This is a final pre-publication version of the paper, which has been published in the *Review of Law & Economics*. It is virtually identical to the published version.

Lasting Legal Legacies: Early English Legal Ideas and Later Caselaw Development During the Industrial Revolution*

Peter Grajzl† Peter Murrell‡

January 2, 2022

Abstract

We explore English legal evolution by empirically investigating the relevance of late-medieval and early-modern legal ideas for caselaw development during the Industrial Revolution, an era of unprecedented societal change. To ascertain the prevalence of specific legal ideas in pre-1765 case reports, we draw on existing topic model estimates. We measure the relevance of those ideas for subsequent caselaw development using post-1764 citations to the pre-1765 cases. We show that deliberations on court cases heard between 1765 and 1870 systematically invoked a broad range of preexisting legal ideas. Strikingly, the strongest effects are exhibited by Coke-style analysis and precedent-based thought. A key legacy of early English caselaw therefore lay in bestowing modes of reasoning. The reason why a subset of preexisting legal ideas does not exert a detectable effect is that those ideas were generally no longer key to post-1764 legal disputes. Our approach to investigating legal development could be applied in many other contexts.

Keywords: English caselaw, legal development, Industrial Revolution, legal history, machine learning

JEL classifications: K10, K30, N43, P10, C81

*We are grateful to Martin Schmidt for sharing his data. For helpful comments and suggestions, we thank participants at the annual meeting of the French Association of Law and Economics, conference of the Online Workshop on the Computational Analysis of Law (OWCAL), seminars at Tel Aviv University and Washington and Lee University, as well as Sofia Amaral-Garcia, Geoffrey Hodgson, Michael Livermore, Christopher Engel (our editor), and an anonymous reviewer.

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

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

1. Introduction

The history of English law is, by and large, a history of caselaw. Generated through myriad judicial decisions rendered in the courts of law, caselaw was a key source of English law already by the middle ages. Statute law, in contrast, became a prominent source of English law relatively late, with the expansion of Parliament's lawmaking activity during the 18th century. Yet even then, caselaw remained critically important. Consequently, the many legal developments that took place during the Industrial Revolution occurred against the backdrop of centuries of accumulated court decisions.

Despite the path-dependence inherent in the development of caselaw, there are many reasons to expect discontinuity in legal development in the late 18th and 19th centuries. Importantly, "[t]he social and economic changes that thrust Britain forward from her early Georgian to her late Victorian condition were the greatest in her history" (Cornish et al. 2019: 1). Industrialization, urbanization, and population growth brought a host of substantively new legal problems concerning accidents, pollution, public health, and labor, to name just a few. Commerce was expanding rapidly both nationally and overseas, with corresponding developments requiring innovations in contracting and resolution of debt. Central and local governments were increasing their reach, with courts not always embracing the resulting changes.

At the same time, the courts and the judiciary were themselves in transition in the late 18th and 19th centuries. With the enhanced role of Parliament and the understanding that courts are usually able to address issues only ex post, legislation had become a much more important source of law. In addition, the courts, especially the Chancery, found themselves plagued by delays, high litigation costs, and venality, thus becoming objects of much public scorn. Spurred by the ideas of Bentham, who regarded the historically-conditioned common law as chaotic and advocated codification, considerations of legal reform became a permanent feature of public discourse (Judson 1910, Lobban 2000). All these developments and influences would have lessened the immediate applicability of prior accumulated legal wisdom for caselaw development in the Industrial Revolution. Thus, a first question we address in this paper is whether preexisting legal ideas were much used at all in caselaw development during the Industrial Revolution.

Having reached the unsurprising affirmative when answering this question, we then examine which specific sets of pre-industrial legal ideas were used especially prominently, and which less intensively, in court deliberations during the late 18th and 19th centuries. And, for those preexisting ideas that were no longer explicitly referred to in the industrial era, we examine whether that was a consequence of the fundamentally altered world or, alternatively, merely a reflection of the widespread acceptance of those ideas, which would have rendered explicit citations unnecessary. To date, these fundamental questions about England's legal development have not been subject to comprehensive scrutiny in a quantitative framework. Without doubt, the central obstacle has been the absence of suitable data. We assemble such a dataset, thereby providing the very first systematic empirical analysis of the legal legacy of early English caselaw for caselaw development in the Industrial Revolution.

Our dataset is derived from The English Reports (Renton 1900-1932) (ER, in short), a collection of reports on cases heard in the high courts of England before the middle of the 19th century. Our observations are pre-1765 cases, and we regress a measure of the use of each of these cases in the post-1764 period on measures of the different legal ideas present in each of the cases. To identify the legal ideas present in the pre-1765 cases, we draw on the work of Grajzl and Murrell (2021a; henceforth GM) who use the corpus of reports of cases heard before 1765 and estimate a topic model. As an unsupervised machine-learning technique, topic modeling offers a route to discovery of salient themes and features of legal reasoning that extend beyond those already known to readers of the ER or identifiable using supervised machine-learning methods. The GM estimates pinpoint the relative extent to which each of the 52,949 case reports features each of the 100 estimated topics, with each topic reflecting a distinctive aspect of English legal thought. The resultant measures of the prevalence of the different legal ideas featured in pre-1765 caselaw are our focal explanatory variables.

We measure the employment of the pre-1765 legal ideas in subsequent caselaw development using the number of post-1764 citations to each of the 52,949 pre-1765 cases. This is our dependent variable. A citation represents "a latent judgment" of legal professionals (judges and counsel) about the precedential relevance of the cited case to the citing case (Fowler et al. 2007, Cross 2010). Citations to legal documents such as judicial opinions or, in our context, case reports are thus a quantitative measure of the influence of those legal documents on the development of the law (Posner 2000; Cross 2010, 2012, Nelson and Hinkle 2018, Landes et al. 1998, Kosma 1998). All else equal, more frequently cited legal documents have more direct relevance for subsequent legal development.

We observe the post-1764 citation counts in two separate time periods: 1765-1815 and 1816-1870. Our dataset, in which the unit of observation is a pre-1765 report, is therefore a two-period panel, a structure that allows us to control for the independent effect of time on citations. We use a negative binomial model, regressing the number of citations to a specific case report on our measures of the presence of early legal ideas in the case reports. To address multiple-hypothesis-testing concerns, we control for the false discovery rate (FDR), an approach focused on choosing an acceptable expected proportion of incorrect rejections of the null hypotheses.

Our results show that deliberations during court cases heard during the Industrial Revolution directly referenced ideas from virtually every major area of medieval and early-modern law and legal thought. As one would naturally expect, comparatively large detectable effects are especially important for those major legal themes present in the pre-1765 corpus that have a relatively modern tenor, such as those pertaining to markets and organizations and debt, as well as those themes with perennial relevance, such as those on families and inheritance. Consistently, we find weaker effects for the sets of pre-1765 ideas that have a less modern tenor, such as real property. These are both areas of law for which industrialization raised legal issues that were of an entirely new character.

But our estimates also uncover surprises. For example, from all the broad areas of law that GM identified, legal ideas comprising ecclesiastical issues exhibit the strongest positive association with post-1764 citations, an unexpected finding given that the immense struggles over

church governance of the 16th and 17th centuries had now long passed. And we find considerable heterogeneity of effects across sets of related legal ideas. For example, even though ideas related to procedure usually have a comparatively modest effect, one specific procedural topic, on equitable relief, exhibits a strong positive association with post-1764 citations.

Strikingly, two early legal topics that exert an especially noteworthy effect on later caselaw are one on precedent-based reasoning and another capturing the distinctive style of legal analysis attributed to Edward Coke. Our estimates thus show that early English legal development not only imparted substantive law to the later era, but was also relevant via the effect of modes of reasoning.

Finally, we develop and implement an empirical methodology to investigate why certain sets of pre-1765 legal ideas are cited less after 1764. We do not find any evidence that non-appearance of post-1764 cites to specific pre-1765 ideas is a consequence of these ideas being no longer explicitly referred to simply because they were so widely accepted (see, e.g., Landes and Posner 1976, Posner 2000). This result confirms the interpretation of our findings given above: the set of the pre-1765 legal ideas that do not exert a detectable influence on post-1764 caselaw development were no longer substantively relevant to legal disputes arising during the Industrial Revolution.

Our paper thereby makes two primary contributions, one substantive and one methodological. On the substantive end, we advance the scholarship on the history of England's institutional development. With regard to law, the predominant focus of the empirically-oriented subset of this literature has been legislation (see, e.g., Hoppit 1996, 2017), some of which, such as estate acts, was an outcome of a Chancery-like petition process and evolved to correct pre-existing deficiencies in court practice (Bogart and Richardson 2009, 2010, 2011). In contrast, the ocean of caselaw has received hardly any quantitative empirical attention. In a recent contribution, Grajzl and Murrell (2021a, 2021b) investigate English caselaw developments in the centuries prior to the Industrial Revolution. We utilize the GM estimates to, first, characterize which specific legal ideas embodied in English caselaw before the Industrial Revolution were especially relevant to caselaw developments during the Industrial Revolution and, second, explore the reasons why some legal ideas were comparatively less relevant.

On the methodological front, we offer an example of how estimates obtained from computational text analysis—in our application topic modeling—can be productively utilized as input into conventional regression analysis. Consequently, we have a new, quantitative way of approaching what some have characterized as perhaps "the most important question facing judicial scholars: What explains the development of the law?" (Hansford and Spriggs 2006: 15). An additional methodological contribution is that we were able to use the unique characteristics of the GM dataset to distinguish between non-citation because of broad acceptance of an area of law and non-citation because an area of law was no longer useful. Thus, we address a methodological problem that has long plagued the literature using citation data as a dependent variable.

We proceed as follows. In Section 2 we provide contextual background on the history of English law. Section 3 introduces our data and variables. In Section 4 we develop our empirical approach. Sections 5 through 7 present and discuss our results. Section 8 concludes and considers research that could build on that documented here.

2. A Primer on English legal history

In this section we provide a brief overview of selected aspects of English legal history in the centuries prior to and during the Industrial Revolution. Our review is intended to provide the minimum necessary background for those readers with no or limited prior knowledge of the history of English law. Given our focus on caselaw and its associated legal ideas, we highlight especially those historical features and developments that are relevant to the activity of the courts, as opposed to legislation or regulation per se.[1]

2.1. Early English Common Law and Equity

English common law emerged in the course of the 12th century. During the reign of Henry II (1154-1189), the King's court was already a focal point of royal administration. Unlike alternative modes for resolution of disputes, the dispensation of justice in the royal courts provided an effective process, a written record, and enforcement backed by the king. Accordingly, the common law of the realm, administered by the royal courts, was by the mid-13th century a "fully fledged juristic entity, with its own specialist practitioners, its own technical language and literature" (Baker 2019: 34). The legal profession, and in particular the judiciary, soon became an immensely powerful actor that would critically shape legal, political and economic developments in the centuries to come.

The central common law courts, located in Westminster, were the King's Bench, the Court of the Common Pleas, and the Court of the Exchequer. In an era when courts were not organized in modern-day hierarchies, these courts were both courts of first instance and, to the extent that judgments could be challenged, courts of appeal. The system of assizes, whereby justices of the central courts travelled across the country to hold sessions, facilitated access to royal justice outside of London.

From its inception to the 19th century, the common law relied on an extensive and elaborate system of writs, a strict and rigid scheme of procedural rules that conditioned the ways in which a complaint could be brought to court. Over time, legal practitioners invented a variety of formulae that prescribed the corresponding forms of legal action, such as that of trespass, on the case, and in assumpsit. Conditional on the appropriate writ, an elaborate process of pleading revealed the essence of a dispute and the factual details relevant for the jury. Pleading was the cornerstone of

[1] English legal history has been the subject of voluminous scholarship. In the interest of brevity, we purposefully omit in-text citations to specific references that this section is based on. Our overview draws especially on Baker (2019) and Cornish et al. (2019), two recent, comprehensive, and complementary works on the subject. Baker (2019) surveys the history of the characteristics and institutions of English law, with a focus on the evolution of common law prior to the second part of the 18th century. Cornish et al. (2019), in contrast, center on the English legal developments from the start of industrialization to mid-20th century. For further recent references on aspects of English legal history highlighted in this section, see Duxbury (2008), Lobban (1991), Lieberman (1989), Brooks (1998), and Harris (2009). Older pertinent treatises of English legal history include Maitland and Montague (1978 [1915]), Plucknett (1948), Kiralfy (1962), Harding (1973), Jenkins (1938), and Allen (1964). Holdsworth's *A History of English Law*, published between 1903 and 1966, comprises 17 volumes. A recent ongoing comprehensive attempt on the subject is *The Oxford History of the Laws of England*, currently consisting of six volumes, but at the time of writing containing no volume covering the period 1559-1819.

early common law and had to be mastered by any budding lawyer. Accordingly, early reports of court cases centered on pleading and provided little, if any, information about judgments per se.

An emphasis on reasoning about points of law emerged after written pleadings replaced oral pleadings in the 15th and 16th century and when improvements occurred in the quality of available law reports. The introduction of written pleadings enabled the disputing parties to clarify the applicable point of issue before the trial. The focus of litigation thus naturally shifted onto the court's decision on the agreed-upon disputed issue, thereby increasing the value of referring to judicial decisions made in similar cases in the past. A distinctive emphasis on precedent-based reasoning, a core aspect of the English legal tradition, thus arose gradually, with the modern-day notion of a binding precedent (*stare decisis*) solidifying only sometime in the late 18th or 19th centuries. At the same time, the growing importance of legislation after the end of the 17th century meant that the courts increasingly concerned themselves with, and developed doctrines for, the interpretation of legislative acts.

The highly formalistic culture of the common-law courts and their rigid adherence to due process exposed the need for alternative remedies and simpler procedures. Equity, as administered by the Court of Chancery, provided a separate, complementary area of law that relied on less strict pleading rules, used an inquisitorial procedure, and provided relief in the form of decrees and injunctions. From the early 16th century onwards, judges in Chancery were typically common-law-educated lawyers. It is perhaps also for this reason that, from the mid-17th century onwards, Chancery cases were regularly reported on and precedent became as important in equity as it was already becoming in the common law. After the 15th century, Chancery's business and jurisdiction expanded. Not unsurprisingly, the relationship between Chancery and the common-law courts was not always an amicable one, although it became more harmonious as the 17th century proceeded.[2]

Much of the early common law was the law of real-property, founded on a feudal understanding of lordship and tenure. Gradually, the law developed a distinction between legal title and beneficial ownership, leading to development of the law of uses. Legal action based on disputes arising out of contractual relations were initially possible under the actions of covenant and debt, then later trespass was used and, eventually, assumpsit.

Jurisdictions changed over time, with the gradual accretion of authority by the common-law courts and Chancery. Various local courts specializing in merchant matters declined in popularity or were legislated out of existence in the 16th century. The Admiralty court, which had much business in commercial matters, was vanquished by the common-law lawyers in the 17th century. The jurisdiction of ecclesiastical courts diminished over time, with common law courts increasingly acquiring jurisdiction over tithes and, in the 16th century, over the tort of defamation. Nevertheless, the ecclesiastical courts retained authority into the 19th century over many family matters and wills. Criminal procedure differed substantially from the procedure used in the private-law domain. With the exception of rare cases featuring important questions of law, which were

[2] The height of hostility was the well-known 17th-century clash between Lord Ellesmere and Sir Edward Coke, at the time chief justice of the King's Bench. Coke's dismissal, Ellesmere's death, and the appointment of Sir Francis Bacon as Lord Chancellor marked the start of a more harmonious relationship between Chancery and common law judiciary.

debated in the central courts, adjudication of criminal cases started and ended in the localities, and thus were rarely reported on.

2.2. Caselaw Developments During the Industrial Revolution

Industrialization, urbanization, and population growth in the late 18th and 19th centuries gave rise to a multitude of new legal issues that spurred developments in many areas of law. Earlier, upon the demise of feudalism, the complex set of inheritance arrangements known as strict settlement had become a means of preserving large family estates. The act of devising the corresponding trusts was a major source of income for the legal profession. Land was increasingly used to produce industrial inputs. Mortgaging of land already enabled access to credit. However, a mortgage became an especially productive financial instrument once the courts ruled that deposit of a title deed with the lender was sufficient to create an equitable mortgage. Financing of new housing now often occurred with the help of building societies and the courts had to address questions pertaining to their contracting rights. Compulsory purchases promoted the building of roads, canals, and railways. The accompanying land valuation issues were, however, the domain of the courts.

