

Characterizing a Legal-Intellectual Culture: Bacon, Coke, and Seventeenth-Century England

Peter Grajzl†

Peter Murrell‡

January 5, 2020

Abstract

A characterization of the ideas of Francis Bacon and Edward Coke, two preeminent English lawyer-scholars, provides insights into the nature of the legal-intellectual culture of early seventeenth-century England. This emerging culture remains underexplored, even though it immediately preceded and provided essential input into the 'culture-of-growth', the eighteenth-century cultural paradigm viewed as a catalyst for England's historically unprecedented technological advance and economic growth. To develop insights, we employ a methodology not previously used in this context, applying structural topic modeling to a large corpus comprising the works of both Bacon and Coke. Estimated topics span legal, political, scientific, and methodological themes. Legal topics evidence an advanced structure of common-law thought, straddling ostensibly disparate areas of the law. Interconnections between topics reveal a distinctive approach to the pursuit of knowledge, embodying Bacon's epistemology and Coke's legal methodology. A key similarity between Bacon and Coke overshadows their differences: both sought to build reliable knowledge based on generalizing from particulars. The resulting methodological paradigm can be understood as reflecting a legacy of common-law thought and constituting a key contribution to the era's emerging legal-intellectual culture. More generally, our analysis illustrates how machine-learning applied to primary texts can aid in exploration of culture.

Keywords: culture; Bacon; Coke; England; 17th century; machine-learning

JEL classifications: B31, Z10, N73, K10, P10

* For helpful comments and discussion we thank John Baker, Michael Livermore, Masha Medvedeva, Molly Roberts, Steve Sheppard, and participants at the Conference on Data Science and Law at ETH Zürich.

† Department of Economics, The Williams School of Commerce, Economics, and Politics, Washington and Lee University, Lexington, VA 24450, USA and CESifo, Munich. Email: grajzl@wlu.edu

‡ Department of Economics, University of Maryland, College Park, MD 20742, USA. Email: pmurrell@umd.edu

1. Introduction

For those interested in the cultural determinants of development, the story of England's early rise raises the question of what were the principal features of English culture during the critical 17th century. Certainly some have argued that it is features of early seventeenth-century English culture that made England increasingly distinctive. This perspective is vividly captured in Wootton's (2015) characterization of the seventeenth-century scientific revolution: "...let us take for a moment a typical well-educated European in 1600 – we will take someone from England, but it would make no significant difference if it were someone from any other European country as, in 1600, they all share the same intellectual culture... Within a few years change was in the air... But now let us jump far ahead. Let us take an educated Englishman a century and a quarter later, in 1733, the year of the publication of Voltaire's *Letters Concerning the English Nation* [whose message] was that England had a distinctive scientific culture: what was true of an educated Englishman in 1733 would not be true of a Frenchman, an Italian, a German or even a Dutchman." Indeed the message of Voltaire's book was of a distinctive culture in general, in all of its social, political, and economic manifestations. For example, the legal and political sphere had been increasingly permeated by common-law thinking, an idiosyncratic product of many centuries of legal evolution (Baker 2019). Thus, "[b]y the seventeenth century, England had developed a political culture completely comfortable with sophisticated legal concepts... [A] legal disposition of mind... was being increasingly brought to bear upon political and constitutional problems. It conditioned men's thought and language and ultimately their actions" (Nenner 1977: x).

What were the principal features and constitutive elements of the legal-intellectual culture that emerged in early seventeenth-century England? We address this question by undertaking a quantitative, machine-learning analysis of the writings of two authors, Francis Bacon (1561-1626) and Edward Coke (1552-1634). We take this approach for five reasons. First, the fundamental contributions of Bacon and Coke came at a time when the underpinnings of the later seventeenth-century legal-intellectual culture were emerging, as made clear by Nenner (1977), Wootton (2015) and others (Cromartie 2006, Hill 1997). Second, as the foremost lawyer-scholars of their age, Bacon and Coke were immensely important in the development of those underpinnings. Third, the existing literature largely characterizes the ideas of Bacon and Coke as conflicting, or at the very least orthogonal. If our analysis finds commonalities between these supposed opposites, then the common features of their thought can be viewed as reflecting important aspects of the broader legal-intellectual culture.

Fourth, the most reliable way to estimate the features of a past culture is by going back to original data from its time, not by using secondary sources. Such data are scarce. Arguably, texts are the only cultural data that exist in sufficient volume to apply a quantitative approach. Fifth, both Coke and Bacon were prolific scholars. Given the substantial breadth and scope of the authors' opuses, it is hardly surprising that there is a dearth of analyses that compare the two using conventional methods. Recent advances in computational textual analysis, however, have opened new pathways to such analysis, creating possibilities for the production of new, macroscopic

evidence based on the application of machine-learning techniques for analysis of large volumes of text.

Our paper thus makes two key contributions to scholarship in economic history and economics more generally. First, our focus on Bacon, Coke, and early seventeenth-century England identifies core elements of a comparatively underexplored emerging culture, which provided a foundation for the subsequent set of ideas and beliefs that stressed the application of scientific methodology to productive use and invention. That eighteenth-century "culture of growth" (Mokyr 2016) has been viewed as instrumental in England's economic ascent. Our analysis is the first to provide quantitative insight into the main features of the immediately preceding culture, one that was distinctly legal-intellectual in character. Thus, understanding the ideas of Bacon and Coke, two preeminent lawyer-scholars, facilitates comprehension of the full breadth of the intellectual origins of England's economic rise, a paradigmatic example of economic development that has motivated a voluminous literature (e.g., North and Weingast 1989, Mathias 1969, Floud and McCloskey 1994, McCloskey 2006, 2010, 2016, Mokyr 2009, Allen 2009, Clark 2009, Hayek 1960, Moore 1966, North et al. 2009, Acemoglu and Robinson 2012).

Second, our paper illustrates how machine-learning can be used productively in the exploration of culture. The study of culture is now more prominent in economics than ever before (Alesina and Giuliano 2015, Guiso et al. 2006, 2016; Barro and McCleary 2003, Nunn 2012, Algan and Cahuc 2010, Tabellini 2008, Bowles and Gintis 2011). Yet unlike humanities scholars (e.g., Newman and Block 2006), economists and social scientists more generally have only recently begun to use the quantitative analysis of text data to investigate the features of salient cultural ideas prevailing at specific points in time (Lucas et al. 2015, Blaydes et al. 2018, Grajzl and Murrell 2019). Within economics in particular, the emphasis to date has been on the 'big data' aspect of text, together with the application of supervised models to create new variables for use in causality-centered regression analyses (see, e.g., Dittmar and Seabold 2016). Our analysis has a different focus: the estimation of an unsupervised model to create a picture of a specific culture that existed at one point in time. This is an approach that has received much less emphasis in the new world of big data and text as input (Gentzkow et al. 2019). But, importantly, it shows the potential of the new techniques to revisit that most traditional of the economic historian's tasks, to characterize the details of a time and place—especially the ideas that then existed. Our approach to the use of machine-learning applied to primary text sources could be fruitfully followed in many contexts to examine the shared and divergent elements of culture and intellectual ideas co-existing within a given era.¹

To set the stage for the analysis, Section 2 introduces Bacon and Coke, their backgrounds, and their professional and personal rivalries. Section 3 describes the machine-learning technique that we use—the structural topic model (STM)—and the data. Topic modeling is particularly

¹ As Gutmann et al. (2018: 283): emphasize: "Textual data in various forms can provide insight into what past economic actors thought.... Textual corpora provide economic historians with a new quantitative approach to questions sometimes addressed in a more narrative style."

suitable for finding the broad themes present in a large corpus. The algorithms used to estimate topic models are unsupervised. Thus, while the researcher interprets the estimated themes, the estimation of those themes is not influenced by the researcher's preconceived notions. As we elaborate in Section 3, STM extends the workhorse Latent Dirichlet Allocation (LDA) model (Blei et al. 2003) in a number of ways, improving the interpretability of the estimated topics and facilitating more reliable estimation of the effects of covariates on topic prevalences (Roberts et al. 2014).

Turning to the heart of our analysis, Section 4 presents estimates of the 25 topics that machine-learning identifies in the corpus of Bacon's and Coke's writings. These topics are a summary estimate of the core ideas present in English legal-intellectual culture around the beginning of the 17th century. In Section 5, we use the power of machine-learning to investigate the shared foundations of different cultural topics and identify their connectedness. The resulting analysis identifies the central cultural nexuses and their linkages, as evidenced in Bacon's and Coke's works. We detect a shared and perhaps unexpectedly deep, theoretical structure in their legal deliberations, with applications cutting across conventional legal subjects. In Section 6, we use STM to provide a first quantitative assessment of the differences and similarities in the emphases present in the writings of Bacon and Coke. While Bacon and Coke differ in their emphasis on particular topics, we find, in contrast to the existing literature, that the similarities between them are as striking as their differences. Importantly, they share a fundamental methodological approach that can be interpreted as a central feature of the era's emerging legal-intellectual thought. Section 7 summarizes our findings more generally and concludes.

2. The Rivals in Their Time

2.1. The Background

In the late 16th and early 17th centuries, English culture was flowering. That period saw the first performances of Shakespeare's plays, the publication of the King James Bible, and early discoveries in medicine and science. It was also an era of competing visions about English institutions. Parliament and the common-law judiciary were increasingly challenging the power of the monarch. In religion, doctrinal and organizational controversies abounded, as debate became more open and the common-law courts challenged the jurisdiction of the ecclesiastical courts.

The common law was establishing its place at the center of English culture (Nenner 1977; Hill 1997; Cromartie 2006). Moreover, the legal profession was becoming an important political and intellectual force. Interactions between monarch and the legal profession were increasingly marked by conflicting conceptualizations of the law and divergent legal philosophies (Berman 1994, Friedrich 1958). For James I especially, law was based on human reason, grounded in the divine right of kings. For common-law lawyers, reason was to be found in both local custom and the accumulated decisions of generations of learned men in a process akin to trial and error.²

² This trial-and-error style was similar to the use of systematic observation and experimentation that had taken root among the practicing artisans, mechanics, and workers of Elizabethan London (Harkness 2007, Hill 1997: 16).

Francis Bacon (1561-1626), best known as the father of the modern scientific methodology, and Edward Coke (1552-1634), widely recognized as the greatest English lawyer, occupied center stage both intellectually and politically.³

Bacon's family was wealthy and politically well-connected. Yet Francis was the youngest of five sons in an age of primogeniture: he needed a profession and he took up the law. In his quest for status and wealth, he gravitated toward service to the monarch. Coke's father had lower social status, but was a successful lawyer in a rich county. As the oldest son, Coke had both financial security and the ambition to follow his father in the law.

Despite these differences in family circumstances, Coke and Bacon were typical of those who entered the country's ruling elite. Both hailed from solid Puritan backgrounds. Both were educated in the humanist tradition and the classics, including rhetoric. Both studied at Cambridge where they were exposed to philosophy and science. Both were thoroughly trained in law at the Inns, England's 'third university', and both served as practicing lawyers. Both had aspirations to attain the highest professional and political status. They both succeeded.

Young Bacon was a quintessential common-law lawyer (Coquillette 1992). As a member of Parliament, he opposed government-granted monopolies and subsidies to business. He spoke against the crown's encroachment on the right of the Commons to set taxes. He argued that royal grants should be subject to a test of compatibility with judicial and statutory law, in a way in which "Sir Edward Coke would have been pleased, not just by Bacon's conclusion, but by the technical, and exhaustive, use of common law precedent on which it was based" (Coquillette 1992: 26). Indeed, his views on taxation were one reason why Elizabeth snubbed Bacon in his early attempts to secure higher office.

Bacon's fortunes improved with the ascent to the throne of James I. Upon impressing James with his intellect and after moderating his public statements on matters involving the monarch, Bacon swiftly climbed the professional ladder. He served first as Solicitor General, then as Attorney General, and eventually as Lord Chancellor. After his impeachment for corruption in 1621, he turned nearly exclusively to the study of natural philosophy.

Coke's career followed a different path. Numerous portrayals of Coke as a fearless combatant against royal prerogative often elide the fact that he spent his early years in service of the crown. As Solicitor General and later Attorney General, Coke was absolutely "ferocious" in his prosecutions against enemies of the monarch, to the extent that "even his contemporaries were occasionally disgusted" (Holdsworth 1938: 114). Once he assumed the position of Chief Justice of the Court of Common Pleas, however, Coke systematically and persistently voiced his convictions about the supremacy of the law and the corresponding constraints on the government.⁴

³ In Holdsworth's (1938: 134) words: "What Shakespeare has been to literature, what Bacon has been to philosophy, what the translators of the authorized version of the Bible have been to religion, Coke has been to the public and private law of England."

⁴ We present here the story that is standard in the historical literature. In a recent historical account, Baker (2017) suggests that Coke was already starting to develop his views on the supremacy of the common law while serving as Attorney General. The story that we present is nevertheless the "typical historian's verdict" (Baker 2017: 357).

He continued in this vein throughout the rest of his life, becoming a leader of the Parliamentary opposition after his dismissal from the bench.

2.2. *The Rivalry*

Bacon and Coke had a lifelong professional and personal rivalry. They competed for the hand of the same woman, they vied for top government positions, they tussled over the superiority of the courts that they each headed, and they were opposing lawyers in landmark cases (Burch 1928, Coquillette 1992, Hart 2003). Neither showed any appreciation of the other's intellectual pursuits, except in one instance when Bacon paid, not unalloyed, tribute to Coke's Reports (Gardner 1916, Hollond 1947). Coke, on the other hand, labeled Bacon's work on the foundation of the scientific method as "sheer folly" (Lyon and Block 1930).

As lawyers, Bacon and Coke were often adversaries.⁵ In Slade's case, which fundamentally altered contract law, Bacon and Coke appeared on different sides and their respective advocacy embodied widely different methods and views. Coke's side won the case and Coke's published report completely ignored Bacon's "learned and persuasive arguments" (Baker 1971: 53; Coquillette 1992: 129, 136).

After he became close to James, Bacon was influential in the decision to transfer Coke from the Court of the Common Pleas to the King's Bench, to reduce Coke's "capacity for harm" (Holdsworth 1935: 335). Personal antipathies between the two "no doubt worked as well to exacerbate tensions and hardened positions" (Hart 2003: 103). In 1616, after another dispute, Coke was dismissed from the King's Bench, with Bacon drafting the letter of dismissal. Coke struck back in 1621. As one of the leaders of the Parliamentary opposition, he led the impeachment of Bacon on corruption charges. Through their rivalry, "Bacon and Coke destroyed each other professionally" (Coquillette 2004: 315).

2.3. *The Literature*

Given the rivalry, it is hardly surprising that the characterizations of Bacon and Coke in the literature are dominated by comments on their differences and contradictions. As Gest (1909: 505) remarked more than a century ago: "It is indeed hard to estimate correctly...those mighty men who then occupied the center of the stage. Everyone who reads the fascinating Elizabethan story becomes insensibly a Baconian or a Cokean, a partisan of one or the other of those wonderful men."

For some, Coke as "a technical, skillful and learned lawyer...has never had a superior" (Burch 1928: 10), while Bacon was no more than "the lawyer without law" (Rogers 1931: 31). For others, even Holdsworth (1935), a great admirer of Coke, Coke had "a credulity which is as medieval as his law" and his use of history was "often unhistorical", while viewing Bacon as the greatest jurist of the day. Wheeler (1983) contrasts "Coke's antiquarian empiricism" with Bacon's

⁵ One notable exception was Calvin's case, still influential in citizenship law, where there was much similarity in the positions of Coke, as the Chief Justice of Common Pleas, and Bacon as the King's Solicitor General (Hart 2003: 88).

"rationalist quality that startles us with its modernity." For Berman (1994), Coke "articulated no systematic philosophy" and his law had "no theory at all". For Coquillette (2004: 312), "the two men were...philosophical opposites, with very different juristic and ideological beliefs".

Differences in views on the substantive application of law naturally follow. For Helgerson (2004), "Bacon favored the king; Coke the law". Cromartie (1995) characterizes Bacon's maxims as "authoritarian". Bacon is even viewed as not in the common-law fold, as a "civilian" (Kelley 1993) for whom "his attempted codification of English law was on the French model" (Gaukroger 2006). In contrast, Coke is depicted as "the greatest oracle of our municipal jurisprudence" and the man "who afforded a bright example of judicial independence" (Burch 1928: 507).

One can find notable exceptions to this characterization of the literature, but these are rare. For Coquillette (1992: 27), it is necessary to separate "Bacon's actual views from his advocate's arguments". Applying this lens, Coquillette argues that Bacon's early work "had the trappings of a parliamentary manifesto, and actually read very much like some of Edward Coke's later opinions". Similarly, de Montpensier (1968: 455) maintains that "Coke and Bacon shared the same views about the foundations and sources of law, the position of the law and the crown, the relations between the courts and Parliament" and that the differences between them should be understood in light of their varied professional circumstances.

2.4. The Task

There are thus two opposing views in the literature on Bacon and Coke, scattered over more than a century of scholarship. There is the near-consensus that Coke and Bacon differed greatly in both general philosophical outlook, legal methods, and substantive application of law. And there is the minority view, expressed by a few scholars, that the two authors shared much in common. We investigate the validity of these contrasting positions using machine-learning. In a departure from the existing literature, we do so in an analysis that is explicitly comparative and distinctly quantitative in character.

3. Methods and Data

3.1. Structural Topic Modeling

In order to examine the works of Bacon and Coke, we estimate a topic model. As a complement to conventional textual analysis, topic models are particularly suitable for analyses of large textual corpora when the principal goal of the analysis is to provide a macroscopic guide to the themes emphasized in a corpus. With the emergence of 'big data' and a growing interest in text-as-data methods, the use of topic models has become increasingly common across a broad range of academic disciplines. The Latent Dirichlet Allocation (LDA) model (Blei et al. 2003) in particular has been fruitfully applied by both social scientists and humanities scholars (Grimmer and Stewart 2013: 283-285; Hansen and McMahon 2016, Hansen et al. 2018, Mohr and Bogdanov 2013). Introducing topic modeling to economic historians, Wehrheim (2019) provides a recent overview of research pertinent to the field and offers an application documenting trends in economic-history scholarship over time.

Topic models belong to a class of generative probability models that require a researcher to postulate a model of the data generating process and then use the data to determine the most likely values for the parameters within the model. To estimate the parameter values, topic models view texts as 'bags of words'. An unsupervised machine-learning algorithm then exploits the co-occurrence of words across documents to identify 'topics' (Blei 2012). It is important to understand that topic modeling does not simply depend on word counts: topic estimation is driven by correlations of word use across documents. Thus, despite the loss of much syntax that occurs by viewing texts as a bag-of-words, the semantics of documents can still be reflected, at least partially, in the resultant estimates (Reich et al. 2015).

The resultant topics are formally conceptualized as probability distributions over the corpus vocabulary. Documents (chunks of text) are modeled as mixtures of topics. The name of each topic is assigned by the researcher after scrutiny of the words most closely associated with the topic and study of the documents that feature a given topic particularly prominently. The topics themselves, however, are solely a product of model estimation. In particular, they are not obtained by matching words and documents to concrete thematic issues that are specified by the researcher prior to estimation (as would be the case in a supervised estimation).

We use the structural topic model (STM; Roberts et al. 2014, 2016a 2016b), the formal statistical structure of which we present in Appendix A.⁶ (All appendices are available online as part of the electronic supplementary material for this paper.) Unlike LDA, STM integrates document-level metadata directly into the estimation of topics and allows topic prevalences to be correlated across documents even when conditioning on the values of the metadata. Intuitively, with documents conceptualized as mixtures of topics, the prevalence of a particular topic will tend to vary across documents because different documents can originate with different authors, or can reflect different time periods, or are intended for different audiences. Thus, rather than estimate topics under the assumption that the corpus documents are fully interchangeable, as implied by LDA, one would like the assumed data generating process to allow topic prevalences to vary with document characteristics. This is exactly what STM does, thereby enabling the researcher to use document-level variables in the estimation of topics and then subsequently to assess the relationship between these variables and topical prevalence.

Roberts et al. (2014: Online Appendix) demonstrate a series of advantages of STM over LDA. The first set of advantages arises because STM tends to improve estimates of the topics themselves. Specifically, simulations show that, in comparison with LDA, STM produces topic estimates that tend to be more exclusive (estimated topics can be better distinguished from each other) and sometimes even more coherent (estimated topics are more internally consistent). STM-identified topics are in general therefore more easily interpretable than LDA-identified topics. In addition, STM-based topic estimates tend to be more robust to the presence of rare words compared to the estimates produced by LDA.

⁶ See <https://www.structuraltopicmodel.com> for a list of published applications of STM.

The second set of advantages of the STM arises in the context of the estimation of relationships between topic prevalences and metadata covariates. Estimating these relationships is intrinsic in the structure of STM, but in LDA they would be estimated using a two-stage process where an LDA (without incorporating covariate information) is followed by a conventional regression of LDA-estimated topic prevalences on document-level covariates. Simulated and real data show that STM's full integration of metadata yields both more accurate and more efficient estimates of the covariate effects than does LDA's two-step process.⁷

3.2. The Corpus of Works

Our corpus merges many works of the two authors. The works of Bacon include all digitized, machine-readable works that could be identified. Our sources for Bacon's works are established repositories of digitized old documents, such as Project Gutenberg (n.d.), Hathi Trust (n.d.), Internet Archive (n.d.), and The Text Creation Partnership for Early English Books Online (2014). Our source of Coke's works is Sheppard (2003), available in electronic format at the Liberty Fund's Online Library of Liberty. Sheppard's anthology is a comprehensive, machine-readable collection of Coke's writings and speeches.

The longer works of each author were broken up into smaller documents in a manual process that used natural breaks in the text. The result was a corpus of 432 text documents of varying length, containing 1,320,262 words, an average of 3,056 per document. Table 1 lists the works included in the corpus. There are more documents from Bacon in the corpus than from Coke. Given an adequately large number of topics to be estimated (see Section 4), this feature of the data should not bias the estimated topics in favor of Bacon's works. The reason is that, in estimating the topics, STM does not rely on word frequencies alone, but also leverages the correlation of word use across documents and explicitly allows for topic prevalence to vary with document-level information such as authorship. Thus, as we demonstrate below, despite the fact that Bacon's works are more prevalent in the corpus than Coke's, a number of topics are featured highly in the documents of both Bacon and Coke, with a further set of topics dominated by either Coke or Bacon.

The documents were processed in a series of steps that converted the chaotic orthography of late sixteenth- and early seventeenth-century English into standard modern orthography and to translate Latin words into English. Appendix B summarizes the details of the processing and also provides evidence in favor of considerable success in standardizing the orthography across the corpus. The corresponding analysis also indicates that potential OCR issues are not a concern for our analysis.

⁷ Specifically, Roberts et al. (2014: Online Appendix) demonstrate that the STM-estimated effects always closely match the true effects, while the two-stage LDA approach often produces estimates featuring incorrect signs. Furthermore, by virtue of incorporating metadata information into topic estimation, STM estimates of covariate effects have smaller confidence intervals than those of LDA estimates. Using permutation analysis, Roberts et al. (2014: Online Appendix) also show that STM-based incorporation of metadata into topic estimation does not introduce spurious relationships between topic prevalences and metadata covariates.

The resultant corpus was imported into R using the `stm` package. To prepare the corpus for estimation, further text processing was implemented using R's `textProcessor` function. All words were converted to lower case. The Porter stemming algorithm was applied. Standard English stop words (natural language words with little meaning, such as 'and', 'the', 'a', 'an'), numbers, and punctuation were removed. The resulting dataset consists of 432 text documents and 215,556 word tokens.

3.3. Metadata

The last step in organizing the data was to assign values of metadata variables to each document. We coded four metavariables: authorship (Bacon or Coke), intended audience (lawyers, politicians, historians, methodologists, philosophers, or scientists), form of finished work (essay, case report, apothegm, book-length tome, letter, or speech), and year of completion. Table 2 provides the document frequencies within each cell of the metadata variables.

The authorship and the form of finished work were readily ascertained. The coding of intended audience was based on scrutiny of each document while simultaneously taking into account existing scholarship on Bacon and Coke. The stated intentions of the authors or the substance of the documents were very important for this coding. For example, the prefaces of Coke's *Reports* were obviously not intended to bolster the legal record but were rather a guide to the methodology that Coke used in writing the reports. Similarly, the letters of Bacon could be easily divided into ones that were in the political or legal sphere and ones advancing his philosophy. Note that for intended audience, the monarch was classified as a politician, since communications with the monarch invariably concerned matters of state.

The year of completion was coded by taking into account any available information about the lives of the two authors. For the vast majority of documents in the corpus (69 percent), we were able to ascertain the exact year of publication. When no additional information about the timing of a work was available, for works published in the author's lifetime we took the year of the work's publication as the year of the work's completion. However, some of Bacon's and Coke's works were published only posthumously or the exact year of their completion is unknown. For those documents (31 percent), we conducted a thorough search of available sources about the author's work and life in order to identify the earliest possible and the latest possible year of completion. To fix a single year for completion, for each document we separately drew a random integer from the interval of possible years. For those works for which the year had to be estimated, the mean absolute value of the error would be less than five years. None of our substantive conclusions would change with a more precise estimate of dating. Appendix C provides full details on the dating process.

4. Estimating and Interpreting Topics

The first decision to be made in estimating an STM is on the number of topics. Because there exists no standard approach on this decision (Roberts et al. 2014, 2016b), we estimated a series of models, with the number of topics varying between 15 and 30. We examined measures

of their goodness-of-fit such as held-out likelihood and size of residuals (Wallach et al. 2009, Taddy 2012, Roberts et al. 2016b). We then compared the set of models that fit the data especially well by using the models' scores for average semantic coherence (indicative of the internal consistency of the topics) and exclusivity (indicative of the extent to which topics in the model can be distinguished from each other). We thereby identified the subset of models on the semantic coherence-exclusivity frontier (Roberts et al. 2014). We inspected the cohesiveness and exclusivity of the topics for these models using our own judgment. This process led to the selection of a model with 25 topics. We verified that all of our substantive findings were robust to small variations in the number of topics.

After naming the topics in a process to be described below, we grouped them into seven general themes to organize our initial findings. Table 3 provides a first, broad overview by listing names for the general themes, the topic names, and a measure of the relative importance of each author for each topic. The table arrays the topics on a natural continuum beginning with law in the abstract and then proceeding via substantive law, to politics, science, and the scientific method.

The organization into themes is based on our a priori conception of which topics belong in standard categories. However, one of the properties of STM is that it is not constrained by a priori notions: it can find unexpected patterns in the data. Thus, when such patterns are examined in the following section, we show how STM offers a slightly different organization of the topics than that in Table 3, improving upon our a priori conceptions and, importantly, offering novel insights about the corpus. For example, STM shows quite clearly that topics can be placed on a circular continuum rather than a linear one: there is a strong connection between the first and the last topics in Table 3.

Table 4 lists the words most strongly associated with each topic. These words are a product of the estimation, and are therefore word stems (e.g., 'judg'). Choosing topic names involves examining in which documents a given topic is most prominent and which words are most strongly associated with each topic. Table 4 provides two lists of the top 30 words for each topic. The highest probability ('Highest prob') words are those most common for a given topic, but are also non-exclusive—they might be the highest probability words for several topics. 'FREX' words are used more frequently in documents highly associated with a topic.⁸ For brevity, we use the shorthand of 'ranked highly' or 'top' when referring either to documents that feature a given topic prominently or to words ranked high by either of the two above criteria.

The assignment of names to topics is central in the analysis because our general conclusions rest upon being able to interpret the content of each topic. Thus, it is notable that we could easily identify the ideas underlying each topic. Moreover, our topic names resonate strongly with concepts in the legal, historical, and traditional text-analysis literature. The paragraphs below present examples of the arguments used to choose topic names. We stress that the focus is on

⁸ We restrict the FREX words to those used with some frequency in order not to focus on unusual words that are used once in a document. Our choice of FREX words is characterized by a frequency to exclusivity ratio of 0.25 (see Roberts et al. 2016b).

examples because space limitations prevent the inclusion of the voluminous evidence we considered in justifying the choice of each topic name.

4.1. *Legal Scholarship*

Both Bacon and Coke are prominent in the topics that capture different dimensions of legal scholarship. The first topic focuses on Understanding Law. (We capitalize the topic names, to easily identify them.) 'Law' is the most used word, while 'book', 'student', 'reader', 'professor', 'treatis', 'commentari', and 'inn' are all highly ranked in this topic and not elsewhere. A majority of the top-twenty documents are Coke's prefaces or conclusions. Coke's concern in these documents is to point out that the purpose of his Reports is not simply for readers to learn about the law, but "to understand what the true sense and sentence of the Lawes then standing is".⁹ For Coke, the texts are intended to be useful even for the highest levels of the legal profession: "... But forasmuch as if a man should spend his whole life in the study of these Lawes, yet he might still add somewhat to his understanding of them".¹⁰ Similarly, for Bacon: "Concerning the Lawes of England: They commend themselves, best to them, that understand them".¹¹

The second topic reflects Jurisprudence, or analytical legal theory, again spanning various areas of substantive law. The most prominent words are 'law', 'statut', and 'case', while among the FREX words are 'cesti' (beneficiary), 'entail', 'covin' (fraud), and 'proviso' (clause), all with specialized meanings in distinct areas of law. The highest-ranked document for this topic is Bacon's "Reading upon the Statute of Uses", a scholarly analysis of a controversial 16th century property-law statute. A number of Bacon's *Maxims* are also highly ranked, each maxim intended to be an analytical statement of the principles of law relevant to widely varying substantive areas. Coke also figures prominently. The tenth-ranked document is Coke's report on Heydon's case, a landmark of statutory interpretation, in which Coke advised "...that for the sure and true interpretation of all statutes in general...restrictive or enlarging of the Common Law, four things are to be discerned and considered".

Disambiguating Law focuses less on general principles and more on clarifying specific legal doctrines. 'Ambigu' appears as a highly-ranked FREX word here, and not elsewhere, reflecting this topic's focus on removing ambiguities in legal rules. Among the most prominent documents are several of Bacon's *Maxims*, which were set down "...so that the uncertainty of law...be somewhat the more settled and corrected."¹² For example, Bacon's twenty-third maxim is ranked fifth: "Hidden ambiguity of words may be supplied by proof; for an ambiguity arising from facts is removed by proof of the facts." Coke is also prominent. In the eighth ranked document he advises: "...to set down in conveyances everything in certainty and particularity, for

⁹ In the preface to Coke's fourth reports.

¹⁰ In the preface to Coke's third reports.

¹¹ In "A Proposition Touching the Compiling and Amendment of the Lawes of England"

¹² In the preface to Bacon's *Maxims*.

Certainty is the mother of quietness and repose, and uncertainty the cause of variance and contentions."¹³

4.2. *Private Law*

One topic addresses Property Rights in general. Among the words distinctive to this topic are 'vest', 'escheat' (reverting of property to the crown), 'purchas', 'beneficiari', 'advowson' (the right to present a candidate for a church office), 'recoveri', and 'escuag' (the military service consequent on a knight's land use rights). The topic is dominated by property in land, but although property in land usually provides the context, the discussion in the top documents is often conducted in terms of more general property rules and rights. For example, one such document expounds on the rights of tenants and landlords over a rented house, another focuses on how Magna Carta specifies the rights and obligations of those who control properties.

A closely associated topic is the one that focuses on Land Inheritance Law. Inspection of the associated words is sufficient to establish its name. 'Heir', 'land', 'shall', 'son', 'inherit', 'father', 'law', 'blood', 'die', and 'issu' are the highest probability words. The top documents have a similar cast. The one non-legal document is Bacon's essay on parents and children that concludes with "Younger brothers are commonly fortunate, but seldom or never where the elder are disinherited". Many of the top documents are from Coke's commentary on Littleton, but there are also case reports: Shelley's case provides a rule on the inheritance of estates and Lord de la Warre's case clarified laws of inheritance.

4.3. *Private-Public Law*

In seventeenth-century England, the institutionalization of society proceeded apace, with the boundaries between private and public gradually becoming delineated. Not surprisingly then, a number of topics span both public and private law. Religion, Law, & Truth reflects the most pressing of these areas of contention. The top FREX words speak collectively to the substantive issues: 'preach', 'controversi', 'church', 'uniti'. But 'law' is also among highest probability words and fifteen of the top twenty documents are about law, with 'true' or its variants prominent in those documents. One top document is Lord Cromwell's Case, on slander in a highly politicized religious setting.

The mixture of religion and the ascertaining of truth in general, often through the use of the law, is best indicated for this topic by uses of the word 'reveal', one of the top FREX words for this topic and only this topic. Coke states that the revelation of the truth is the purpose of his reports.¹⁴ In a highly-ranked section from his great methodological work *The Advancement of Learning*, Bacon wrote: "...I note this deficiency, that there hath not been...sufficiently inquired and handled the true limits and use of reason in spiritual things...to search and mine into that which

¹³ In the first volume of Coke's Institutes.

¹⁴ The tenth report's preface states: "This part containeth a true and just Report...to avoid that, which venerable Verity [truth] doth blush at for fear, that is, that she which is the Foundation of Justice should not be hidden and unknown. Neither is she pleased, when once she is found out and revealed to be called into argument and question'd again, as if she were not in Verity indeed".

is not revealed". This hints at Bacon's epistemology of eliminative induction, for which he is most renowned. Similarly, in top documents, Coke clearly states his modes of reasoning. For example, in the preface to the sixth volume of his reports, Coke comments "That if the ancient Laws of this noble Island, had not excelled all others...some of the several Conquerors and Governors thereof...would (as every of them might) have altered or changed [these laws]." This is an example of a general evolutionary mode of argumentation, which was characteristic of the common law at that time (Grajzl and Murrell 2016).

Interestingly, Religion, Law, and Truth and Epistemology, which is introduced below, are the only two topics for which 'true' is a highly rated word. Moreover, a majority of the documents most highly associated with Religion, Law, and Truth have the two authors focusing on the structure of arguments that should be applied to religion and to doctrinal and church-organization issues.¹⁵ These are often, but not always, legal arguments. At the current level of analysis, therefore, it is simply not clear to what extent Religion, Law, and Truth is a topic about the application of law to religion versus a topic about methods of argumentation that happen to be often applied to religion and law. Fortunately, STM has techniques in its toolbox that will cast light on this issue. We make use of those techniques in the following section.

Defendant Rights focuses on law that establishes the rights of defendants in criminal trials and the defenses possible in civil suits. Thus words prominent in this topic are 'imprison', 'magnacarta', 'habea', 'corpus', and 'bail', but also 'assumpsit' and 'loan', which are relevant to civil matters. All the top five documents are Coke's case reports. Semayne's Case concerns when an official can legally break into a house; Vaux's Case clarifies double jeopardy; and William Aldred's Case elucidates when an act constitutes a nuisance, containing the memorable dictum that "one ought not to be of so delicate nosed, that he cannot endure the scent of hoggs". Bacon's *Maxims* are also highly ranked, the fifth on the impossibility defense; the seventh on motive and extent of damage in criminal and civil cases; and the twenty-second on duress.

The next topic also covers heterogeneous substantive areas. The top documents are all legal ones, but there is no 'law' among the most common words. This paradox suggests that the topic concerns matters that legal authorities often wanted to comment on, but for which formal law was less important. There are many FREX words that refer to market relationships: 'debt', 'franchis', 'pay', 'sale', 'chattel' (movable property), and 'rate'. Bacon's *The Use of the Law* is prominent, examining the varied ways to obtain property in goods, including mechanisms other than markets. The third ranked document is Coke's report on market-overt, explaining how custom regulates ownership of stolen goods. Bacon's essay on usury uses moral, legal, and economic arguments. Highly ranked documents consider taxes imposed on the beneficiaries of drainage projects and the giving and taking of property for religious purposes. This topic, then, is about the

¹⁵ It is not surprising that the form of argumentation should be important to the two authors in these substantive areas. Religion was at the fulcrum of English politics throughout the two centuries in which Bacon and Coke lived. Debate about religious issues was becoming more open and more intertwined with the law as the common-law courts wrested jurisdiction from the ecclesiastical courts. But such debate was risky, with blasphemy still a capital offense; therefore focusing on details of the method of argumentation would be a safer strategy than articulating one's convictions.

variety of institutional and cultural mechanisms that could support exchange. We use the rather modern name, Regulation of Exchange.

As noted above, the separation between the public and private spheres was central in seventeenth-century political struggles. Public & Private Authority captures elements of these struggles. It is all Coke, with the highest ranked documents all legal cases. Among these are rulings on when the state can make a monopoly (Case of Monopolies); on when restrictions on trade are lawful (The Chamberlain of London's Case); on the powers of town governments (James Bagg's Case); on the legal authority of an arbitrator (Vynior's Case); on the creation of an official office to benefit a private citizen (Walter Chute's Case); on the power of private bodies to implement taxes for use of public works (The Case of the Isle of Ely); and on whether the King can dispense with the law for particular individuals (The Case of Non-Obstante).

4.4. Public Law

We label the next topic Legal Jurisdiction; indeed, 'jurisdict' is among the top FREX words. It focuses on specifying which institutions and organizations have which powers and obligations. The highest-ranked document is Langdale's case concerning inter-court jurisdiction; second is the case of Praemunire, concerning the sphere of ecclesiastical and temporal courts (and ultimately the authority of foreign powers within the Kingdom). Ranked next is Fuller's case in which "It was resolved when there is any question concerning what power or jurisdiction belongs to Ecclesiastical Judges...the determination of this belongs to the Judges of the Common Law...". Ranked fourth is Calvin's case, concerning the jurisdiction of English law for Scottish citizens, the two countries being separate nations with a common monarch. Other highly-ranked documents address jurisdictional disputes between the Court of Common Pleas and the Court of the Admiralty, between the Crown and courts, and between different local officials.

Next is a topic on the Criminal Justice System. The top five documents are all from the third part of Coke's *Institutes*, which examines "...De malo, viz. of High Treason, and other Pleas of the Crowne, and Criminall Causes...". Bacon is also prominent, his "Cases of Treason" covering a broad range of crimes, together with the organization of the courts dealing with them. This is consistent with the types of words that are ranked highly: 'indict', 'court', 'punish', 'steward', 'constabl', 'accus', 'offenc', 'gaol', 'arraign', etc. In other top-ranked documents, Coke reports on the case of Floyd and Barker, which concerned the conduct of criminal proceedings, and Bacon's *The Use of the Law* lays out the purpose and logic of different adjudicatory and law enforcement institutions. This topic is clearly about both criminal law and criminal procedure.

Constitutional Law is almost entirely a product of Coke. Prominent are elements of the fourth part of the *Institutes*, reviewing such topics as the powers of the courts, the law and custom of parliament, the monarch's rights in religion, Parliamentary control over taxation, rules on elections, stipulations on who can sit in Parliament, etc. Consistently, there are a number of distinctive words that are highly associated with this topic and not with others: for example, 'tunnag', 'subsidi', 'frankpledg', 'imposit', 'poundag', and 'alnag' all refer to issues of taxation.

In contrast, King, Law, & Nation is dominated by Bacon. A large proportion of the FREX words refer to aspects of nationhood or monarchy (e.g., 'union', 'papist', 'kingdom', 'nation', 'sovereignti', 'territori'). The highest probability word is 'law' but the topic goes beyond law. A top-ranked document discusses the way in which a King should approach decisions.¹⁶ Another is a letter to one of the King's favorites advising what (extra-legal) powers the recipient has and how he should conduct himself.¹⁷ This topic reflects both Bacon's legal training and his background in Renaissance humanism, with its emphasis on the connection between improved knowledge and good government.

4.5. Politics

On the continuum of themes, we now move from law into politics. Legal terminology is absent from the most-used words, while royal and aristocratic terms are prominent. King & Court focuses on the monarch's role in government and inter-personal relations in his court. Most of the documents that are strongly associated with this topic are Bacon's letters, most referring to court politics. Sometimes the politics is of a very personal kind, with Bacon revealing his most unctuous self in pleading for personal favors.

The top documents for Dynastic Politics are Bacon's histories of the reigns of the three great Tudor monarchs and letters from Bacon that use lessons from history to advise others. Coke is not absent here: three of the top twenty documents are his. These have a similar tenor: all three use historical examples to elucidate the author's points on a variety of political and legal questions.

The next topic is easily labeled by looking at the top ten FREX words: 'spain', 'lopez', 'spaniard', 'invas', 'squir', 'spanish', 'portug', 'cathol', 'war', 'turk'. This is Foreign Relations, and it is all Bacon. In the documents most associated with this topic, there is discussion of a war with Spain, an analysis of just wars, a deliberation of what makes kingdoms great (arms), ruminations on the appropriateness of a holy war, observations on the foreign causes of internal problems, and a speech in parliament on the necessity of raising taxes to increase the external strength of England.

Civic Knowledge is the topic in which Bacon's background in Renaissance humanism comes to the fore. All top-twenty documents are his essays, except one, a section of *The Advancement of Learning* that is exclusively about "civil knowledge". The words particularly associated with the topic are primarily of two kinds, mental dispositions—'envi', 'felic', 'beware', 'anger', 'dissimul', 'virtu', 'secreci', 'discontent', 'cun', 'perturb', 'malum'—and historical figures used to motivate the discussion—'tacitus', 'caesar', 'cicero', 'demosthen', 'solomon', 'sulla', 'machiavelli', 'seneca', 'tiberius'. This topic focuses on discussing the states of mind that are relevant in civic action and on the process of detecting them in others. It is not a methodological topic, but rather an application of ideas on human nature to the civic world.

¹⁶ "An Essay of a King".

¹⁷ "A Letter Of Advice Written By Sr Francis Bacon To The Duke Of Buckingham."

4.6. Science

This theme solely reflects Bacon's science (or natural philosophy). There are five topics, four of which have little interest for this paper, but the first is thematically related to those discussed before. This topic is Human Nature, where Bacon attempts a natural philosophy of psychology. Among the highest-probability words are 'man', 'imagin', and 'natur'. FREX words include 'dream', 'belief', and 'magic'. In the large variety of documents that are associated with this topic, the common thread is the workings of the human mind, the erroneous patterns of thought that must be excised, and the qualities of human thinking that are to be admired. Bacon's prayers ask for the "unlocking of the gates of sense, and the kindling of a greater natural light" to create a "mind, thoroughly cleansed and purged from fancy and vanities". The essay on adversity is a disquisition on how human nature is formed, "for prosperity doth best discover vice, but adversity doth best discover virtue". Bacon's most enigmatic work, *The New Atlantis*, features this topic prominently, with its inclusion of many comments on the exercise of the mind: "let us know ourselves", "every man reform his own ways", "the knowledge of causes", "full of piety and humanity", and "ye shall also understand". Over the centuries, there have been many interpretations of the *New Atlantis*. Here, we emphasize its discourses on Human Nature because STM finds this topic prominently in that work.¹⁸

The remaining science topics are all related to Bacon's writings on the philosophical study of nature and the physical universe. All names for these topics are self-evidently justifiable from both the highest probability and FREX words. We name these topics Botany, Pharmacology, Physics (Air & Sound), and Physics (Energy). A fuller discussion of the content of these topics would have little relevance to the substance of this paper.

4.7. Methodology

The final theme has three topics addressing methodology, Bacon's most renowned contribution to knowledge. The documents most highly associated with the first of these topics are heterogeneous. Several top documents are sections from the *Wisdom of the Ancients* and from *History Natural and Experimental of Life and Death* in which Bacon attempts to derive scientific lessons from the lives of biblical, mythological, Greek, and Roman characters. There is also a section of *The Advancement of Learning*, where Bacon interprets the contributions of historical figures. A disquisition on Queen Elizabeth attributes her success to a contested succession because "Princes brought up in Regal houses, to hope of succession not uncertain, are often depraved with soft and licentious breeding, and become immoderate in their reign".¹⁹ Thus, Bacon uses historical and mythological stories to generate quasi-scientific insights, consistent with his broad philosophy that all experience should be used to advance natural philosophy. We name this topic Extracting Meaning.

¹⁸ This is the topic that is most prominent in *The New Atlantis*. The topic has no connection to the organization of science, the interpretation most often associated with that work (Sargent 1996).

¹⁹ In "The Felicity Of Queen Elizabeth".

The last two topics constitute two elements of the set of ideas that Bacon's later followers came to refer to as the "Baconian program" (Mokyr 2005, 2010, 2016).²⁰ The key to understanding the first of these two, Probing for Facts, is that there are two words that are used with a high probability by this topic—'ask' and 'answer'—and that are highly specific to this topic, as indicated by the FREX words. This is consistent with the common thread that we find in the disparate collection of highly-ranked documents. Several are sections of the *Apothegms*, a collection of isolated short stories that often describe one person asking and another answering. In the top documents, there are three charges made by Bacon as prosecutor. Bacon asks why a person is guilty and answers with evidence. There is also one case from Coke's reports, Sir Stephen Procter's Case, discussing the status of the law when judges are divided in how they answer the question on somebody's guilt. Thus, STM identifies a key aspect of Bacon's scientific method that encourages asking questions and finding answers wherever they can be found, an aspect of Bacon's thought highlighted by many scholars (see e.g., Peltonen 1996a: 17; Mahlerbe 1996, Mokyr 2005: 289, 304).

The label of the final topic, Epistemology, follows immediately from its most highly associated words: 'philosophi', 'knowledg', 'method', 'system', 'logic', 'inquiri', 'discoveri', 'experi', 'axiom', 'theori', etc. The most prominent documents are sections from the *Novum Organum* and *The Advancement of Learning*, the foremost works cataloging Bacon's methodology. For example, the most highly associated document is a section of the latter where Bacon argues that "...the induction which the logicians speak of...whereby the principles of sciences may be pretended to be invented...is utterly vicious and incompetent...For he that shall attentively observe how the mind doth gather this excellent dew of knowledge...shall find that the mind of herself by nature doth manage and act an induction much better than they describe it. For to conclude upon an enumeration of particulars, without instance contradictory, is no conclusion, but a conjecture...".

4.8. Insights into Early Seventeenth-Century Legal-Intellectual Culture

The estimates from STM provide an integrated overview of the works of Bacon and Coke. This overview provides a picture of the intellectual relationship between the two authors that is different from current characterizations in the literature. The analysis places their works on an equal footing and produces a quantitative assessment. It generates novel insights into the contributions of Bacon and Coke to seventeenth-century ideas and culture.

Thirteen of the 25 topics are legal ones. Of the thirteen, only six would fit standard area-of-legal-application categories within modern classification schemes (Property Rights; Land Inheritance Law; Criminal Justice System; Constitutional Law; Jurisprudence; King, Law, & Nation).²¹ Three legal topics are on legal scholarship (Understanding Law; Jurisprudence; Disambiguating Law), with two of these, Understanding Law and Disambiguating Law, hardly to

²⁰ The naming of these two topics is the same as that in Grajzl and Murrell (2019), which focuses only on Bacon but uses the same corpus of Bacon's works as in this paper.

²¹ For example, that used for the *Journal of Economic Literature* (<https://www.aeaweb.org/jel/guide/jel.php>) or by LexisNexis (<https://www.lexisnexis.com/LegalNewsRoom/lexis-hub/b/legaltoolbox/posts/area-of-law-research>)

be expected from the existing literature. Most importantly, four legal topics cross area divides in ways that are suggestive of a deep hidden structure to legal reasoning: Defendant Rights, Regulation of Exchange, Public & Private Authority, and Legal Jurisdiction. These topics are conceptual rather than substantive. For example, Defendant Rights ranges across many areas of the substantive application of law, capturing general conceptual ideas concerning which defenses are admissible or desirable, and when. Similarly, the ideas embodied in Legal Jurisdiction are used to discuss disputes between courts, between nations, between church and state, between the Crown and the law, and between varieties of legal officials. In early seventeenth-century English culture, the ideas in the common law structured debates about conflicts of all kinds.

Thus, our STM estimates show commonalities between different subjects that do not appear in standard treatments. Often, the estimates do more than simply reflect the ostensible focus of a work—for example, crime or contract: they reveal a commonality between writings that span different areas of the application of law. Coke is the primary user of these topics, but Bacon also uses them. This suggests a commonality in their deployment of legal reasoning. It was always a dream of Bacon, never accomplished, to systematize the common law in a way that would reveal its deeper structures and logic. In the end, Coke's unwieldy and rambling *Institutes* and *Reports* were the closest anybody in the 17th century came to explicitly accomplishing this goal, as Bacon grudgingly acknowledged (Cromartie 1999: 86). Nonetheless, our estimates suggest a sophisticated structure within seventeenth-century common-law reasoning, but one that was implicit rather than explicit, and perhaps not even understood clearly by the lawyers themselves.

Figure 1 illustrates the relative importance of the various topics in the whole corpus. We note that the topics that capture more general concepts, methodology, or modes of understanding, tend to rank above average in importance, contrasting with the topics that center on domains of application. This suggests that these more general topics are used across a wide range of documents, and therefore a wide range of subject areas.

Finally, it is important to identify topics that could have been identified by STM, but were, in fact, not. This is part of the process of defining the set of ideas that exists in a specific milieu: in delineating the features of a distinct culture, it is necessary to identify what views and ideas were absent from the core components of that culture. Here we simply focus on those ideas that have been prominently attributed to Bacon and Coke in the existing literature but which we do not find in the STM topics, meaning that they were not emphasized by the authors. For example, Coke has sometimes been characterized as a proponent of *laissez-faire* and an ally of commercial interests, most famously by Heckscher (1935) and Hill (1965). Our 25-topic STM identifies no topics that could be associated with these ideas. As in Grajzl and Murrell (2019), there is also no topic that could be interpreted as capturing utilitarianism nor one that focuses on the centralized organization of science. These are ideas that were adopted by eighteenth-century Baconians and often attributed to Bacon himself. However our estimates do not provide evidence in favor of this

attribution: these ideas are not emphasized in the corpus.²² As we demonstrate in Appendix D, these findings are fully robust to estimating a 100-topic STM that offers an overview of the corpus at a much higher level of granularity.

5. Exploring the Connections Between the Cultural Topics

In this section, we use STM to assess the relationships between the estimated topics, harnessing the power of machine-learning to identify commonalities among the cultural ideas featured in different topics. These commonalities would not be readily apparent to a human reader in the sense that they reflect minutely detailed patterns in word usage both within and across documents, together with variations in such patterns that are associated with the values of meta-variables. In particular, as we demonstrate below, an examination of topic connectedness illuminates an important facet of Religion, Law, & Truth: it is a methodological topic drawing on examples from religion and law, as opposed to a topic about either religion or law per se.

We first examine document-level correlations between topics. Documents are mixtures of topics. Co-occurrence of two topics at the document level is evidence of complementarity in the use of topics. It shows that the two topics aid each other in expressing a specific set of ideas, indicating a shared conceptual foundation. We then examine vocabulary overlap between topics. Doing so allows us to assess the degree to which topics share a common semantic foundation.

Figure 2 provides a visualization of all positive topic correlations. The thickness of a link indicates the strength of the corresponding pairwise correlation. Among the featured correlations, the average correlation coefficient (ρ) equals 0.079.²³ (The complete set of correlation coefficients between topic pairs is summarized in Table E1 in Appendix E.) The positive topic correlations featured in Figure 2 can be visualized as comprising four elements: the nexus of legal topics (top-left); the nexus featuring a mixture of scientific and social-scientific topics predominantly used by Bacon (to the right); a three-topic nexus consisting of two core methodological topics, Understanding Law and Epistemology, and an associated one, Religion, Law, & Truth (in the center); and the completely disconnected Dynastic Politics. In the following, we examine each of the three connected parts of the nexus.

5.1. A Common Law Culture

The legal nexus comprises all legal topics featured in the corpus, with the exception of King, Law, & Nation. Most topics in the legal nexus are used by both Bacon and Coke. In particular, at the heart of the legal nexus is Jurisprudence, which is used equally by the two authors. Jurisprudence is correlated with six other topics, including Disambiguating Law ($\rho = 0.225$), also

²² The literature on Bacon's influence most often emphasizes four contributions. As in Grajzl and Murrell (2019), STM finds two of these, the inductive logic of interpreting the world (Epistemology) and the focus on cataloging the world (Probing for Facts). See, for example, Peltonen (1996), Rossi (1996), and Malherbe (1996). STM does not find any emphasis on the utilitarian value of produced knowledge or on centralized organizational arrangements for scientific investigation. See, for example, Rossi (1968), Gaukroger (2001), Mokyr (2005, 2010), Sargent (1996), and Harkness (2007).

²³ The data-generating model implies that if all data were random, the correlation between topics would be -0.0417. Figure 2 captures the 11.3% of highest-valued correlations.

a more-theoretical legal topic used equally by both. We view this as direct evidence that Bacon and Coke shared much legal theorizing.

Several other patterns stand out within the legal nexus. The cluster of correlations among Defendant Rights, Criminal Justice System, and Legal Jurisdiction illuminates the co-occurrence of ideas that form the foundation of modern-day criminal law and justice ($\rho = 0.105$ between Defendant Rights and Criminal Justice System; $\rho = 0.103$ between Criminal Justice System and Legal Jurisdiction; $\rho = 0.063$ between Defendant Rights and Legal Jurisdiction). Property Rights, a fundamental private law topic, is strongly correlated with Land Inheritance Law ($\rho = 0.458$), as one would expect, but also with both Disambiguating Law ($\rho = 0.167$) and Jurisprudence ($\rho = 0.111$). The latter pattern is consistent with the old common-law theme that the legal rights of Englishmen were viewed as inherited property, and indeed the connection between Property Rights and Constitutional Law ($\rho = 0.078$) shows how the two were intertwined. Public & Private Authority and Regulation of Exchange are especially likely to co-occur ($\rho = 0.177$). They are featured prominently in Coke's work. This is evidence of an early culture of regulation, while counter to the view that Coke was an advocate of *laissez-faire*.²⁴ Finally, the correlation among Constitutional Law, Property Rights, and the methodological Understanding Law foreshadows ideas and scholarship on the rule of law. ($\rho = 0.077$ between Constitutional Law and Property Rights; $\rho = 0.058$ between Property Rights and Understanding Law; $\rho = 0.117$ between Constitutional Law and Understanding Law.)

5.2. *A Culture of the Methodology for Expanding Knowledge*

The legal nexus is connected with the scientific and social-scientific nexus through Understanding Law, which is dominated by Coke, via Religion, Law, & Truth, used by both authors, and Bacon's Epistemology.²⁵ ($\rho = 0.161$ between Understanding Law and Religion, Law, & Truth; $\rho = 0.058$ between Religion, Law, & Truth and Epistemology.) The STM correlations indicate that Religion, Law, & Truth has many characteristics of—and in fact may be thought of as—a methodological topic.

Given the centrality of the methodological nexus, we investigate the connectedness between Religion, Law, & Truth, Epistemology, and Understanding Law by analyzing overlaps in vocabulary use, which indicate common semantic foundations. Figures 3-5 present a series of plots featuring the top 80 FREX words for each chosen topic pair. In each of the plots, the size of

²⁴ See Heckscher (1935) and Hill (1965) for such a characterization of Coke. In contrast, Malament (1967) rejects the characterization of Coke as proponent of *laissez-faire*.

²⁵ Grajzl and Murrell (2019) examine the genesis of Bacon's scientific methodology and provide quantitative evidence of the common-law origins of Bacon's epistemological thought. The results in Figure 2 are consistent with that evidence in that use of Epistemology co-occurs with the use of Religion, Law, & Truth. In Figure 2, Epistemology is also connected to Civic Knowledge, a finding suggestive of the influence of Renaissance humanism on Bacon's epistemological ideas (see Gaukroger 2001). Analysis of overlapping vocabulary between topic pairs reveals that the connection between Epistemology and Religion, Law, & Truth is stronger than the connection between Epistemology and Civic Knowledge. This suggests that any influence of Bacon's background in Renaissance humanism on his epistemological reasoning was weaker than the influence of his immersion in common law. Detailed evidence supporting this point is available upon request from the authors.

any displayed word is proportional to the word's use within the featured topic-pair. The horizontal position of a word measures the difference in the probabilities that associate a word with each topic, normalized by the maximum difference that occurs in the set of 80 words (Roberts et al. 2016b, fn. 20). The horizontal position of each word thus conveys how common a word is in one topic versus the other topic.²⁶ Words located close to the vertical dashed line are ones shared equally by the two topics.

The shared semantic foundations of Religion, Law, & Truth, Epistemology, and Understanding Law are clearly identifiable. While Epistemology focuses on 'natur' and Religion, Law, & Truth on 'church' and 'god', both use 'reason', 'true', 'cause', 'understand' and words employed in logical reasoning such as 'yet', 'upon', 'matter' and 'without' (Figure 3). Similarly, while Understanding Law is concerned with 'law' and 'king', it shares with Religion, Law, & Truth 'learn', 'know', 'reason', 'judgment', 'true', together with 'yet', 'upon', and 'without' (Figure 4). In the same vein, Epistemology and Understanding Law share the emphasis on 'learn', 'understand', 'true', 'reason', 'caus', 'know', 'observ', and 'find', as well as the logical words 'upon', 'see', 'therefore', 'yet', and 'matter' (Figure 5). Consistent with the evidence from topic correlations, Epistemology and Understanding Law capture two seemingly disparate, but methodologically congruent, approaches to understanding the world, while the applied methodology of Religion, Law, & Truth links them. The set of ideas in these three topics captures elements of a culture reflecting the methodology used in expanding knowledge.

5.3. *A Culture of Inquiry, Scientific and Social Scientific*

The scientific and social-scientific nexus features a collage of topics on natural science, methodology, politics, and even psychology, all strongly dominated by Bacon. Coke's work hardly figures here. This is easily explained: Coke writings concentrate on law. Given that our primary interest is in identifying shared elements of the ideas of Bacon and Coke, we comment on this nexus only very briefly.

Notably, a cluster of natural science topics is correlated with two methodological topics, Epistemology and Extracting Meaning. Extracting Meaning, a topic using old texts to pursue new inquiries, co-occurs with a political theme (Foreign Relations; $\rho = 0.063$) and with Civic Knowledge ($\rho = 0.116$), a topic highlighting Bacon's background in renaissance humanism with its spirit of inquiry into the practice of government. Also included in this cluster is Probing for Facts, a broad methodological topic emphasizing the need to ask questions to generate facts about the world. In sum, the connections between methodological topics and practical scientific investigation are indicative of a culture stressing the importance of inquiring into how the world works and thereby building reliable knowledge. This is a culture of empirical inquiry, in great contrast to the deductive approach in interpreting existing facts that was characteristic of the prevailing Aristotelian paradigm.

²⁶ In contrast, the vertical position of a word is random and carries no substantive interpretation.

6. Identifying Differences and Similarities Between Bacon and Coke

The previous section has shown that a machine-learning analysis of the Bacon-Coke corpus provides evidence of considerable consistency between the ideas of these two intellectual giants. But this insight leads to something of a paradox because the existing literature that utilizes conventional text analysis has invariably emphasized differences. In this section, we leverage STM's use of metavariables to explore the differences and similarities between the topical emphases of the two intellectuals. We show that although Bacon and Coke certainly differ in their emphases on specific ideas, there are also important similarities between them. These similarities appear especially when estimates of topic prevalence are conditioned on the intended audience and, above all, on the time at which the works were produced.