As industrialization took off in the 19th century, so too did the risks faced by the members of society. Accidents were acknowledged as a social problem and the courts had to concern themselves with the associated disruptions. Given the largely unchartered legal territory, the emerging caselaw of accidents necessitated many refinements. For example, as the courts settled on the doctrine of negligence as an approach to ex-post compensation for harm, they faced the challenge of clarifying the notion of duty of care. The possibility of awarding damages stimulated a debate about the appropriate magnitude and scope of damage awards. The use of vicarious liability, whereby principals could be sued for agents' actions, required the delineation of the domain of relationships to which it could and should apply. The judges had to adopt a stance toward possible exemptions from liability. Once safety legislation increased in scope, the judiciary needed to decide whether industrial claims could be founded solely on the employer's breach of a statute.

Industrialization caused pollution and public health issues. In the early industrial era, these problems were addressed via the old common law of nuisance. As it became clear that the courts lacked adequate expertise and could only act after the harm had already occurred, environmental and public health issues came to be addressed through legislation and regulation. The courts, however, did not always welcome the expansion of judicial-like powers awarded to the newly founded administrative bodies. In the context of the associated litigation, the courts strove to protect private property rights.

The late 18th and 19th centuries saw a surge in theorizing about contract law that was founded on a substantial body of precedent. In this period, the attention of the courts gradually shifted away from the older practice of rectifying elements of unfairness in a given bargain toward insisting on fulfillment of the agreed-upon terms. Accordingly, courts were increasingly inclined to discuss contract formation using terms of offer and acceptance. Important doctrinal differences arose between common law and equity with regard to contract. Common law, with its priority accorded

to procedure, focused on damages and emphasized privity of contract. Equity, in contrast, resolved contractual disputes by rescinding or rectifying agreements and granting injunctions. By placing a comparatively greater emphasis on non-written elements of a contract, equity was also better positioned than common law to discover parties' intentions.

Use of debt became more and more common. The 19th century saw a repeal of usury laws, abolishment of imprisonment for debt, and increasing use of secured credit. The flourishing economy witnessed the birth and the demise of many businesses, and thus bankruptcy was a central legal issue. The law of corporations was beginning to take shape. The early 18th century saw an increase in the number of unincorporated joint stock companies. The Bubble Act, a response to the 1720 stock market crash, prohibited the raising of transferrable stock without explicit charter or statute. The law, however, was not consistently enforced, at least until the early 1800s, and was eventually repealed in 1825. The modern limited liability company finally arose in the mid-19th century.

In the 19th century the judiciary revisited the role of law in promoting competition. Early common law judges, including Coke, viewed monopolies as restraints on trade, but were very much focused on monopolies created by the Crown. During the industrial era, judges at first viewed many private collusive arrangements as no more than enforceable contracts. Guild rules on employment did fall out of favor with the courts by the early 19th century, yet the judiciary still did not regard the law as an appropriate tool for active competition policy. The late 18th century and the 19th century also featured the emergence of the modern patent system. The courts played an important role in clarifying rules for what constituted a patentable invention.

As industrial activity became organized in factories and workshops, labor relations started to take on the modern form of wage-based permanent employment. Factories relied heavily on child labor. The issue of workday length was at the heart of a heated controversy. Trade unions were gaining ground. The judiciary, however, was generally hostile to the movement: acts of collective action were viewed as limiting the freedom of (labor) contracting.

Emerging family law focused on the relation between husbands and wives, and parents and children. Religious considerations were key in shaping the law on marriage and its dissolution. Ecclesiastical courts could not terminate a marriage, but rather only nullify it or order a suspension of the obligation to cohabit. Social changes, however, led to consideration of the option of a full divorce. Yet even before judicial divorce was a possibility, the secular courts had to sort out many legal nuances. For example, once separation via negotiation became a reality, the courts had to establish the rules for the interpretation of separation agreements, such as whether the separated wife could enter contracts. The Divorce Court, a specialized secular tribunal, was introduced in the last quarter of the 19th century.

In the domain of poverty alleviation and education, private charities occupied a central role. A much-debated legal issue was the use of charitable trusts, and in particular whether there could be deviations from the testator's or donor's presumed intention. At the same time, the poor laws, which included rules on geographic settlement, led to tensions between parishes, culminating in legal disputes about funding for the relief of poverty. The courts, for example, disapproved of the

practice of using charitable endowments to reduce the poor rate, a property tax earmarked for poor relief. Eventually, with the expansion of the workhouse model, much of the administration of poverty alleviation shifted from individual parishes to unions of parishes that were better able to absorb the growing financial burden of providing poor relief.

As in the earlier era, criminal law developments were less affected by the reasoning processes that had shaped the development of private law. The central courts continued to hear the cases that featured the most perplexing legal issues, such as the incorporation of mental elements or negations of *mens rea* in the context of a homicide. At the same time, despite increasing complexity of substantive criminal law, the judiciary vehemently opposed any attempts at codification.

Legal developments during industrialization of course did not occur in a vacuum. By the very nature of caselaw, earlier cases and corresponding legal ideas certainly must have played a role. But how important overall were early legal ideas? And which particular sets of earlier legal ideas would one expect to be especially relevant in the age of industrialization? Below, we offer some specific conjectures on this matter. To this end, we first clarify how we empirically measure the presence of different pre-industrial legal ideas and the relevance of those ideas in subsequent legal development.

3. The Variables: Data Sources and Some Conjectures about Anticipated Effects

3.1. Data sources

The source of our data is a digitized version of the ER (Renton 1900-1932), a compilation of 129,042 reports of decisions rendered in the English courts of law between the early 13th century and the mid-19th century. The cases covered by the ER are neither the universe nor a random sample of cases heard in the English courts. Instead, they constitute a specific selection of all cases considered in the English superior courts. Grajzl and Murrell (2021a: Appendix A) discuss the history of the ER and the mechanisms underpinning the choice of the reported cases that would eventually be included in the final edition. Reporters were especially keen on preserving a record of those cases highlighting aspects of law that were deemed novel or unsettled. Consequently, the ER should not be viewed as reflecting English legal development in its broadest sense. Nevertheless, the ER provide an immensely valuable record of those cases from the late medieval period to late 19th century that came to be used by lawyers as the basis for both legal precedent and legal education.[3] As such, the ER represents the key collection of case reports that practicing lawyers and judges viewed as providing insights into the nature of pre-19th-century English caselaw development. No comprehensive exploration of English legal history could be conducted without resorting to the ER. No alternative legal-historical corpus of comparable breadth and depth is available. And no additional legal corpus exists that could supplement the ER to any degree.

Figure 1 shows the distribution of the case reports included in our analysis by the year in which reported cases were heard in court. The reports on cases heard prior to the mid-16th century are scant, but coverage improves significantly thereafter. The mid-17th century decline in the

[3] It is perhaps for this reason that the editors of the ER in fact viewed their publication as reflective of the "Complete Verbatim Re-issue of the Decisions of the English Courts prior to 1866".

volume of reports is undoubtedly a consequence of the Civil War and the Interregnum: these events reduced both the amount of court activity and the systematic reporting of that activity. The less pronounced relative decline in the volume of reports in the second half of the 18th century most likely primarily reflects an overall decrease in the volume of litigation during that era, a trend that was accompanied by a contraction in the size of the legal profession. By the end of the 18th century, the demand for the services of the courts and legal business began to rise due to the demographic, economic, and social changes brought about by the Industrial Revolution (Brooks 1998: Ch. 4, 5). Throughout the era under consideration, the process of selection of the types of cases included in the ER may have varied in ways that remain elusive to modern-day researchers. Accordingly, our analysis can only be interpreted as reflecting English legal development at large under the assumption that this case selection has no confounding impact on our estimates. Obviously, this is an assumption we cannot test. Therefore, our less ambitious goal is to leverage the breadth of information in the ER to explore the extent to which different pre-industrial legal ideas were referred to in later court case deliberations. A finding that particular pre-industrial cases emphasizing specific legal ideas were frequently cited during industrialization would constitute direct evidence that the corresponding early legal ideas left a lasting legacy on subsequent caselaw development.

3.2. Focal Explanatory Variables: Measuring the Presence of Pre-1765 Legal Ideas

The units of observation in our regression analysis are the 52,949 pre-1765 case reports. Our focal explanatory variables, measured separately for each of these reports, capture key legal ideas that had been accumulated in the caselaw prior to the Industrial Revolution. We draw on the estimates generated by Grajzl and Murrell (2021a; henceforth GM). Utilizing the definitive version of the ER (Renton 1900-1932), GM assemble a corpus amenable to computational text analysis and then estimate a 100-topic structural topic model (Roberts et al. 2014, 2016) that characterizes the development of English caselaw and its associated legal ideas prior to 1765, the approximate onset of the Industrial Revolution.[4,5]

In topic modeling, the researcher specifies a model of the data generating process and then estimates the model's most likely parameter values using the corpus as data. With documents (here,

[4] For details on GM's pre-processing of the corpus and justification of STM modeling choices, including the number of estimated topics, see Grajzl and Murrell (2021a).

[5] Strictly speaking, we use estimates obtained on the basis of a structural topic model that is identical to the GM model in all but one respect: to forestall endogeneity concerns given the outcome variable used here (citations), we exclude the citation count for each case report from the set of metavariables that enter the topic prevalence equation of the structural topic model (STM). We have verified that the structural topic model estimates, reported by GM, do not change as a result. (Evidence in support of this point is available on request.) As in GM, we continue to model the topic prevalence equation in the underlying STM as a function of reporter name, identity of the adjudicating court, and the ER volume number. The inclusion of these metavariables leverages an important advantage of STM over earlier topic-model variants: the incorporation of metadata in the estimation improves the semantic coherence and exclusivity, and thus overall interpretability, of the estimated topics (Roberts et al. 2014: Online Appendix). In our setting, the use of a broad set of metavariables in the estimation of the underlying STM is especially warranted. Precisely-estimated topics mitigate the measurement error in the focal explanatory variables that we derive from the STM (see the remainder of Section 3.2) and that we then use to estimate the influence of pre-1765 legal ideas on caselaw development during the Industrial Revolution.

case reports) conceptualized as bags of words, an unsupervised machine-learning algorithm leverages the co-occurrence of word-use across documents to identify 'topics' (Blei 2012), each of which captures a specific salient emphasis within the corpus. Importantly, the topics themselves are solely the product of estimation; in this sense, topic modeling is an exercise in discovery. However, the interpretation of the substance of each topic, culminating in the assignment of topic names, is left to the researcher.[6]

The ER, and thus the GM corpus, span multiple centuries during which the English language certainly changed. Nevertheless, topic modeling is an especially suitable methodological approach for generating a macroscopic digest of the ER texts for three primary reasons. First, the system of writs and bills, which rigidly conditioned how legal action could be brought before a court, enforced continuity and relative stability in the use of language. Second, because topic modeling identifies topics based on co-appearance of words in documents, different substantive meanings of words, whether over time or across different reporters, would be reflected in the estimates as appearing in distinct topics.[7] Third, in pre-processing the corpus, GM eliminated Law French reports, standardized the English orthography, and translated Latin.

The topics generated by this exercise in machine-learning are probability distributions over word use. For readers unfamiliar with topic modeling, we provide three illustrative examples of the topics uncovered by the GM estimates.[8] One topic uses the following word-stems relatively heavily: 'case', 'reason', 'law', 'opinion', 'determin', 'think', 'object', 'question', 'differ', 'court', and 'cite'. The reports that use this topic intensively contain debates about which caselaw is relevant to the particular dispute. For example, in one case featuring this topic prominently a party argued "that there were many precedents in Tremayne, where such express words are omitted; and besides this is omitted in Ven. 110, which is a precedent of a case directly in point for the substance likewise of this mandamus". The same party then later "cited the case of The King and Sympson, Mich. 11 Geo. 1, where this point was determined. But the Chief Justice said, that in the case of The King and Harwood, Trin. the same year, that case was denied to be law; and he took it clearly that it was not" (Rex v Ward, 1 Barnardiston KB 411, 94 ER 277). GM name this topic Precedent.

[6] To name the estimated topics, one examines the key words associated with the topic and conducts a close reading of the documents that feature the topic most prominently.

[7] On the ability of topic-modeling to see through polysemy because meanings are embodied in combinations of word usage, not in single words, see DiMaggio et al. (2013). For example, in the context of an early topic that GM name Royal Patents & Tenures (see Grajzl and Murrell 2021a, Appendix E), 'patent' refers to the monarch's granting of an appointment, e.g. "A scire facias was brought to reverse a patent, in which the case was; King Charles the Second, anno 12, of his reign, grants the office of searcher to Martin, durante beneplacito…" (Rex v Kemp, Skinner 446, 90 ER 198). In contrast, in a case that prominently features, a comparatively recent topic, Publishing & Copyright (see Grajzl and Murrell 2021a, Appendix E), the use of 'patent' is related narrowly to publishing, e.g. "In action upon the statute against the defendant for printing and publishing an almanack, to their damage; special verdict finds that the usage of printing hath been regulated by the King, &c. and that he by patent granted them the sole printing of almanacks, and of the Common-Prayer Books…" (Corporation of Stationers v Seymor, 3 Keble 792, 84 ER 1015). For this example, the GM-estimated topic model thus clearly distinguishes between the nuances of meaning by identifying two distinct topics.

[8] For a detailed elaboration on the naming of each of the 100 topics, see Grajzl and Murrell (2021a, Appendix E).

Another topic is characterized by the key words 'coven', 'assign', 'breach', 'break', 'assigne', 'perform', 'accept', 'assur', 'accord', 'parti', 'indentur', and 'lessor'. The reports in which this topic appears extensively center on establishing and clarifying whether breach of contract has occurred. For instance, a report featuring this topic prominently states: "The breach assigned was, that the defendant lessor covenanted that it shall be lawful for the plaintiff, being lessee, quietly to enjoy the land; and that the lessor himself ousted him. And it was thereupon demurred; for this illegal ouster is no breach of the covenant" (Corus v Anonymous, Croke and Elizabeth 544, 78 ER 791). GM assign to this topic the name Identifying Contractual Breach.

The key words for the last illustrative topic include 'roll', 'coke', 'case', 'abr', 'elizabeth', 'report', 'inst', and 'law'. Many of the reports that feature this topic prominently are from Coke reporting on the activities of the King's Bench. The reports contain the distinctive style of reporting attributed to Coke who is widely credited for having revolutionized the practice of case reporting.[9] Unlike his predecessors, Coke insisted on a careful scrutiny of pertinent points of law and drew on a variety of past decisions and historic sources, while including many remarks and asides. For example, one of Coke's reports featuring this topic prominently states: "In this case two points were adjudged: 1. That (h) a descent of a copyhold in fee shall not toll the entry of him who has right to the copyhold (A), which agrees with the resolution in Brown's and the other cases before. 2. That where the custom of the manor of Allesley, in the county of Warwick, was, that copyhold lands may be granted to any person in feodo simplici, that a grant to one and his heirs of his (i) body is within the custom: for be it a fee-simple conditional, or an estate tail, it is within the Custom (B). So he may grant for life or for years by the same custom, for an estate in fee-simple includes all (C); and it is a maxim in law, cui licet (D) quod majus, non debet quod minus est non licere (D). (h) Ante 22 a. Cro. Jac. 36. (A) The doctrine of descent cast does not affect copyhold or customary estates, because the freehold is in the lord, ace. Poph. 33. 35. March. 6. Goodtitle d. Falkner v Morse, 3 T. R. 368. Doe d. Cook v Danvers, 7 East, 299. 3 Smith, 291. Gilb. Ten. 160. Adams Eject. 42. Vin. Abr. Descent N. pl. 1. Com. Dig. Copyhold E. Bac. Abr. Copyhold B. Ante 22 a. p. 322. n. (Q)…" (Gravenor v Todd, 4 Coke Report 23a, 76 ER 922). This topic is named Coke-Style Reporting.[10]

The names of all 100 estimated topics are listed in the first column of Table 1. Following GM, we classify the estimated topic into 15 broad themes, with each theme depicting a particular area of law and legal thought that is featured prominently during the pre-1765 era.

In topic modeling, documents are mixtures of topics. Thus, GM's estimated model allows the researcher to pinpoint the extent to which each of the 52,949 reports on pre-1765 cases feature each of the estimated 100 topics. Some case reports might heavily emphasize a specific topic (e.g. Identifying Contractual Breach), implying that the corresponding reports devote a large amount of

[9] See, for example, Coquillette (2004: 314), Boyer (1997), Plucknett (1948: 265), Harding (1973: 200), Allen (1964: 208-209), Holdsworth (1938: 122), and Lewis (1932: 234-238).