A caveat is in order before proceeding with the analysis. Topic modeling, with its conceptualization of text as a bag of words and reliance on a macroscopic lens to view the corpus, is a blunt tool for uncovering fine semantic similarities and distinctions between the ideas of different authors. For example, it could be that two authors reach opposite conclusions about a subject while using very similar words so that STM finds no difference in topic prevalences.²⁷ Nevertheless, given the existing literature's predominant narrative of the conflicts between the ideas of Bacon and Coke, quantitative evidence of similarities in the authors' emphases on specific topics—as well as in the pattern of the evolution of those emphases—would alone constitute an important finding.

At the same time, STM offers scope for a comparatively more nuanced analysis of differences and similarities between the ideas of Bacon and Coke than some of the alternative approaches for comparing textual corpora, such as the use of Kullback-Leibler/Jensen-Shannon divergence measures or cosine similarity (see, e.g., Gallagher et al. 2018, Lu et al. 2018, Gomaa and Fahmy 2013). Much like STM, these alternative approaches rely on a bag-of-words representation of text. However, unlike STM, they infer differences and similarities between texts based on word frequencies alone, without identifying coherent topics, without leveraging the correlation of words across documents, and without allowing for the possibility of the influence of metadata covariates on the prevalence of specific ideas.

6.1. Unconditional Analysis

Figure 6 shows how the proportions of topic use in the documents written by Coke differ from the same proportions for Bacon. For example, the proportion of Understanding Law in Coke

²⁷ There is an important lacuna in the methodological literature on STM that is relevant at this point. There is no agreement on whether divergent opinions on a given subject will inevitably result in differences across authors in the prevalences of the topics relevant to that subject, or whether authors could have similar topic prevalences and yet still feature substantively divergent opinions. The former seems to be the dominant view in practical applications of STM. Examples can be found in Tingley (2017), Lynam (2016), Farrell (2015), Reich et al. (2015), and Tvinnereim and Fløttum (2015). If we took this point of view, then the results presented in the ensuing section would imply that Bacon and Coke had substantively similar views on many aspects of law, for example. But given the methodological lacuna, we do not jump to this stronger conclusion.

documents is nearly 0.1 higher than in Bacon documents and the 95% confidence interval for the estimated difference in proportions lies between 0.04 and 0.14.

In some respects, Figure 6 shows how the dominant view in the literature might have arisen: overall there are considerable differences in topical emphases between Bacon and Coke. For most topics, there is no overlap between the confidence intervals and the zero-line, the line indicating that both use a topic equally. When examining the whole corpus the two authors look quite different—one is a scientist-methodologist and the other is the lawyer. Using Figure 6, the only way in which the dominant view could be subject to some scrutiny would come from a focus on Jurisprudence, Disambiguating Law, and Religion, Law, & Truth: the differences between Bacon and Coke are not statistically significant at the 95% level. But this is only three of 25 topics.

In ascertaining similarities or differences between the two authors' emphases, Figure 6 has a problem: it reflects the whole corpus. We do know that Bacon and Coke had very different careers, both over time and in the different audiences they would need to address. Perhaps the unconditional differences highlighted in Figure 6 simply reflect the broad outlines of their careers. For a more precise comparison, it is necessary to compare the two when they are writing for similar audiences or at a particular point in time.

6.2. Conditioning on Target Audience and Communication Form Eliminates Some Differences But Amplifies Others

There exists only a limited number of cells of the meta-data variables where both Bacon and Coke contribute a non-trivial number of documents to our corpus (see Table 2). These situations especially involve the targeting of lawyers, and to a lesser extent also methodologists and politicians. Also, the two authors both contributed speeches.²⁸ In interpreting the results produced by conditioning on the values of meta-variables, it is important to keep in mind that it would be perfectly natural for both Bacon and Coke to ignore certain topics when addressing specific audiences (e.g. Pharmacology or Botany to lawyers or politicians). Thus, the absence of a statistically significant difference in the authors' emphases on such topics does not imply anything about similarities between the authors' topical uses; it would simply reflect a general lack of relevance of the specific topic in a given situation. In the ensuing discussion we therefore purposefully ignore such topics.

We first investigate the consequences of conditioning on documents addressed to lawyers (Figure 7). Note that, in this case, Coke contributes more documents than Bacon and the total number of considered documents (205) is adequately large. Moreover, relative to unconditional analysis, conditioning on lawyers somewhat improves the balance in the documents contributed by the authors (see Table 2). Conditioning on lawyers, in contrast to the unconditional analysis (Figure 6), shows that it is Bacon, 'the lawyer without law'—not Coke, 'the greatest oracle of our

²⁸ The only other communication form that both Bacon and Coke used extensively was essays, but this category includes such a heterogeneous collection of documents that conditioning on essays does not yield any new insights beyond reducing some of the unconditional differences by a small amount.

municipal jurisprudence'—who places more emphasis on Jurisprudence and Disambiguating Law. Bacon and Coke also differ notably in their emphases on affairs conducted at the nation's highest level. Coke, an important constitutional actor, emphasizes Constitutional Law whereas Bacon, deeply embedded in court politics, emphasizes the matters that directly touched on the monarch's current decisions and place in the country (King & Court; King, Law, & Nation). This is evidence that the two authors thought very differently about the place of the monarch in England's constitutional politics.²⁹ However, in most areas of the concrete application of the law (Land Inheritance Law; Defendant Rights; Regulation of Exchange; Criminal Justice System), conditioning on lawyers eliminates statistically significant differences in the authors' emphases on topics that both used when addressing legal audiences. That is, with respect to their emphases on a range of substantive areas of law, Bacon and Coke were certainly not as dissimilar as one might have expected them to be based on the depictions in the existing literature utilizing conventional textual analysis.

Conditioning on documents addressed to politicians and methodologists, as well as conditioning on speeches, decreases the number of eligible documents and, in the case of conditioning on politicians and methodologists, increases the imbalance in the proportion of documents attributable to each of the authors. This tends to increase the confidence intervals for the estimates of differences in the authors' topical emphases. Nevertheless, conditioning on methodologists (Figure 8) amplifies the divergence between the authors' uses of their own methods for pursuit of knowledge: Epistemology for Bacon and Understanding Law for Coke. In this context, Religion, Law, & Truth, which is used by both authors, is more important for Coke than Bacon. Similarly, conditioning on documents directed at politicians (Figure 9) preserves the unconditional differences between the authors with respect to their emphasis on topics such as Defendant Rights, Constitutional Law, and King & Court. These differences perhaps reflect the kind of politicians being addressed: Coke addresses Parliament, while Bacon produces essays on strategy in interpersonal relations or private letters to the monarch or courtiers.

When conditioning on speeches (Figure 10), we see that Coke in comparison with Bacon emphasized Defendant Rights. It is possible that this finding points to the authors' differential willingness to engage in strategic communication of their ideas at a time when political disfavor could be very costly. Coke, who "stubbornly fought to limit the king's prerogative powers" (Berman 1994: 1674) was willing to publicly articulate his convictions about the importance of defendant rights. Bacon, in contrast, carefully distinguished between his public and private utterances, as can be seen by contrasting Figures 6 and 10. Elements of his writings might have been public lies, covering private truths (Kuran 1995).

In sum, conditioning on targeted audience and communication form eliminates some estimated differences in the topical emphases of Bacon and Coke, but amplifies others. The

²⁹ This point harks back to the second, methodological, paragraph of Section 6. In an area in which it is beyond dispute that Bacon and Coke had profound differences, these differences do clearly result in STM estimating two separate topics in the same subject area (King, Law, & Nation and Constitutional Law).

resulting analysis, however, does not incorporate the potentially important role of time: after all, the intellectual and professional pursuits of the two authors followed different timelines. A quantitative examination of their changing topical emphases over time thereby offers a unique analytical perspective for analysis of the conception and the evolution of thought of the two lawyer-intellectuals. As we illustrate below, it is precisely this exercise that allows us to uncover profound similarities between Bacon and Coke.

6.3. Conditioning on Time Reveals a Fundamental Similarity Between Bacon and Coke

Figure 11 has 25 elements, each depicting the change over time in the use of one of the topics by both authors. Bacon's timelines are represented by solid lines while Coke's are dashed. There are three lines for each, the estimated mean topic proportion over time and the 95% confidence bounds for the estimated mean. (Where only two lines appear for an author, the lower 95% confidence bound lies completely below zero.) With the exception of the earliest period, Bacon's documents are more abundant than Coke's, and thus the confidence intervals around the estimated mean topic proportions tend to be smaller for Bacon than for Coke. Yet this imbalance in featured documents should not affect the conclusions we draw in this subsection, as our interest here is in ascertaining the broad patterns in the evolution of the two authors' changes in topical emphases over their respective lifetimes. We do not focus on assessing differences in the emphases of the two authors on specific topics at any particular point in time.

The young Bacon is the quintessential common-lawyer, using standard topics when applying the law concretely. He contributes in such areas as Land Inheritance Law, Defendant Rights, the Criminal Justice System, and the Regulation of Exchange. The theoretical cast to his writings begins early: Jurisprudence and Disambiguating Law are used heavily even at the beginning of his career. Indeed, Bacon's emphasis on these two topics fades only slowly during his life, whereas emphases on concrete applications of law decline earlier. These applications of law are replaced by scientific topics, beginning in mid-career and rising steeply in importance thereafter.

The timelines provide insight into the sources of Bacon's most celebrated intellectual contribution, his work on the methodology of science. From the very beginning, Bacon was using his Epistemology and his use of this topic increased throughout his life. Thus, it was not the turn to science that spurred the development of Epistemology: it was already present from the very beginning when Bacon focused on law. This endorses the conclusion, already mooted in Section 5, that Bacon's epistemology was a product of his education in the common-law and his early experiences at the center of the legal profession. This conclusion is buttressed by the observation that Bacon used the proto-methodological topic, Religion, Law, & Truth, in his earliest contributions and that this topic stayed with him throughout his life, even as his writings focused increasingly on science and social science.

Bacon's other great methodological contribution, Probing for Facts, shows a very different trajectory from Epistemology. It is not present in the young Bacon, when he was focused on legal subjects. But its use begins before the turn to science, indicating that there is no evidence that

Probing for Facts was a product of Bacon's later-life concern with science. Probing for Facts seems to be more *sui generis* Bacon than is Epistemology.

It is easy to see from the topics emphasized in Coke's early writings how he could be cast as a typical medieval legal thinker. His most important topics are standard concrete applications of law—Defendant Rights, Regulation of Exchange, Legal Jurisdiction, and Public & Private Authority. There is no theoretical cast to his writings, in contrast to Bacon. For the young Coke, Understanding Law is a much less important topic than it would later become, Disambiguating Law does not appear, and Jurisprudence is overshadowed by topics reflecting applications of law. But there is one exception to this characterization, Religion, Law, & Truth, which is important in the first years of Coke's writings.

As the years pass, Coke moves to more general, more conceptual modes of deliberation. He places more emphasis on methodology (Understanding Law), on more theoretical aspects of law (Disambiguating Law), and on law at the highest level (Property Rights and Constitutional Law). One conjecture that might explain this broadening is Coke's change from being a servant of the monarch to a Chief Judge who guarded his independence jealously. But this reasoning is inconsistent with the observation that Coke's use of Legal Jurisdiction and Public & Private Authority declines greatly over time: these would be exactly the legal themes emphasized by a Chief Judge who was highly protective of his territory. A hypothesis that is more consistent with the topic timelines is that Coke was gradually developing his own broader theoretical approach to legal reasoning. In terms of the language of others, he was moving away from having "no theory at all" (Berman 1994) while simply dealing in an unsystematic set of rules (Siegel 1981) to having a systematic general perspective on legal reasoning.

Conditioning on time, therefore, reveals a fundamental congruity between the ideas of Bacon and Coke. At the beginning of their careers, they could each be characterized as traditional lawyers in the common-law mold. They worked on similar topics in apparently complementary ways. In the earlier years, Bacon had a more theoretical bent to his writings. But the later Coke did employ a coherent methodological approach, one that had many similarities with Bacon's better-known methodology, as Section 5 has already shown.

The evolution of the ideas of Bacon and Coke exhibits a common thread. Bacon moved from an emphasis on a conceptual understanding of common-law reasoning to a general scientific methodology. Coke began by immersing himself in the details of the common law but then moved on to develop a broader legal theorizing. Both moved from the particular to the general, but eventually in disparate fields of inquiry. Therefore, one cannot say, as the literature often does, that Bacon was modern and Coke medieval. They were both grappling with the crucially important issue of how to derive truthful statements from a set of apparently disparate facts. Both used the inductive approach that had developed for centuries as common-lawyers wrestled with generalizing from particulars. This was the intellectual culture that these two intellectual giants both imbibed and advanced.

7. Conclusion

The period spanning the late 16th and early 17th centuries was pivotal for England. At the dawn of an era of fundamental institutional changes and technological advances, elements of a distinct intellectual culture were taking shape, a culture that would leave a lasting imprint on the modern world. We have estimated the features of one aspect of this emerging culture by analyzing the works of Francis Bacon and Edward Coke, two lawyer-scholars who did much to contribute to the new intellectual environment. We thereby offer the first quantitative analysis of the legal-intellectual ideas that immediately preceded, and provided key input into, the subsequent, eighteenth-century culture that is believed to have spurred pioneering inventive activity and sustained technological progress (Mokyr 2016). Our insights into early seventeenth-century English culture therefore provide a key step in understanding the cultural origins of England's rise, a topic that has stimulated a voluminous body of research on comparative economic development, and in economic history in general.

Among the many detailed cultural and historical observations that we make, the following three substantive conclusions deserve particular emphasis. First, common-law thinking, as exemplified in the work of both Bacon and Coke, had by the early seventeenth century already developed a deep, theoretical structure. Applications cut across conventional legal subjects found in standard legal-classification schemes. That is, early seventeenth century English common-law thought was not simply an atheoretical collection of cases and facts, as has been the standard characterization.

Second, the interconnected nexuses of ideas in the corpus suggest a commitment to the systematic derivation of truthful statements from facts, combined with an emphasis on practical inquiry as a means of understanding the broader world. Looking forward, these components of early seventeenth-century English legal-intellectual culture provided an important input into the much broader culture, one that would eventually facilitate lasting economic progress. Looking backward, these components of culture are intimately linked to the use of common law-style reasoning, which had been in development for many centuries.

Third, while Bacon and Coke differed in their emphasis on specific subjects, their works evidence a methodological commonality that has not been discussed in the literature to date. Bacon's much celebrated epistemology is closely related to Coke's supposedly absent theoretical approach to understanding the law. Both derive their approach to understanding the world using the inductive reasoning inherent in common-law thought. Their shared methodology emphasizes practical inquiry and the building of reliable knowledge. The presence of this methodology in the works of both authors suggests that it is a central element of seventeenth-century English legal-intellectual culture.

Finally, we highlight a methodological point. Our analysis illustrates how machine-learning, and in particular structural topic modeling applied to original texts, can be productively utilized to investigate the features of a culture. Given the recent substantial interest among

economists in the study of culture, our approach suggests a novel quantitative approach to ascertaining the core characteristics of a culture prevailing in a specific time and place.

References

- Acemoglu, Daron and James A. Robinson. 2012. *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York, NY: Crown Publishers.
- Alesina, Alberto and Paola Giuliano. 2015. "Culture and Institutions." *Journal of Economic Literature*, 53:4, 898-944.
- Algan, Yann and Pierre Cahuc. 2010. "Inherited Trust and Growth." *American Economic Review*, 100:5, 2060-2092.
- Allen, Robert C. 2009. *The British Industrial Revolution in Global Perspective*. Cambridge, UK: Cambridge University Press.
- Baker, John. 2019. *An Introduction to English Legal History*. Fifth edition. Oxford, UK: Oxford University Press.
- Baker, John. 2017. *The Reinvention of Magna Carta 1216-1616*. Cambridge, UK: Cambridge University Press.
- Baker, J.H. 1971. "New Light on Slade's Case." *The Cambridge Law Journal*, 29:1, 51-67.
- Moore, Barrington Jr. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston, MA: Beacon Press.
- Barro, Robert J. and Rachel M. McCleary. 2003. "Religion and Economic Growth Across Countries." *American Sociological Review*, 68:5, 760-781.
- Berman, Harold J. 1994. "The Origins of Historical Jurisprudence: Coke, Selden, Hale." *Yale Law Journal*, 103:7, 1651-1738.
- Blaydes, Lisa, Justin Grimm, and Alison McQueen. 2018. "Mirrors for Princes and Sultans: Advice on the Art of Governance in the Medieval Christian and Islamic Worlds." *Journal of Politics*, 80:4, 1150-1167.
- Blei, David M. 2012. "Probabilistic Topic Models." *Communications of the ACM*, 55:4, 77-84.
- Blei, David M. and John D. Lafferty. 2007. "A Correlated Topic Model of Science." *The Annals of Applied Statistics*, 1:1, 17-35.
- Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. "Latent Dirichlet Allocation." *Journal of Machine Learning Research*, 3, 993-1022.
- Bowles, Samuel and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton, NJ: Princeton University Press.
- Burch, Charles N. 1928. "The Rivals." *Virginia Law Review*, 14:7, 507-525.
- Clark, Gregory. 2009. *A Farewell to Alms: A Brief Economic History of the World*. Princeton, NJ: Princeton University Press.
- Coquillette, Daniel R. 1992. *Francis Bacon*. Stanford, CA: Stanford University Press.
- Coquillette, Daniel R. 2004. *The Anglo-American Legal Heritage, Introductory Materials*, Second Edition, Durham, NC: Carolina Academic Press.
- Cromartie, Alan. 1995. *Sir Matthew Hale, 1609-1676: Law, Religion and Natural Philosophy*. Cambridge, UK: Cambridge University Press.

- Cromartie, Alan. 1999. "The Constitutionalist Revolution: The Transformation of Political Culture in Early Stuart England." *Past & Present*, 163, 76-120.
- Cromartie, Alan. 2006. *The Constitutionalist Revolution: An Essay on the History of England, 1450-1642*. Cambridge, UK: Cambridge University Press.
- De Montpensier, Roy L.S. 1968. "Bacon as Lawyer and Jurist." *ARSP: Archiv für Rechts- und Sozialphilosophie*, 54:4, 449-483.
- Dittmar, Jeremiah and Skipper Seabold. 2016. "Media, Markets, and Radical Ideas: Evidence from the Protestant Reformation." Unpublished manuscript.
- Farrell, Justin. 2015. "Corporate Funding and Ideological Polarisation About Climate Change." *Proceedings of the National Academy of Sciences*, 113:1, 92-97.
- Floud, Roderick and Deirdre N. McCloskey (eds.) 1994. *The Economic History of Britain since 1700, Volume 1: 1700-1860*. Cambridge, UK: Cambridge University Press.
- Friedrich, Carl Joachim. 1958. *The Philosophy of Law in Historical Perspective*. Chicago, IL: The University of Chicago Press.
- Gardner, William A. 1916. "Bacon's Lost Rules and Decisions." *American Law Review*, 50, 536-553.
- Gaukroger, Stephen. 2001. *Francis Bacon and the Transformation of Early-Modern Philosophy*. Cambridge, UK: Cambridge University Press.
- Gaukroger, Stephen. 2006. *The Emergence of a Scientific Culture. Science and the Shaping of Modernity, 1210-1685*. Oxford, UK: Oxford University Press.
- Gest, John M. 1909. "The Writings of Sir Edward Coke." *The Yale Law Journal*, 18:7, 504-532.
- Gentzkow, Matthew , Bryan T. Kelly, and Matt Taddy. 2019. "Text as Data." *Journal of Economic Literature*, forthcoming.
- Gallagher, Ryan J. Andrew J. Reagan, Christopher M. Danforth, and Peter Sheridan Dodds. 2018. "Divergent Discourse Between Protests and Counter-Protests: #BlackLivesMatter and #AllLivesMatter." *PLoS ONE*, 13:4, e0195644. Available at: <https://doi.org/10.1371/journal.pone.0195644>.
- Gomaa, Wael H. and Aly A. Fahmy. 2013. "A Survey of Text Similarity Approaches." *International Journal of Computer Applications*, 68:13, 13-18.
- Grajzl, Peter and Peter Murrell. 2016. "A Darwinian Theory of Institutional Development Two Centuries Before Darwin." *Journal of Economic Behavior and Organization*, 131(A) November: 346-372.
- Grajzl, Peter and Peter Murrell. 2019. "Toward Understanding 17th Century English Culture: A Structural Topic Model of Francis Bacon's Ideas." *Journal of Comparative Economics*, 47:1, 111-135.
- Grimmer, Justin and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis*, 21:3, 267-297.
- Guiso, Luigi, Paola Sapienza and Luigi Zingales. 2006. "Does Culture Affect Economic Outcomes?" *Journal of Economic Perspectives*, 20:2, 23-48.

- Guiso, Luigi, Paola Sapienza, Paola, and Luigi Zingales. 2016. "Long Term Persistence." *Journal of the European Economic Association*, 14:6, 1401-1436.
- Gutmann, Myron P., Emily Klancher Merchant, and Evan Roberts. 2018. "'Big Data' in Economic History." *The Journal of Economic History*, 75:4, 268-299.
- Hansen, Stephen and Michael McMahon. 2016. "Shocking Language: Understanding the Macroeconomic Effects of Central Bank Communication." *Journal of International Economics*, 99, S114-S133.
- Hansen, Stephen, Michael McMahon, and Andrea Prat. 2018. "Transparency and Deliberation Within the FOMC: A Computational Linguistics Approach." *Quarterly Journal of Economics*, 2:1, 801-870.
- Harkness, Deborah E. 2007. *The Jewel House: Elizabethan London and the Scientific Revolution*. New Haven, CT: Yale University Press.
- Hart, James S. Jr. 2003. *The Rule of Law, 1603-1660: Crowns, Courts and Judges*. Harlow, UK: Pearson Longman.
- Hathi Trust. n.d. <https://www.hathitrust.org>. Accessed January-May 2016.
- Hayek, Friedrich A. 1960. *The Constitution of Liberty*. Chicago, IL: The University of Chicago Press.
- Heckscher, Eli F. 1935. *Mercantilism*. London, UK: Routledge.
- Helgerson, Richard. 2004. "Writing the Law." In: Boyer, Allen D. (ed.), *Law, Liberty, and Parliament: Selected Essays on the Writings of Sir Edward Coke*. Indianapolis, IN: Liberty Fund, 26-69.
- Hill, Christopher. 1965. *Puritanism and Revolution: Studies in Interpretation of the English Revolution of the 17th Century*. New York, NY: Schocken Books.
- Hill, Christopher. 1997. *Intellectual Origins of the English Revolution Revisited*. New York, NY: Oxford University Press.
- Holdsworth, William S. 1935. "Sir Edward Coke." *The Cambridge Law Journal*, 5:3, 332-346.
- Holdsworth, William S. 1938. *Some Makers of English Law*. London, UK: Cambridge University Press.
- Hollond, H.A. 1947. "English Legal Authors Before 1700." *The Cambridge Law Journal*, 9:3, 292-329.
- Internet Archive. n.d. <https://archive.org>. Accessed January-May 2016.
- Kelley, Donald R. 1993. "Elizabethan Political Thought." In: Pocock, John G.A., Gordon J. Schochet, and Lois Schwoerer (eds.), *The Varieties of British Political Thought, 1500-1800*, Cambridge, UK: Cambridge University Press, 47-79.
- Kuran, Timur. 1995. *Private Truths, Public Lies. The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.
- Lichbach, Mark I. and Alan S. Zuckerman (eds.). 2009. *Comparative Politics. Rationality, Culture, and Structure*. New York, NY: Cambridge University Press, 134-161.

- Lu, Jinghui, Maeve Henchion, and Brian MacNamee. 2018. "Extending Jensen Shannon Divergence to Compare Multiple Corpora." *Proceedings of the 25th Irish Conference on Artificial Intelligence and Cognitive Science*. Available at: <http://hdl.handle.net/10197/10517>.
- Lucas, Christopher, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. "Computer-Assisted Text Analysis for Comparative Politics." *Political Analysis*, 23:2, 254-277.
- Lynam, Timothy. 2016. "Exploring Social Representations of Adapting to Climate Change Using Topic Modeling and Bayesian Networks." *Ecology and Society*, 21:4, Article 16.
- Lyon, Hastings and Herman Block. 1929. *Edward Coke, Oracle of the Law*. Boston, MA: Houghton Mifflin Co.
- Mahlerbe, Michel. 1996. "Bacon's Method of Science." In: Peltonen, Markku (ed.), *The Cambridge Companion to Bacon*. Cambridge, UK: Cambridge University Press, 75-98.
- Malament, Barbara. 1967. "The 'Economic Liberalism' of Sir Edward Coke." *Yale Law Journal*, 76:7, 1321-1358.
- Mathias, Peter. 1969. *The First Industrial Nation: An Economic History of Britain, 1700-1914*. London, UK: Methuen.
- McCloskey, Deirdre Nansen. 2006. *The Bourgeois Virtues: Ethics for an Age of Commerce*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre Nansen. 2010. *Bourgeois Dignity: Why Economics Can't Explain the Modern World*. Chicago, IL: University of Chicago Press.
- McCloskey, Deirdre Nansen. 2016. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. Chicago, IL: University of Chicago Press.
- Mohr, John W. and Amin Ghaziani. 2014. "Problems and Prospects of Measurement in the Study of Culture." *Theory and Society*, 43, 3-4, 225-246.
- Mohr, John W. and Petko Bogdanov. 2013. "Introduction—Topic Models: What They Are and Why They Matter." *Poetics*, 41, 545-569.
- Mokyr, Joel. 2005. "The Intellectual Origins of Modern Economic Growth." *Journal of Economic History*, 65:2, 285-381.
- Mokyr, Joel. 2009. *The Lever of Riches: Technological Creativity and Economic Progress*. New York, NY: Oxford University Press.
- Mokyr, Joel. 2010. "The Contribution of Economic History To The Study of Innovation and Technical Change: 1750-1914." In: Bronwyn H. Halland and Nathan Rosenberg (Eds.), *Handbook of the Economics of Innovation, Vol. 1*, Burlington, MA: Academic Press, 11-50.
- Mokyr, Joel. 2016. *A Culture of Growth: The Origins of the Modern Economy*. Princeton, NJ: Princeton University Press.
- Nenner, Howard. 1977. *By Colour of Law: Legal Culture and Constitutional Politics in England, 1660-1689*. Chicago: University of Chicago Press.

- Newman, David J. and Sharon Block. 2006. "Probabilistic Topic Decomposition of an Eighteenth-Century American Newspaper." *Journal of the American Society for Information Science and Technology*, 57:6, 653-767.
- North, Douglass C. and Barry Weingast. 1989. "Constitutions and Commitment: The Evolution of Institutions Governing Public Choice in Seventeenth-Century England." *Journal of Economic History*, 49:4, 803-832.
- North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. New York, NY: Cambridge University Press.
- Nunn, Nathan. 2012. "Culture and the Historical Process." *Economic History of Developing Regions*, 27:S1, 108-126.
- Peltonen, Markku. 1996a. "Introduction." In: Peltonen, Markku (ed.), *The Cambridge Companion to Bacon*. Cambridge, UK: Cambridge University Press, 1-24.
- Project Gutenberg. n.d. <https://www.gutenberg.org>. Accessed January-May 2016.
- Reich, Justin, Dustin Tingley, Jetson Leder-Luis, Margaret E. Roberts, Brandon M. Stewart. 2015. "Computer-Assisted Reading and Discovery for Student-Generated Text in Massive Open Online Courses." *Journal of Learning Analytics*, 2:1, 156-184.
- Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2016b. "stm: R Package for Structural Topic Models." *Journal of Statistical Software*, forthcoming.
- Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoidi. 2016a. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association*, 111:515, 988-1003.
- Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science*, 58:4, 1064-1082.
- Rogers, James G. 1928. "Bacon and Coke: A Contrast in the Law Career." *Rocky Mountain Law Review*, 1:1, 30-34.
- Sargent, Rose-Mary. 1996. "Bacon as an Advocate for Cooperative Scientific Research." In: Peltonen, Markku (ed.), *The Cambridge Companion to Bacon*. Cambridge, UK: Cambridge University Press, 146-171.
- Sheppard, Steve (ed.), 2003. *Selected Writings and Speeches of Sir Edward Coke*. Indianapolis, IN: Liberty Fund.
- Siegel, Stephen A. 1981. "The Aristotelian Basis of English Law, 1450-1800." *New York University Law Review*, 56:1, 18-59.
- Tabellini, Guido. 2008. "Institutions and Culture." *Journal of the European Economic Association*, 6:2-3, 255-294.
- Taddy, Matthew A. 2012. "On Estimation and Selection for Topic Models." *Proceedings of the 15th International Conference on Artificial Intelligence and Statistics*, 1184-1193.

- The Text Creation Partnership for Early English Books Online. 2014. <http://www.textcreationpartnership.org>. Accessed January-May 2016.
- Tingley, Dustin. 2017. "Rising Power on the Mind." *International Organization*, 71(S1): S165-S188.
- Tvinnereim, Endre and Kjersti Fløttum. 2015. "Explaining Topic Prevalence in Answers to Open-Ended Survey Questions About Climate Change." *Nature Climate Change*, 5: 744-8.
- Wallach, Hanna M., Iain Murray, Ruslan Salakhutdinov, and David Mimno. 2009. "Evaluation Methods for Topic Models." *ICML '09 Proceedings of the 26th Annual International Conference on Machine Learning*, 1105-1112.
- Wehrheim, Lino. 2019. "Economic History Goes Digital: Topic Modeling the Journal of Economic History." *Cliometrica*, 13:1, 83-125.
- Wheeler, Harvey. 1983. "The Invention of Modern Empiricism: Juridical Foundations of Francis Bacon's Philosophy of Science." *Law Library Journal*, 76, 78-120.
- Wootton, David. 2015. *The Invention of Science: A New History of the Scientific Revolution*. Epub Edition. New York, NY: Harper.