[10] For five topics, we chose slightly different names than those in GM publication, to better reflect their substantive content. We renamed Interacting in Court into Decisional Logic, Coke Reporting into Coke-Style Reporting, Keble Reporting into Keble-Style Reporting, Modern Reporting into Modern-Style Reporting, and Coke's Procedural Rulings into Coke-Style Procedural Rulings.

content to that topic. Other reports might feature many topics, with no particular topic clearly dominating. The topic proportions, one for each of the 100 GM-estimated topics, therefore vary across the 52,949 reports of pre-1765 cases. These topic proportions constitute our focal explanatory variables.

The third and the fourth columns of Table 1 show the descriptive statistics for each of the 100 topic-proportion variables, with variable names corresponding to those used by GM. The topic that is featured most prominently in the reports is Precedent (with a corpus-wide topic prevalence of 3.81 percent). A number of prominent topics are on court procedure (e.g. Motions, Decisional Logic, Correct Pleas, Writs of Errors, Procedural Rulings on Actions, Procedural Bills). Indeed, the procedure theme features the largest number of topics and is on average the most prevalent theme in the corpus, with as much as 28.2 percent of all reports devoted to procedure. As stressed in our overview of English legal history, an emphasis on procedural rigor and consistency was already at the heart of the operation of English courts by the late medieval period.

3.3. The Outcome Variable: Citations as a Measure of Influence

As our dependent variable, we use counts of the citations to the pre-1765 case reports that appear in the post-1764 case reports. Intuitively, if a legal idea that was developed during the pre-1765 era exerted an influence on caselaw development after 1764, then the pre-1765 case reports that feature the idea prominently should be cited relatively more frequently in the post-1764 case reports. The number of citations to a case is a quantitative measure of the precedential relevance of the case, and thus of the legal ideas featured prominently in the case.

In general, legal cases, judicial opinions, or case reports tend to be cited because they resolve a particular legal question, pose a novel legal challenge, represent an advance in legal reasoning, or otherwise help clarify elements of a legal doctrine (see, e.g., Choi et al. 2010, Posner 2000). As such, citations to a case or report are an indicator of the influence of the cited case or report on subsequent legal development (see, e.g., Landes and Posner 1976, Landes et al. 1998, Posner 2000, Choi et al. 2010, Cross and Spriggs 2010, Black and Spriggs 2013, Nelson and Hinkle 2018, Fix and Fairbanks 2020).

The ER do not come with citation counts for individual reports. Therefore, to obtain a citation count for each of the 52,949 reports on cases heard prior to 1765, we drew on, and augmented, the database of citations produced by Schmidt (2016). The main challenge with correctly identifying citations in the ER is the plethora of abbreviations used for reporter names, a key element of a citation that additionally consists of report volume and page number. To address this problem, Schmidt (2016) first compiled a comprehensive list of abbreviations for each reporter name. He then used the resulting reporter name abbreviation lists to convert all reporter abbreviations featured in the reports into a uniform format, one for each reporter. The use of a common format for each reporter name is crucial for ensuring a high degree of accuracy in identifying instances of citation.[11] The reported degree of accuracy of Schmidt's (2016: 65-70) citation extraction algorithm is high: Schmidt's manual analysis of a random sample of reports revealed that more than 80

[11] For operational details of this process, see Ch. 3 and Appendix A in Schmidt (2016).

percent of citations appearing in the reports were correctly identified by his citations-identifying computational algorithm. Additionally there was a very low rate of the recording of citations that do not actually appear in the actual reports (less than 2 percent). Notably, Schmidt (2016) did not find any systematic error patterns in the application of his algorithm for identifying citations. Among the different citations in the sample manually examined by Schmidt, more than one half of citations appear in the context of arguments made by counsel and a quarter in the context of arguments made by judges. The remaining citations can be attributed to the reporters themselves and the report editors.

Schmidt (2016), however, focused on the 18th century and beyond, and therefore his database needed to be supplemented for the years before 1700. The missing information was primarily in the dates of cases. Murrell (2021) filled in missing dates using three procedures: visual inspection of the pertinent pages of the volumes of the ER, using regularities in the relationship between the pages of volumes and dates, and finally a randomization procedure within the years covered by four smaller sets of very early reports. This process naturally led to some checking and correcting of errors in the correspondence between citing and cited cases. The discovered errors were few.

Because our interest is in measuring the influence of legal ideas on the subsequent development of caselaw, we include in our citation count the citations appearing in all post-1764 cases. Furthermore, we do not distinguish between string citations and citations that involve an in-depth discussion of a past case, or between positive and negative treatments of a past case. From the substantive standpoint, there is little reason to discriminate between different types of citations on some a priori normative grounds. In particular, even string citations and citations to prior erroneous decisions still clearly contextualize legal discussion, and thereby provide evidence of influence of a past case on the ongoing development of caselaw (see, e.g., Cross 2010: 521-522, 529; 2012: 724; Landes et al. 1998: 273; Posner 2000: 385; Fix and Fairbanks 2020: 813).[12]

For each of the 52,949 reports on cases heard prior to 1765, we observe the post-1764 citation count twice: during the years 1765-1815 and during the years 1816-1870. We choose the year 1815 as the end year of the early post-1764 sub-period for two major reasons. First, the number of reports included in the ER drops sharply after 1865, the year typically viewed as marking the final year of the reports.[13] The choice of the year 1815 as the end year for the early post-1764 sub-period thus

[12] Inspection of random samples of the reports reveals that the citations featured in the ER are overwhelmingly instances of positive citations that draw analogies to prior similar cases. As in modern-day legal opinions (see, e.g., Cross and Spriggs 2010), instances of negative citations in the ER seem to be exceedingly rare. Based on his manual analysis of a random sample of reports, Schmidt (2016: 68) estimates the presence of negative citations to be less than one percent. Murrell's (2020) analysis of his own random sample suggests that the share of negative citations is less than three percent, with a caveat that even when instances of negative citations could be identified, "the adjective 'negative' is perhaps too strong to capture the tenor of the report".

[13] The number of the reports equals 1097 in 1863, 848 in 1864, 779 in 1865, 205 in 1866, and then drops to 41 in 1867. The editors of the ER (Renton 1900-1932) viewed their publication as covering "the Decisions of the English Courts prior to 1866". Veeder (1901) likewise views 1865 as the last year of the reports. Precise reasons for the availability of a small number of reports on cases heard after 1865 have been lost in history. Given their presence in our digitized database, we include the post-1865 reports in our count of citations for completeness.

splits the 1765-1865 era into two sub-periods of comparable length. Second, the year 1815 marks the end of the Napoleonic wars, a decisive moment in English history.

Our dataset, in which the unit of observation is a pre-1765 report, is therefore a two-period panel. As we clarify below, the panel structure of the data allows us control for the impact of more general legal, societal, and reporting changes that arose during the Industrial Revolution and that might have shaped legal development independent of the effect of specific legal ideas that arose during the pre-1765 era. The split of the post-1764 era into two periods gives rise to the most parsimonious data structure that facilitates controlling for such factors.

Table 2 shows the distribution of the incidence of citations to pre-1765 reports during the entire 1765-1870 period as well as separately for the early (1765-1815) and the late (1816-1870) post-1764 sub-periods. The vast majority (74 percent) of the pre-1765 reports are never cited. 99 percent of the reports are cited less than a dozen times. The mean number of citations per pre-1765 case that was accumulated during the entire 1765-1870 period is less than one. The overall incidence of citations is somewhat higher in the later (1816-1870) period than in the earlier (1765-1815) one. This pattern coincides with, and likely reflects, the overall steady increase in the incidence of case reports, especially in the 19th century (see Figure 1).

3.4. Anticipated Effects: Some Basic Conjectures

Which specific sets of pre-1765 legal ideas would we anticipate to be most relevant for post-1764 legal development, as captured by citations to the pre-1765 cases that feature those sets of ideas prominently? For many areas of law, traditional legal-historical research provides only limited guidance on this issue: the absence of systematic data has made it difficult to articulate, or even imagine, specific relationships. The conjectures we offer are therefore a mixture of our own insights and selected observations from the legal-historical literature.

We would certainly expect continued relevance for those sets of early legal ideas that were most likely to grow in importance in the new world of industrialization, commercial growth, and population expansion. From the 15 themes cataloging our 100 topics, we thus expect to observe a comparatively strong positive association in our regressions for topics relating to markets and organizations, debt, and personal property. We would also expect to uncover a strong positive association for legal ideas pertinent to families and inheritance. These two perennially relevant legal domains perhaps even increased in importance as family wealth increased in the new era of economic growth, and at the onset of industrialization there was a comparatively developed, applicable body of caselaw, for example, on the settlement of family estates and design of wills.

Contract is one area of law for which the legal-historical scholarship has investigated continuity between the pre-industrial and industrial periods. Some scholars have argued that the pertinent law experienced a rather radical transformation in the late 18th century, stimulated by the widespread growth of industrial capitalism and an emphasis on laissez-faire ideology (Atiyah 1979, Horwitz 1977). Cornish et al. (2019: 196) summarize that view by stating that "an older idea of contracts" that had relied on "relationships generated by reliance or receipt of benefit" is said to have given way to "a new idea" under which the courts began to consider executory promises and

started awarding expectation damages in the event of breach. Under this view, therefore, preexisting legal ideas on contracts would not have been especially relevant to subsequent caselaw developments.

Other scholars have expressed skepticism about this sudden-transformation theory, noting that many elements of the 19th-century law on contracts had already been present in English law during a much earlier period (Baker 1980, Simpson 1979, Holden 1951, Ibbetson 2013, Liebermann 1989: 99, 111). Proponents of this alternative view believed that there was considerable "continuity between the law of the 1700s and 1800s" (Cornish et al. 2019: 197). Thus, the overall relevance of preexisting contracting ideas on subsequent legal developments would have been greater than afforded under the sudden-transformation perspective. In the area of contract law, the literature thus does not offer clear cut predictions on continuity.

Ex ante, we would expect relatively less relevance during industrialization for early legal ideas on real property and torts. Intuitively, with the ending of feudalism and new usages of land as an industrial resource, preexisting legal ideas on real property would have naturally lost much relevance. Similarly, while ideas on torts were featured in early legal discourse (e.g. defamation, nuisance), the overall scope of relevant legal issues was comparatively narrow and not readily applicable in the industrializing world where entirely new type of accidents became prominent. With rapid growth in urbanization, the old law of nuisance, for example, was soon viewed as inadequate for addressing the type and scale of new legal problems (Cornish et al. 2019: 157-158). In the same vein, it would also be natural to assume that earlier ecclesiastical law would not have had much relevance in the later period. By 1765, England's two centuries of religious struggles were over. The place of Protestant Dissenters and Catholics had been defined by legislation and the state was firmly in control of the Church of England.

We close this discussion of conjectures with two points. First, the above conjectures apply to themes, that is general areas of law. Given the 100 distinct topics captured by our explanatory variables, it would be intractable to list ex ante conjectures about the relative magnitudes of effects for each of the individual topics. But many of our themes are collections of very heterogeneous topics. Indeed, as we clarify in the discussion of our results below, the data do show that, within themes, there is considerable heterogeneity across the constituent topics with respect to empirically detectable effects. Second, we assume that an already well-developed area of law is still cited, even when it is so well accepted that it is embedded deeply in the minds of the legal profession. This, in fact, is a controversial point in the literature that uses citations as a dependent variable (see, e.g., Landes and Posner 1976, Posner 2000). Fortunately, in contrast to most studies that use citations as a dependent variable, we are able to test this assumption and find that our results are not consistent with the view that non-citation arises from complete familiarity (see Section 7).

4. Methods

4.1. Empirical Model

Our dependent variable, the count of the number of citations, is characterized by overdispersion. We therefore utilize a negative binomial model.[14] Accordingly, we posit the following model of the conditional mean of the number of citations, y_{jt} to pre-1765 case j that were accrued in post-1764 time period t:

$$\mathrm{E}\left[y_{jt} \middle| Z_{ij}, \lambda_t, \kappa_\tau \right] = \exp\left\{ \beta_0 + \sum_i \beta_i Z_{ij} + \lambda_t + \kappa_\tau + \lambda_t \kappa_\tau \right\}. \quad (1)$$

β_0 is the regression constant. Z_{ij} is the proportion of case report j that is devoted to topic i. (Recall that every report is a mixture of topics.) The β_i's are the coefficients of primary interest. λ_t and κ_τ are fixed effects, to be discussed below.

We investigate the relation between the prevalence within a case of the pre-1765 topics and the post-1764 citations to that case. For each topic, a detectable positive relationship (positive β_i) is an indication of direct relevance of the corresponding pre-1765 ideas for post-1764 legal development. But we are also interested in the variation across topics of the strength of the estimated connections. Because the topics vary considerably in their prevalence across pre-1765 reports (sometimes by a factor of 40, see Table 1), ease of comparison across topics of the numerical values of the estimated coefficients is facilitated by using standardized topic proportion variables. Each Z_{ij} in expression (1) has a mean equal to zero and a standard deviation equal to one. Thus, each β_i reflects the effect of a one-standard-deviation increase in topic prevalence. Therefore direct comparisons between the estimated effects provide straightforward insights into whether elevated prominence of a legal idea in the pre-industrial era translates into a relatively important use of that idea in court case deliberations after 1764.

Topic modeling constrains the sum of all (non-standardized) topic proportions to be one for any single case report. Thus, for each case report j a weighted sum of standardized topic proportions equals zero (see the Appendix, especially expression (A13)). Therefore, omission of one topic proportion (Z_{ij}) is unavoidable for identification; dropping the regression constant would not allow one to identify the coefficients on all 100 standardized topic proportion variables.[15] Thus,

---

[14] None of our findings change if we instead use a quasi-maximum likelihood (QML) Poisson approach. (Estimates based on the QML Poisson approach are available upon request.) When the functional form for overdispersion is correct, the negative binomial estimator will be more efficient than the QML Poisson estimator, even though the latter is still consistent. The negative binomial approach, however, is less robust to distributional misspecification than is the QML Poisson approach.

[15] This is a critical analytical point, showing that, in a setting where non-standardized explanatory variables sum to a non-zero constant (for us, the value one), the typical intuition about models with non-standardized explanatory variables does not carry over to models with standardized explanatory variables. The difference follows from the fact that when a set of non-standardized variables sum a non-zero constant, that set of variables plus the constant (and no smaller set) are linearly dependent, while, in contrast, the corresponding set of standardized variables alone (without the constant) are themselves linearly dependent. Thus, when the sum of non-standardized explanatory variables equals a non-zero constant, there is a fundamental distinction when it comes to options concerning parameter identification between (a) a model with non-standardized explanatory variables and (b) a model with the corresponding standardized

the number of standardized topic proportions included in our estimated model equals 99, one less than the total number of GM-estimated topics.

We chose Motions for the omitted topic. This is part of procedure, the largest theme in the pre-1765 corpus, and Motions is the most prevalent of the topics in the procedure theme. Of course, many legal ideas will affect future legal developments simply through their embodiment in professional practices, norms, and routines entrenched within the legal community, even if those ideas are not especially relevant to the substantive content of the cases to which they are applied. Procedure, in particular, encompasses the broad thematic element of law that is critical to the functioning of the courts, and is thus almost bound to have such an effect, especially given the tremendous importance of the writ system. However, we want to go beyond this influence that arises more or less autonomously due to the way lawyers frame their work. We want to focus on the influence of past caselaw on subsequent developments that arises because of the substantive social, economic, political, and intellectual features of the legal ideas embodied in that past caselaw. Therefore, we compare all estimated effects to that of Motions.

4.2. Controls

We include two different groups of time fixed effects as well as their interactions. λ_t is the fixed effect for the time period during the post-1764 era when a case was cited, either an early time period (1765-1815) or a later time period (1816-1870). κ_τ is the fixed effect for the specific time period of the pre-1765 era during which the reported case was heard in court. We split the pre-1765 era into eleven non-overlapping time periods.[16] $\lambda_t \kappa_\tau$ is the full set of interactions between the two groups of fixed effects.[17]

The inclusion of fixed effects for the time period during the pre-1765 era controls for the age of a case and therefore for (plausibly non-linear) depreciation of its precedential value (Landes and Posner 1976, Black and Spriggs 2013).[18] The inclusion of the fixed effects for 1765-1815 and 1816-1870 serves two distinct purposes. On the one hand, these fixed effects absorb the variation in citations to pre-1765 reports that arises simply because the overall volume of case reports, and thus the likelihood of citation of pre-1765 cases, is much higher in the later time period than the earlier one (see Figure 1). At the same time, these fixed effects absorb any variation that is due to the fact that the relevance of early legal ideas might have changed over the course of the Industrial Revolution. Conditional on the nature of societal changes and promulgated legislation, legal ideas

explanatory variables. Under (a), the coefficients on all non-standardized explanatory variables (in our case, 100 of them) could be identified simply by dropping the regression constant, even if the estimated coefficients themselves from such a model would not have a natural interpretation that corresponds to a comparative-statics scenario that could plausibly be observed in the real world. But under (b), the coefficients on all standardized explanatory variables (in our case, 100 of them) cannot be identified by simply omitting the regression constant.

[16] These are: up to and including 1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1764.

[17] This approach is therefore equivalent to including a full set of 22 indicators, one for each cell defined by the Cartesian product of the eleven time periods when the reported case was heard in court and the two time periods when it was cited.

[18] None of our qualitative findings change if we instead control for the age of the case using a quadratic polynomial in year of the case.

embedded in caselaw can become either more or less pertinent over time (see, e.g., Hansford and Spriggs 2006: 24). Thus, there will be an independent effect of time on the number of citations. The interaction of the pre-1765 era fixed effects with the post-1764 fixed effects absorbs any changes in the impact of a case's age that arise because the relevance of early legal ideas changes over the course of the Industrial Revolution.[19] These interactions also absorb any variation that is due to the fact that some pre-1765 cases are further in time from the era in which the number of citing reports is most voluminous (after 1815).

Importantly, we do not include court and reporter fixed effects, even though we know the identity of the adjudicating court and the reporter for each case. The reason is that, in our context, both the court and the reporter are endogenous to the case report's substantive content. Historically, English high courts competed over jurisdiction for specific legal cases (fees were a source of revenue). Relatively early on, each court became considerably specialized with respect to particular legal issues. For example uses, and afterwards trusts, were the domain of Chancery, while contentious criminal cases were usually heard at the King's Bench (Baker 2019: 270, 328, 544). Similarly, particular reporters were attracted to specific legal issues and therefore focused their reporting on those cases that featured these issues especially prominently (see, e.g., Helmholz 2015). Consequently, adjudication venue and the identity of the reporter for a particular case themselves depend on the substantive content of the case and are, in this sense, endogenous outcomes. Our interest lies in detecting the overall relation between the topical content of pre-1765 legal ideas and post-1764 development of caselaw, and not in the net effect of the former on the latter within courts or within reporters. Therefore, indicators for adjudicating court and reporter should not be used as controls: their inclusion would bias estimates of the effects on which we are focused. For the same reason, we also do not control for the length of the report as the length of a reporter's depiction of the case will depend on the substantive features of the case. [20,21]

4.3. Dispelling Multicollinearity Concerns

Given that there are 99 explanatory variables of direct interest, multicollinearity might be a cause for concern. Indeed, because topics are distributions over a vocabulary, some pairs of topics will naturally be correlated across documents.[22] There exists no standard approach to testing for multicollinearity in a negative binomial regression model. However, in order to ensure that we obtained quantitative insight into whether multicollinearity could be affecting our analysis, we

[19] The inclusion of the interactions of the pre-1765 era fixed effects with the fixed effect for the 1816-1870 period into the estimated model has virtually no impact on the point estimates of the IRRs for the individual topics, but marginally increases the precision of the estimates relative to the model when those interactions are excluded.

[20] Suppose, for simplicity, that the expected number of post-1764 citations, y^E, equals (i) $y^E=aZ+bC+cR+dL$, where Z is (standardized) topic proportion, C court, R reporter, L report length, and a, b, c, and d are parameters. We know that $C=eZ+u_1$, $R=fZ+u_2$, and $L=gZ+u_3$, where e, f, and g are parameters and u_1, u_2, and u_3 are error terms. Then, (ii) $y^E=hZ+u$, where $h\equiv a+be+cf+dg$ and u is a composite error term. Our interest lies in estimating $h=\mathrm{d}y^E/\mathrm{d}Z$ based on (ii) rather than $a=\partial y^E/\partial Z$ based on (i).

[21] Reporting on issues pertaining to Rulings on the Calendar, for example, will on average necessarily give rise to shorter reports than reporting on cases deliberating on issues about the Clarification of Legislative Acts.

[22] In our data, for example, the topics Transfer of Ownership Rights and Length & Expiry of Leases are positively correlated, implying that they tend to co-occur across the reports.

estimated an analogous OLS regression in which the outcome variable was defined as the logged number of citations plus one. We examined the variance inflation factors (VIFs) for the resulting regression. The mean and the maximum VIF of our focal explanatory variables (the standardized topic proportions) were 1.36 and 2.12, respectively, which are notably below the standard threshold value of ten that would be indicative of multicollinearity problems (Wooldridge 2013: 98). In sum, multicollinearity is not a concern for our empirical analysis.

4.4. Approach to Statistical Inference and Multiple Hypothesis Testing

We base statistical inference on heteroscedasticity-robust standard errors clustered at the level of case reporter. Clustering at the level of reporter is appropriate on two counts. First, because controlling for reporter fixed effects would be inappropriate in our setting (see Section 4.3), plausible reporter-specific common shocks remain unabsorbed. Clustering at the case reporter level addresses these concerns (Cameron and Trivedi 2015). Second, the genesis of the ER certainly entailed elements of selection between law reporters and, conditional on the choice of reporter, selection of particular reports authored by the chosen reporters (Grajzl and Murrell 2021a: Appendix A). Consequently, clustering at the case reporter level is justified on the grounds of sample design (Abadie et al. 2017). There are 104 reporter-based clusters in our data.[23]

After estimating the model, we test 99 coefficients for significance (one per included topic).[24] Given the large number of tests that we perform, we must address concerns about false positives that arise in multiple-hypothesis-testing settings.

Two distinct approaches to correcting for multiple hypothesis testing have been proposed in the literature (see, e.g., Dudoit and van der Laan 2008, Efron 2010). Under the family-wise error rate (FWER) approach, the researcher controls the probability of incorrectly rejecting at least one true null hypothesis. Under the false discovery rate (FDR) approach, the researcher instead controls the expected proportion of falsely rejected null hypotheses. Relative to the FWER, FDR has greater power in detecting true effects. Under FDR, however, one is less confident that all of the detected rejections are correct. Accordingly, the use of FDR methods is especially suitable in settings such as ours, where the benefit of detecting true positives plausibly exceeds the cost of false positives.

We use the Benjamini and Yekutieli (2001) FDR method that allows for arbitrary correlation between the uncorrected *p*-values (those relevant to conventional hypothesis testing). We set the FDR at the conventional value of five percent, implying that, in our analysis, the expected proportion of rejected null hypotheses that are incorrectly rejected will not exceed 0.05.[25]

[23] Case reporters who reported on less than ten cases have been grouped into a separate, other reporters, category.

[24] Recall that one theme, the procedure theme (in the regression with themes), or one topic, Motions (in the regression with topics), are always omitted for purposes of identification.

[25] We also explored the use of multiple FWER methods. Holm's (1979) FWER method in particular, much like Benjamini and Yekutieli's (2001) FDR method, allows for arbitrary correlation between the uncorrected *p*-values. Applying Holm's five percent FWER instead of Benjamini and Yekutieli's five percent FDR correction reduces the number of detectable effects by about a third. We also experimented with the FWER procedure developed by Romano and Wolf (2005a, 2005b, 2016), an approach that entails greater statistical power than earlier FWER procedures, including Holm's (1979). The corresponding bootstrap-based estimates did not always converge. When they did,

4.5. Interpretation of the Estimates

We close this section with three remarks concerning the interpretation of our estimates. First, we report the point estimates in the form of incidence rate ratios (IRR), where the IRR_i corresponding to a β_i in (1) equals $\exp(\beta_i)$. The value of $(IRR_i-1)\times 100$ is the percentage change in the expected number of citations when the prevalence of topic i increases by one standard deviation, holding all else equal. As we demonstrate in the Appendix, given that we always omit the topic Motions, the reported IRR for a particular standardized topic proportion captures the ratio of the IRR for the corresponding non-standardized topic to the IRR for the non-standardized topic Motions.[26] The reader should thus keep in mind that all of our estimated effects are relative to the effect of Motions.[27]

Second, as clarified above, all topic prevalences enter the estimated model in a standardized form. The relative magnitudes of the estimated IRRs are thus immediately informative of the relative strength of the connection between different pre-1765 legal ideas and caselaw development during the Industrial Revolution.

Finally, our estimates of the later effects of earlier legal ideas provide a picture of the legacy of early English law for caselaw development in the Industrial Revolution that has never been provided before in such a systematic quantitative fashion. However, it is important to stress that post-1764 cases could also rely on pre-1765 ideas indirectly, without citing the relevant pre-1765 cases, perhaps instead referring to recent post-1764 cases that themselves cite or otherwise incorporate pre-1765 legal ideas. This type of dependence will, by construction, not be captured by our estimates. In this sense, our estimated effects should not be interpreted as capturing the full causal influence of pre-1765 legal ideas on post-1764 caselaw development.

5. Core Results and Interpretations

Our estimation results are summarized in Table 3. The topic-proportion variables are first ordered according to whether their estimated effects are statistically significant or not. Then, within each of these two sets (significant, not significant), the topic-proportion variables have been ordered on the basis of the magnitude of $|IRR-1|$. In what follows we make use of the resulting estimates to address three broad questions informative of the legacy of early legal ideas for legal development during industrialization.

5.1. Were Pre-1765 Legal Ideas Relevant?

We first test the null hypothesis that the effects of all 99 topics included in the estimated model are indistinguishable from the effect of Motions (the omitted topic). The null is readily rejected: $\chi^2=150937$, with p-value<0.0001. This is a first indication that substantive legal ideas accumulated

however, our findings using the Romano and Wolf (2005a, 2005b, 2016) FWER method were very similar to those obtained under the preferred FDR approach.

[26] The ratio of the two IRRs is further adjusted for the extent of variability of the focal topic; see the Appendix.

[27] While identification constraints prevent us from ascertaining the effect of the topic Motions per se, descriptive evidence indicates that case reports featuring the topic Motions prominently certainly were cited. The mean number of citations to reports at or above the 90th percentile based on the prevalence of the topic Motions is 0.1746. The maximum number of citations for the corresponding subsample of the reports is 23.

prior to 1765 mattered for the development of caselaw during the Industrial Revolution: if none of the effects of legal ideas were important, then a joint test of all 99 effects relative to Motions would not reject the null hypothesis.

The test of joint significance, however, does not tell us how many of the early ideas exerted an effect on later caselaw developments. To answer this question, we turn to the testing of hypotheses on the effect of each topic-proportion variable. After the FDR adjustment, 71 out of 99 topics evidence a statistically significant effect (see Table 3). These significant effects are distributed across all themes: at least one topic within each of the 15 themes exerts a detectable effect. This finding indicates that caselaw developments during the Industrial Revolution drew on a broad range of caselaw that had developed during the earlier era.

5.2. Which Sets of Pre-1765 Legal Ideas (Themes) Were More vs. Less Relevant?

To obtain insight into which sets of pre-1765 themes were most relevant for post-1764 legal development, we next compute average IRRs by themes. To address the fact that some topics do not exert statistically detectable effects, we set the IRR for the corresponding topics equal to one. We then order the themes by the size of average IRR, here considered as a convenient summary statistic.

Figure 2 summarizes the results. Next to each theme name in Figure 2, the labels [exp. large or exp. small] indicate our conjectures (see Section 3.4) about the expected size of the effects of a theme's constituent topics. The resulting visualization thus allows us to readily assess our conjectures. Next to the bar indicating the average size of IRR's for each theme, we also provide information on the share of the topics comprising the theme that exert a statistically significant effect. That information is one indicator of within-theme heterogeneity with respect to effects of specific legal ideas. In what follows we highlight a subset of the insights that can be gleaned from Figure 2.

As anticipated, the topics comprising the themes personal property, inheritance and debt all exert comparatively large effects. Also as anticipated, real property and torts exhibit comparatively small effects, on average. However, families, which we had expected to exert a comparatively strong effect on post-1764 citations, is ranked only midway in terms of the strength of its effect.

The contract theme has a strong positive effect, and seven out of eight of its constituent topics are statistically significant. Overall, our evidence thus provides support for the view of English contract-law development as exhibiting considerable continuity, as opposed to featuring intrinsic novelty with the start of industrialization (see Section 3.4). Our analysis thereby casts a new, quantitative light on a long-standing debate among legal historians (see, e.g., Cornish et al. 2019: 196-197; Lobban 1991).

Strikingly, the theme exhibiting the strongest effect is ecclesiastical. This is a finding that we did not expect given that England's great religious struggles were far gone when our later period begins. However, in retrospect, these results are entirely explicable. The effect of the topic Temporal & Spiritual Jurisdiction is indeed not especially large (see Table 3) and this is plausibly a reflection of the fact that the boundary between state and church was relatively settled. However,

the topics Tithes and Ecclesiastical Appointments do exert comparatively sizeable effects. Ex post, this finding can be interpreted by recognizing that the courts still needed to address many mundane ecclesiastical issues:[28] Who was exempt from tithes and which income generating activities were exempt? What property rights over parish property did the local deacon have? Could a canon hold two church positions? What were the rights of village grandees vis-à-vis the church in local church appointments? Thus, early ideas pertinent to the ecclesiastical theme very much continued to be relevant to later legal disputes because the courts were responsible for resolving those disputes that the church did not have the authority to solve by fiat.

Finally, the theme for which the constituent topics exhibit the weakest effect on average is jurisdiction, the only theme for which the mean IRR is smaller than one. This result is partly driven by the effect of Equitable Relief, the only topic that exerts a negative effect (relative to Motions) (see Table 3). More generally, however, our findings provide evidence that early legal ideas on jurisdiction, a core subject occupying early common-law and equity courts, were on average not especially relevant after the start of industrialization.

Figure 2 visualizes our results on the basis of average topic effects by theme. Within themes, however, there is considerable heterogeneity. Some topics exhibit especially large effects. We highlight those topics, which constitute the key legacy of early English law, in the next section. And within all but two themes (ecclesiastical and personal property), there are topics that do not exert a statistically detectable effect on post-1764 citations. In Section 7, we return to further investigation of the topics that do not have a statistically detectable effect.

5.3. Which Particular Pre-1765 Legal Ideas Were Especially Relevant?

Based on the magnitude of its IRR, the topic that exerts the largest detectable effect on citations is, interestingly, Coke-Style Reporting. A one-standard-deviation increase in the prevalence of Coke-Style Reporting is associated with a 58.4 percent increase in the number of citations relative to the effect of Motions. In interpreting this result, it is important to emphasize that, even though the prevalence of the topic Coke-Style Reporting is obviously especially high in Coke's own reports, that topic is also featured prominently in the reports of other reporters. It is therefore not a topic that simply reflects one person's work but rather a topic that captures the style of that person's work. It is a style that other reporters used, presumably inspired by Coke, even if they were usually unable to quite measure up to him in terms of quality.[29] Notably, the effect of Coke-Style Reporting has the largest detectable effect even when we include an additional control, a Coke fixed effect, which captures the influence that Coke himself exerted on the development of

[28] For examples of such cases, see Mullett (1939). Perhaps the difference between our predictions and our findings is due to what Young (2000: 862) refers to as the marginalization of ecclesiastical history from the mainstream of historiography.

[29] There are 410 reports authored by reporters other than Coke for which the prevalence of Coke-Style Reporting exceeds the mean prevalence of Coke-Style Reporting for the 661 Coke-authored reports in our corpus. Coke's style "was followed by many inferior imitators, not a few of whom, falling into the very vices which he condemned, would have incurred his robust disdain" (Allen 1964: 209).

English caselaw solely through his own reports.[30] The effect of Coke-Style Reporting on citations therefore does not occur exclusively, or even primarily, through the effect of Coke himself. Rather, the detected effect of Coke-Style Reporting captures a general mode of reasoning about legal cases that substantially originated with Coke, but later, via dissemination of pertinent reports and legal education, became part of the methodology adopted by many members of the profession.[31]

Our findings thereby cast quantitative light on the venerable legal-historical debate about which particular pre-19th century lawyers had the most effect on subsequent legal development (see, e.g., Holdsworth 1938, Coquillette 2004: Ch. 10; Berman 1994). While our analysis does not settle this debate, our results on Coke-Style Reporting provide compelling quantitative evidence in supporting the view of those who have considered Coke as "the most important" among the "many Makers of English Law" (Holdsworth 1938: 113). Going beyond this, our results could be viewed as showing evidence that one individual can really leave a lasting legacy on the development of law.