Table 1: Works Included in the Corpus

Name or type of work	Author	Document count
<i>New Atlantis</i>	Bacon	5
<i>Novum Organum</i>	Bacon	17
<i>The Advancement of Learning</i>	Bacon	12
<i>The Use of the Law</i>	Bacon	5
<i>A Collection of Apothegms, New and Old</i>	Bacon	6
<i>History Natural and Experimental of Life and Death</i>	Bacon	9
<i>Sylva Sylvarum</i>	Bacon	10
<i>The History of The Reign of King Henry the Seventh</i>	Bacon	6
<i>The Natural and Experimental History of Winds</i>	Bacon	7
<i>The Wisdom of the Ancients</i>	Bacon	5
<i>The Elements of the Common Laws of England</i>	Bacon	28
Bacon's other writings (essays, letters, speeches, case reports)	Bacon	172
<i>The Reports, Part One</i>	Coke	2
<i>The Reports, Part Two</i>	Coke	4
<i>The Reports, Part Three</i>	Coke	3
<i>The Reports, Part Four</i>	Coke	5
<i>The Reports, Part Five</i>	Coke	9
<i>The Reports, Part Six</i>	Coke	2
<i>The Reports, Part Seven</i>	Coke	4
<i>The Reports, Part Eight</i>	Coke	4
<i>The Reports, Part Nine</i>	Coke	4
<i>The Reports, Part Ten</i>	Coke	3
<i>The Reports, Part Eleven</i>	Coke	5
<i>The Reports, Part Twelve</i>	Coke	22
<i>The Reports, Part Thirteen</i>	Coke	2
<i>The First Part of the Institutes, or a Commentary upon Littleton</i>	Coke	32
<i>The Second Part of the Institutes</i>	Coke	17
<i>The Third Part of the Institutes</i>	Coke	9
<i>The Fourth Part of the Institutes</i>	Coke	11
Coke's other writings (essays, speeches)	Coke	12
Total number of documents for Bacon		282
Total number of documents for Coke		150
Total number of documents in the corpus		432

Note: The names of some of the works have been edited for rendering in modern English.

Table 2: Frequency of Documents by Author, Form of Finished Work, Intended Audience, and Time Period of Completion

Panel A: Documents by form of finished work			
Form of finished work	Bacon	Coke	Total
Essay	95	84	179
Case report	2	56	58
Apothegm	34	0	34
Speech	14	10	24
Tome	76	0	76
Letter	61	0	61
Total	282	150	432

Panel B: Documents by intended audience			
Intended audience	Bacon	Coke	Total
Lawyers	74	131	205
Politicians	58	7	65
Historians	14	3	17
Philosophers	50	0	50
Methodologists	39	9	48
Scientists	47	0	7
Total	282	150	432

Panel C: Documents by time period of completion			
Time period	Bacon	Coke	Total
Prior to 1590	0	6	6
1590-1599	52	11	63
1600-1609	67	38	105
1610-1619	63	39	102
1620 or later	100	56	156
Total	282	150	432

Note: Bacon lived 1561-1626. Coke lived 1552-1634.

Table 3: Themes, Topics, and Authorship

Themes and topics	% top 50 documents by Bacon
<i>Legal scholarship</i>	
Understanding Law	14
Jurisprudence	54
Disambiguating Law	50
<i>Private law</i>	
Property Rights	28
Land Inheritance Law	30
<i>Private/public law</i>	
Religion, Law, & Truth	50
Defendant Rights	28
Regulation of Exchange	22
Public & Private Authority	10
<i>Public law</i>	
Legal Jurisdiction	10
Criminal Justice System	34
Constitutional Law	0
King, Law, & Nation	82
<i>Politics</i>	
King & Court	96
Dynastic Politics	84
Foreign Relations	92
Civic Knowledge	100
<i>Science</i>	
Human Nature	96
Botany	100
Pharmacology	100
Physics, Air & Sound	92
Physics, Energy	94
<i>Methodology</i>	
Extracting Meaning	100
Probing For Facts	92
Epistemology	100

Note: % top 50 documents by Bacon refers to the percent of the 50 documents featuring a given topic most prominently that are authored by Bacon (as opposed to Coke).

Table 4: Topics and Top Words for the Estimated 25-Topic STM

Understanding Law

Highest Prob: law, king, time, say, case, great, court, justic, book, man, mani, shall, may, will, author, one, make, year, learn, ancient, first, england, common, reign, part, commonlaw, call, statut, observ, judg

FREX: prefac, student, sage, conqueror, demurr, reader, professor, treatis, commentari, inn, dom, justinian, forest, herein, institut, client, compil, confessor, reign, reverend, conquest, sergeant, greek, advoc, cite, edit, publish, lawyer, judici, cautious

Jurisprudence

Highest Prob: use, statut, shall, law, feoff, make, upon, case, will, land, estat, seiz, may, word, therefor, act, take, yet, say, life, first, can, time, right, possess, fine, good, heir, tenant, give

FREX: cesti, feoff, entail, covin, lesse, feme, lessor, remaind, statut, use, disseise, remit, proviso, seiz, leas, formedon, levi, fraud, bargain, convey, atturn, remitt, conting, estat, tenanc, asset, surrend, trust, revers, stranger

Disambiguating Law

Highest Prob: shall, land, grant, king, word, deed, pass, one, name, man, may, make, take, call, yet, say, law, year, upon, non, wit, can, signifi, give, time, place, therefor, rule, rend, hold

FREX: domesday, ambigu, praecip, deed, estov, dale, meadow, acr, widow, aver, quarantin, revoc, signifi, revok, pastur, arbitra, falsiti, terra, liveri, injust, papyrus, detractor, moietii, habendum, oblig, rei, promontori, date, style, oblige

Property Rights

Highest Prob: heir, shall, say, case, land, edward, son, purchas, make, estat, law, use, take, man, can, bodi, shelley, tail, issu, word, therefor, male, execut, may, life, yet, tenant, will, first, hold

FREX: shelley, vest, male, tail, ventur, escheat, homag, recoveri, purchas, recoveror, elder, advowson, indentur, forasmuch, haer, fol, divest, beneficiari, reenter, sue, leas, escuag, remaind, bastard, heir, ancestor, jane, fealti, feesimpl, donor

Land Inheritance Law

Highest Prob: heir, land, shall, son, inherit, father, law, blood, die, issu, part, man, brother, old, mother, seiz, make, feesimpl, without, descend, author, enter, case, descent, can, littleton, daughter, call, fee, purchas

FREX: inherit, brother, blood, father, mother, uncl, cousin, descent, heir, section, sister, littleton, descend, son, feesimpl, issu, daughter, lineal, attaint, soccag, attaind, die, guardian, haered, collater, freehold, sex, maxim, albeit, parent

Religion, Law, & Truth

Highest Prob: god, church, man, say, law, will, may, shall, great, time, good, thing, upon, bishop, know, case, matter, word, part, can, one, holi, first, work, yet, true, see, mani, religion, publish

FREX: preach, controversi, church, uniti, preacher, holi, christ, ghost, etern, liturgi, worship, libel, publish, scriptur, ministri, sermon, pastor, reveal, apostl, saviour, atheist, profan, primit, prayer, scandal, baptism, chariti, god, sin, censur

Defendant Rights

Highest Prob: law, say, shall, man, king, will, may, case, upon, hous, caus, give, edward, make, justic, can, take, without, imprison, judg, reason, non, defend, good, parti, great, one, time, arrest, action

FREX: arrest, etcetera, loan, corpus, imprison, pillag, habea, april, detain, verdict, grievanc, num, bail, breve, magnacarta, atia, prison, martial, warrant, forethink, barg, passeng, assumpsit, march, june, malic, excus, indict, stanford, felon

Regulation of Exchange

Highest Prob: shall, good, land, man, may, make, custom, debt, king, will, take, upon, one, say, court, prescript, time, can, everi, pay, day, call, manor, lord, sheriff, parti, within, give, justic, use

FREX: swan, prescript, debt, pledg, eyr, tourn, usuri, executor, franchis, prescrib, legaci, leet, chattel, creditor, properti, owner, pay, sewer, copyhold, manor, circuit, rate, forfeit, bank, satisfact, sale, recogniz, shop, market, sheriff

Public & Private Authority

Highest Prob: say, king, shall, make, case, act, aforesaid, grant, henri, may, law, word, edward, time, can, one, caus, without, hospit, incorpor, good, give, hous, upon, plaintiff, well, borough, within, patent, resolv

FREX: hospit, borough, colleg, sutton, censor, disfranchis, aforesaid, monopoli, mayor, incorpor, burgess, plymouth, bag, jame, licenc, art, card, chantri, commonalti, thoma, patent, governor, ordin, trade, dispens, physic, sole, rectori, partnership, corpor

Legal Jurisdiction

Highest Prob: king, law, say, england, court, shall, case, one, henri, upon, allegi, subject, edward, may, justic, judg, natur, can, make, appear, bear, plea, writ, caus, protect, statut, prohibit, cap, act, within

FREX: allegi, protect, prohibit, tith, praemunir, calvin, pendent, fol, ecclesiast, alien, modus, gascoin, obedi, plea, jurisdict, cannon, cogniz, lieg, extra, municip, surmis, born, deniz, cap, spiritu, bracton, lib, dominion, parson, headway

Criminal Justice System

Highest Prob: king, treason, act, shall, statut, law, case, say, justic, man, feloni, make, high, indict, upon, may, lord, take, person, offenc, word, parliament, realm, within, court, judgement, punish, give, death, can

FREX: treason, mispris, indict, feloni, peer, steward, constabl, heresi, accus, offenc, gaol, overt, heret, convict, counterfeit, guilti, arraign, sorceri, accessari, abjur, punish, purvien, petti, offend, conspiraci, murder, roy, oier, witchcraft, traitor

Constitutional Law

Highest Prob: king, parliament, lord, hold, common, say, shall, make, writ, edward, henri, come, act, year, call, statut, per, great, law, may, time, custom, realm, baron, grant, one, england, quoth, take, appear

FREX: tunnag, subsidi, baroni, frankpledg, parliament, imposit, poundag, wool, concilium, summon, men, baron, elig, abbot, burgess, assent, holiday, johann, roll, chapter, alnag, relief, comitatus, ibidem, comrad, prorog, diem, print, realli, testa

King, Law, & Nation

Highest Prob: law, shall, will, may, king, majesti, england, one, natur, kingdom, subject, make, time, first, therefor, can, scotland, unto, peopl, part, sever, yet, great, person, now, point, question, take, upon, parliament

FREX: plantat, scotland, union, speaker, papist, kingdom, britain, nation, unit, vote, sovereignti, gascoign, majesti, seminari, recus, monarchi, submiss, jesuit, abus, committe, pope, born, provinc, territori, hostil, scottish, foreign, nobil, ireland, excommun

King & Court

Highest Prob: majesti, will, shall, may, good, time, lordship, lord, upon, great, think, make, can, self, king, now, know, give, god, well, honour, yet, take, man, thing, letter, matter, much, place, though

FREX: lordship, majesti, humbl, gracious, essex, hope, pleas, solicitor, acquaint, revenu, wish, letter, thank, honour, honest, care, con, glad, happi, remembr, beseech, crave, unworthi, duti, messag, secretari, vouchsaf, advic, callisthen, presum

Dynastic Politics

Highest Prob: king, upon, make, great, will, time, shall, come, may, man, part, take, person, one, peopl, war, yet, also, unto, good, well, think, lord, much, give, princ, england, can, duke, two

FREX: maximilian, perkin, ferdinando, treati, flander, duke, castill, charl, ambassador, rebel, york, margaret, ladi, plantagenet, fillip, britain, french, lovel, earl, bruge, stanley, ambassag, succour, sanctuari, napl, burgundi, duchess, thousand, clifford, calai

Foreign Relations

Highest Prob: upon, great, war, make, man, spain, will, time, state, king, shall, yet, say, may, one, part, law, take, england, princ, come, peopl, now, like, though, much, true, mani, never, nation

FREX: spain, lopez, spaniard, invas, squir, spanish, portug, cathol, war, turk, antonio, duel, enterpris, pollio, manuel, nation, valour, christendom, design, indi, armi, quarrel, confeder, leagu, lowcountri, germani, libel, palatin, navi, sundri

Civic Knowledge

Highest Prob: man, good, will, make, great, say, upon, may, thing, time, one, can, much, natur, well, shall, like, see, yet, person, mind, take, mani, use, virtu, therefor, come, first, part, fortun

FREX: envi, tacitus, caesar, felic, cicero, fortun, bewar, faction, anger, dissimul, demosthen, solomon, sulla, reprehens, virtu, secreci, discours, convers, lover, precept, machiavelli, poverti, discontent, cun, seneca, perturb, tiberius, malum, busi, proverb

Human Nature

Highest Prob: man, shall, will, upon, make, great, may, thing, one, come, say, think, also, work, take, give, imagin, good, day, time, place, let, see, thou, yet, like, god, mani, can, natur

FREX: thi, ointment, imagin, witch, galleri, bensalem, inventor, thou, boat, belief, bead, magic, pillar, travel, jew, miracl, wart, earthquak, front, room, plagu, perfum, herald, sick, delug, scroll, hebrew, dream, blue, remnant

Botany

Highest Prob: tree, will, upon, plant, fruit, make, earth, may, water, put, grow, ground, caus, come, root, also, forth, like, flower, great, herb, seed, leav, set, see, bear, kind, one, therefor, sun

FREX: sap, plum, bough, cherri, pear, stalk, graft, herb, oak, tree, holli, moss, peach, dung, fig, elm, radish, plant, cucumb, blossom, onion, appl, lettuc, earli, rosemari, flower, fruit, mushroom, colewort, mistleto

Pharmacology

Highest Prob: bodi, spirit, will, part, water, may, make, thing, also, see, upon, long, caus, air, heat, great, man, much, therefor, like, take, live, use, time, good, natur, cold, littl, motion, one

FREX: liquor, tooth, intener, oil, nitr, aliment, stomach, amber, explic, opiat, putrefact, beer, spirit, saffron, consumpt, decoct, refriger, bottl, opium, sweat, sugar, cool, liver, indur, diet, flesh, dri, repar, malaciss, purger

Physics, Air & Sound

Highest Prob: wind, sound, air, will, make, one, may, blow, thing, great, motion, water, bodi, part, upon, see, like, likewis, much, also, caus, sail, come, place, shall, man, sea, yet, can, two

FREX: wind, string, echo, sail, sound, tone, brass, rain, percuss, blow, pipe, mast, south, nois, east, bell, concav, articul, cloud, north, dram, lute, engend, audibl, air, presag, gale, nurseri, metal, diapason

Physics, Energy

Highest Prob: motion, heat, bodi, natur, instanc, will, water, air, must, also, may, fire, can, place, let, part, first, one, like, cold, substanc, yet, flame, great, observ, differ, spirit, power, appear, shall

FREX: magnet, similar, instanc, howev, expans, anim, rapid, heat, investig, latent, predomin, liquid, tendenc, exist, excit, ebb, ray, heterogen, homogen, expand, conspicu, migrat, perpendicular, ignit, flame, format, motion, concret, subjoin, exclus

Extracting Meaning

Highest Prob: man, thing, year, life, may, live, age, natur, also, one, long, yet, hundr, great, say, shall, will, time, see, seem, mani, unto, god, make, like, mind, old, well, can, neither

FREX: nineti, jupit, prometheus, pan, eighti, proserpina, fabl, orpheus, parabl, page, miss, nemesi, siren, typhon, muse, pentheus, sphinx, perseus, youth, giant, bacchus, cupid, palla, cere, vicissitud, hundr, moreov, septenti, eleg, atalanta

Probing for Facts

Highest Prob: say, will, lord, man, one, upon, king, shall, great, come, answer, make, give, take, may, sir, can, time, first, day, know, like, queen, thing, tell, ask, much, think, overburi, see

FREX: overburi, impoison, somerset, weston, bacon, ask, poison, vespasian, diogen, slander, aristippus, peacham, madam, pompey, aggrav, injunct, pillow, tell, ladi, gentleman, forgiv, seaman, displac, chanceri, cardin, captain, sir, raleigh, answer, alexand

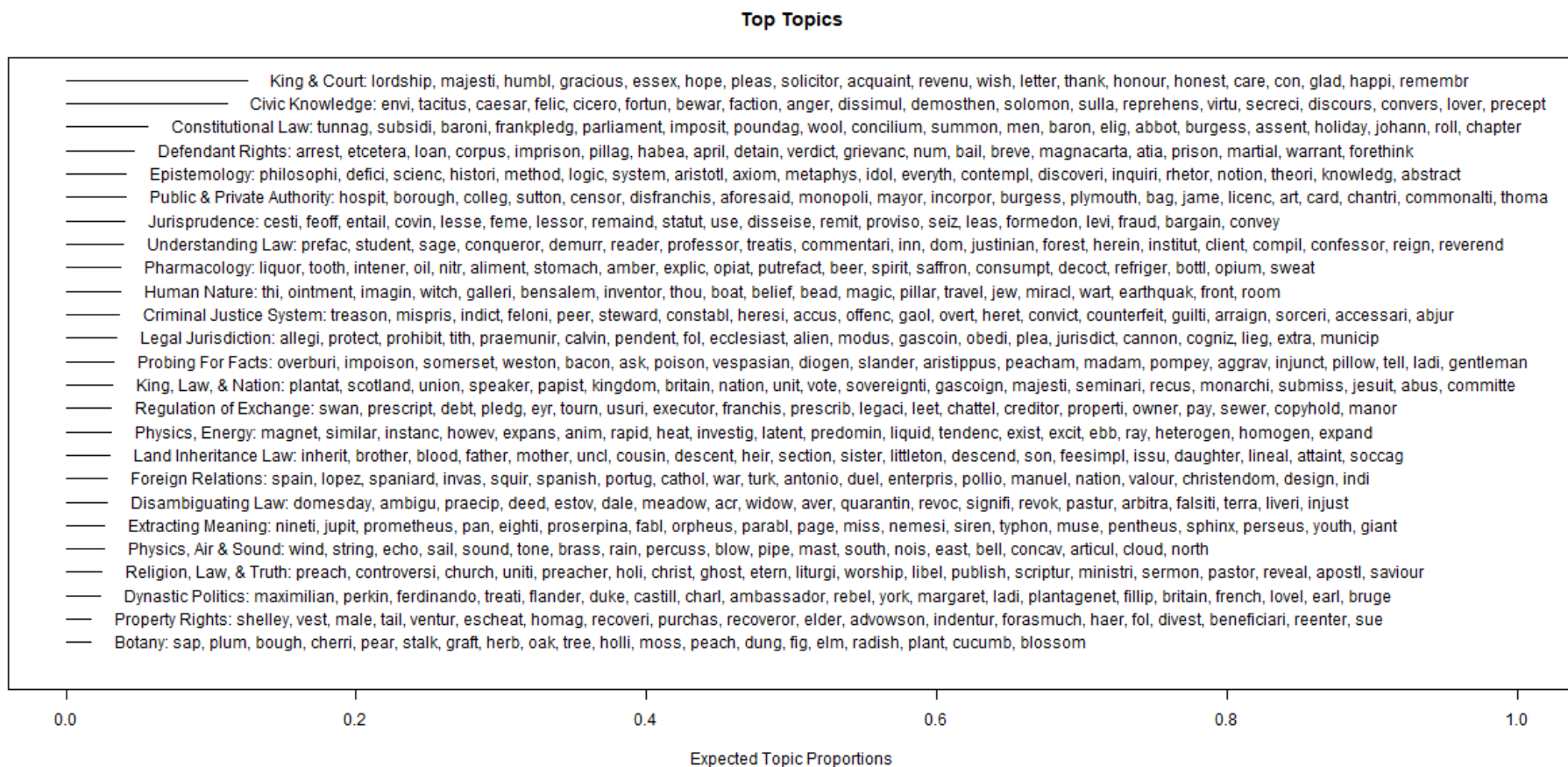
Epistemology

Highest Prob: natur, man, thing, will, may, philosophi, knowledg, histori, use, part, one, scienc, great, mind, can, make, upon, yet, shall, experi, therefor, learn, matter, first, much, caus, particular, work, true, find

FREX: philosophi, defici, scienc, histori, method, logic, system, aristotl, axiom, metaphys, idol, everyth, contempl, discoveri, inquiri, rhetor, notion, theori, knowledg, abstract, mathemat, theolog, mankind, principl, poesi, deduc, sophist, plato, intellectu, subtleti

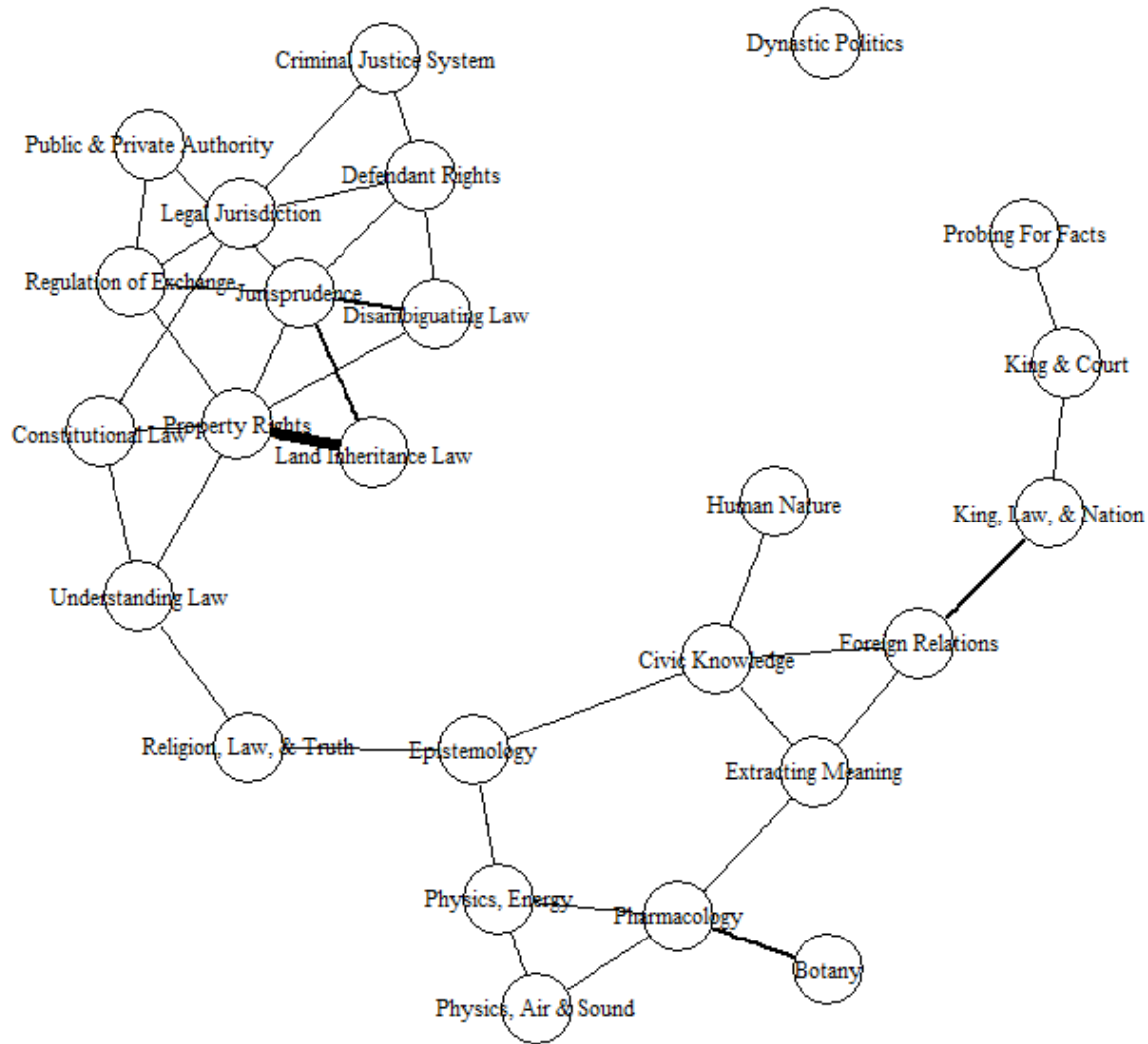
Note: The 'words' listed in the table are those used by STM after reducing all original text words to their stemmed form. Thus, for example, 'possess' could reflect an original usage of possess, possesses, possession, possessing, possessed, possessions, etc.

Figure 1: Expected Values for Topic Proportions in the Bacon and Coke Corpus



Note: For each topic, the figure displays the assigned names and the top fifteen FREX words. The size of the bars to the left of each topic is proportional to the probability that a random word drawn from the whole corpus has been generated by that particular topic.

Figure 2: Positive Topic Correlations



Note: The relative thickness of links reflects the magnitude of (positive) correlations between respective topic pairs.

Figure 5: Vocabulary Use Across Topic Pairs: Epistemology vs. Understanding Law

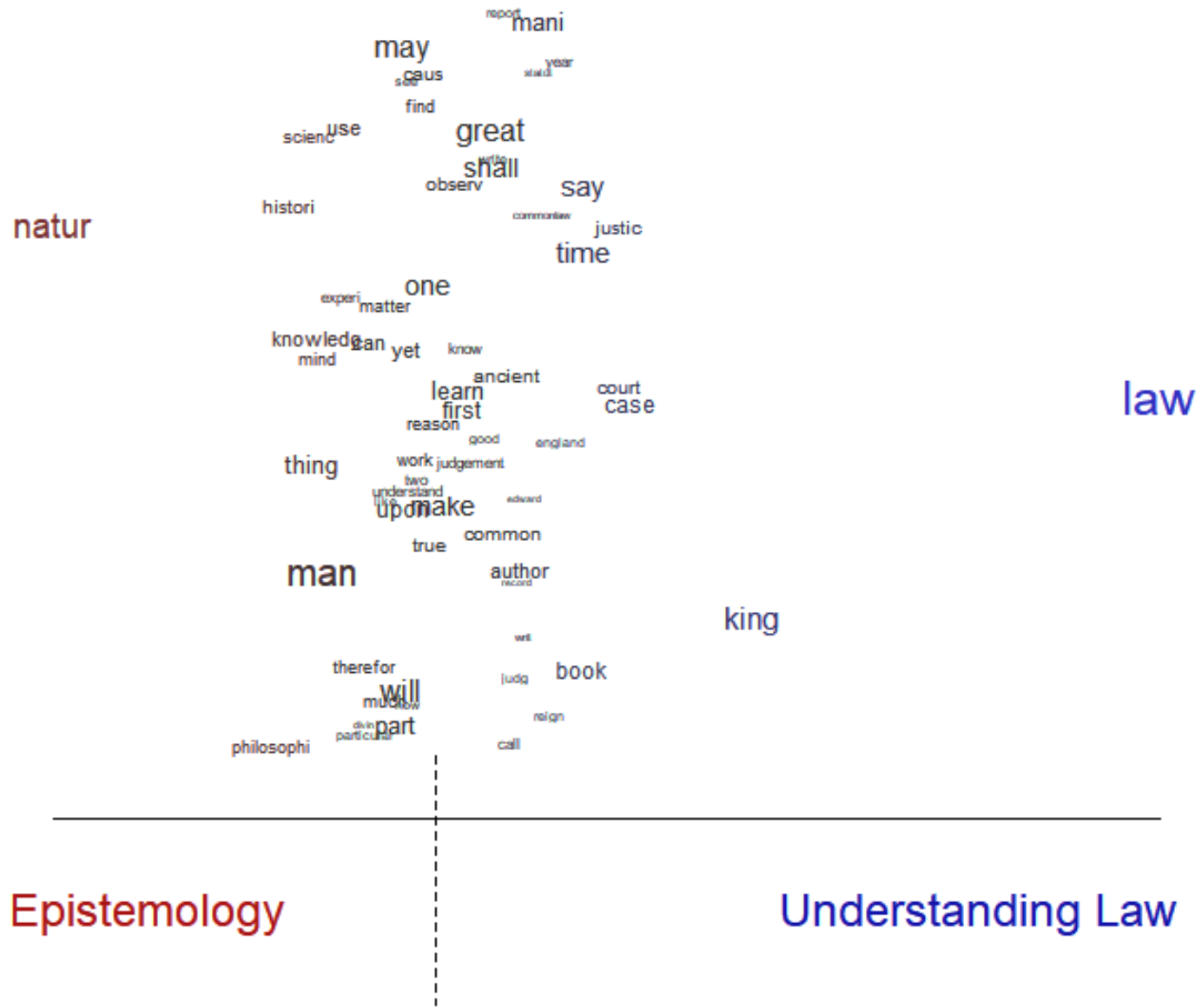


Figure 6: Use of the 25 Topics: Bacon versus Coke

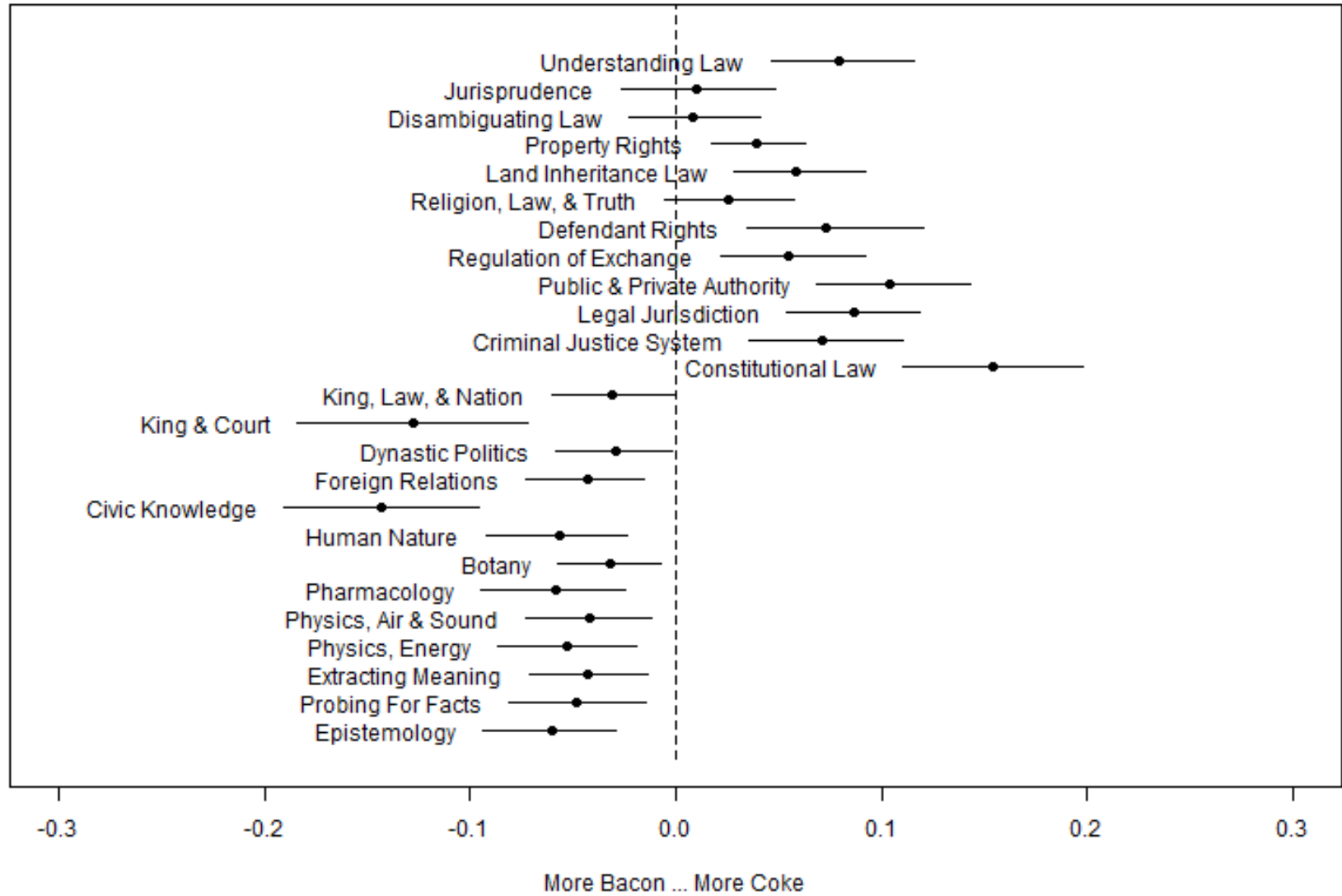


Figure 7: Use of the 25 Topics: Bacon versus Coke When Addressing Lawyers

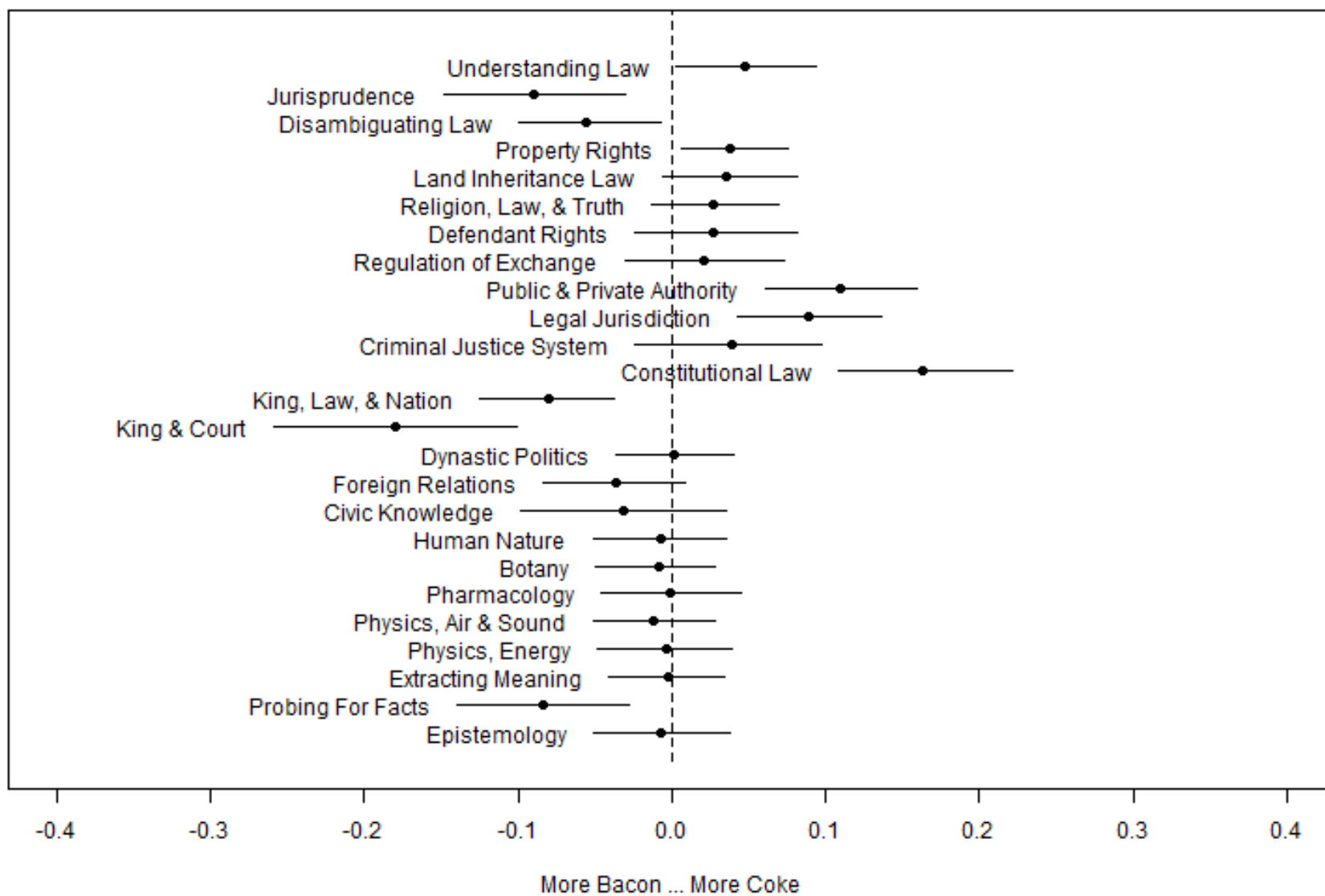


Figure 8: Use of the 25 Topics: Bacon versus Coke When Addressing Methodologists

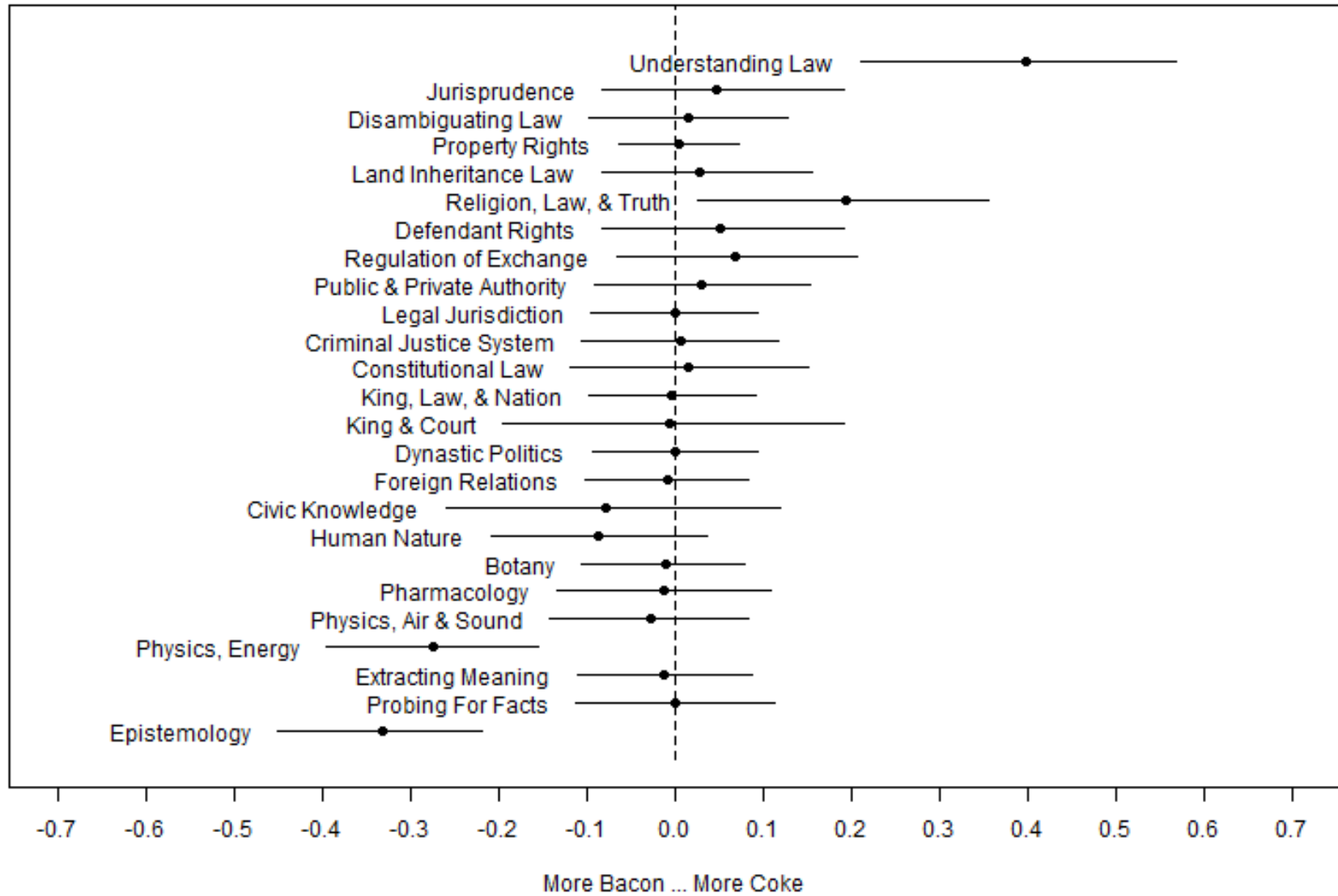


Figure 9: Use of the 25 Topics: Bacon versus Coke When Addressing Politicians

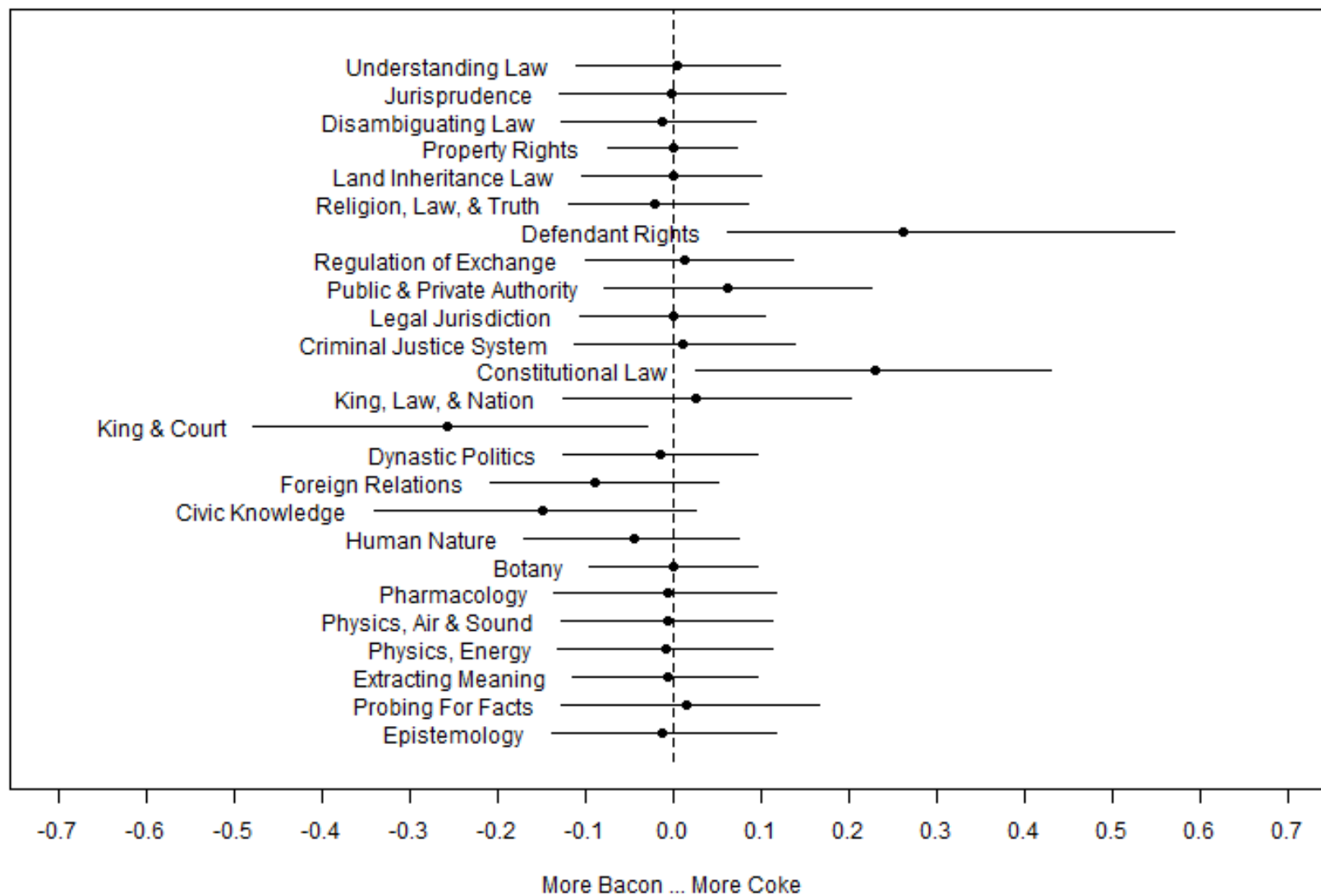


Figure 10: Use of the 25 Topics: Bacon versus Coke When Making Speeches

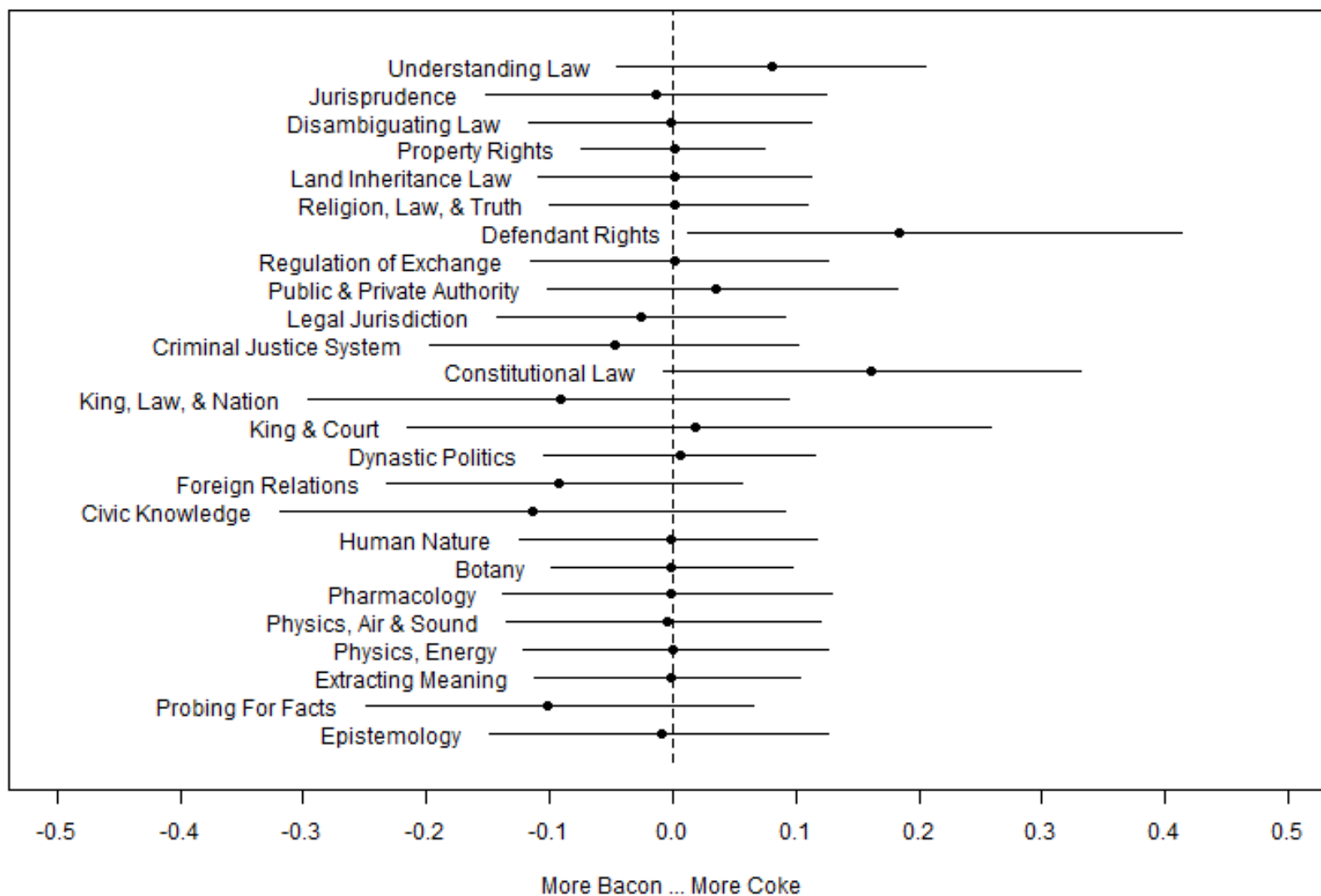
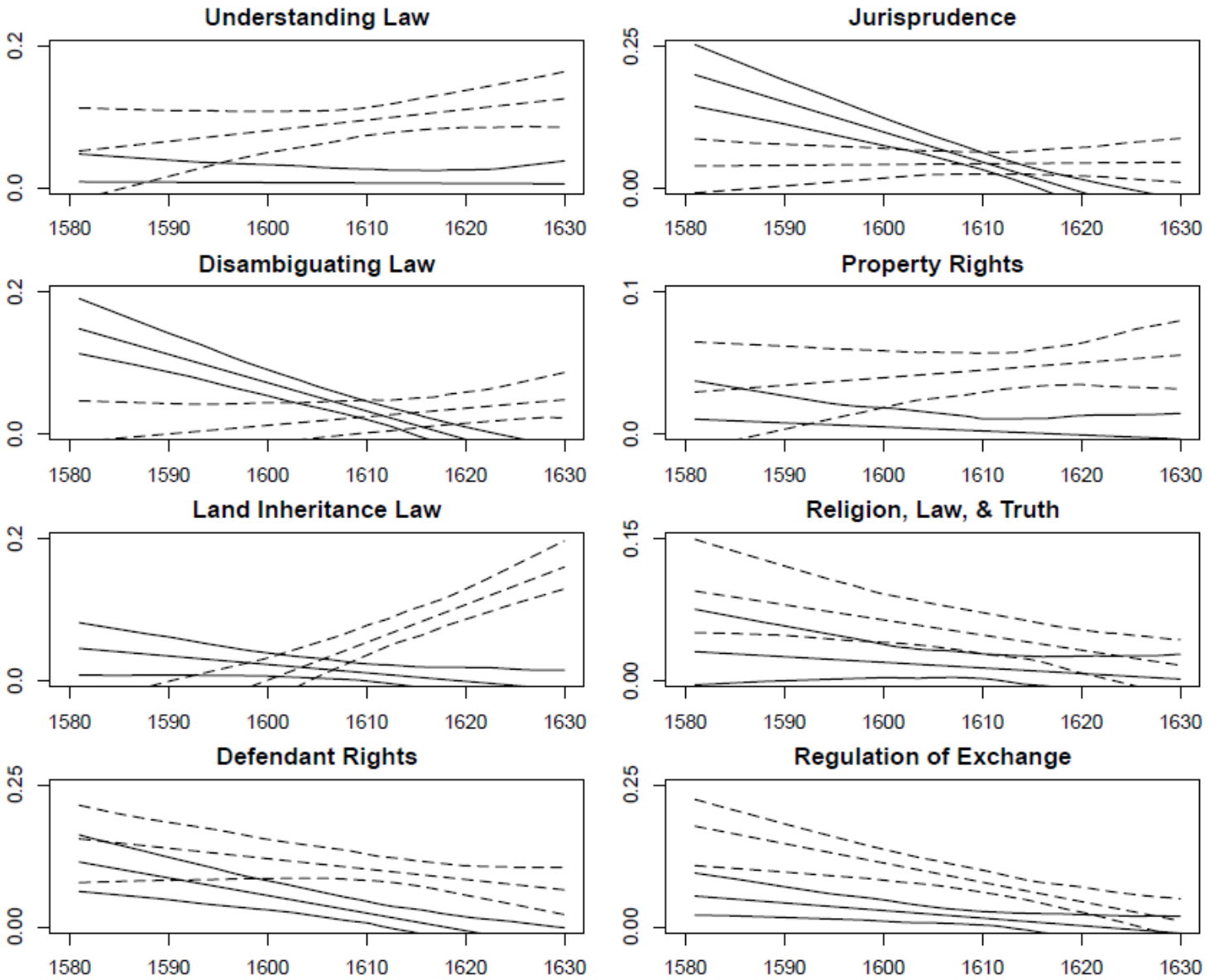
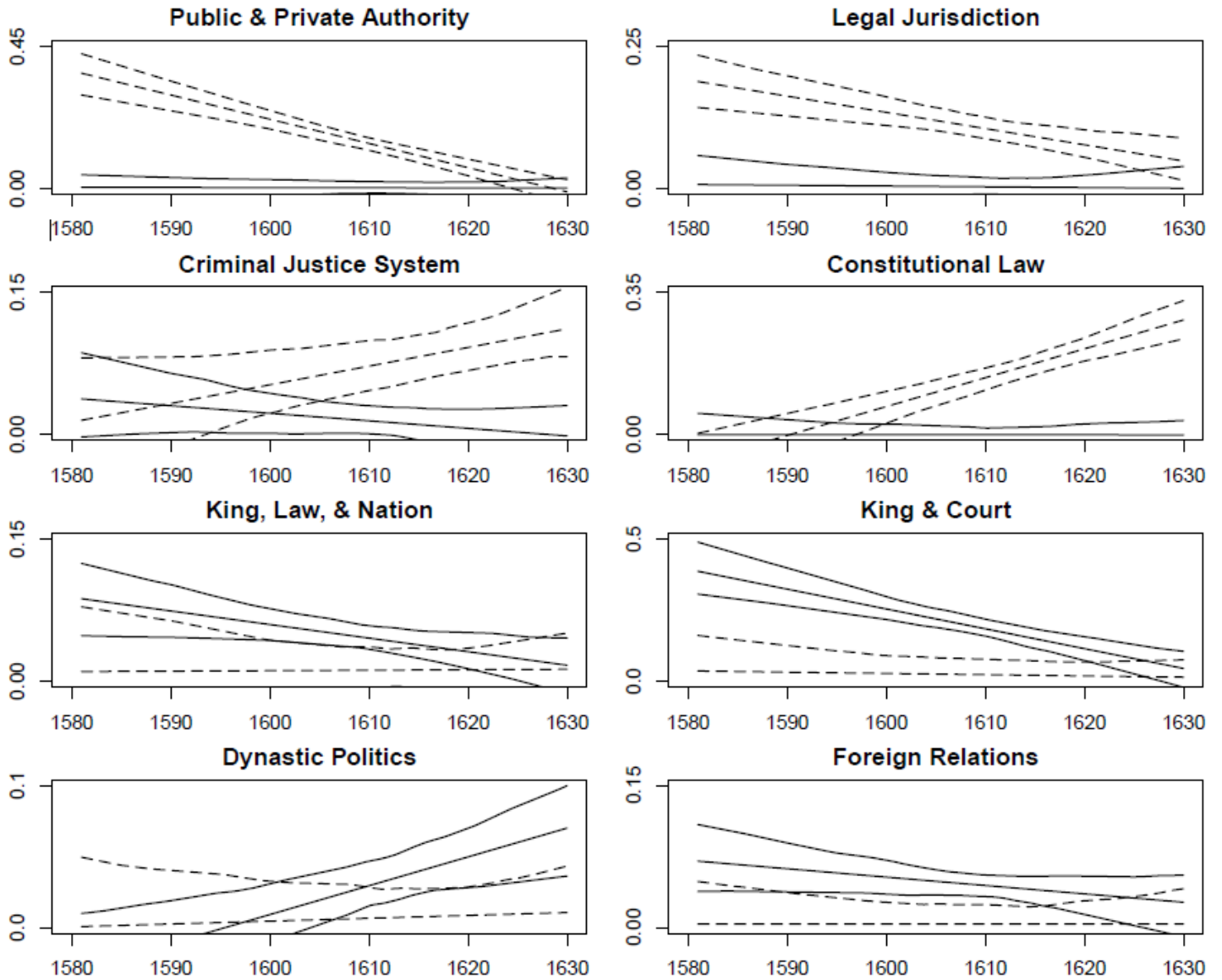


Figure 11, Panel A:
Variations Over Time in the Use of the 25 Topics by Bacon and Coke



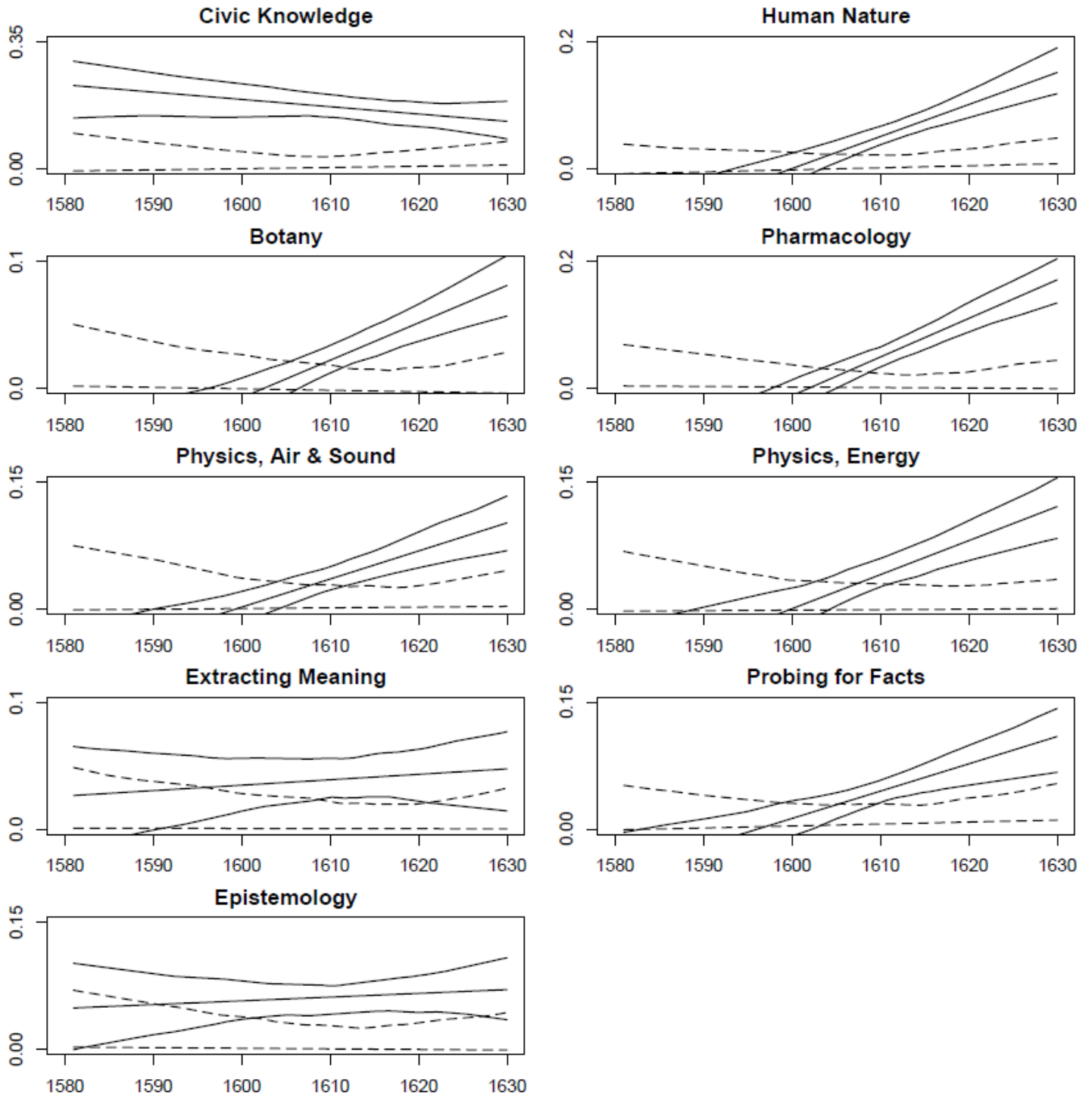
Note: Estimated mean values for topic usage and 95% confidence intervals. Solid lines are for Bacon, dashed lines for Coke.