The topic featuring the second largest significant effect is Precedent.[32] A one-standard-deviation increase in the prevalence of Precedent in a case report is associated with a 39.2 percent increase in the number of citations. Precedent-based reasoning by definition involves use of past cases and the ideas embodied in those cases to bolster the strength of a particular line of argument. The key words associated with the topic Precedent and the pre-1765 case reports that feature this topic prominently are clearly indicative of this type of reasoning (see Section 3.2). Therefore, our estimates show that precedent-based thought from the late medieval and early modern era had an important effect even before past cases came to be viewed as binding for future decisions (see, e.g., Duxbury 2008). This is strong evidence of the deep historical roots of English legal development and a finding that could not be readily ascertained on the basis of traditional legal-historical analysis.

These results on Precedent thus resonate with the findings of recent scholarship on the determinants of modern-day legal citations. According to this literature, the incorporation of citations in a judicial opinion increases the extent to which an opinion is perceived as well-reasoned and persuasive. Judicial opinions that incorporate more citations—and hence rely heavily on precedent-based reasoning—tend themselves to be comparatively more cited (see, e.g., Szmer et al. 2020, Nelson and Hinkle 2018, Cross 2010, Hansford and Spriggs 2006). Our analysis, however, is the first to provide systematic quantitative evidence of the legacy of early (pre-industrial) precedent-based reasoning for subsequent legal development. And, notably, our

[30] Plucknett (1948: 265) notes that Coke's reporting is "dominated by Coke's personality, and derives authority from him….anything that Coke wrote, be it case or comment, was received with the highest respect". The inclusion of a dummy for Coke's report decreases the IRR for Coke-Style Reporting to 1.4488. The estimated coefficient on the Coke indicator dummy is highly statistically significant and the associated IRR equals 3.1932. Full results on this point are available upon request.

[31] Notably, our estimates reveal that another topic connected to Coke, Coke-Style Procedural Rulings, exerts the seventh largest effect (see Table 2).

[32] The effects of Coke-Style Reporting and Precedent are statistically indistinguishable (p-value for the corresponding χ^2-test of equality of effects equals 0.1647).

evidence is produced in a different form than that on modern courts. Because our explanatory variable reflects the words in cases, we show that early cases incorporating precedent-based thought tend to be cited more heavily in later years. Our results do not simply reflect judges conveniently citing a case that includes many citations. Rather they show judges citing cases that emphasize the very idea of precedent. To our knowledge, no similar connection between legal thought during and prior to industrialization has been articulated in the existing legal-historical scholarship.

Overall, our results on the importance of Coke-Style Reporting and Precedent in deliberations on post-1764 cases reveal a central insight: a key legacy of early English caselaw was as much in bestowing upon subsequent generations modes of reasoning as it was in imparting substantive law. This insight could not have been anticipated via the study of traditional legal-historical research. As such, it demonstrates the value of systematic, data-driven empirical inquiry.

Finally, Publishing & Copyright, an early intellectual-property topic within the theme markets and organizations, exhibits the fifth largest effect. Relative to the effect of Motions, a one-standard-deviation increase in the prevalence of Publishing & Copyright in a case report is associated with a 27.4 percent increase in the post-1764 citation count. This high rank of Publishing & Copyright is noteworthy in light of prevailing legal-historical narratives. In 1774, a rather unanticipated judicial decision restricted the scope for subsequent judicial lawmaking in the domain of publishing and copyright law; thereafter, common law considered a published book to be public property (Baker 2019: 486). Much of the late 18th and 19th century advancement of copyright law thus occurred via legislation, a development that would have normally narrowed the scope for continued relevance of early legal ideas promulgated by the courts. Yet our estimates show that pre-industrial legal ideas on Publishing & Copyright were during industrialization nevertheless used relatively extensively. The rise of legislation pertinent to publishing and the associated new forms of property did not extinguish the relevance of earlier legal ideas in this legal domain.

6. The Context of Use of Pre-1765 Legal Ideas: Illustrative examples

In what way were the legal ideas that featured prominently during the pre-1765 era directly used in caselaw developments that took place during the Industrial Revolution? Our quantitative analysis is silent on this point and a systematic analysis of this issue extends beyond the scope of the present paper. Yet providing some examples of the context of the use of pre-1765 legal ideas in subsequent caselaw development seems important. In our data, the cited and the citing case reports are often centuries apart. Some readers might therefore be skeptical about the ability of our quantitative inquiry to clearly identify the legacy of specific pre-1765 legal ideas on post-1764 caselaw, as reflected in the ER. For these readers, specific narratives of the influence of preexisting ideas on the subsequent development of caselaw should increase confidence in the ability of our empirical approach to truly capture substantive aspects of the process of legal development.

We provide three illustrative examples. We highlight the post-1764 use of pre-1765 legal ideas associated with three topics, each reflecting a different theme and each exerting a large detectable effect: Procedural Bills (in the procedure theme), Assumpsit (contract), and Competing Land Claims (real property).

The first example highlights the relevance of preexisting legal ideas on subsequent caselaw development in a closely related legal domain. Legal ideas about Competing Land Claims are featured in a 1554 case (Panel v Moor, 1 Plowden 91, 75 ER 145) concerning an instance of disseizin (an unlawful disposition of real property; in this particular case, a cellar under a parson's church). The report on the case emphasizes that "one may not appear in an assize as bailiff to a corporation, without warrant in writing", a ruling clarifying that a bailiff must possess a formal deed. This case is cited in a report on a 1843 case (Fishmongers' Company v Robertson, 5 Manning and Granger 131, 134 ER 510) about a multi-party contractual arrangement concerning the draining and allotment of certain slob and waste land. In describing a judge's view on the case, the report states: "As to the necessity of the contract being by deed, in order to bind the plaintiff, [the judge] cited…Panel v Moore (Plowd. 91)…". Therefore, pre-1765 ideas pertinent to the resolution of competing land claims were relevant to post-1764 caselaw development in a closely connected area of law.

The second example showcases the importance of preexisting legal ideas in one specific substantive domain for subsequent caselaw development in a very different domain. Ideas about Assumpsit are featured prominently in a 1609 case (Bradley v Toder, Croke Jac 228, 79 ER 198). In that case, the court clarified that when A promises to pay B a sum of money if B marries a person that A would like to see married (in this case, A's cousin), a formal notice of marriage is not necessary for the obligation of payment to arise. The court resolved that "in the like action verbatim, and no notice alledged…it was good enough; for it is a necessary intendment, that when after marriage he requested the payment of the money, that notice was given of the marriage". This case is cited in a report on a 1796 case (Reynolds v Davies, 1 Bosanquet and Puller 625, 126 ER 1100) featuring a dispute about appropriate means of indorsement of a promissory note. Invoking the legal ideas featured in the 1609 case, the court in 1796 ruled that "Notice…is not necessary: but if it be, the averment of the maker's liability and promise to pay, being followed up by an averment of a special request to pay, amounts to an allegation of notice, Bradley v Toder, Cro. Jac. 228." Thus, pre-1765 legal ideas about implied promises in one legal domain, family matters, were used in post-1764 caselaw developments concerning promises in a substantively distinct legal domain, the negotiability of notes.

The third, and final, example illustrates the legacy of preexisting ideas about court procedure on later considerations about procedure. Ideas about Procedural Bills were the subject of a 1754 case (Baldwin v Mackown, 3 Atkyns 817, 26 ER 1267) in which "a supplementary bill" had been brought against a defendant "who was no party to the original bill, to answer the matters charged in the original bill". In that case, in which the reporter omits discussion of the underlying substance, the Lord Chancellor allowed the defendant's demurrer (objection). This case is cited in a 1839 case (Semple v Price 10, Simons 238, 59 ER 604) in which the object of the original bill "was to charge…the surviving trustee of the Plaintiffs marriage settlement, with a breach of trust in selling out a sum of stock, part of the settled property". In the process, the plaintiff filed a supplementary bill, which was demurred to by the defendant. A participant in the case, likely the defendant's counsel, "in support of the demurrer…said that the filing of the supplemental bill was an attempt

to evade that order. He also cited Baldwin v Mackown; (3 Atk. 817)". Thus, pre-1765 legal ideas about procedural aspects of the use of bills were used to assert a point made by one of the parties in a post-1764 case featuring analogous procedural considerations.

7. Which Pre-1765 Legal Ideas Were Less Cited and Why?

Given that an overwhelming majority of topics exerts a detectable effect, it is informative to examine which types of topics do not exert an effect that can be statistically distinguished from that of Motions, the benchmark topic. Those topics capture aspects of caselaw and associated legal ideas from the late medieval and early modern era that were, in this sense, comparatively less relevant for the development of caselaw during the Industrial Revolution.

Among the topics that do not exert a statistically detectable effect is Employment of Apprentices & Servants, the only topic of this kind in the contract theme. Considerations concerning labor relations were undoubtedly at the forefront of many legal debates during the Industrial Revolution (Cornish et al. 2019: Ch. 4). Legal historians have shown that some of the resultant developments, for example on restraints on trade, were facilitated by preexisting legal ideas (Blake 1960).[33] However, our results suggest that, even though the caselaw developed during the Industrial Revolution drew on many early contracting-specific ideas, those pertaining specifically to employment were not particularly important.

Similarly, within the families theme, the only topic that does not exert a detectable effect is Minors & Guardians. Thus, even though issues pertaining to the well-being of children certainly continued to arise during the Industrial Revolution, earlier ideas on the related legal issues were not particularly relevant to those developments. Further examples of topics that do not exert a detectable effect include Actionable Defamation within the tort theme, Statute Applicability within the sources of law theme, Regulating Commerce within the markets and organizations theme, as well as several topics within the real-property and procedure themes.

Given our methodology, it is theoretically possible that the ideas we have referred to in the paragraphs immediately above were actually relevant to subsequent legal development, but simply never cited. To investigate this possibility, we must understand why some legal ideas, such as those noted above, were less cited during the Industrial Revolution even though these ideas had been an important part of prior discourse among legal practitioners. Is it that those legal ideas became so widely accepted within the legal profession by the late 18th century that, even though they were relevant, legal professionals using them in the later era no longer felt the need to cite specific cases (hypothesis H1)? Or is it that, as we had conjectured in Section 3.4, those legal ideas were generally less applicable in the late 18th and 19th century (hypothesis H2)? Our empirical approach, laid out in Section 4, cannot directly distinguish between these hypotheses. In order to examine these hypotheses, we therefore utilize information from outside our existing framework.

[33] In their discussion of the law of labor relations during the late 18th and 19th century, Cornish et al. (2019: 278) for example posit: "The turning of employment from a serf-like status, with fixed and subjugate conditions, into a 'free contractual relationship'—[a] celebrated indicator of a progressive society—was not something which followed on from industrial capitalism. Rather, the law in the centuries preceding the industrial revolution had already developed in such a way as to permit many employers to organize their workers in ways largely free from outside interference".

The essence of our approach rests on using a measure of the degree to which the corresponding legal ideas had been settled in pre-1765 caselaw. If hypothesis H1 is true, then we would expect a measure of acceptance of legal ideas to be negatively related with whether the ideas exert a detectable effect. On the other hand, if hypothesis H2 is true, we would not expect to observe a systematic relationship between the two variables.

To produce a measure of acceptance of legal ideas, we use data generated by Grajzl and Murrell (2021a; GM). GM produce timelines showing the evolution of the topic proportions for each of the 100 topics over the years 1550-1764. Using a model of the diffusion of legal ideas and drawing on examples from legal history, GM demonstrate that the integral over time of a topic timeline—the cumulative prevalence—captures the relative extent to which the topic's ideas have been accepted by a specific point in time. In contrast, topic prevalence at any specific time reflects the relative amount of change in the adherence to the corresponding ideas at that point in time. Conceptually, these notions are exactly the same as the distinction in epidemiological models between the cumulative number of infections by a particular point in time, as represented by the height of the S-curve of the spread of disease, and the incidence of infection at that point in time, captured by the slope of the S-curve.

Thus, for a given process of diffusion of specific legal ideas, as represented by the corresponding S-curve, the change in the height of the S-curve over a time interval just before the onset of the Industrial Revolution is inversely related to the extent to which the pertinent legal ideas have been accepted prior to the start of the Industrial Revolution. Figure 3 illustrates this point graphically. Part (i) of Figure 3 represents the scenario where early legal ideas were largely accepted prior to the start of the Industrial Revolution. The change in the cumulative prevalence of those ideas (i.e. the change in the height of the S-curve) over the time period 1700-1750 is relatively small: hence the ratio a/b is close to one, with a and b reflecting the heights of the S-curve in 1700 and 1750, respectively. In contrast, part (ii) of Figure 3 corresponds to the scenario where early legal ideas were still very much under contention on the eve of the Industrial Revolution. Thus, the change in the cumulative prevalence of those ideas from 1700 to 1750 is relatively large: c/d is close to zero, with c and d reflecting the heights of the S-curve in 1700 and 1750.

Drawing on the above insights and the GM estimates, we construct the following empirical measure of the extent to which the specific set of legal ideas in topic i were accepted by the onset of the Industrial Revolution:

$$A_i = \frac{\sum_{y=1550}^{1700} T_{i,y}}{\sum_{y=1550}^{1750} T_{i,y}}. \tag{2}$$

$T_{i,y}$ is the topic proportion of topic i in year y. The numerator in (2) measures the cumulative prevalence of topic i in the corpus between 1550 and 1700. The denominator measures the

cumulative prevalence of topic i between 1550 and 1750.[34] A value of A_i close to one indicates that the ideas associated with topic i have been, to a significant extent, already accepted by the onset of the Industrial Revolution. In contrast, a value of A_i close to zero indicates that the pertinent ideas have only recently entered professional discourse and therefore have not diffused widely by the start of the Industrial Revolution.

Importantly, A_i is a measure of the extent of acceptance of ideas in topic i and not of the incidence of topic i in the corpus. To stress this point, Table 4 shows the distribution of all 100 topics based on (i) the relative size of A_i and (ii) the relative report-level prevalence of topic i as captured by the mean reported in Table 1. The value of the χ^2 statistic for the test of the null of independence between the two measures equals 3.21, with p-value equal to 0.071. The p-value for Fisher's exact test is 0.109. We thus find at best limited evidence of a positive association between topic acceptance and topic prevalence. Indeed, Precedent, the most prevalent topic in the pre-1765 corpus, features a comparatively low value of A_i: in early 18th century, precedent-based reasoning was still very much in contention (Grajzl and Murrell 2021b).

Given our measure of acceptance of legal ideas by the Industrial Revolution, we then estimate a simple bivariate linear probability model. The dependent variable is a binary indicator equal to one if topic i exerts a significant effect as indicated by the results in Table 3. The explanatory variable is the standardized ratio of ($A_i/A_{Motions}$), where $A_{Motions}$ is the value of expression (2) for the omitted topic (Motions). The estimation uses 99 observations, one for each topic in model (1).

A finding of a negative association between the two variables would be consistent with hypothesis H1 above. This is the theory that less cited topics correspond to those legal ideas, embodied in the early caselaw, that were relatively settled by the beginning of the Industrial Revolution. In contrast, absence of a systematic association between the two variables would suggest that the lack of a detectable effect of specific topics is consistent with the argument that the legal ideas associated with those topics were no longer important in late 18th- and 19th-century legal disputes.

Figure 4 summarizes the results for the resulting regression, showing the corresponding line of best fit. We find no evidence of a systematic relationship between whether a topic exerts a detectable effect and our measure of the extent to which the topic has been settled by the start of the Industrial Revolution. (There is a 0.971 p-value for the test of the null that the slope of the regression line is equal to zero.) We thus do not find evidence that the set of legal ideas that had become generally accepted by 1765 were no longer cited because they were so well accepted. Our findings are consistent with the interpretation that the set of pre-1765 legal ideas that do not exert

[34] Our choice of the years 1550 and 1750 as the end-points of the intervals used for computation of the elements of (2) reflects the specifics of GM's data. The GM estimates of $T_{i,y}$ for $y<1550$ are generally less precise than the estimates of $T_{i,y}$ for $y\geq1550$ because of the scarcity of reports on pre-1550 cases. The GM estimates of $T_{i,y}$ for $y\in(1750, 1764]$, where 1764 is the last year of reported cases included in the GM corpus, are naturally less reliable than the estimates for $y\in[1550, 1750]$ because spline-based methods for the computation of $T_{i,y}$ near the end-point of available data are sensitive to outliers.

a detectable effect on post-1764 caselaw development lost relevance amidst the many social, economic, and legislative changes occurring during the Industrial Revolution.