Figure 11, Panel B:
Variations Over Time in the Use of the 25 Topics by Bacon and Coke



Note: Estimated mean values for topic usage and 95% confidence intervals. Solid lines are for Bacon, dashed lines for Coke.

Figure 11, Panel C:
Variations Over Time in the Use of the 25 Topics by Bacon and Coke



Note: Estimated mean values for topic usage and 95% confidence intervals. Solid lines are for Bacon, dashed lines for Coke.

Electronic Supplementary Material

to

Characterizing a Legal-Intellectual Culture: Bacon, Coke, and Seventeenth-Century England

(Cliometrica)

Peter Grajzl[†]

Peter Murrell[‡]

[†] (Corresponding author) Department of Economics, The Williams School of Commerce, Economics, and Politics, Washington and Lee University, Lexington, VA 24450, USA and CESifo, Munich, Germany. Email: grajzlp@wlu.edu

[‡] Department of Economics, University of Maryland, College Park, MD 20742, USA. Email: pmurrell@umd.edu

Appendix A: The Structural Topic Model (STM)

A1. The Data Generating Process

The generative process of the STM (see Roberts et al. 2014, 2016a, 2016b) has the following general structure.¹ The corpus consists of D documents indexed by d . The data generating process views a document, d , as a collection of N_d initially empty positions, each of which is to be filled with a word. The first step in the process of filling a position involves a choice of a topic, k , from a fixed number of available topics, K . That step uses a K -dimensional vector of the parameters of a distribution that generates one of the topics $k \in \{1, \dots, K\}$ for each position in the document. This is the topic-prevalence vector, which lists the probabilities that each of the K topics will be assigned to an empty position. Then, given the chosen topic, k , a word is chosen from a corpus-level vocabulary, the elements of which are indexed by $v \in \{1, \dots, V\}$. This choice is determined by a topic-specific V -dimensional vector specifying the probabilities that each element of the vocabulary will be chosen to fill an empty position given that topic.²

Early versions of topic models viewed the topic-prevalence vector as drawn from the same distribution for all documents. In STM, however, the topic-prevalence vector is a function of variables characterizing documents, referred to as metadata (to be distinguished from the text that constitutes the corpus). Incorporation of document-level variables improves the identification of topics and allows the researcher to estimate the relationship between topic prevalence and document characteristics (Roberts et al. 2014).

We now describe the process of filling a word-position in a document in more formal terms. The generation of a document-specific topic-prevalence vector uses the metadata. Let the metadata be given by a matrix \mathbf{X} , each row (denoted \mathbf{x}_d) listing the values of all metadata covariates for document d . Then, the topic-prevalence vector for document d , $\boldsymbol{\theta}_d$, is drawn from a logistic-normal distribution with parameters that are a function of the covariate values:

$$\boldsymbol{\theta}_d \sim \text{LogisticNormal}(\boldsymbol{\Gamma}\mathbf{x}_d', \boldsymbol{\Sigma}). \quad (\text{A1})$$

$\boldsymbol{\Gamma}$ is a matrix of coefficients relating covariate values to mean topic-prevalence. $\boldsymbol{\Sigma}$ is a general variance-covariance matrix that allows for the possibility of non-zero correlations across documents in topic-prevalence vectors. Given a topic-prevalence vector, one specific topic, z_{dn} , is associated with empty position $n \in \{1, \dots, N_d\}$ in document d through the following process:

$$z_{dn} \sim \text{Multinomial}(\boldsymbol{\theta}_d), \quad (\text{A2})$$

where the k^{th} element of z_{dn} is unity and all other elements are zero when topic k is chosen.

The choice of a specific vocabulary word, v , is modeled as a function of two parameters: one capturing the baseline importance of that word across all documents, m_v , and one capturing

¹ Our exposition of the STM here closely follows the exposition in Grajzl and Murrell (2019).

² These probabilities can also be allowed to vary in a systematic way across documents. We do not use this feature of STM.

the importance of the word given the topic k , κ_{kv} . Transforming the sum of these two parameters into probabilities for use in a multinomial distribution via a logistic transformation, one obtains:

$$\beta_{dkv}|z_{dn} \propto \exp(m_v + \kappa_{kv}), \quad (\text{A3})$$

where β_{dkv} is the probability of choosing vocabulary word v to fill a position in document d given topic k . Then a specific word, w_{dn} , is chosen from the overall corpus vocabulary to fill position n in document d , as follows:

$$w_{dn} \sim \text{Multinomial}(\beta_{dk1}, \dots, \beta_{dkV}). \quad (\text{A4})$$

A2. Estimation and Output

The data to be used in estimation are the metadata matrix, \mathbf{X} , and all words in all documents, w_{dn} , where $n = 1, \dots, N_d$ and $d = 1, \dots, D$. Given these data, one estimates Γ , θ_d , m_v , κ_{kv} , and β_{dkv} by maximizing the posterior likelihood that the observed data were generated via the above data-generating process. Computationally, the maximization problem is solved using an iterative approximation-based, variational expectation-maximization algorithm available in R's `stm` package (Roberts et al. 2016a, 2016b). To address problems due to non-convexity, we rely on the spectral initialization approach advocated by Roberts et al. (2016b, 2016c).

The key output of the K -topic STM are, first and foremost, the K topics, conceptualized as probability distributions over corpus vocabulary and reflecting the estimates of β_{dkv} for each $k \in \{1, \dots, K\}$. STM output also includes the estimated document-level topical prevalence, reflecting the estimates of θ_d for each document $d \in \{1, \dots, D\}$. We make use of both sets of these estimates in Section 4 when identifying and naming the topics. In Section 5, we use the estimates of β_{dkv} to examine overlaps in the usage of words within specified topic pairs. Finally, the output of STM produces a matrix of document-level topic correlations, reflecting the estimate of $\theta_d' \theta_d$. We investigate topic correlations in Section 5 also.

A3. Analysis of the Effect of Metadata Covariates on Topical Prevalence

The distinguishing feature of the STM vis-à-vis earlier topic models is the possibility of examining the effect of metadata covariates on document-level topical prevalence. Analytically, we proceed as follows. Upon estimating the K -topic STM, we investigate the relationship between document-level topical prevalence, reflecting the estimates of θ_d , and metadata covariates, stacked in matrix \mathbf{X} , thereby obtaining estimates of the pertinent column of Γ . To estimate the resulting regressions featuring the STM-estimated document-level proportions devoted to a topic as the dependent variable and metadata covariates as explanatory variables, we make use of the `estimateEffect` function available in R's `stm` package (see Roberts et al. 2016b, 2018). To summarize the results, we plot mean differences in estimated topic proportions for different values of pertinent covariates, sometimes further conditioning the resulting differences on additional covariate values. We display the point estimates and the corresponding 95-percent confidence intervals. We report the results of this analysis in Section 6.

Appendix B: Initial Processing of the Corpus

Prior to importation into R, the corpus of Bacon's and Coke's works was processed in a number of stages. All operations were carried out using Python programs written by the authors.

The orthography of late 16th and early 17th century English was converted into standard modern orthography using the database available with MorphAdorner (2013). This database contains translations between spelling variants and standardized spelling for words common in 16th and 17th century English. In the process of constructing our corpus many corrections and additions were made to this database, resulting in translations being available for 361,161 spelling variants.

Given standardized spelling, word inflections were removed by converting all words to their lexical roots, again using databases available with MorphAdorner. Corrections and additions were made to these databases, resulting in the availability of translations between standard spellings and lexical roots for 468,040 words.

The next steps entailed the use of a comprehensive list of standard English spellings, names, abbreviations, and acronyms. The construction of this list initially relied on databases from Moby Word Lists (2002). With additions made during the processing of the corpus, the resultant word list contained 385,098 standard English words. If, at this stage, a word in the corpus matched a word in the English dictionary, then it was left in the corpus and omitted from all subsequent steps of this initial processing.

Bacon and Coke occasionally used Latin. Thus, a very large proportion of the words that did not match standard English ones were Latin words. The variant of Latin that was common during Bacon's and Coke's time involved the use of many, often idiosyncratic, accents on letters that do not appear in classical Latin. Therefore, accents were removed from all words that did not appear in the English dictionary, replacing accented letters with their nearest ASCII equivalent.

Those words remaining in the corpus that were not in the English word list were then matched against a word list of 1,034,156 Latin spellings (including a large number of inflections of Latin root words). This Latin spelling list was obtained from Whitaker's Words (2006). Any word in the Latin spelling list, but not in the English word list, was stemmed using a version of the Schinke Latin stemming algorithm (Schinke et al. 1996). Through stemming, the variant forms of a word were transformed into standard forms giving the inflectionless form of a word.

Once the Latin words in the corpus were in their standard forms, they were translated into English. The Latin-English dictionary relied initially on the database available at Whitaker's Words (2006). Then, if a word in the corpus was simultaneously not in the English word list, within the Latin word list, and not in the Latin words in this initial Latin-English dictionary, a program searched for the word in the online Lewis and Short Latin Dictionary (Perseus n.d.). The Latin words and their online translations were then added to the original Latin-English dictionary. The resultant dictionary database contained 59,458 Latin to English translations. The ensuing corpus was imported into R for further processing, as described in Section 3.2.

Prior to importation into R, we also examined the corpus for the overall quality of the initial processing as outlined above. To this end, we computationally examined each of the 432 processed corpus documents, additionally purged of punctuation, for the presence of non-numeric strings that could not be matched to any word in the English dictionary (see above). Any 'leftover entries' of this type are either non-translated foreign (especially Latin) words or character strings indicative of OCR problems. A comparison of every non-numeric character string in the corpus to the list of words in the English dictionary indicated a negligible presence of such entries: the average percent of leftover entries per document, weighted by the number of all character strings in a document, was equal to 1.05 percent. This is evidence, first, that our initial processing was successful at standardizing the variants of English orthography across the corpus and, second, that any OCR issues should not be a concern for our analysis.

Appendix C: Coding the Year of Completed Work

The coding of the year of completion of corpus documents proceeded as follows. For each work associated with documents in the corpus, we first attempted to identify the exact year of publication. As long as the publication year of a specific work was known and publication occurred during the author's lifetime, we took the year of publication as the year of the work's completion. Based on this criterion, we were readily able to ascertain the year of completion for 215 out of 282 Bacon documents and for 81 of 150 Coke documents. In the case of some case reports, speeches, and letters, we were often able to further refine the estimated year of completion by identifying the year of the specific subject matter (e.g. political issue, or personal correspondence) addressed by the author in his writing.

In the case of 67 Bacon documents (24 percent of all Bacon documents) and 68 Coke documents (45 percent of all Coke documents), the pertinent works were published posthumously or the year of completion is simply not known. For each of those 136 works, we carefully examined all available historical sources to identify a time interval during which a specific work was completed with a very high probability. We then placed a value on the year of completion of such works by drawing a random integer from the time interval.

For the 67 Bacon documents (46 essays, ten tomes, seven letters, and four speeches) for which the exact year of completion could not be ascertained, the average width of the time interval used for estimating the year of completion equals 6.4 years, with standard deviation equal to 7 years. Thus, even if assuming that the true year of completion for each document is distributed uniformly over the estimated time interval, the mean absolute value of the error would be smaller than five years.

All of the 69 Coke documents for which the exact year of completion could not be readily ascertained are from his *Institutes*, a comprehensive series of Coke's legal treatises widely recognized as constituting a foundational work for the common law. Scholars of Coke's opus have not been able to determine when exactly Coke completed his *Institutes*. Sheppard (2003b: 574) states that Coke published the first part of the *Institutes* as early as 1608, while "portions of [the rest of the *Institutes*] seem to have been completed in the later 1610s and 1620s, after [Coke's] dismissal as chief justice [in 1616]". Through personal correspondence with Professor Sheppard (June 2, 2017), we established that the reference to year 1608 is in fact a typo, and that the correct year of publication for the first part of the *Institutes* is 1628. This is fully consistent with Baker (1972: 78-179) who comments that the *Institutes* were "written for the most part before 1628". In our private correspondence with the pre-eminent legal historian of this period, Sir John Baker (May 9, 2017), he noted that "it fair to assume that [Coke] would not have had the time to write [the *Institutes*] before 1616 or after 1627...I believe they were all finished in the 1620s, though only the first part made it into print in his lifetime". Accordingly, we chose as the start of the interval used for estimation of the year of completion of the *Institutes* 1616 and as the end of the interval 1627, a width of 12 years.

Appendix D: Investigating a 100-Topic STM

D1. The Goal and Central Findings

This appendix reports findings from the robustness check mentioned in Section 4.8. As made clear in Section 4, our 25-topic STM features no emphasis on

- (1) the utilitarian promise of science;
- (2) the centralized organization of the scientific quest;
- (3) a laissez-faire stance toward commerce.

Ideas (1) and (2) have been often attributed to Bacon. (3) has been ascribed to Coke.

To examine whether an emphasis on ideas (1)-(3) emerges in a model featuring a notably larger number of topics, we estimated first a 50-topic and then a 100-topic STM. Increasing the number of topics from 25 to 50 and 100 decreases model-fit as measured by held-out likelihood (see Figure D1). In contrast, increasing the number of topics initially increases and eventually decreases the size of residuals, as one would perhaps expect to be the case for an over-fitted model. Similarly, while the average coherence of the estimated topics decreases, the average topic exclusivity actually increases as the number of topics grows from 25 to 100. The relative increase in the average exclusivity measure (4.2 percent), however, is much smaller than the relative decrease in the average coherence measure (39.5 percent). Based on these diagnostic measures, as well as our own inspection of the cohesiveness and exclusiveness of the estimated topics in the 50-topic and 100-topic models, both the 50-topic and 100-topic STMs are clearly inferior to our chosen 25-topic STM. Nonetheless, the estimation of STM models featuring 50 or 100 topics allows us to explore the corpus at a much greater level of topic granularity. Thus, although we do not believe it is methodologically advisable to base important insights on a 50-topic or a 100-topic model estimated on a corpus of 432 documents featuring about 1.3 million words (see Section 3.2), we do believe that an examination of such models provides a suitable robustness test for conclusions that highlight the absence of emphases on specific subjects, where those conclusions have been previously derived from a superior model involving a smaller number of topics.

The essence of the robustness test is an examination of whether there are any topics that could plausibly be designated as capturing ideas (1)-(3). We first investigated the higher probability and FREX words of the 50-topic STM. We found no indication at all of an emphasis on ideas (1)-(3). We therefore proceeded with a scrutiny of a 100-topic STM. Panel D2 lists the highest probability and FREX words for all 100 topics in the resulting STM. A close inspection of the resultant word lists reveals that in the overwhelming majority of topics, there is yet again no trace whatsoever of ideas reflecting (1)-(3). Importantly, we found no word lists that were readily suggestive of topics corresponding to ideas (1)-(3). This in itself is evidence that the 100-topic model provides confirmation of the conclusions reached in the paper: that (1)-(3) are not emphases to be found in the corpus. However, to be absolutely sure we investigated further, examining in more detail any topic for which the corresponding word lists contained any hint at all of these ideas. We identified six such topics: 3, 16, 46, 49, 79, and 81.

Figure D3 shows the relative importance of these six topics in the corpus based on expected topic proportions. The figure demonstrates that only topic 79 is ranked in the top quarter of all topics, and even topic 79 is ranked only 23rd. That is, none of these six topics reflects a major emphasis in the corpus.

For each of the six topics, we then identified the top twenty documents that feature the given topic most prominently. Following the process for examining the topical content described in Section 4, we then carefully examined each of those documents. Importantly, due to the overall diminished coherence of the estimated topics in the 100-topic STM relative to our main 25-topic STM (see Figure D1), identifying the precise content of a topic in the 100-topic STM was often challenging, if not infeasible. This in itself is evidence that estimating a 100-topic STM is not advisable. Nevertheless, we were able to ascertain beyond any reasonable doubt whether a given topic features an emphasis on any of the ideas (1)-(3).

As we clarify in the discussion below, we find no evidence of the existence of topics centered on any of the ideas (1)-(3). Of course, conventional textual analysis based on a close reading of some documents does reveal a limited presence of notions that could be interpreted as consistent with ideas (1), (2), or (3). This finding is hardly surprising given that scholars adopting conventional textual analysis have attributed specific notions to Bacon or Coke precisely based on their own (or other scholars') close reading and interpretation of selected passages written by either of the two authors. However, based on our own careful reading (reported below), we must emphasize the following three points. First, these notions, highlighted by scholars on the basis of conventional textual analysis, are not systematically stated across the set of documents that feature a given topic prominently. Second, such notions are confined to particular passages contained in only a handful of documents. And third, these notions are often directly countered in the very same documents or in other documents that feature the same topic prominently. That is, the combination of careful scrutiny of topic estimates of the 100-topic STM and close reading of the documents featuring these topics prominently lead to the conclusion that ideas (1)-(3) cannot be considered as constituting any kind of emphasis within the corpus. In other words, our findings reported in Section 4.8 are fully robust to analysis of the corpus at a notably higher level of granularity than that implied by a 25-topic STM.

D2. Detailed Analysis of Estimated Topics from the 100-Topic STM

Topic 3 is about learning and science (natural and humanistic) in general, hence keywords include 'learn', 'man', 'histori', 'natur'. All top documents featuring this topic prominently are contributed by Bacon. The documents include portions of Bacon's most celebrated methodological works such as *The Advancement of Learning* and *Novum Organum*, but also *New Atlantis*, rather heterogeneous private letters (where Bacon often introduces the recipient of his letter to his new work on learning), political speeches (e.g. upon a calling of Sir William Jones to serve as a Chief Justice of Ireland), a historical treatise on the history of Great Britain, an opinion on the disposition of a charity (where Bacon argues that funds should be used as "a beneficence towards the relief of

the poor"; hence 'benefic' and 'beggar' among keywords), and even an essay on the imminent union of Scotland and England.

A close reading of these documents reveals a very modest and inconsistent presence of ideas that later followers could have interpreted as implying that Bacon, among others, discussed the organization of science (understood very broadly). For example, in a passage in *The Advancement of Learning* Bacon notes: "...amongst so many great foundations of colleges in Europe, I find strange that they are all dedicated to professions, and none left free to arts and sciences at large". In the same document, Bacon also suggests that salaries of lecturers of "arts" and "professions" are too low and argues in favor of scientific cooperation in Europe. Yet these same ideas are entirely absent from other top ranked documents. In a similar vein, in another top ranked document (from *New Atlantis*), Bacon describes Salomon's House, often interpreted by Bacon's later followers as embodying his vision of a modern research facility. Yet Salomon's House is not featured prominently across the top ranked documents for topic 3. In fact, it is not even found among the keywords most associated with this topic. Only three of the top-20 documents for topic 3 contain any reference to the organization of science, meaning that in this topic there is no evidence of Bacon's emphasis on such organization.

In selected passages of a few top ranked documents, Bacon also articulates his view on the role and benefits of science. None of these documents, however, entails a perspective that is readily interpreted as utilitarian. In fact, in a passage from *The Advancement of Learning*, rather than adopting a narrowly utilitarian lens, Bacon lists a large number of different benefits from learning. It benefits individuals: "the beneficence of learning...give fortune to particular persons" and it provides inherent pleasure: "Again, for the pleasure and delight of knowledge and learning, it far surpasseth all other in nature....But of knowledge there is no satiety, but satisfaction and appetite are perpetually interchangeable; and, therefore, appeareth to be good in itself simply, without fallacy or accident...But it is a pleasure incomparable, for the mind of man to be settled, landed, and fortified in the certainty of truth." In the same document Bacon notes that knowledge provides a stepping stone towards the highest level of understanding, which is to be valued in itself: "But yet the commandment of knowledge is yet higher than the commandment over the will; for it is a commandment over the reason, belief, and understanding of man, which is the highest part of the mind, and giveth law to the will itself."

In another highly ranked document, a parliamentary speech, Bacon hopes to draw "his majesty into a serious consideration of the mineral treasures of his own territories, and the practical discoveries of them by way of [his] philosophical theory". Bacon views "recovering all such drowned mineral work" as "the most probable means to relieve all the poor thereof without any other stock or benevolence, than that which divine bounty should confer on their own industries and honest labours". Thus, Bacon argues, "by this unchangeable way...have I proposed to erect the academical fabric of this island's Salomon's House, modelled in my New Atlantis. And I can hope...that my midnight studies to make our countries flourish and outvy European neighbours in mysterious and beneficent arts, have not so ingratfully affected the whole intellects, that you will

delay or resist his majesty's desires, and my humble petition in this benevolent, yea, magnificent affair." Note, however, that Bacon's argument here is focused on the specific pursuit of mining as a source of wealth and certainly not on the scientific quest per se.

In sum, while conventional textual analysis of documents featuring this topic reveals traces of ideas (1) and (2) in selected passages of a small subset of these documents, the topic itself does not entail an emphasis on these ideas and certainly could not be labelled as reflecting such ideas. Four of the top ranked documents could be labeled as referring to the uses of science. But those documents also contain references to uses that are distinctly non-utilitarian in nature.

Topic 16 reflects a variety of legal issues. All top documents are contributed by Coke. There is some emphasis on the broad subject of royal prerogative and individual rights, which were both clear points of emphasis in Coke's career and are also captured by our 25-topic STM (e.g. via the topics Constitutional Law; Legal Jurisdiction; Public & Private Authority). A close reading of the top-20 documents reveals that only five documents bear any direct relevance to commerce and trade.

In the top ranked among these five documents, *The Case of Monopolies*, Coke opposes the creation of a monopoly on the sale of playing cards (hence 'card', 'monopoli' among keywords), concluding his commentary with: "And note, Reader, and well observe the glorious preamble and pretense of this odious monopoly... Monopolies are things against the Lawes of this Realm, and therefore expressly commands that No Suitor presume to move him to grant any of them". Similarly, in *Petition of Grievances, Privileges of Parliament and Impeachments*, which features a variety of legal issues involving the exercise of the king's powers (Coke among others notes "I had rather live under severe laws than under any man's discretion"), Coke again argues against monopolies. Some have interpreted such statements as implying a devotion to laissez-faire. But in the document ranked third (*The Chamberlain of London's Case*), Coke reported approvingly on a case that allowed the City of London to maintain the rule "that if any citizen, freeman, or stranger within the said city, put any broad cloth to sale within the City of London before it be brought to Blackwell-Hall to be viewed and searched, so that it may appear to be saleable", then the cloth would be "forfeited". This ruling was based, at least partially, on an old case that confirmed the king's power to regulate trade by requiring incoming ships to unload only at particular locations, which Coke mentions as an important precedent.

The third-ranked document is *The Case of the Tailors of Habits &c. of Ipswich*. In this case, Coke objected when a guild, the Merchant Taylors, required that a fully apprenticed craftsman must obtain the guild's permission in order to exercise a trade. Coke argued that such restraints were "against the freedom and liberty of the subject, and are a means of extortion...or of oppression of young tradesmen, by the old and rich of the same trade". Some scholars, most notably Heckscher (1935), have interpreted Coke's position in this case as supportive of laissez-faire. This interpretation, however, has been contested. Malament (1967: 1321-1322), in particular, notes that, in arguing the case, Coke in fact did not object to the inherently anti-laissez-faire Statute

of the Artificers but merely "the possibility that a man qualified to work under the act might be prevented from doing so" (ibid.).

Thus, the estimated topic itself is not a topic on laissez-faire and thus not about idea (3). Only five of the top-20 documents are in any way about commerce, and therefore conceivably relevant to laissez-faire. Moreover, in those five documents there are several statements that refer to the balancing of private and public interests, with the king free to intervene where the latter could be pursued: "Upon all which and divers records which we had seen, it appeared to us, that the King cannot at his pleasure put any imposition upon any merchandize to be imported into this kingdom, or exported, unlesse it be for advancement of trade and traffic, which is the life of every island, pro bono publico."

Much like Topic 16, Topic 46 is about a broad range of issues without an evident central thread, though with some emphasis on ideas about the king's powers. Top documents are contributed by both Coke and (to a lesser extent) Bacon. Among the top documents by Coke are again *Petition of Grievances*, *Privileges of Parliament and Impeachments* and *The Case of the Tailors of Habits &c. of Ipswich* (hence 'monopoli' among keywords). The remaining Coke documents revolve around issues as diverse as decision-making in common-law courts; legal jurisdiction; the limits of royal revenue; and whether a king can personally decide legal cases. The top ranked case is a charge by Bacon, the attorney general, against a defendant who allegedly attempted to poison the king (hence 'impoison' and 'injunct' among key words). In sum, this topic decidedly does not emphasize any of ideas (1)-(3).

Topic 49 revolves around experimentation and the methodology of conducting science ('experiment', 'discoveri', 'scienc', 'natur' are among keywords). All top documents are from Bacon's works. Among the top ten documents are portions of *Novum Organum* and *The Natural and Experimental History of Winds*. Keywords most associated with this topic are not evidently indicative of ideas (1) or (2).

A close reading of the top ranked documents reveals passages where Bacon does hint at the utilitarian promise of science. In a passage from *Novum Organum*, for example, Bacon notes: "For, though we do not deny that by transferring the experiments from one art to another..., many new experiments may be discovered tending to benefit society and mankind". Yet in another portion of the same work that features this topic prominently, Bacon adopts a distinctly non-utilitarian stance toward science: "Truth, therefore, and utility, are here perfectly identical, and the effects are of more value as pledges of truth than from the benefit they confer on men." A further passage suggests yet another reason for applying the scientific method, the conquest of nature: "We must next, however, proceed to the supports and corrections of induction, and thence to concretes, the latent process, and latent conformations...order that, like good and faithful guardians, we may yield up their fortune to mankind upon the emancipation and majority of their understanding; from which must necessarily follow an improvement of their estate, and an increase of their power over nature". For Bacon, the conquest of nature, or the power over it, had much more connection to religious dogma than to utilitarian ideas, since the purpose of the conquest of

nature was to restore man to his prelapsarian state. In sum, the utilitarian promise of science certainly appears in the documents that feature this topic prominently, but it is not the predominant perspective emphasized by Bacon. This topic therefore does not center on ideas (1) or (2). Given the mix of subjects in the top documents, it is a medley of ideas about doing science in general.

Topic 79 is featured prominently in Bacon's essays on diverse subjects, including marriage, parents and children, and imperial dynasties (hence 'famili', 'child' among keywords). Only one among the top documents is Coke's (a preface to one of the Institutes, where Coke talks about several legal issues, including marriage; hence 'wife' among keywords).

Two of the top documents (ranked first and 20th) are portions of the *New Atlantis*; hence 'bensalem' and 'salomon' among keywords. A close reading of these two documents reveals that, in the document ranked first, Bacon mentions Salomon's House as "the noblest foundation (as we think) that ever was upon the earth; and the lantern of this kingdom". Bacon, however, does not in any way elaborate on the notion of organization of science. Similarly, in the passages of the document ranked 20th, Bacon describes Salomon's House, but never explicitly connects it to any ideas about organization of the scientific quest. Neither idea (1) nor (2) are thus central to, or otherwise emphasized in the context of this topic. With the top documents in the topic featuring such themes as love, adversity, the domains of affairs to which single or married men are best matched, and the lives of emperors, this topic does not even emphasize science.

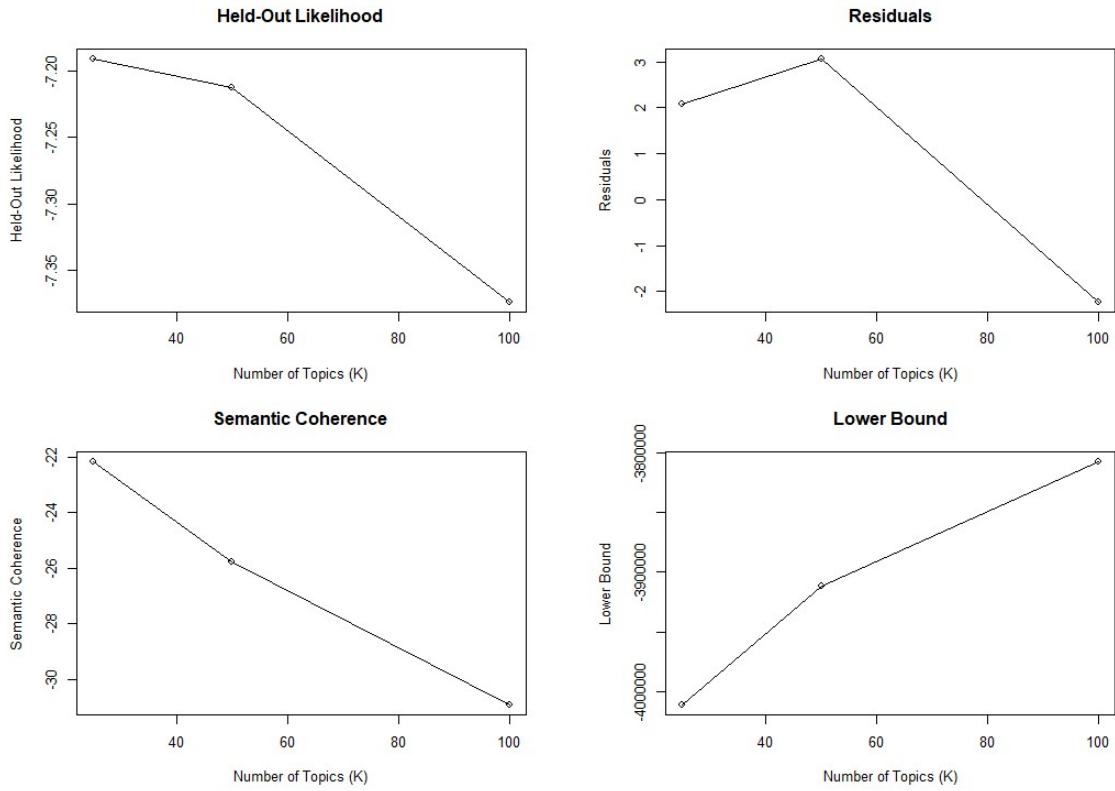
Topic 81, the final topic we investigated closely, is again dominated by Bacon. The topic focuses on elements of science and scientific reasoning (keywords include 'system', 'scienc', 'idol', 'metaphys', 'abstract', 'notion', 'method', 'axiom') and in part on the promise of science for mankind (hence 'human', 'mankind', 'power' among the keywords). The top documents are all portions of *Novum Organum*. A close reading of these documents exposes passages where Bacon does indicate the possibility of a utilitarian promise of science, stating e.g. that "the benefits derived from inventions may extend to mankind in general". However, in the very same document alternative perspectives are equally prominent. Bacon, for example, endorses the scientific quest with the express purpose of conquering nature: "...if one were to endeavor to renew and enlarge the power and empire of mankind in general over the universe, such ambition (if it may be so termed) is both more sound and more noble than [to enlarge one's power in one's country or to enlarge the power and empire of one's country over mankind]. Now the empire of man over things is founded on the arts and sciences alone, for nature is only to be commanded by obeying her". Such pursuit, according to Bacon, is in fact in full congruence with God's will (with its goal of attainment of the prelapsarian state): "Only let mankind regain their rights over nature, assigned to them by the gift of God, and obtain that power, whose exercise will be governed by right reason and true religion". That is, a utilitarian perspective on the scientific quest (idea (1)) is by no means central to this topic.

At the same time, in a passage of a highly ranked document from *Novum Organum* that features this topic prominently, Bacon argues in favor of decentralization of the scientific quest: "If, therefore, there be any one who is more disposed and prepared for mechanical art, and

ingenious in discovering effects, than in the mere management of experiment, we allow him to employ his industry in gathering many of the fruits of our history and tables in this way, and applying them to effects, receiving them as interest till he can obtain the principal." This is a comment distinctly not about centralized organization of science (idea (2)). In sum, topic 81 focuses neither on idea (1) nor on idea (2). This conclusion is reinforced by the observation that documents featuring this topic refer to such subjects as human psychology, "the happy union of the Kingdomes of England, and Scotland", Tudor history, and "collecting the rules and grounds, dispersed throughout the body of the [English] lawes".

Figure D1: Diagnostics measures for 25-, 50-, and 100-topic STM

Diagnostic Values by Number of Topics



Panel D2: Topics and keywords for 100-topic STM

Topic 1

Highest Prob: parliament, king, common, lord, writ, henri, edward, hous, make, hold, act, baron, say, call, shall, time, one, may, royal, year, richard, men, knight, cap, can, law, england, power, elect, place

FREX: elig, parliament, holiday, prorog, roll, prorogu, summon, baneret, burgess, print, elect, speaker, chevali, banner, gilbert, judicator, westminst, legitim, baroni, baron, quad, monday, men, concilium, knight, citizen, adjourn, expensur, parsonag, testa

Topic 2

Highest Prob: sound, air, make, water, will, string, one, give, great, upon, voic, caus, bodi, tone, motion, nois, note, may, strike, hear, percuss, dram, part, wind, shall, weigh, man, two, grain, equal

FREX: tone, dram, percuss, string, articul, nois, concav, tripl, pipe, sound, diapason, grain, bell, interior, weigh, bullet, loud, barrel, trunk, lute, music, voic, strike, overweigh, exterior, balanc, air, outsid, sauncer, base

Topic 3

Highest Prob: learn, man, histori, natur, may, great, time, work, will, part, make, use, knowledg, shall, see, one, scienc, therefor, can, well, state, like, yet, good, upon, thing, kind, find, likewis, place

FREX: inventor, histori, learn, profici, defici, narrat, scienc, illumin, lectur, extant, contempl, poesi, repres, antiqu, virgil, event, benefic, magnific, stori, beggar, degener, immort, impostur, journal, socrat, convert, astronomi, memor, imag, imit

Topic 4

Highest Prob: custom, prescript, time, man, manor, say, prescrib, may, make, lord, shall, statut, titl, writ, commonlaw, can, law, will, use, copyhold, right, within, common, take, year, therefor, limit, good, one, general

FREX: prescript, prescrib, copyhold, cautious, custom, manor, usag, forest, ancestor, estatut, perci, titl, eir, interrupt, widowhood, estov, limit, estrai, alleg, agn, memori, deodand, impound, merton, leet, littleton, pai, mill, engag, acquit

Topic 5

Highest Prob: treason, king, act, statut, high, case, say, parliament, make, shall, law, word, lord, within, man, justic, declar, realm, take, petit, death, great, feloni, compass, offenc, judgement, money, hold, mispris, indict

FREX: treason, mispris, overt, compass, counterfeit, roy, petit, signior, specifi, tiel, high, regnant, class, conspir, companion, preambl, ceo, maria, purvien, edmond, levi, act, feloni, imagin, facto, coin, conspiraci, pur, declar, offenc

Topic 6

Highest Prob: son, land, shall, heir, blood, father, die, brother, old, issu, seiz, case, young, enter, inherit, law, possess, feesimpl, ventur, sister, make, entri, man, daughter, without, descend, descent, whole, half, condit

FREX: son, ventur, brother, sister, blood, actual, entri, father, daughter, feesimpl, puisn, young, issu, elder, descent, old, die, tail, parcel, seiz, corona, section, half, enter, descend, inherit, attaint, possess, heir, reenter

Topic 7

Highest Prob: good, swan, may, properti, shall, one, executor, king, man, will, take, debt, say, owner, mark, make, pay, administr, time, case, natur, sale, legaci, market, within, can, henri, use, shop, law

FREX: swan, legaci, cygnet, properti, executor, shop, market, fera, owner, steal, administr, stray, mark, intest, game, estray, dorset, executorship, cupboard, goldsmith, privilegi, partridg, sale, swim, debt, workman, gile, thame, scriven, thief

Topic 8

Highest Prob: law, case, book, great, publish, man, say, may, mani, good, will, know, learn, year, king, court, can, time, ancient, upon, justic, shall, read, write, find, make, call, part, true, common

FREX: libel, publish, client, reader, apprentic, sergeant, chariti, sage, commentari, pieti, verita, professor, invect, prisot, prefac, inn, antiqu, scribe, treatis, versus, magistr, compos, abridg, student, incarn, mawd, nova, justinian, councillor, vener

Topic 9

Highest Prob: lordship, majesti, will, may, good, shall, lord, letter, humbl, know, upon, servic, can, self, time, think, place, toward, make, pleas, hope, now, great, much, god, man, year, mean, king, well

FREX: lordship, humbl, solicitor, hope, salisbury, glad, attorney, majesti, pleas, pray, letter, servic, thank, toward, gracious, assur, commend, signif, unworthi, accept, happi, recommend, gift, favour, oblat, presum, sovereign, wish, chancellor, mistress

Topic 10

Highest Prob: majesti, law, friend, unto, great, shall, judg, young, time, upon, may, purpos, offic, place, peac, self, good, yet, must, govern, though, honour, man, also, present, much, roman, judgement, mani, make

FREX: friend, kinsfolk, friendship, acquaint, murrey, toga, roman, senat, dedic, sacr, young, clemenc, loath, amend, upright, gracious, humbl, purpos, bestow, acta, dear, matthew, cluster, youth, departur, choir, banquet, season, majesti, distast

Topic 11

Highest Prob: say, land, estat, purchas, feesimpl, hold, name, shall, make, king, man, grant, can, leas, take, custom, fee, lord, statut, without, may, year, heir, will, case, person, good, tenant, book, law

FREX: copyhold, feesimpl, baptism, simplex, heydon, leas, purchas, ware, heredita, waiv, forest, surnam, tenement, capac, formedon, nullo, intrus, dirt, capabl, dean, fee, bastard, disagre, warden, perdur, copi, forg, estat, valuabl, jane

Topic 12

Highest Prob: water, bodi, will, part, make, caus, see, upon, spirit, great, may, also, like, air, therefor, cold, come, natur, shall, time, put, heat, nourish, man, motion, take, turn, littl, well, thing

FREX: tooth, sweat, expel, tickl, purger, version, indur, stomach, rhubarb, dens, clarifi, gut, rheum, clay, crush, tangibl, egg, nail, coal, almond, compress, staunch, purg, peubl, urin, water, cold, bone, plaster, moistur

Topic 13

Highest Prob: hous, say, case, may, sheriff, shall, break, law, defend, good, time, pillag, king, man, action, plaintiff, make, execut, upon, writ, non, one, without, reason, will, suit, liberti, door, per, great

FREX: bivalv, door, existen, hous, easement, window, sheriff, pillag, adj, capia, denial, egress, messuagium, septemb, rood, idem, request, break, swine, totum, pig, process, messuag, magnum, nuisanc, distrain, plaintiff, defend, facia, stop

Topic 14

Highest Prob: man, will, good, thing, say, busi, speech, may, think, know, well, let, first, shall, can, make, time, one, much, give, use, anoth, counsel, come, take, speak, turn, wise, show, find

FREX: busi, speech, discours, cun, secreci, simul, bold, occas, dissembl, counsel, advantag, wise, surpris, pack, hasti, gather, wisdom, jest, amiss, friend, often, tedious, absurd, wors, convers, walk, ridicul, trick, tale, turn

Topic 15

Highest Prob: spirit, part, bodi, will, water, make, see, heat, air, also, upon, great, may, flame, oil, therefor, forth, long, caus, natur, motion, wine, canon, putrefact, time, much, one, put, man, first

FREX: beer, bottl, putrefact, acceler, canon, appl, matur, oil, puls, shine, oili, commix, milk, clarif, wax, malaciss, flame, liquor, putrifi, putrefi, lee, candl, digest, dew, conservatori, drink, moistur, clarifi, turret, bran

Topic 16

Highest Prob: say, make, shall, law, grant, trade, act, edward, card, king, ordin, within, good, case, use, may, realm, year, statut, one, subject, plaintiff, queen, henri, parliament, upon, penalti, without, bring, cap

FREX: card, trade, monopoli, ordin, tailor, penalti, apprentic, play, forbid, bylaw, artific, commonwealth, ralph, monopol, imprisonetur, chamberlain, ipswich, deputi, dice, recreat, patente, mysteri, sole, warden, dispens, factor, exercis, traffic, cloth, penal

Topic 17

Highest Prob: learn, say, make, man, king, princ, time, great, upon, excel, will, speech, give, alexand, use, caesar, one, yet, name, thing, well, shall, virtu, can, may, see, unto, roman, god, take

FREX: alexand, trajan, callisthen, antoninus, caesar, learn, dictat, sulla, falinus, eloqu, antipat, greas, vener, diogen, cornelius, philosoph, marcus, theme, xenophon, homer, scholar, lucius, virgil, proprieti, quirir, endu, commodus, literatur, parmenio, emperor

Topic 18

Highest Prob: land, grant, name, true, yet, addit, patent, fals, word, one, pass, unto, thing, refer, demonstr, day, shall, tenur, good, certainti, part, upon, whereof, parish, rule, place, therefor, first, lie, law

FREX: falsiti, denomin, dale, burst, demonstr, addit, tenur, rubi, guilielmi, certainti, annex, gentlewoman, specificat, parish, box, meum, fals, indentura, arra, varianc, plot, multitudin, ici, refer, pictur, appel, date, patent, promontori, represent

Topic 19

Highest Prob: shall, law, may, majesti, scotland, will, england, whether, nation, make, question, sever, one, point, part, britain, yet, parliament, first, time, name, kingdom, crown, subject, case, great, opinion, take, therefor, say

FREX: scotland, britain, nation, pollio, union, intern, infidel, vindic, scot, admiralti, whether, languag, hostil, pirat, border, territori, sovereignti, question, financ, articl, vote, dialect, style, impost, inequ, scottish, merchant, barwick, divid, superscript

Topic 20

Highest Prob: king, court, shall, justic, say, writ, law, hold, may, parliament, statut, take, call, time, common, make, one, man, come, person, great, ancient, counti, england, quoth, act, well, year, word, appear

FREX: tourn, frankpledg, comitatus, concilium, placita, leet, amerc, amercia, abbot, crime, capit, bracton, retourn, novel, assiz, burgh, fleta, testa, glanvill, haec, wiseman, shire, assisa, justitia, bound, appeal, justitiarii, maresch, nous, henceforth

Topic 21

Highest Prob: lopez, king, spain, good, letter, great, majesti, shall, matter, make, mean, take, upon, time, give, man, servic, first, may, say, antonio, unto, will, manuel, write, also, practic, natur, confess, come

FREX: lopez, antonio, manuel, spain, intellig, emanuel, portug, mendoza, count, secretari, advertis, evas, intercept, calai, confess, constantinopl, passport, mke, suspicion, antwerp, confer, detest, letter, retinu, cun, reveal, excess, doctor, fidel, credenc

Topic 22

Highest Prob: man, yet, shall, reign, will, may, thing, time, make, state, happi, princ, queen, upon, see, take, can, life, one, king, age, self, love, long, first, kingdom, felic, great, peopl, mind

FREX: con, felic, unwis, popish, happi, hap, miseri, regal, fame, rare, priest, prais, faction, ged, prin, reign, aswel, apell, henryth, dom, tune, glori, scan, debar, blemish, blot, ous, aim, helm, woe

Topic 23

Highest Prob: say, will, one, man, answer, upon, come, great, king, lord, sir, take, make, shall, queen, ask, give, tell, can, may, like, know, time, much, hous, think, use, see, now, friend

FREX: bacon, aristippus, diogen, cardin, madam, ask, raleigh, pompey, seaman, captain, nichola, crassus, owl, pillow, yes, pyrhus, surgeon, alexand, antigonus, uva, harbing, hannib, ladi, peasant, sir, answer, boot, tell, athenian, fabius

Topic 24

Highest Prob: bodi, caus, spirit, will, see, water, heat, great, air, part, may, also, upon, draw, make, creatur, man, therefor, good, come, like, cold, thing, putrefact, can, one, put, time, forth, moistur

FREX: amber, putrefact, sugar, concoct, shower, moistur, congeal, portend, moon, fragil, smell, liquefi, rub, percept, vivif, honey, visual, attract, tangibl, winter, summer, excrement, basen, tobacco, liquor, worm, feather, moist, bird, rain

Topic 25

Highest Prob: law, inherit, shall, heir, may, statut, book, next, say, uncl, man, case, author, writ, great, make, issu, can, cousin, littleton, commonlaw, purchas, call, word, without, observ, father, common, lineal, intend

FREX: cousin, lineal, uncl, collater, creation, charta, littleton, disparag, infant, inherit, guardian, regist, haered, chivalri, oust, descent, warranti, duchess, parent, next, estoppel, merton, lament, shame, maxim, ascent, section, propinqu, transversali, nobil

Topic 26

Highest Prob: god, thou, thi, man, shall, work, time, will, holi, creatur, earth, christ, day, good, word, spirit, yet, world, natur, make, can, unto, thing, lord, law, one, may, great, first, power

FREX: thi, thou, ghost, etern, christ, psalm, jesus, holi, thyself, mediat, bide, sabbath, righteous, advers, god, creation, mirac, angel, everlast, uniti, creatur, translat, vale, prais, carol, redempt, zion, godhead, revel, regener

Topic 27

Highest Prob: tree, will, upon, plant, fruit, make, earth, may, water, put, grow, root, ground, forth, come, caus, also, herb, seed, like, flower, great, set, leav, kind, much, see, one, therefor, bear

FREX: sap, bough, plum, herb, moss, holli, dung, fig, cherri, peach, tree, cucumb, mushroom, plant, graft, oak, pear, lettuc, blossom, radish, stalk, escul, colewort, mistleto, flower, seed, root, vine, earli, fir

Topic 28

Highest Prob: king, law, can, make, shall, henri, subject, grant, may, parliament, will, good, act, judg, power, offic, edward, justic, take, statut, pardon, give, proclam, prerog, case, say, offenc, elizabeth, non, accord

FREX: recogniz, proclam, dispens, pardon, benevol, prerog, sole, insepar, bridg, malum, walter, recus, forfeitur, subordin, penal, discret, imprison, confer, offenc, subject, henri, patent, weal, levi, administ, engross, redress, offic, king, repair

Topic 29

Highest Prob: man, upon, will, imagin, may, shall, thing, spirit, make, one, think, bodi, work, take, great, like, use, natur, say, part, also, can, anoth, time, much, experi, must, good, tell, therefor

FREX: ointment, imagin, bead, belief, wart, witch, perfum, plagu, dissimul, magic, transmiss, sympathi, juggler, brain, solitud, cramp, wolf, immateri, emiss, friendship, ingredi, infect, card, suspicion, agu, anoint, smell, weapon, jaundic, unburi

Topic 30

Highest Prob: lord, will, time, say, self, queen, good, shall, well, make, think, can, know, love, lordship, Essex, upon, man, majesti, take, may, yet, mani, matter, though, ever, part, great, one, tell

FREX: Essex, madam, mat, con, wri, ber, cion, remem, sati, dam, love, lover, forsak, lordship, queen, pub, tell, ved, othersid, ire, fortun, pinion, ous, bee, rememb, test, lord, ireland, rack, axletre

Topic 31

Highest Prob: make, law, shall, will, edward, non, yet, say, act, claus, year, bind, can, one, take, stand, may, revoc, repeal, man, contract, good, back, without, william, parti, give, quoth, well, word

FREX: revoc, arbit, revok, duress, countermand, rug, detractor, repeal, arbitra, menac, claus, unpunish, irrevoc, contract, derogatori, layout, award, wild, facto, oblig, oblige, william, scribe, attorn, abid, parol, immutauit, maxim, exchang, aunt

Topic 32

Highest Prob: protect, king, law, shall, man, cast, upon, writ, action, may, allow, can, case, make, one, necess, delay, like, life, day, plaintiff, time, come, good, servic, take, subject, first, without, defend

FREX: protect, passeng, barg, delay, facia, headway, cast, disallow, praemunir, ferryman, necess, reattach, excus, allow, default, tempest, justifi, assail, surcharg, sue, trespass, inspect, font, debtor, safeguard, custodian, mous, standf, venir, repeal

Topic 33

Highest Prob: law, will, may, time, make, shall, say, king, great, good, man, upon, mani, reason, case, author, first, work, majesti, yet, learn, one, take, rule, thing, know, new, common, part, judg

FREX: innov, compil, lawyer, ego, recompile, bail, theme, ang, frankalmoign, edit, obj, resp, student, digest, jura, legitim, noos, languag, mainpriz, edgar, regius, incertainti, edict, dux, abridg, decis, quotat, discret, jump, obsolet

Topic 34

Highest Prob: king, upon, will, great, make, time, shall, princ, take, come, also, man, one, year, part, can, peopl, may, perkin, earl, yet, two, much, well, give, think, good, say, unto, though

FREX: perkin, castill, ferdinando, fillip, katherin, aragon, scottish, treati, exet, empson, arthur, cornishmen, warwick, arch, truce, ladi, fox, bray, cornwal, fatherlaw, ambassag, dudley, amiti, earl, napl, suffolk, ble, chamberlain, lewi, ken

Topic 35

Highest Prob: will, metal, make, silver, iron, ounce, brass, gold, incorpor, glass, may, use, stone, lead, bodi, whether, copper, inquir, colour, first, half, mixtur, weight, tin, like, fire, well, sever, compound, melt

FREX: ounce, brass, copper, tin, metal, calcin, silver, addita, iron, quicksilv, tough, cheap, fixat, calaminar, vitriol, malleabl, stuff, melt, gold, incorpor, vitrifi, mixtur, glass, milki, compound, vitrif, luster, flint, tinctur, volatil

Topic 36

Highest Prob: heir, land, shall, part, inherit, man, mother, father, word, blood, law, feesimpl, feoff, say, without, grant, son, case, make, escheat, die, estat, can, issu, give, wife, pass, fee, hold, purchas

FREX: mother, escheat, heir, haer, feesimpl, inherit, successor, duplic, blood, father, ventr, male, releas, femal, feoff, wife, haered, sandi, fee, everywher, attaind, hang, intrus, pregnant, donat, convey, polit, mesn, habendum, ancestor

Topic 37

Highest Prob: shall, fine, make, levi, king, say, land, statut, law, covin, will, rend, case, year, lesse, without, appear, may, bar, debtor, lessor, act, tenant, fraud, seiz, good, hold, husband, debt, yet

FREX: covin, debtor, lessor, widow, levi, pledg, fraud, lesse, quarantin, dower, injust, fine, asset, warranti, vex, encroach, estov, ind, maker, smith, ceux, husband, rend, bar, avowri, proclam, fermor, strife, licenc, owe

Topic 38

Highest Prob: king, shall, justic, case, franchis, parliament, say, come, may, writ, accord, lord, claim, jame, eyr, upon, parti, time, make, grant, law, place, year, hold, charter, great, resolv, can, act, counti

FREX: franchis, eyr, petition, disabl, claim, warranto, michaelma, jame, summon, mich, repleui, quo, bailiff, charter, gloucest, forthwith, triniti, prelat, griev, adjudg, venir, ancestor, articl, circuit, cambridg, counti, justic, novum, allocandi, proclam

Topic 39

Highest Prob: law, shall, feoff, yet, act, make, die, upon, use, therefor, condit, land, grant, seiz, will, life, take, man, heir, rend, now, case, purchas, elizabeth, right, edward, good, leas, statut, one

FREX: disseise, remit, feme, dissent, discontinu, remitt, lesse, feoff, lach, atturn, survivor, discontinue, advowson, companion, asset, impropri, devis, pursuanc, precip, leas, moieti, rob, demis, dagger, condit, attorney, liveri, accessari, accept, monitor

Topic 40

Highest Prob: land, heir, shall, tenant, make, man, upon, call, use, life, estat, can, king, servic, entail, may, seiz, hold, will, give, statut, fine, good, lord, one, convey, inherit, parti, first, writ

FREX: entail, tenant, tenur, soccag, convey, escheat, wardship, attaind, forfeit, liveri, remaind, revers, land, knight, leas, outlawri, acr, coven, ken, servic, recoveri, attaint, heir, prius, owner, lack, appoint, dispos, capit, debt

Topic 41

Highest Prob: court, say, prohibit, king, upon, plea, shall, law, case, judg, caus, ecclesiast, tith, statut, edward, commonlaw, justic, kill, common, may, jurisdict, modus, give, act, henri, yet, spiritu, one, grant, hold

FREX: tith, prohibit, pendent, ecclesiast, praemunir, modus, plea, jurisdict, parson, spiritu, demurr, cogniz, kill, consult, sue, examen, lege, libel, bennet, belong, court, schism, senat, placitum, disregard, tempor, alibi, banc, modi, forum

Topic 42

Highest Prob: will, shall, unto, may, law, pope, king, majesti, justic, one, make, god, therefor, upon, know, subject, time, kingdom, come, judg, year, queen, must, execut, papist, much, now, receiv, can, abus

FREX: papist, pope, excommun, jesuit, cathol, recus, seminari, bull, invas, abus, rebellion, traitor, upright, denounc, bribe, toler, hell, partial, popish, depos, timber, purveyor, proud, specifi, idl, reform, wench, sovereign, complot, enforc

Topic 43

Highest Prob: man, say, shall, will, death, law, upon, make, think, take, give, see, may, one, like, can, thing, court, yet, though, great, die, time, day, duel, speak, much, good, come, lord

FREX: duel, vespasian, combat, csar, reveng, mourn, generous, dionysius, antiochus, wright, spill, galba, caesar, cicero, nero, contum, forgiv, homicid, fight, archbishop, valiant, legion, terribl, fortitud, depos, mutini, perfidi, censur, insidi, celebr

Topic 44

Highest Prob: author, law, great, shall, time, work, sir, book, mani, honour, institut, part, read, learn, first, observ, king, write, may, can, man, yet, upon, three, littleton, good, one, opinion, will, report

FREX: institut, student, vacant, famous, cite, littleton, herein, poster, esquir, studious, sir, commentari, print, match, reverend, shell, surnam, reader, furnish, argent, flourish, publish, author, cambridg, epistl, almighti, daughter, briberi, thoma, read

Topic 45

Highest Prob: heir, case, shall, edward, say, shelly, son, use, execut, tail, male, bodi, land, make, law, estat, issu, take, therefor, recoveri, purchas, life, can, will, yet, word, henri, may, judgement, richard

FREX: shelly, male, tail, recoveri, vest, elder, recoveror, forasmuch, fol, sue, uncl, indentur, execut, heir, descent, daughter, beneficiari, purchas, stat, richard, son, beget, femal, edward, issu, remaind, leas, entri, divest, limit

Topic 46

Highest Prob: will, king, shall, may, lord, great, make, man, say, upon, court, one, justic, time, first, grant, thing, caus, come, give, take, majesti, law, can, three, therefor, much, find, judg, case

FREX: monopoli, mompesson, injunct, weston, february, complain, committe, chancellor, chanceri, bill, decre, grievanc, ransom, etcetera, proctor, stephen, impoison, gile, export, partnership, matto, pottag, rob, riot, pronounc, sole, conceal, whip, poison, examin

Topic 47

Highest Prob: instanc, bodi, natur, motion, water, will, heat, air, place, must, power, also, observ, may, can, spirit, like, great, one, substanc, part, differ, effect, shall, therefor, first, flame, sens, anim, take

FREX: similar, magnet, instanc, rapid, investig, exist, latent, expans, howev, particl, migrat, expand, concret, anim, conspicu, tendenc, contact, connect, format, class, sphere, liquid, classif, imped, ebb, veloc, dens, exert, remark, deviat

Topic 48

Highest Prob: case, say, shall, action, plaintiff, upon, defend, law, debt, good, may, word, parti, make, court, act, king, justic, take, resolv, will, edward, preced, henri, bar, great, plead, one, satisfact, matter

FREX: creditor, bankrupt, plaintiff, satisfact, debt, nov, assumpsit, sedit, justif, defend, action, payment, justifi, rate, indebt, commission, outlawri, rye, bar, oblig, wage, regist, clark, vicar, preced, insuffici, wager, coke, sale, demurr

Topic 49

Highest Prob: natur, histori, will, may, man, shall, experi, one, thing, must, great, yet, hope, use, matter, can, work, particular, part, scienc, let, upon, mind, new, make, philosophi, observ, discoveri, think, well

FREX: experiment, instinct, discoveri, histori, instaur, axiom, mechan, induct, intermedi, prodigi, hitherto, mode, scienc, afford, mankind, investig, arrang, compil, pueril, philosophi, deduc, theori, remark, copious, result, hope, experi, uninterrupted, doubtless, cultiv

Topic 50

Highest Prob: shall, king, take, say, statut, person, act, case, make, indict, upon, word, feloni, man, save, poison, coron, within, part, give, counti, one, great, record, yet, justic, author, thing, wife, sorceri

FREX: sorceri, witchcraft, charm, coron, buggeri, enchant, oier, garrison, conjur, certiorari, sorcer, constabularius, castl, appris, accessari, termin, host, devil, mid, weston, poison, attaind, dower, middlesex, mirror, saul, christiani, fortilegi, sodom, constabl

Topic 51

Highest Prob: lord, overburi, will, impoison, shall, poison, somerset, upon, first, man, murder, may, must, make, can, secret, king, give, time, great, act, three, part, say, proof, tower, natur, self, mean, ladi

FREX: overburi, somerset, impoison, empoison, poison, displac, weston, tower, hatr, murder, ladi, abet, proof, trap, secret, northampton, lieuten, malic, gradus, dispatch, friendship, actor, intox, packet, jargon, abett, divert, peer, devic, countess

Topic 52

Highest Prob: year, man, live, hundr, age, nourish, life, great, eighti, one, also, old, upon, long, may, time, seven, nineti, four, creatur, mani, five, will, seventi, twenti, yet, well, three, woman, thirti

FREX: eighti, nineti, seventi, bait, hundr, fifti, nourish, graft, thirti, rhetorician, affabl, aaron, marcus, augustus, weep, fulfil, live, tiberius, flood, sixti, pragmat, simeon, xenophan, stout, apostl, caius, devout, seven, age, year

Topic 53

Highest Prob: church, man, will, shall, say, thing, may, god, good, great, upon, time, one, bishop, make, yet, part, well, matter, person, first, law, can, take, much, think, govern, state, speak, caus

FREX: usuri, preach, church, preacher, controversi, uniti, liturgi, baptism, pastor, borrow, ministri, zeal, reform, blasphemi, bishop, primit, minist, worship, benefic, ceremoni, usur, profan, prayer, partial, erkenwald, plural, contumaci, synod, holi, christ

Topic 54

Highest Prob: sound, will, make, bodi, air, see, one, echo, hear, may, much, like, man, shall, great, light, speak, upon, pass, visibl, voic, string, well, come, time, motion, better, likewis, therefor, stand

FREX: echo, visibl, audibl, medium, pipe, sound, lute, string, reflect, bell, vial, delat, sight, repercuss, distanc, music, imit, harp, speci, voic, line, refract, exquisit, cranni, articul, purl, concav, belli, pupil, appetit

Topic 55

Highest Prob: king, law, england, allegi, say, bear, subject, natur, one, henri, alien, case, kingdom, shall, make, edward, can, man, land, appear, within, obedi, may, cap, come, realm, sever, book, fol, yet

FREX: allegi, calvin, alien, obedi, gascoïn, fol, municip, born, deniz, dominion, lib, gersey, lieg, cilicia, ergo, denizen, extra, dukedom, local, scot, normandi, englishman, bear, trusti, isl, england, irishman, nichola, birth, scotland