Our analysis in this section also speaks to a broader methodological debate about the use of citations as a quantitative means of tracing the influence of legal ideas, and provides a new method of contributing to this debate. A frequently-articulated concern about reliance on citations as a measure of influence has been that the (relative) absence of citations to a past case or judicial opinion poses an interpretative challenge about the true influence of that case or opinion (see, e.g., Landes and Posner 1976, Landes et al. 1998, Posner 2000, Cross 2010, 2012). Does the lack of citations reflect the irrelevance of the specific case or opinion or does the low citation rate simply reflect the fact that a case or opinion has effectively settled the law to such an extent that it is not in anybody's interest to refer to the pertinent issues? Drawing on the evolution of English caselaw as an example, our analysis offers a straightforward methodological approach that is able to empirically distinguish between the alternative hypotheses for why citation rates might be low for specific sets of ideas. Given the increasing reliance on citations in empirical legal scholarship, our approach could be used in many other contexts where it is important to understand why certain legal ideas and associated legal documents have attracted relatively few citations.

8. Concluding Remarks

How English caselaw evolved over the centuries remains incompletely understood. The Industrial Revolution, in particular, resulted in unprecedented societal change and created a host of new legal problems. The question of which preexisting legal ideas were of more or of less relevance for legal development in the late 18th and 19th centuries is therefore a highly pertinent one, but also one that remains empirically underexplored. To address it, we have used a dataset derived from the definitive collection of reports on cases heard in the English high courts from medieval times to the mid-19th century. Methodologically, we have combined existing data produced by an unsupervised machine-learning analysis of text with conventional regression-based techniques.

Consistent with our expectations, our estimates show that among early legal ideas that were invoked most in court deliberation during industrialization were those of comparatively most modern tenor: for example, on markets and organizations, personal property, and debt. However, our analysis also reveals a surprisingly strong relevance (on average) of ideas comprising other themes, such as the ecclesiastical. Moreover, preexisting legal ideas exhibiting detectable effects on subsequent caselaw development are nested within all major pre-industrial legal themes. Indeed, the two specific sets of late-medieval and early-modern legal ideas that exert the strongest effect are emphasis on precedent and the distinctive style of legal analysis associated with Edward Coke. These findings indicate that a central legal legacy of early English legal thought is to be found as much in the bestowed modes of legal reasoning as in imparted substantive law.

Finally, a little more than a quarter of the identified preexisting legal ideas do not exert a detectable effect on subsequent caselaw development. Our test of why this is so shows that, as we conjectured in our baseline analysis, the corresponding legal ideas were simply no longer key to the legal disputes arising in the Industrial Revolution. We do not find evidence in favor of the

alternative hypothesis, that the applicable legal ideas had been so widely accepted by the late 18th century that they were no longer explicitly referred to in pertinent cases.

Our analysis leads to further questions on English caselaw development. For example, were preexisting ideas applicable to a particular domain of law (e.g. contracts) or legal thought (e.g. precedent-based reasoning) predominantly relevant to subsequent caselaw development in the same domain? Or were early legal ideas relevant to subsequent caselaw development in multiple legal domains, as when preexisting rulings on the obligations of carriers and innkeepers stimulated later developments in the tort law of negligence (Cornish et al. 2019: 460-461)? If such cross-fertilization was prevalent, which preexisting legal ideas were especially important in spreading across domains? We hope to tackle such questions in the future. The methods developed here and our findings constitute just one step in generating a quantitative understanding of the process of development of English caselaw, which remains an intriguing and empirically underexplored research avenue. Indeed, our approach can be viewed as providing a productive methodological template for gaining an empirical understanding of the development of caselaw in general.

References

Abadie, A., Athey, S., Imbens, G.W., and Wooldridge, J.M. (2017). When should you adjust standard errors for clustering? NBER Working Paper No. 24003.

Allen, C.K. (1964). Law in the Making. Seventh edition. Oxford, UK: Clarendon Press.

Ash, E., MacLeod, W.B., and Naidu, S. (2019). The language of contract: Promises and power in union collective bargaining agreements. Paper presented at the 23rd Annual Conference of the Society for Institutional and Organizational Economics.

Atiyah, P.S. (1979). The Rise and Fall of Freedom of Contract. Oxford, UK: Clarendon Press.

Baker, J.H. (1980). The Rise and Fall of Freedom of Contract by P.S. Atiyah. The Modern Law Review 43(4): 467-469.

Baker, J.H. (2019). An Introduction to English Legal History. Fifth edition. Oxford, UK: Oxford University Press.

Benjamini, Y. and Yekutieli, D. (2001). The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics 29(4): 1165-1188.

Berman, H.J. (1994). The origins of historical jurisprudence: Coke, Selden, Hale. Yale Law Journal 103(7): 1651-1738.

Black, R.C. and Spriggs II, J.F. (2013). The citation and depreciation of U.S. Supreme Court precedent. Journal of Empirical Legal Studies 10(2): 325-358.

Blake, H.M. (1960). Employee agreements not to compete. Harvard Law Review 73(4): 625-691.

Blei, D.M. (2012). Probabilistic topic models. Communications of the ACM 55(4): 77-84.

Bogart, D. and Richardson, G. (2009). Making property productive: Reorganizing rights to real and equitable estates in Britain, 1660 to 1830. European Review of Economic History 13(1): 3-30.

Bogart, D. and Richardson, G. (2010). Estate acts, 1600 to 1830: A new source for British history. Research in Economic History 27: 1-50.

Bogart, D. and Richardson, G. (2011). Property rights and Parliament in industrializing Britain. Journal of Law and Economics 54(2): 241-274.

Boyer, A.D. (1997). Sir Edward Coke, Ciceronianus: Classical rhetoric and the common law tradition. International Journal for the Semiotics of Law 10(28): 3-36.

Brooks, C.W. (1998). Lawyers, Litigation and English Society Since 1450. London, UK: The Hambledon Press.

Cameron, A.C. and Miller, D.L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources 50(2): 317-372.

Choi, S.J., Gulati, G.M., and Posner, E.A. (2010). Professionals or politicians: The uncertain empirical case for an elected rather than appointed judiciary. Journal of Law, Economics, and Organization 26(2): 290-336.

Coquillette, D.R. (2004). The Anglo-American Legal Heritage, Introductory Materials, Second Edition, Durham, NC: Carolina Academic Press.

Cornish, W., Banks, S., Mitchell, C., Mitchell, P., and Probert, R. (2019). Law and Society in England 1750-1950. Second Edition. Oxford, UK: Hart Publishing.

Cross, F.B. (2012). The ideology of Supreme Court opinions and citations. Iowa Law Review 97(3): 693-752.

Cross, F.B., Spriggs II, J.F., Johnson, T.R., and Wahlbeck, P.J. (2010). Citations in the U.S. Supreme Court: An empirical study of their use and significance. University of Illinois Law Review 2010(2): 489-576.

DiMaggio, P., Nag, M., and Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. Poetics 41(6): 570-606.

Dudoit, S. and van der Laan, M.J. (2008). Multiple Testing Procedures with Applications to Genomics. New York, NY: Springer.

Duxbury, N. (2008). The Nature and Authority of Precedent. Cambridge, UK: Cambridge University Press.

Efron, B. (2010). Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing and Prediction. Cambridge, UK: Cambridge University Press.

Fix, M.P. and Fairbanks, B.R. (2020). The effect of opinion readability on the impact of U.S. Supreme Court precedents in state high courts. Social Science Quarterly 101(2): 811-824.

Fowler, J.H., Johnson, T.R., Spriggs II, J.F., Jeon, S., and Wahlbeck, P.J. (2007). Network analysis and the law: Measuring the legal importance of precedents at the U.S. Supreme Court. Political Analysis 15(3): 324-346.

Grajzl, P. and Murrell, P. (2021a). A machine-learning history of English caselaw and legal ideas prior to the Industrial Revolution I: Generating and interpreting the estimates. Journal of Institutional Economics 17(1): 1-19.

Grajzl, P. and Murrell, P. (2021b). A machine-learning history of English caselaw and legal ideas prior to the Industrial Revolution II: Applications. Journal of Institutional Economics 17(2): 201-216.

Hansford, T.G. and Spriggs II, J.F. (2006). The Politics of Precedent on the U.S. Supreme Court. Princeton, NJ: Princeton University Press.

Harding, A. (1976). A Social History of English Law. Gloucester, MA: Peter Smith.

Harris, R. (2000). Industrializing English Law. Entrepreneurship and Business Organization, 1720-1844. Cambridge, UK: Cambridge University Press.

Harris, R. (2004). Government and the economy, 1688-1850. In: Floud, R. and Johnson, P. (eds.), The Cambridge Economic History of Modern Britain, Volume I: Industrialization, 1700-1860. Cambridge, UK: Cambridge University Press, 204-237.

Holden, J.M. (1951). Bills of exchange during the seventeenth century. Law Quarterly Review 67(2): 230-248.

Holdsworth, W.S. (1938). Some Makers of English Law. London, UK: Cambridge University Press.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6(2): 65-70.

Hoppit, J. (1996). Patterns of parliamentary legislation, 1660-1800. The Historical Journal 39(1): 109-131.

Hoppit, J. (2017). Britain's Political Economies: Parliament and Economic Life, 1660-1800. Cambridge, UK: Cambridge University Press.

Horwitz, M.J. (1977). The Transformation of American Law, 1780-1860. Cambridge, MA: Harvard University Press.

Ibbetson, D. (2013). The Assessment of contractual damages at common law in the late sixteenth century. In: Dyson, M. and Ibbetson, D. (eds.), Law and Legal Process: Substantive Law and Procedure in English Legal History. Cambridge, UK: Cambridge University Press, 126-147.

Jenks, E. (1938). A Short History of English Law. Fifth Edition. London, UK: Methuen & Co. Ltd.

Judson, F.N. (1910) A modern view of the law reforms of Jeremy Bentham. Columbia Law Review 10(1): 41-54.

Kiralfy, A.K.R. (1962). Potter's Historical Introduction to English Law and Its Institutions. Fourth Edition. London, UK: Sweet & Maxwell Limited.

Kosma, M.N. (1998). Measuring the influence of Supreme Court justices. The Journal of Legal Studies 27(2): 333-372.

Landes, W.M. and Posner, R.A. (1976). Legal precedent: A theoretical and empirical analysis. The Journal of Law and Economics 19(2): 249-307.

Landes, W.M., Lessig, L., and Solimine, M.E. (1998). Judicial influence: A citation analysis of Federal Courts of Appeals judges. The Journal of Legal Studies 27(2): 271-332.

Lewis, T.E. (1932). The history of judicial precedent, part IV. Law Quarterly Review 48(2): 230-247.

Lieberman, D. (1989). The Province of Legislation Determined. Legal Theory in Eighteenth Century Britain. Cambridge, UK: Cambridge University Press.

Lobban, M. (1991). The Common Law and English Jurisprudence, 1760-1850. Oxford, UK: Oxford University Press.

Lobban, M. (2000) Henry Brougham and law reform. The English Historical Review 115(464): 1184-1215.

Maitland, F.W. and Montague, F.C. (1978 [1915]). A Sketch of English Legal History. New York, NY: G.P. Putnam's Sons.

Mullett, C.F. (1939). The legal position of the English Protestant dissenters, 1767-1812. Virginia Law Review 25(6): 671-697.

Murrell, P. (2021). Did the independence of judges reduce legal development in England, 1600-1800? Journal of Law and Economics 64(3): 539-565.

Nelson, M.J. and Hinkle, R.K. (2018). Crafting the law: How opinion content influences legal development. Justice System Journal 39(2): 97-122.

Plucknett, T.F.T. (1948). A Concise History of the Common Law. Fourth edition. London, UK: Butterworth & Co.

Posner, R.A. (2000). An economic analysis of the use of citations in the law. American Law and Economics Review 2(2): 381-406.

Renton, A.W. (1900-1932). The English Reports. Great Britain. Parliament. House of Lords. Edinburgh, UK: W. Green & Sons.

Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S.K., Albertson, B., and Rand, D.G. (2014). Structural topic models for open-ended survey responses. American Journal of Political Science 58(4): 1064-1082.

Roberts, M.E., Stewart, B.M., and Airoldi, E.M. (2016). A model of text for experimentation in the social sciences. Journal of the American Statistical Association 111(515): 988-1003.

Romano, J.P. and Wolf, M. (2005a). Exact and approximate stepdown methods for multiple hypothesis testing. Journal of the American Statistical Association 100(469): 94-108.

Romano, J.P. and Wolf, M. (2005b). Stepwise multiple testing as formalized data snooping. Econometrica 73(4): 1237-1282.

Romano, J.P. and Wolf, M. (2016). Efficient computation of adjusted *p*-values for resampling-based stepdown multiple testing. Statistics and Probability Letters 113: 38-40.

Schmidt, M. (2016). Institutional Persistence and Change in England's Common Law: 1700-1865. Unpublished doctoral dissertation. University of Maryland, College Park.

Simpson, A.W.B. (1979). The Horwitz thesis and the history of contracts. The University of Chicago Law Review 46(3): 533-601.

Szmer, J., Christensen, R.K., and Grubbs, S. (2020). What influences the influence of U.S. Courts of Appeals decisions? European Journal of Law and Economics 49(1): 55-81.

Veeder, V.V. (1901). The English Reports, 1292-1865. Harvard Law Review 15(1): 1-25.

Wooldridge, J.M. (2013). Introductory Econometrics, A Modern Approach. Fifth edition. Mason, OH: South-Western, Cengage Learning.

Young, B.W. (2000). Religious history and the eighteenth-century historian. The Historical Journal 43(3): 849-868.

Table 1 (part 1): Topics and report-level descriptive statistics for the topic proportion variables

Topic	Part of theme	Mean	S.D.
Assumpsit	contract	0.0151	0.0494
Bonds	contract	0.0129	0.0309
Contract Interpretation & Validity	contract	0.0056	0.0211
Employment of Apprentices & Servants	contract	0.0049	0.0185
Executable Purchase Agreements	contract	0.0062	0.0223
Identifying Contractual Breach	contract	0.0064	0.0236
Length & Expiry of Leases	contract	0.0083	0.0299
Rental Payments	contract	0.0066	0.0243
Decisions After Conviction	criminal	0.0148	0.0487
Habeas Corpus	criminal	0.0153	0.0492
Indicting for Murder	criminal	0.0054	0.0243
Bankruptcy	debt	0.0048	0.0258
Claims from Financial Instruments	debt	0.0075	0.0207
Mortgages	debt	0.0056	0.0316
Pleadings on Debt	debt	0.0099	0.0305
Prioritizing Claims	debt	0.0083	0.0347
Repaying Debt	debt	0.0159	0.0271
Ecclesiastical Appointments	ecclesiastical	0.0072	0.0374
Temporal & Spiritual Jurisdiction	ecclesiastical	0.0034	0.0218
Tithes	ecclesiastical	0.0070	0.0437
Daughters' Legacies	families	0.0055	0.0292
Geographic Settlement of Children	families	0.0085	0.0531
Marriage Settlement	families	0.0060	0.0293
Minors & Guardians	families	0.0048	0.0220
Rights of Married Women	families	0.0090	0.0289
Contingency in Wills	inheritance	0.0063	0.0320
Disentangling Heirs	inheritance	0.0075	0.0232
Estate Tail	inheritance	0.0058	0.0342
Excluding Beneficiaries of Wills	inheritance	0.0031	0.0184
Execution & Administration of Estates	inheritance	0.0112	0.0356
Implementing Ambiguous Wills	inheritance	0.0089	0.0398
Intestacy	inheritance	0.0062	0.0307
Specifying Inherited Property Rights	inheritance	0.0104	0.0360
Validity of Wills	inheritance	0.0035	0.0377
Equitable Relief	jurisdiction	0.0135	0.0492
Equity Jurisdiction	jurisdiction	0.0109	0.0274
Geographic Jurisdiction of Laws	jurisdiction	0.0048	0.0208
Inferior-Court Jurisdiction	jurisdiction	0.0109	0.0312
Prohibiting Jurisdiction	jurisdiction	0.0137	0.0521
Governance of Private Organizations	markets and organizations	0.0028	0.0232
Municipal Charters	markets and organizations	0.0079	0.0402
Negotiable Bills & Notes	markets and organizations	0.0059	0.0321
Publishing & Copyright	markets and organizations	0.0037	0.0198
Regulating Commerce	markets and organizations	0.0056	0.0209
Restraints on Trade	markets and organizations	0.0051	0.0275
Attorney- & Solicitor-General	multiple	0.0094	0.0303
Coke-Style Reporting	multiple	0.0097	0.0282
Determining Damages & Costs	multiple	0.0087	0.0259
Keble-Style Reporting	multiple	0.0251	0.0695
Modern-Style Reporting	multiple	0.0257	0.0471
Multiparty Cases	multiple	0.0053	0.0132
Non-Translated Latin	multiple	0.0098	0.0374
Revocation	multiple	0.0075	0.0282
Vesey Footnotes	multiple	0.0120	0.0501

Notes: The first column of the table features the names given by Grajzl and Murrell (2021a; GM) to the 100 topics that GM estimated using a corpus of 52,949 reports on cases heard before 1765. The second column provides a grouping of the topics into broader themes. Analytically, case reports are mixtures of topics. The GM estimates pinpoint the proportion of each of the 52,949 pre-1765 case reports that is devoted to each of the 100 topics. The values of each of the corresponding 100 topic proportion variables therefore differ across the reports. Consequently, the third and fourth columns show the mean and standard deviation of the 100 topic proportion variables across the 52,949 case reports.