Topic 56

Highest Prob: king, custom, grant, make, edward, parliament, hold, shall, upon, act, subsidi, statut, merchant, wool, land, say, stranger, may, cloth, great, imposit, take, henri, appear, new, year, common, without, realm, ancient

FREX: tunnag, wool, poundag, imposit, subsidi, alnag, merchant, cloth, penni, draperi, tun, leather, merchandis, sack, dimid, stranger, baroni, canva, trader, prisa, prisag, marc, custom, earthenwar, northern, entreat, impos, alien, wine, butlerag

Topic 57

Highest Prob: king, law, say, great, year, england, shall, time, reign, man, book, one, call, mani, case, first, ancient, justic, make, statut, edward, will, henri, report, caus, court, writ, may, appear, commonlaw

FREX: conqueror, dom, greek, abbot, prefac, confessor, charta, magnacarta, forest, uta, conquest, reader, glanvill, burgo, commentari, easter, center, normandi, hubert, british, tomorrow, haer, inn, archbishop, cancel, pertinentii, reign, barrist, villa, extant

Topic 58

Highest Prob: man, say, good, natur, make, mind, can, knowledg, great, may, will, time, one, upon, fortun, thing, much, see, part, shall, yet, virtu, like, well, god, state, learn, govern, true, first

FREX: atheism, fortun, moral, cato, cicero, felic, precept, perturb, demosthen, superstit, decenc, applic, machiavelli, socrat, knowledg, pedant, sophist, lover, atheist, cultur, caesar, honesti, caution, animi, livi, pursuit, magnus, tacitus, eleg, atticus

Topic 59

Highest Prob: hospit, king, incorpor, say, shall, case, hous, make, sutton, name, word, charter, grant, licenc, give, corpor, can, edward, object, may, henri, without, find, poor, law, one, first, appear, governor, take

FREX: hospit, sutton, incorpor, governor, chaplain, ramsey, licenc, chantri, erect, corpor, founder, foundat, preacher, dean, instructress, charter, explanatori, requisit, reg, nomin, chariti, implead, usher, postea, poor, school, mortmain, essenc, object, establish

Topic 60

Highest Prob: law, king, man, shall, will, say, may, make, caus, edward, imprison, can, parliament, give, without, take, time, great, commit, prison, judg, england, henri, case, must, act, reason, land, justic, come

FREX: etcetera, imprison, april, atia, magnacarta, habeas, fee farm, breve, villain, num, stanford, loan, corpus, bastardi, detain, bail, prison, mandatum, convoc, mittimus, terra, freeman, burgag, bondman, pasch, rogerus, matrimoni, rot, gaol, parium

Topic 61

Highest Prob: law, king, natur, england, shall, say, case, one, subject, may, can, person, will, upon, time, first, reason, therefor, statut, make, kingdom, man, act, word, yet, bear, two, parliament, come, crown

FREX: gascoign, conquest, allegi, born, submiss, normandi, infer, conquer, england, parent, platform, anjou, scottishmen, obeis, alien, signori, monarchi, scotland, privileg, englishmen, guyenn, provinc, post, lieg, duchi, walter, foe, intermeddl, confut, kingdom

Topic 62

Highest Prob: indict, say, arrest, law, pillag, case, shall, upon, justic, find, may, offic, will, kill, king, quoth, warrant, take, make, murder, resolv, night, one, london, give, verdict, matter, defend, judgement, can

FREX: arrest, pillag, forethink, indict, nich, pleint, ridley, allocatur, night, sergeant, verdict, malic, plaint, warrant, cur, watchman, murder, kill, mace, jeopardi, felon, parish, recusaveri, minist, receipt, tire, johann, acquitt, doorkey, constabl

Topic 63

Highest Prob: motion, bodi, thing, water, part, will, heat, may, one, natur, yet, air, let, togeth, great, can, first, desir, place, like, must, congreg, spirit, well, name, come, iron, call, also, make

FREX: congreg, predomin, motion, homogen, ebb, vacuiti, assimil, magnet, dimens, coition, pole, penetr, flow, heterogen, load, excit, connexion, abhor, rotat, compress, pressur, lesser, agit, gun, centaur, curb, numb, diffus, reject, densiti

Topic 64

Highest Prob: shall, law, upon, indict, say, case, justic, may, court, peer, verdict, error, give, trial, find, record, judg, parti, prison, lord, judgement, plead, accus, juri, make, king, steward, tri, writ, bring

FREX: peer, verdict, steward, accus, indict, arraign, trial, judgment, juri, pare, guilti, dacr, prison, juror, error, conspiraci, oye, inquest, bee, mar, conus, formedon, attaind, panel, purvien, scarlet, tri, lessor, tortur, indictor

Topic 65

Highest Prob: land, king, shall, deed, pass, man, wit, name, may, signifi, call, grant, henri, one, word, say, style, make, also, year, time, wood, upon, indent, water, can, hous, first, seal, domesday

FREX: indent, domesday, terra, signifi, aquitain, style, deed, praecip, pastur, soil, dux, centum, ang, wit, testimonium, infam, meadow, ridg, saxon, reddendum, altern, supremum, heath, dominus, haec, tenendum, liveri, stabl, wood, marsh

Topic 66

Highest Prob: great, majesti, king, time, upon, state, man, make, one, spain, say, good, may, shall, mani, england, yet, will, unto, never, part, much, like, peopl, princ, realm, year, war, also, honour

FREX: libel, faction, spain, untruth, lowcountri, sundri, cathol, pag, christendom, guis, turk, avouch, leagu, amiti, indi, cassius, state, confederaci, franc, assay, burleigh, treati, persecut, invas, sedit, jealousy, invad, macedon, counselor, germani

Topic 67

Highest Prob: use, statut, law, feoff, case, upon, shall, make, time, say, will, take, land, remedi, commonlaw, trust, give, therefor, may, reason, good, cesti, possess, consider, conscienc, yet, right, one, action, man

FREX: cesti, feoff, conscienc, trust, subpoena, remedi, chanceri, notic, tenanc, collus, use, priviti, statut, quaer, mortmain, consider, relief, transfer, participl, creditor, definit, opus, praecip, provisor, covin, recontinu, commonlaw, remaind, possess, releas

Topic 68

Highest Prob: parliament, king, come, predict, lord, gentleman, say, per, quoth, john, will, die, mar, queen, common, henri, command, head, place, court, law, modus, hic, hold, make, without, pro, year, self, give

FREX: mic, dit, proxi, kanc, hic, johann, ver, diem, mar, eux, antonius, johan, cur, roper, other, gentleman, salop, sequitur, predict, thomam, roy, shaggi, edm, nich, transcrib, quad, richardum, winton, oxon, pur

Topic 69

Highest Prob: man, use, will, knowledg, one, may, make, therefor, method, thing, great, reason, part, word, natur, much, see, shall, write, well, exercis, scienc, rule, yet, speech, whereof, say, memori, affect, though

FREX: rhetor, method, logic, similitud, aphor, defici, poesi, cipher, hieroglyph, tradit, charact, intellectu, eloqu, mathemat, exercis, knowledg, aristotl, extempor, cogit, faculti, conceit, appendix, scienc, grammar, applic, inquiri, studi, stammer, moral, enfold

Topic 70

Highest Prob: king, upon, make, time, great, shall, will, may, part, lord, war, person, come, man, unto, duke, take, yet, peopl, well, peac, good, think, britain, england, one, much, also, first, french

FREX: maximilian, perkin, flander, britain, duke, ambassador, york, plantagenet, charl, rebel, french, bruge, lovel, clifford, succour, duchess, confeder, stanley, treati, urswick, orlean, margaret, lambert, ambassag, simon, simnel, burgundi, battl, sanctuari, lincoln

Topic 71

Highest Prob: will, man, great, may, shall, upon, honour, can, last, good, matter, like, time, make, place, case, right, everi, person, now, parliament, know, king, suit, undertak, much, particular, self, must, whether

FREX: undertak, suitor, discount, peril, messag, disgrac, merit, adventur, protest, reput, honour, sober, suit, accid, whither, committe, provoc, ignor, hat, invas, embrac, tend, denial, insist, handl, expostulatori, last, indian, farewell, punctual

Topic 72

Highest Prob: god, man, natur, divin, upon, law, part, reason, scriptur, word, true, church, much, thing, make, see, two, matter, say, use, will, first, point, well, knowledg, unto, know, may, accord, can

FREX: anger, scriptur, idolatri, worship, divin, heathen, reluct, inspir, ador, reveal, reced, folk, posit, doctrin, moral, total, angri, holi, sin, summari, infer, illustr, theolog, deduc, egyptian, defici, ark, satisfactori, israelit, handl

Topic 73

Highest Prob: aforesaid, say, borough, caus, mayor, burgess, king, lord, time, jame, shall, chief, bag, day, disfranchis, one, word, thoma, plymouth, year, may, certifi, well, remov, will, within, offic, reason, good, john

FREX: borough, plymouth, disfranchis, aforesaid, burgess, bag, mayor, certifi, commonalti, jame, mayoralti, thoma, remov, twelv, contemptu, inhabit, innkeep, presenc, keeper, citizen, victual, malici, clement, february, alehous, depriv, magistr, perfidi, fals, chief

Topic 74

Highest Prob: heat, natur, bodi, cold, hot, fire, thing, air, will, make, water, flame, instanc, burn, also, degre, warm, yet, glass, first, let, may, one, self, motion, appear, must, tabl, like, upon

FREX: warm, heat, perpendicular, hot, ignit, ray, expans, negat, exclus, cold, flame, reject, sparkl, weather, burn, howev, tube, tenuiti, liquid, attrit, fever, fire, decreas, dung, instanc, apron, anim, snow, glass, cancer

Topic 75

Highest Prob: long, life, year, live, age, man, bodi, thing, also, spirit, great, hundr, old, much, creatur, one, will, may, make, certain, dri, bird, like, air, part, time, short, yet, water, less

FREX: liver, diet, statur, bird, eagl, desicc, goos, wrinkl, tallow, womb, long, dri, slender, live, flesh, youth, emaci, creatur, growth, bald, raven, mountain, curl, hundr, sign, age, fourteen, grey, hair, moistur

Topic 76

Highest Prob: king, will, shall, upon, may, man, give, lord, one, princ, time, law, great, say, make, subject, pope, like, opinion, parliament, matter, can, yet, charg, now, first, take, true, thing, god

FREX: slander, talbot, sediti, pope, benevol, forerun, doctrin, depos, cathol, irrit, papist, tyrant, venom, aggrav, contagion, proscib, major, conspiraci, tyron, menac, clemenc, extenu, incens, sentenc, princ, violat, channel, awak, miser, duel

Topic 77

Highest Prob: majesti, will, shall, may, great, upon, man, take, good, law, like, kingdom, mani, king, make, thing, therefor, say, unto, yet, first, one, work, mean, time, hous, peopl, can, god, now

FREX: plantat, majesti, undertak, project, abus, gracious, sting, miseri, discard, provis, portion, manifold, alderman, poor, harp, crave, annual, ina, kingdom, reliev, almighty, adventur, ireland, habit, mighti, enterpris, stipend, commod, humbl, prize

Topic 78

Highest Prob: wind, air, blow, thing, will, sail, motion, great, part, one, sea, may, east, north, make, south, place, rain, west, water, sun, way, much, cloud, also, certain, name, rise, come, likewis

FREX: wind, sail, south, east, mast, north, rain, blow, cloud, presag, nurseri, gale, engend, west, prognost, moon, storm, tropic, weather, mizon, monit, calm, stormi, vapour, whirl, allay, impuls, whirlwind, optat, subterran

Topic 79

Highest Prob: man, king, make, shall, say, child, will, come, mani, great, time, one, thing, see, upon, take, wife, much, famili, think, father, give, son, may, yet, like, hous, day, desir, ever

FREX: bensalem, herald, propheci, famili, feast, chastiti, child, pace, grape, chast, hebrew, ivi, dream, dinner, jew, solyman, enamel, lad, scroll, mission, lineag, vespasian, unmarri, chair, salomon, parent, marriag, cluster, golden, voyag

Topic 80

Highest Prob: use, statut, feoff, law, word, shall, seiz, estat, case, upon, person, may, possess, make, first, execut, heir, cesti, land, can, three, therefor, right, time, will, save, take, king, titl, life

FREX: cesti, feoff, proviso, use, seiz, conting, remaind, preambl, statut, possess, feme, regress, save, convey, estat, coven, joint, titl, exclud, execut, materi, purview, limit, bargain, infant, parol, waiv, extirp, word, remit

Topic 81

Highest Prob: natur, man, philosophi, scienc, will, thing, can, understand, experi, mind, caus, great, first, system, one, human, upon, even, axiom, therefor, ancient, discoveri, object, method, yet, sens, power, common, rather, notion

FREX: system, idol, metaphys, philosophi, notion, investig, abstract, dogma, anxious, everyth, axiom, unanim, logic, theori, discoveri, scienc, greek, contempl, mankind, howev, exist, basi, method, theolog, steadi, sceptic, henc, intermedi, anticip, generalti

Topic 82

Highest Prob: shall, will, natur, time, law, kingdom, peopl, upon, may, speaker, master, say, great, now, can, state, one, first, see, union, make, yet, therefor, much, take, man, point, come, self, think

FREX: speaker, union, master, provinc, commixtur, barbar, revolt, nobil, popul, lot, unit, desol, surcharg, candidaci, kingdom, famili, nation, scottish, irish, communiti, rebel, ireland, reclaim, peopl, aragon, restrict, canton, yoke, goth, polici

Topic 83

Highest Prob: war, upon, great, spain, will, make, man, state, come, time, yet, may, say, part, shall, take, speak, forc, now, just, sea, spaniard, land, fear, true, though, much, year, armi, king

FREX: spaniard, spanish, palatin, war, navi, spain, enterpris, valour, armi, confeder, germani, fight, turk, parma, defens, invas, syria, veteran, indi, dominion, bohemia, battl, retreat, courag, drake, brave, ship, design, thousand, lacedaemonian

Topic 84

Highest Prob: may, bodi, take, thing, spirit, will, much, use, upon, part, also, good, wine, life, oper, touch, make, like, littl, water, therefor, heat, air, long, meat, juic, espec, drink, oil, three

FREX: saffron, astring, claret, broth, intener, myrrh, aliment, alo, bath, myrtl, draught, stomach, juic, anoint, ale, grain, unction, meat, flux, shirt, quinc, rhubarb, infus, fricat, almond, citron, morn, tragacanth, decoct, vinegar

Topic 85

Highest Prob: king, will, may, shall, make, man, great, upon, can, must, god, much, give, good, yet, time, one, take, thing, law, offic, let, honour, now, self, think, person, kingdom, first, matter

FREX: needless, coloni, beseech, flatter, choic, revenu, numero, courtier, art, cost, favourit, poperi, leak, subsidi, unskil, board, regul, essay, dispatch, june, trade, interpos, rash, neutral, suitor, petit, vaniti, awhile, admonit, heartstr

Topic 86

Highest Prob: counsel, upon, man, one, time, make, king, squir, shall, person, princ, may, matter, come, will, great, take, yet, hand, good, god, give, england, counsellor, queen, part, think, thing, use, natur

FREX: squir, counsellor, ambiti, counsel, dispatch, stabl, saul, secreci, jesuit, vow, fugit, papist, suborn, persuas, confessor, affair, conspir, princ, actor, voyag, bladder, event, retract, pericular, saddl, meti, backward, heathenish, essex, palla

Topic 87

Highest Prob: majesti, lord, will, may, king, upon, time, shall, say, make, court, opinion, think, judg, busi, chancellor, give, well, day, letter, take, thing, much, speak, can, self, good, though, god, true

FREX: peacham, coke, chancellor, majesti, chanceri, compani, busi, paper, shrewsburi, humbl, yesterday, advertis, lake, counselor, afternoon, solicitor, sorri, praemunir, glad, fever, mistrust, subscrib, twofac, cough, auricular, houghton, clamour, accommod, allen, financ

Topic 88

Highest Prob: shall, word, take, grant, rule, rend, upon, acr, therefor, plead, land, intend, yet, heir, reserv, void, law, may, ambigu, intent, make, good, give, case, will, part, three, except, ten, matter

FREX: ambigu, acr, aver, paten, reserv, shill, entail, frank, rend, graunte, feme, ten, dale, grantor, warren, fealti, rule, moiety, replic, pound, precis, plead, implic, void, distress, white, intend, constru, warranti, abat

Topic 89

Highest Prob: thing, man, may, natur, also, say, god, unto, one, will, seem, shall, yet, see, matter, make, fabl, can, time, great, come, jupit, like, take, first, kind, well, much, neither, begin

FREX: prometheus, jupit, proserpina, pan, orpheus, page, typhon, parabl, miss, fabl, pentheus, sphinx, perseus, cupid, siren, nemesi, giant, bacchus, cere, palla, moreov, allegori, muse, atalanta, eleg, thebe, icarus, saturn, juno, hercul

Topic 90

Highest Prob: great, also, come, make, may, man, will, one, shall, let, upon, say, thing, land, divers, place, side, part, good, mani, see, know, hous, take, give, well, day, likewis, like, use

FREX: boat, galleri, plantat, pillar, blue, travel, front, velvet, chariot, garden, carpent, alley, pool, diari, bay, pave, atlanti, ship, statu, america, inund, room, ark, victual, europ, scene, stair, cane, navig, sick

Topic 91

Highest Prob: law, judg, court, upon, justic, say, caus, may, oath, shall, good, make, appear, examin, king, parti, time, peac, take, will, counti, case, england, can, sheriff, everi, stand, bind, see, indict

FREX: cannon, oath, layman, gaol, excommun, habeas, accus, laypeopl, tourn, circuit, linwood, advoc, interrogatori, examin, legat, clerk, eyr, rot, clergi, jurisdict, billa, ecclesiast, judg, infami, swear, annoy, curia, recogniz, grand, matrimoni

Topic 92

Highest Prob: king, man, feloni, treason, shall, constabl, justic, may, peac, offic, law, make, take, person, offenc, upon, case, punish, court, good, realm, call, land, proceed, petti, inquir, within, everi, commit, kill

FREX: constabl, petti, chap, feloni, shire, session, leet, verg, treason, currant, abjur, supremaci, punish, villag, peac, breach, recus, praemunir, offenc, mispris, affray, accessari, malici, kill, victual, jurisdict, oyer, oath, sanctuari, jesuit

Topic 93

Highest Prob: shall, king, heir, law, per, say, hold, land, one, knight, servic, quoth, church, make, take, man, fee, relief, may, baron, homag, see, word, writ, lord, first, ancient, come, england, part

FREX: homag, relief, escuag, custodi, lifetim, knight, baroni, earldom, advowson, patron, ward, lanc, chival, chapter, militari, utrum, loom, glanvill, haered, ind, baron, dic, slaveri, ecclesia, wardship, comrad, seisitus, goe, fee, manual

Topic 94

Highest Prob: king, law, say, parliament, will, shall, may, lord, statut, great, make, realm, let, now, person, hous, majesti, word, can, common, man, give, petit, time, never, take, edward, commiss, like, upon

FREX: loan, comptrol, petit, june, etcetera, imposit, privycouncil, messag, vic, pretext, martial, household, grievanc, enact, unlaw, supremaci, num, magnacarta, project, marin, commiss, parliamentari, lend, march, projector, empson, excis, breve, wale, dudley

Topic 95

Highest Prob: good, man, make, one, thing, will, seem, say, see, may, great, quoth, colour, take, kind, mani, time, therefor, yet, evil, shall, part, natur, knowledg, matter, first, place, come, without, well

FREX: reprehens, gradus, reprehend, malum, blossom, spur, schoolman, alterc, melior, incept, understat, belief, contrariwis, colour, start, luck, evil, felic, angel, deceiv, alchemi, distemp, realiti, stoic, plural, planet, mala, burden, aesop, badg

Topic 96

Highest Prob: say, act, shall, statut, henri, make, king, case, word, may, upon, heresi, colleg, time, imprison, one, caus, come, can, without, land, edward, cap, within, commiss, physic, parliament, london, censor, give

FREX: censor, colleg, heresi, physic, heret, sewer, art, rectori, imprison, tith, disobedi, religi, diocesan, gutter, doctor, dictam, median, gaoler, convict, nemo, maker, daniel, communiti, cessavit, commission, monasteri, warburton, repeal, maidston, convoc

Topic 97

Highest Prob: spirit, man, bodi, thing, life, long, old, part, also, will, live, may, repar, death, touch, much, age, last, tree, can, yet, year, now, natur, repair, good, sleep, way, come, therefor

FREX: repar, opiat, explic, aliment, nitr, opium, consumpt, repair, condens, inanim, diet, intener, vital, sleep, prolong, refriger, eager, spirit, arteri, robust, renov, consubstanti, porch, subordin, perspir, durabl, organ, flight, membran, arefact

Topic 98

Highest Prob: man, natur, part, knowledg, upon, use, bodi, may, mind, make, thing, shall, invent, one, will, great, can, yet, therefor, much, philosophi, like, imagin, particular, true, say, speak, two, find, first

FREX: defici, invent, anatomi, mahomet, medicin, mathemat, variabl, metaphys, vicissitud, calendar, inquiri, philosophi, proposit, faculti, physician, elench, imagin, magic, induct, rhetor, earthquak, knowledg, aristotl, principi, delug, neptun, cure, astronomi, scienc, diseas

Topic 99

Highest Prob: man, great, envi, upon, good, person, make, virtu, will, natur, like, rich, much, can, one, come, certain, may, yet, best, therefor, honour, busi, fame, well, time, say, common, shall, use

FREX: envi, rich, discontent, fame, bold, tacitus, braveri, envious, prais, bewar, proverb, flatter, virtu, deform, eunuch, discount, enrich, personag, ostent, beauti, despis, blister, scorn, vespasianus, mucianus, aesop, sedit, pluto, glorious, melancholi

Topic 100

Highest Prob: thing, natur, concern, inquisit, fold, four, may, histori, three, bodi, condit, see, seem, either, much, order, wherefor, also, earth, experi, consist, sometim, true, air, els, meet, will, interpret, inquir, togeth

FREX: fold, potenti, inquisit, alphabet, subjoin, delta, wherefor, celesti, meteor, artifici, histori, adapt, perspicu, speci, canon, arrog, interpret, interpos, mass, consist, repetit, meet, four, sake, durabl, concern, proposit, els, design, monstrous

Appendix E

Table E1: Pairwise topic correlations for 25-topic STM

	Understanding Law	Jurisprudence	Disambiguating Law	Property Rights	Land Inheritance Law	Religion, Law, & Truth	Defendant Rights	Regulation of Exchange	Public & Private Authority	Legal Jurisdiction	Criminal Justice System	Constitutional Law	King, Law, & Nation	King & Court	Dynastic Politics	Foreign Relations	Civic Knowledge	Human Nature	Botany	Pharmacology	Physics, Air & Sound	Physics, Energy	Extracting Meaning	Probing For Facts	Epistemology
Understanding Law	1.000	-0.012	-0.020	0.058	0.026	0.161	-0.049	0.016	-0.077	-0.036	-0.053	0.117	-0.054	-0.083	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.084	-0.064
Jurisprudence	-0.012	1.000	0.225	0.111	0.208	-0.069	0.106	0.091	0.055	-0.037	-0.030	-0.068	-0.100	-0.163	-0.070	-0.109	-0.160	-0.108	-0.057	-0.086	-0.065	-0.071	-0.088	-0.096	-0.107
Disambiguating Law	-0.020	0.225	1.000	0.167	0.028	-0.047	0.119	0.030	0.026	-0.050	-0.062	-0.003	-0.095	-0.131	-0.058	-0.094	-0.134	-0.081	-0.038	-0.070	-0.053	-0.059	-0.073	-0.080	-0.086
Property Rights	0.058	0.111	0.167	1.000	0.458	-0.038	-0.031	0.045	-0.057	-0.047	-0.039	0.077	-0.097	-0.127	-0.057	-0.094	-0.127	-0.071	-0.058	-0.075	-0.058	-0.064	-0.075	-0.082	-0.089
Land Inheritance Law	0.026	0.208	0.028	0.458	1.000	-0.052	-0.011	-0.001	-0.061	-0.041	-0.002	0.005	-0.094	-0.140	-0.044	-0.087	-0.130	-0.075	-0.058	-0.078	-0.059	-0.066	-0.064	-0.087	-0.095
Religion, Law, & Truth	0.161	-0.069	-0.047	-0.038	-0.052	1.000	-0.050	-0.066	-0.017	-0.039	-0.025	-0.037	-0.063	-0.068	-0.046	-0.032	-0.026	-0.009	-0.048	-0.062	-0.045	-0.048	-0.040	-0.047	0.058
Defendant Rights	-0.049	0.106	0.119	-0.031	-0.011	-0.050	1.000	0.003	0.052	0.063	0.105	-0.002	-0.082	-0.118	-0.067	-0.108	-0.162	-0.114	-0.059	-0.094	-0.059	-0.075	-0.094	-0.072	-0.114
Regulation of Exchange	0.016	0.091	0.030	0.045	-0.001	-0.066	0.003	1.000	0.177	0.056	0.025	0.054	-0.093	-0.144	-0.053	-0.103	-0.121	-0.097	-0.055	-0.081	-0.060	-0.067	-0.083	-0.085	-0.098
Public & Private Authority	-0.077	0.055	0.026	-0.057	-0.061	-0.017	0.052	0.177	1.000	0.024	0.035	-0.036	-0.068	-0.122	-0.044	-0.094	-0.147	-0.095	-0.048	-0.074	-0.053	-0.061	-0.077	-0.059	-0.089
Legal Jurisdiction	-0.036	-0.037	-0.050	-0.047	-0.041	-0.039	0.063	0.056	0.024	1.000	0.103	0.120	-0.052	-0.121	-0.052	-0.082	-0.141	-0.091	-0.050	-0.071	-0.052	-0.059	-0.074	-0.053	-0.088
Criminal Justice System	-0.053	-0.030	-0.062	-0.039	-0.002	-0.025	0.105	0.025	0.035	0.103	1.000	0.008	-0.025	-0.115	-0.057	-0.034	-0.143	-0.086	-0.054	-0.078	-0.058	-0.065	-0.078	-0.005	-0.096
Constitutional Law	0.117	-0.068	-0.003	-0.095	-0.131	-0.058	-0.094	-0.134	-0.081	-0.038	-0.070	-0.053	-0.071	-0.058	-0.094	-0.134	-0.081	-0.038	-0.070	-0.053	-0.059	-0.073	-0.080	-0.086	-0.107
King, Law, & Nation	-0.054	-0.100	-0.095	-0.131	-0.058	-0.094	-0.134	-0.081	-0.038	-0.070	-0.053	-0.071	-0.058	-0.094	-0.134	-0.081	-0.038	-0.070	-0.053	-0.059	-0.073	-0.080	-0.086	-0.107	-0.114
King & Court	-0.083	-0.163	-0.070	-0.109	-0.160	-0.108	-0.057	-0.086	-0.054	-0.078	-0.058	-0.065	-0.078	-0.005	-0.096	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.095
Dynastic Politics	-0.063	-0.070	-0.109	-0.160	-0.108	-0.057	-0.086	-0.054	-0.078	-0.058	-0.065	-0.078	-0.005	-0.096	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.095	-0.107
Foreign Relations	-0.077	-0.109	-0.160	-0.108	-0.057	-0.086	-0.054	-0.078	-0.058	-0.065	-0.078	-0.005	-0.096	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.095	-0.107	-0.114
Civic Knowledge	-0.113	-0.160	-0.108	-0.057	-0.086	-0.054	-0.078	-0.058	-0.065	-0.078	-0.005	-0.096	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.095	-0.107	-0.114	-0.114
Human Nature	-0.096	-0.108	-0.057	-0.086	-0.054	-0.078	-0.058	-0.065	-0.078	-0.005	-0.096	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.095	-0.107	-0.114	-0.114	-0.114
Botany	-0.054	-0.086	-0.065	-0.078	-0.005	-0.096	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.077	-0.113	-0.096	-0.054	-0.083	-0.063	-0.068	-0.074	-0.095	-0.107	-0.114	-0.114
Pharmacology	-0.083	-0.063	-0.068	-0.074	-0.095	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059	-0.089	-0.059
Physics, Air & Sound	-0.063	-0.065	-0.053	-0.058	-0.059	-0.045	-0.059	-0.060	-0.053	-0.052	-0.058	-0.065	-0.062	-0.080	-0.035	-0.046	-0.061	0.006	0.003	0.060	1.000	0.081	-0.029	-0.049	0.010
Physics, Energy	-0.068	-0.071	-0.059	-0.064	-0.066	-0.048	-0.075	-0.067	-0.061	-0.059	-0.065	-0.074	-0.060	-0.090	-0.040	-0.061	-0.081	-0.042	-0.015	0.050	0.081	1.000	-0.045	-0.057	0.163
Extracting Meaning	-0.074	-0.088	-0.073	-0.075	-0.064	-0.040	-0.094	-0.083	-0.077	-0.074	-0.078	-0.090	-0.058	-0.080	0.000	0.063	0.116	0.035	0.010	0.091	-0.029	-0.045	1.000	-0.055	-0.028
Probing For Facts	-0.084	-0.096	-0.080	-0.082	-0.087	-0.047	-0.072	-0.085	-0.059	-0.053	-0.005	-0.092	-0.001	0.193	-0.027	0.014	-0.068	-0.059	-0.048	-0.068	-0.049	-0.057	-0.055	1.000	-0.070
Epistemology	-0.064	-0.107	-0.086	-0.089	-0.095	0.058	-0.114	-0.098	-0.089	-0.088	-0.096	-0.110	-0.064	-0.105	-0.050	-0.062	0.106	0.007	-0.040	-0.012	0.010	0.163	-0.028	-0.070	1.000