Table 1 (part 2): Topics and report-level descriptive statistics for the topic proportion variables

Topic	Part of theme	Mean	S.D.
Vesey Reporting	multiple	0.0051	0.0293
Bailment	personal property	0.0057	0.0256
Ownership of War Bounty	personal property	0.0058	0.0429
Trespass to Goods	personal property	0.0096	0.0371
Dignitaries	politics	0.0046	0.0140
Local Administrative Appointments	politics	0.0062	0.0223
Rights of Public Office	politics	0.0064	0.0246
Royal Patents & Tenures	politics	0.0087	0.0307
Arbitration & Umpires	procedure	0.0067	0.0350
Coke-Style Procedural Rulings	procedure	0.0110	0.0382
Correct Pleas	procedure	0.0258	0.0487
Court Petitions	procedure	0.0048	0.0256
Equity Appeals	procedure	0.0032	0.0298
Evidence Gathering & Admissibility	procedure	0.0136	0.0368
Decisional Logic	procedure	0.0265	0.0284
Jury Procedures & Trials	procedure	0.0156	0.0342
Mistakes in Court Records	procedure	0.0082	0.0253
Motions	procedure	0.0363	0.0898
Procedural Bills	procedure	0.0186	0.0542
Procedural Rulings on Actions	procedure	0.0204	0.0312
Procedural Rulings on Writs	procedure	0.0170	0.0451
Rendering Judgement	procedure	0.0182	0.0395
Reviewing Local Orders	procedure	0.0146	0.0615
Rulings on the Calendar	procedure	0.0183	0.0264
Writs of Error	procedure	0.0234	0.0563
Common-Land Disputes	real property	0.0082	0.0374
Competing Land Claims	real property	0.0060	0.0310
Conveyancing by Fine	real property	0.0046	0.0203
Elizabethan Land Cases	real property	0.0176	0.0471
Equitable Waste	real property	0.0031	0.0175
Implementing Trusts	real property	0.0062	0.0228
Manorial Tenures	real property	0.0074	0.0336
Possession & Title	real property	0.0092	0.0205
Self-Help in Real-Property Disputes	real property	0.0033	0.0289
Shared & Divided Property Rights	real property	0.0160	0.0169
Timing of Property Rights	real property	0.0072	0.0276
Transfer of Ownership Rights	real property	0.0074	0.0305
Tree Law	real property	0.0049	0.0156
Uses	real property	0.0062	0.0289
Clarifying Legislative Acts	sources of law	0.0099	0.0213
Contrasting Cases & Statutes	sources of law	0.0009	0.0042
Precedent	sources of law	0.0381	0.0495
Statute Applicability	sources of law	0.0113	0.0210
Actionable Defamation	torts	0.0136	0.0630
Nuisance	torts	0.0089	0.0369
Wrongful Possession	torts	0.0096	0.0172

Notes: See notes under part 1.

Table 2: The distribution of the number of citations to pre-1765 reports

Citing period 1765-1870			Citing period 1765-1815			Citing period 1816-1870		
Citations	Freq.	Cum. %	Citations	Freq.	Cum. %	Citations	Freq.	Cum. %
0	39,040	73.73	0	43,975	83.05	0	42,465	80.20
1	5,607	84.32	1	4,975	92.45	1	4,993	89.63
2	2,693	89.41	2	1,876	95.99	2	2,049	93.50
3	1,583	92.40	3	910	97.71	3	1,137	95.65
4	1,050	94.38	4	474	98.60	4	693	96.96
5	742	95.78	5	257	99.09	5	400	97.71
6	485	96.70	6	162	99.40	6	322	98.32
7	347	97.35	7	76	99.54	7	203	98.70
8	250	97.82	8	77	99.68	8	153	98.99
9	204	98.21	9	45	99.77	9	98	99.18
10	173	98.54	10	26	99.82	10	72	99.31
11	133	98.79	11	22	99.86	11	71	99.45
12	101	98.98	12	20	99.90	12	47	99.54
13	74	99.12	13	9	99.92	13	35	99.60
14	58	99.23	14	9	99.93	14	29	99.66
15	49	99.32	15	5	99.94	15	26	99.71
16	38	99.39	16	3	99.95	16	17	99.74
17	53	99.49	17	8	99.96	17	15	99.77
18	34	99.56	18	1	99.96	18	19	99.80
19	30	99.61	19	3	99.97	19	18	99.84
20	19	99.65	20	4	99.98	20	5	99.85
>20	186	100.00	>20	12	100.00	>20	82	100.00
Mean	0.93		Mean	0.36		Mean	0.57	
S.D.	3.54		S.D.	1.45		S.D.	2.37	
Min.	0		Min.	0		Min.	0	
Max.	368		Max.	171		Max.	197	
Obs.	52,949		Obs.	52,949		Obs.	52,949	

Notes: The table shows the distribution of the incidence of citations to pre-1765 reports. The left-most set of three columns shows the distribution of citations accrued during the entire post-1764 period (1867-1870). The middle set shows the distribution of citations accrued during the early post-1764 period (1765-1815). The right-most set shows the distribution of citations accrued during the late post-1764 period (1816-1870).

Table 3 (part 1): The estimated effect of pre-1765 legal ideas on post-1764 caselaw development, negative binomial regression results

Topic (standardized)	Part of theme	IRR	Uncorr. *p*	FDR
Coke-Style Reporting	multiple	1.5840	0.00000	S
Precedent	sources of law	1.3916	0.00004	S
Procedural Bills	procedure	1.3265	0.00206	S
Equitable Relief	jurisdiction	0.7255	0.00031	S
Publishing & Copyright	markets and organizations	1.2744	0.00116	S
Implementing Ambiguous Wills	inheritance	1.2669	0.00006	S
Coke-Style Procedural Rulings	procedure	1.2567	0.00386	S
Specifying Inherited Property Rights	inheritance	1.2535	0.00000	S
Assumpsit	contract	1.2530	0.00033	S
Reviewing Local Orders	procedure	1.2481	0.00286	S
Competing Land Claims	real property	1.2431	0.00013	S
Estate Tail	inheritance	1.2394	0.00000	S
Tithes	ecclesiastical	1.2266	0.00018	S
Contingency in Wills	inheritance	1.2242	0.00000	S
Procedural Rulings on Writs	procedure	1.2241	0.00076	S
Pleadings on Debt	debt	1.2187	0.00024	S
Correct Pleas	procedure	1.2180	0.00140	S
Nuisance	torts	1.2138	0.00004	S
Geographic Settlement of Children	families	1.2104	0.00313	S
Habeas Corpus	criminal	1.2047	0.00242	S
Ownership of War Bounty	personal property	1.2030	0.00359	S
Negotiable Bills & Notes	markets and organizations	1.2018	0.00000	S
Non-Translated Latin	multiple	1.2015	0.00028	S
Evidence Gathering & Admissibility	procedure	1.1992	0.00081	S
Bankruptcy	debt	1.1936	0.00000	S
Municipal Charters	markets and organizations	1.1876	0.00078	S
Equity Appeals	procedure	1.1874	0.00002	S
Prioritizing Claims	debt	1.1848	0.00046	S
Clarifying Legislative Acts	sources of law	1.1838	0.00001	S
Contract Interpretation & Validity	contract	1.1834	0.00000	S
Self-Help in Real-Property Disputes	real property	1.1807	0.00035	S
Marriage Settlement	families	1.1804	0.00004	S
Mortgages	debt	1.1790	0.00013	S
Trespass to Goods	personal property	1.1789	0.00046	S
Procedural Rulings on Actions	procedure	1.1758	0.00008	S
Revocation	multiple	1.1718	0.00001	S
Daughters' Legacies	families	1.1716	0.00009	S
Ecclesiastical Appointments	ecclesiastical	1.1694	0.00102	S
Bailment	personal property	1.1690	0.00001	S
Determining Damages & Costs	multiple	1.1677	0.00002	S
Execution & Administration of Estates	inheritance	1.1672	0.00085	S
Common-Land Disputes	real property	1.1649	0.00150	S
Temporal & Spiritual Jurisdiction	ecclesiastical	1.1637	0.00002	S
Intestacy	inheritance	1.1610	0.00407	S
Arbitration & Umpires	procedure	1.1608	0.00196	S
Court Petitions	procedure	1.1590	0.00118	S
Rights of Public Office	politics	1.1527	0.00004	S
Timing of Property Rights	real property	1.1504	0.00020	S
Implementing Trusts	real property	1.1486	0.00019	S
Length & Expiry of Leases	contract	1.1476	0.00029	S

Notes: The table presents the incidence rate ratios (IRR) from a negative binomial regression. The underlying regression utilizes 105,898 observations: there are 52,949 pre-1765 case reports and citations to them are observed in the 1765-1815 period and then again in the 1816-1870 period. The estimated model includes full sets of fixed effects for the time period of when a case was heard in court, the time period when the case was cited, and their interactions. (The estimates of the fixed-effects themselves are not reported.) All topic proportion variables have been standardized so that they have means equal to zero and standard deviations equal to one. The omitted topic is Motions. The reported uncorrected *p*-values are based on standard errors clustered at the case reporter level. In the last column, S indicates that the effect is statistically significant at the five percent FDR. The FDR correction is based on the Benjamini and Yekutieli (2005) method. The subset of topics with statistically significant effects are listed first, followed by the subset that are not statistically significant. Within the two subsets, topics are ordered on the basis of effect size, that is the absolute value of the difference between the IRR and one.

Table 3 (part 2): The estimated effect of pre-1765 legal ideas on post-1764 caselaw development, negative binomial regression results

Topic (standardized)	Part of theme	IRR	Uncorr. *p*	FDR
Executable Purchase Agreements	contract	1.1437	0.00046	S
Transfer of Ownership Rights	real property	1.1437	0.00101	S
Rental Payments	contract	1.1400	0.00001	S
Rights of Married Women	families	1.1363	0.00099	S
Royal Patents & Tenures	politics	1.1340	0.00372	S
Governance of Private Organizations	markets and organizations	1.1336	0.00002	S
Identifying Contractual Breach	contract	1.1316	0.00014	S
Uses	real property	1.1287	0.00205	S
Manorial Tenures	real property	1.1277	0.00491	S
Indicting for Murder	criminal	1.1273	0.00079	S
Inferior-Court Jurisdiction	jurisdiction	1.1265	0.00573	S
Bonds	contract	1.1256	0.00390	S
Restraints on Trade	markets and organizations	1.1253	0.00025	S
Possession & Title	real property	1.1027	0.00064	S
Repaying Debt	debt	1.1024	0.00428	S
Local Administrative Appointments	politics	1.1023	0.00041	S
Geographic Jurisdiction of Laws	jurisdiction	1.0994	0.00044	S
Conveyancing by Fine	real property	1.0853	0.00209	S
Excluding Beneficiaries of Wills	inheritance	1.0847	0.00231	S
Wrongful Possession	torts	1.0766	0.00253	S
Contrasting Cases & Statutes	sources of law	1.0317	0.00000	S
Vesey Footnotes	multiple	1.5493	0.01382	
Modern-Style Reporting	multiple	1.4013	0.01902	
Keble-Style Reporting	multiple	0.7558	0.00756	
Actionable Defamation	torts	1.2290	0.00850	
Decisions After Conviction	criminal	1.1723	0.00768	
Prohibiting Jurisdiction	jurisdiction	1.1712	0.01661	
Dignitaries	politics	1.1688	0.00847	
Writs of Error	procedure	1.1687	0.02459	
Jury Procedures & Trials	procedure	1.1378	0.00787	
Mistakes in Court Records	procedure	1.1337	0.00879	
Vesey Reporting	multiple	1.1185	0.16887	
Validity of Wills	inheritance	0.8982	0.19661	
Attorney- & Solicitor-General	multiple	1.0980	0.04991	
Rulings on the Calendar	procedure	1.0974	0.02864	
Multiparty Cases	multiple	1.0967	0.03588	
Elizabethan Land Cases	real property	1.0946	0.30644	
Regulating Commerce	markets and organizations	1.0811	0.01117	
Minors & Guardians	families	1.0766	0.01053	
Decisional Logic	procedure	1.0739	0.23975	
Equitable Waste	real property	1.0715	0.04378	
Disentangling Heirs	inheritance	1.0711	0.02873	
Equity Jurisdiction	jurisdiction	1.0630	0.32144	
Employment of Apprentices & Servants	contract	1.0584	0.02025	
Claims from Financial Instruments	debt	1.0533	0.19912	
Statute Applicability	sources of law	1.0496	0.12100	
Tree Law	real property	1.0390	0.07871	
Shared & Divided Property Rights	real property	1.0309	0.18296	
Rendering Judgement	procedure	1.0237	0.82281	

Notes: See notes following part 1 of the table.

Table 4: The relation between pre-1765 topic prevalence and acceptance of corresponding ideas by the Industrial Revolution

	Topic prevalence		
Topic acceptance	At or below median	Above median	Total
At or below median	30	21	51
Above median	20	29	49
Total	50	50	100

Notes: The table shows the distribution of the 100 pre-1765 topics based on the extent of acceptance by the Industrial Revolution as measured by A_i (defined in expression (2)) relative to the corresponding median value (0.7785) and average report-level prevalence in the pre-1765 corpus relative to the corresponding median value (0.0081).

Figure 1: The distribution of the reports in the corpus with respect to the timing of heard cases

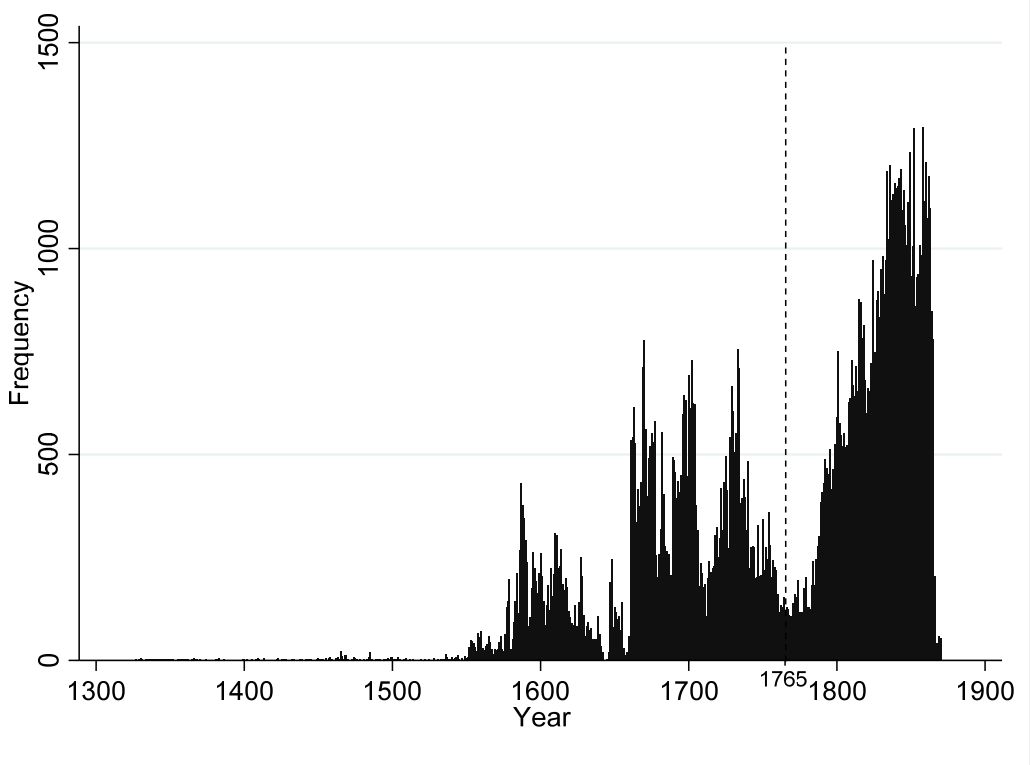
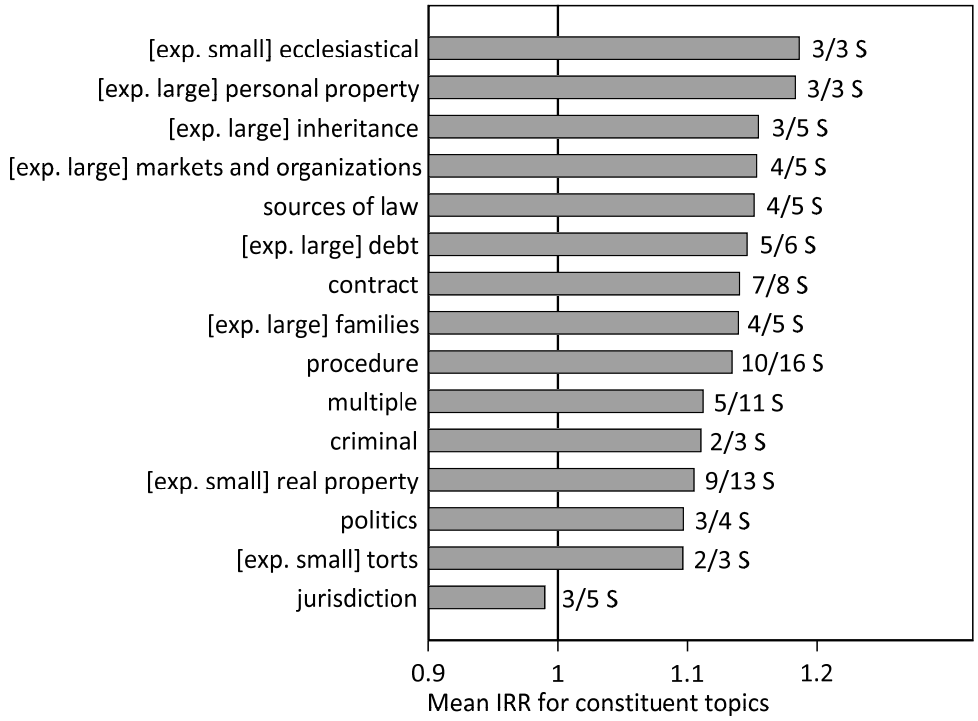
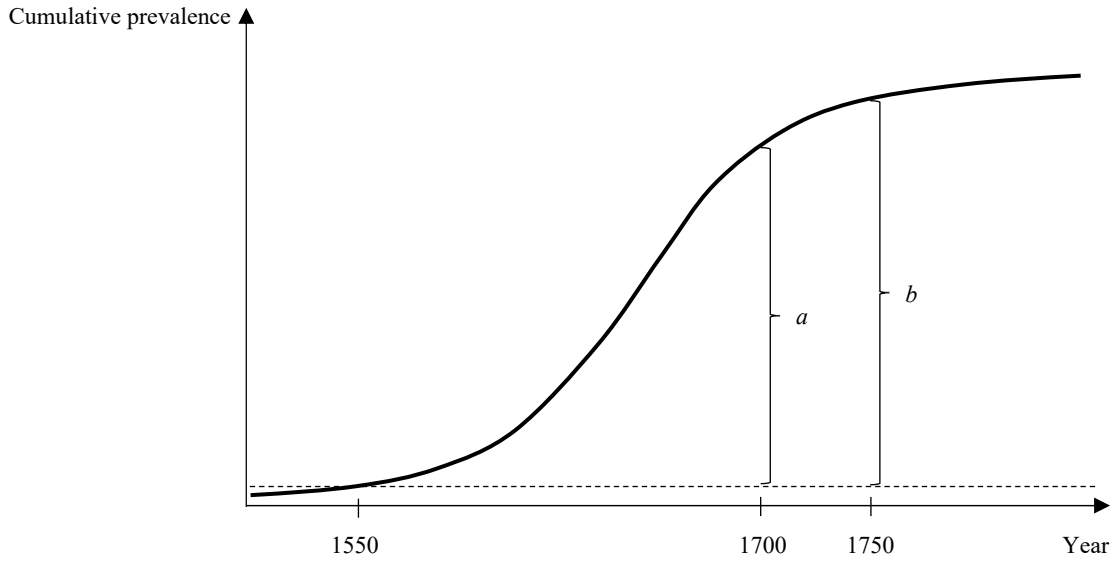


Figure 2: Summary of average effects by theme

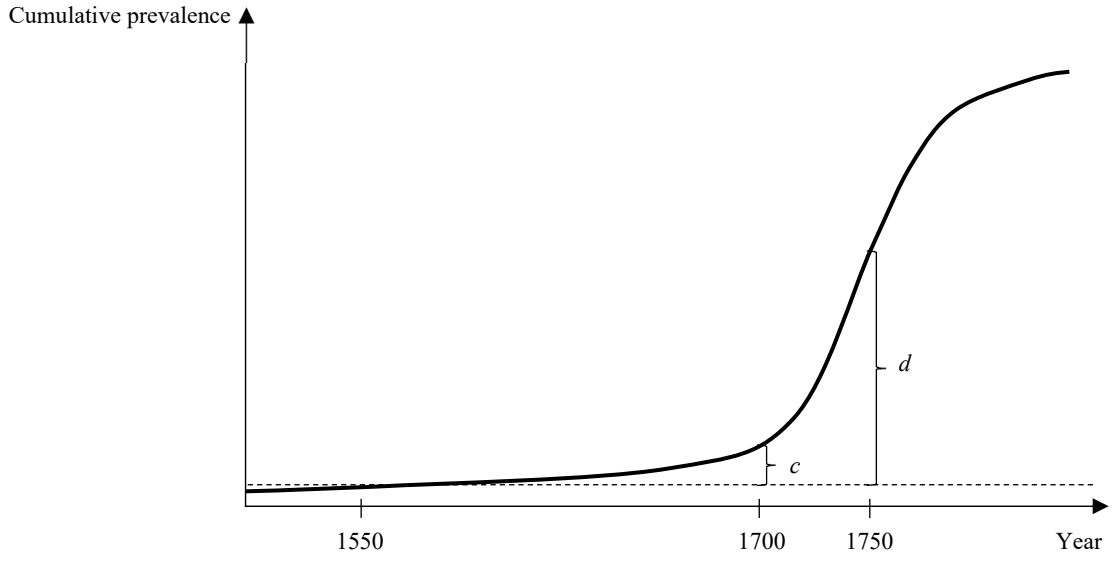


Notes: For each theme, mean IRR is computed as a simple average of IRRs of constituent topics, where IRR for statistically insignificant topics (see Table 3) is set equal to one. [exp. large] and [exp. small] respectively indicate that, based on our conjectures in Section 3.4, we would have anticipated the constituent topics to exert comparatively large or small effects on post-1764 citations. Adjacent to the bar indicating the size of mean IRR for each theme, we list the number of topics comprising the theme that exert a statistically significant effect (denoted S in Table 3) as a share of all topics comprising the theme.

Figure 3: An illustration of the construction of the measure of acceptance of an idea by 1750

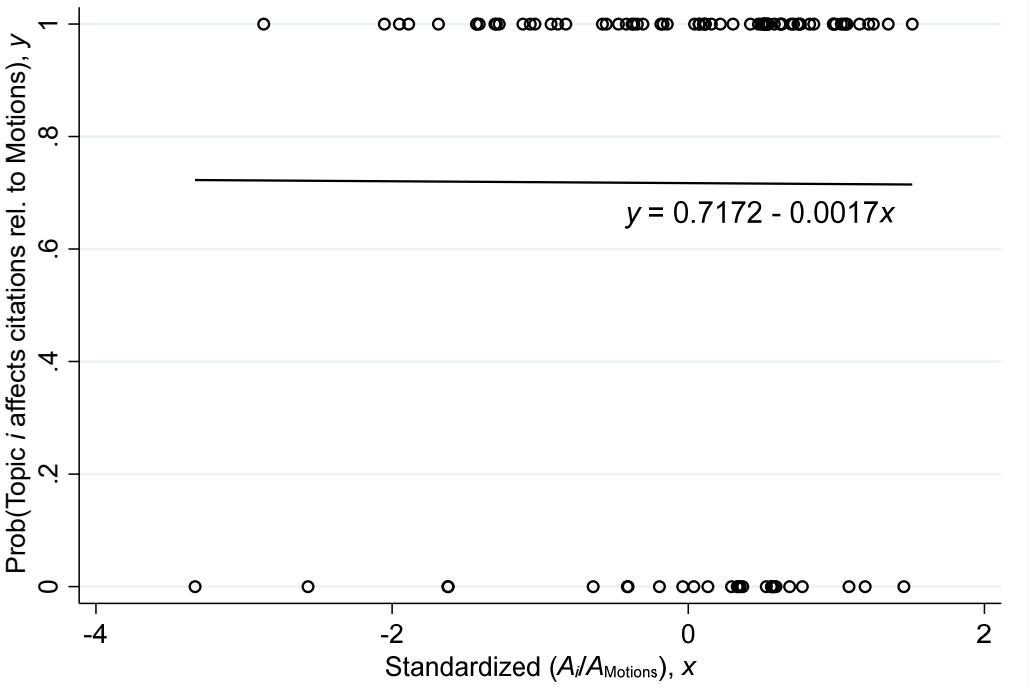


(i) Early legal ideas largely accepted by the onset of the Industrial Revolution



(ii) Early legal ideas still in contention at the onset of the Industrial Revolution

Figure 4: A test of why certain topics do not exert a detectable effect on citations



Notes: Each circle in the figure depicts a particular topic *i* from the set of 99 topics included in the estimation of model (1). The figure then shows the relationship between (i) the value of the binary indicator equal to one if topic *i* exerts a detectable effect on post-1764 citations based on our results in Table 3 and (ii) the standardized value of ($A_i/A_{Motions}$), where A_i is defined by expression (2) and $A_{Motions}$ is the value of expression (2) for the topic Motions (the omitted topic). The featured line is the OLS-estimated line of best fit.

Appendix

Each pre-1765 report is a mixture of 100 topics. We begin with the following model for the conditional mean of citation counts, as implied by the negative binomial model:

$$E\left[y \middle| T_1,...,T_{100} \right] = \exp\left\{ \alpha_0 + \sum_{i=1}^{100} \alpha_i T_i \right\}. \tag{A1}$$

In (A1), y is the number of citations to a particular report in a specific post-1764 period. For ease of exposition, we drop all subscripts denoting observations. Without loss of generality, we also suppress time-period fixed effects. T_i is the proportion of topic i in a given report, bounded from below by zero and from above by one. α_i's are the corresponding parameters.

(A1) implies:

$$\ln E\left[y \middle| T_1,...,T_{100} \right] = \alpha_0 + \sum_{i=1}^{100} \alpha_i T_i \,. \tag{A2}$$

For any report, the sum of topic proportions is one:

$$\sum_{i=1}^{100} T_i = 1 \,. \tag{A3}$$

Thus, model (A2) is not identified because of perfect collinearity. One topic proportion variable must be omitted. Suppose, for illustrative purposes only, that the omitted topic proportion variable is T_{100}. Then, solving (A3) for T_{100} and substituting the resulting expression into (A2) gives:

$$\ln E\left[y \middle| T_1,...,T_{100} \right] = \alpha_0 + \alpha_{100} + \sum_{i=1}^{99} \left(\alpha_i - \alpha_{100} \right) T_i \,. \tag{A4}$$

Averaging (A4) across all reports yields:

$$\overline{\ln E\left[y \middle| T_1,...,T_{100} \right]} = \alpha_0 + \alpha_{100} + \sum_{i=1}^{99} \left(\alpha_i - \alpha_{100} \right) \overline{T}_i \,, \tag{A5}$$

where $\bar{X}$ denotes the report-level average for the variable X. Subtracting (A5) from (A4) and rearranging terms:

$$\ln E\left[y \middle| T_1,...,T_{100} \right] = \overline{\ln E\left[y \middle| T_1,...,T_{100} \right]} + \sum_{i=1}^{99} \left(\alpha_i - \alpha_{100} \right) \frac{\left(T_i - \overline{T}_i \right)}{\sigma_i} \sigma_i \,, \tag{A6}$$

where σ_i is the report-level standard deviation for T_i. Thus, (A6) can be expressed as

$$\ln E\left[y \middle| Z_1,...,Z_{99} \right] = \beta_0 + \sum_{i=1}^{99} \beta_i Z_i \,, \tag{A7}$$

where

$$\beta_0 \equiv \overline{\ln E\left[\, y \middle| T_1, ..., T_{100} \,\right]} \tag{A8}$$

and

$$\beta_i \equiv \left(\alpha_i - \alpha_{100}\right)\sigma_i \tag{A9}$$

$$Z_i \equiv \frac{T_i - \overline{T}_i}{\sigma_i} \tag{A10}$$

for $i \in \{1,2,\ldots,99\}$. That is, Z_i in (A7) is the standardized version of topic proportion i with mean equal to zero and standard deviation equal to one.

(A7) is the empirical model (1) for the conditional mean of citation counts that we estimate. Note that expression (A7) omits one standardized topic proportion, Z_{100}. The reason is that a model otherwise analogous to (A7) but with all $Z_1,\ldots,Z_{100}$ included among the explanatory variables is not identified, even if the constant β_0 is dropped. To see this, average (A3) across the reports to obtain:

$$\sum_{i=1}^{100} \overline{T}_i = 1. \tag{A11}$$

Then, using (A3) to express T_{100}, (A11) to express $\overline{T}_{100}$, and inserting the resulting expressions into (A10) for i=100 yields:

$$Z_{100} = \frac{T_{100} - \overline{T}_{100}}{\sigma_{100}} = \frac{1 - \sum_{i=1}^{99} T_i - 1 + \sum_{i=1}^{99} \overline{T}_i}{\sigma_{100}} = \frac{-\sum_{i=1}^{99}\left(T_i - \overline{T}_i\right)}{\sigma_{100}} = \frac{-\sum_{i=1}^{99} \frac{\left(T_i - \overline{T}_i\right)}{\sigma_i}\sigma_i}{\sigma_{100}}. \tag{A12}$$

Thus, (A12) simplifies to

$$Z_{100} = -\sum_{i=1}^{99} Z_i \frac{\sigma_i}{\sigma_{100}}. \tag{A13}$$

That is, a model with all of $Z_1,\ldots,Z_{100}$ included among the explanatory variables, with or without the regression constant, is not identified.

How should one interpret the parameters β_i, $i \in \{1,2,\ldots,99\}$, from expression (A7)? A one-unit change in the standardized proportion for topic k from the baseline value of Z_k to the value Z_k+1 (that is, a one-standard deviation change in T_k) gives rise to the following change in the expected incidence of citations:

$$\ln E\left[\, y \middle| Z_1, ..., Z_k + 1, ..., Z_{99} \,\right] - \ln E\left[\, y \middle| Z_1, ..., Z_k, ..., Z_{99} \,\right] = \beta_k. \tag{A14}$$

Substituting (A9) into the right-hand side of (A14) and exponentiating the resulting expression implies that the corresponding incidence rate ratio (IRR) equals:

$$\frac{E\left[y\middle|Z_1,...,Z_k+1,...,Z_{99}\right]}{E\left[y\middle|Z_1,...,Z_k,...,Z_{99}\right]}=e^{\beta_k}=e^{(\alpha_k-\alpha_{100})\sigma_k}. \tag{A15}$$

This implies:

$$\frac{E\left[y\middle|Z_1,...,Z_k+1,...,Z_{99}\right]}{E\left[y\middle|Z_1,...,Z_k,...,Z_{99}\right]}=\left(\frac{e^{\alpha_k}}{e^{\alpha_{100}}}\right)^{\sigma_k}. \tag{A16}$$

That is, the IRR associated with a one-standard deviation change in the prevalence of topic *k* equals the ratio of the IRR associated with a unit-change in the (non-standardized) proportion of topic *k* to the IRR associated with a unit-change in the (non-standardized) proportion of the omitted topic, adjusted by the extent of the variability of the prevalence of topic *k*.