Political Analysis (2012) 20:329–350
doi:10.1093/pan/mps019

The Genealogy of Law

Tom S. Clark
Department of Political Science, Emory University, 327 Tarbutton Hall, 1555 Dickey Drive, Atlanta, GA 30322
e-mail: tom.clark@emory.edu (corresponding author)

Benjamin E. Lauderdale
Methodology Institute, London School of Economics, Columbia House, Houghton Street, London, WC2A 2AE, UK
e-mail: b.e.lauderdale@lse.ac.uk

Edited by Jonathan N. Katz

Many theories of judicial politics have at their core the concepts of legal significance, doctrinal development and evolution, and the dynamics of precedent. Despite rigorous theoretical conceptualization, these concepts remain empirically elusive. We propose the use of a genealogical model (or "family tree") to describe the Court's construction of precedent over time. We describe statistical assumptions that allow us to estimate this kind of structure using an original data set of citation counts between Supreme Court majority opinions. The genealogical model of doctrinal development provides a parsimonious description of the dependencies between opinions, while generating measures of legal significance and other related quantities. We employ these measures to evaluate the robustness of a recent finding concerning the relationship between ideological homogeneity within majority coalitions and the legal impact of Court decisions.

1 Introduction

Contemporary research on judicial politics has moved away from conceiving of law as a potential constraint on judges' ability to pursue their political goals and instead recognizes that law and politics are inextricably linked (for a review, see Lax 2011). Law is the material with which judges work, and politics is a component of what they try to do with law. The literature has witnessed the development of increasingly sophisticated theoretical models of how judges shape the content of judge made law (e.g., Kornhauser 1992b; Lax 2007; Gennaioli and Shleifer 2007; Clark and Carrubba 2012)—which we call, for convenience, doctrine. Unfortunately, the empirical literature has struggled to keep pace with these theoretical developments. While scholars have made some progress in measuring the content of judicial decisions (Spriggs and Hansford 2000; McGuire and Vanberg 2005; Clark and Lauderdale 2010; Sulam 2011), less has been achieved in developing statistical models of the fundamental feature of judge-made law—how lines of doctrine develop and evolve through a series of cases. In this article, we offer a step forward in that direction.

We introduce a statistical model of legal doctrine, which we call the *genealogy of law*. Our directed network model makes use of citation data among Supreme Court precedents to parsimoniously describe the structural relationships among cases. As contrasted with existing network analyses of Supreme Court precedent (e.g., Fowler and Jeon 2008; Fowler et al. 2007), we (1) employ richer information about citations (not only the fact of citation between opinions but the extent to which one case relies on another) and (2) build into our model a number of features of doctrine, chief among which is the idea that doctrine evolves through time. Our model captures in a quantitative way the substantive features of law that previously have been studied only through

Authors' note: We thank Brandon Bartels, Barry Friedman, John Kastellec, Drew Linzer, and Jeff Staton for helpful comments and suggestions. We also thank Josh Strayhorn for helpful research assistance. This research was supported by the National Science Foundation (SES-0961058).

laborious, subjective, and qualitative evaluation by legal experts. As such, it provides a rich descriptive tool for systematically and parsimoniously capturing the substantive, legal links among cases in a line of doctrine.

However, our model does more than simply provide a descriptive representation of doctrine. As scholars continue to develop dynamic models of legal evolution through a series of cases (e.g., Kornhauser 1992a; Bueno de Mesquita and Stephenson 2002; Gennaioli and Shleifer 2007), our model can be used to empirically evaluate predictions about how judges sequentially build law, when they continue to refine an area of doctrine, and when they decide to eschew a previous line of argumentation and start a new line of doctrine. Further, our model yields estimates of quantities of interest in a variety of theories that are more closely connected to their underlying substantive features. In particular, our model allows us to generate estimates of a case's significance in terms of its impact on subsequent legal development, a useful alternative to existing proxies like media coverage that tell us more about immediate political salience than legal significance.

The article proceeds as follows. Section 2 provides an overview of the data and methods currently available for modeling legal doctrine and situates our approach in the past research. Section 3 describes our data, modeling assumptions, functional form, estimation strategy, and results. In Section 4, we use our model to generate an original measure of a case's legal significance that is based on its influence on doctrine. Section 5 reports the results of our replication of a recent study of the determinants of consequential law-making by the Supreme Court. In Section 6, we conclude and discuss other potential applications of our model to citation count data.

2 Characterizing Legal Doctrine

To develop our model, we begin by considering what substantive students of the law—namely, lawyers and judges—do when they study doctrine. To study doctrine, lawyers read court opinions, tracing a line of argumentation through a series of cases. Citations are the most important source of information in this process, as they direct the reader to know which cases relate to each other, and which cases build from which others. Indeed, American legal training and argumentation is based on the "case method," which centers around the identification and application of the set of doctrinally relevant cases (Levi 1949; Patterson 1951). Within a given opinion, many other cases (precedents) will be cited, though some will be cited more often than others within that opinion. Our goal is to develop a statistical model of one piece of information that lawyers and judges seek to distill from a collection of opinions: how a line of reasoning develops through a series of related cases. We refer to this process as *doctrinal development*. While recent scholarship has made use of patterns of citation among Supreme Court cases to measure which Supreme Court precedents are frequently cited (e.g., Fowler et al. 2007; Fowler and Jeon 2008), those studies are ill-equipped to capture doctrinal structure and evolution. Specifically, we need a model that identifies the logical connections among cases and the sequential structure of doctrinal development.

2.1 *Data for Measuring Doctrine*

We begin by considering the quantifiable manifestations of doctrine that one could use to develop a statistical model. Ultimately, observable data about legal doctrine derives almost exclusively from the texts of opinions. While data on the justices' dispositional voting and which opinions the justices "join" (which opinions they endorse in a case) can provide a great deal of information about the justices' preferences, opinion texts are the most important source of information about doctrine and its evolution. It is this kind of information to which lawyers, judges, legal academics, and (increasingly) political scientists turn when they seek to study the content of the law. Opinion texts can be analyzed at several levels of abstraction, each of which involves different trade-offs between automation and careful definition of quantities of interest. At one extreme are methods that directly use the text as data, almost always exploiting relative word frequencies in different opinions as the observed data (e.g., McGuire and Vanberg 2005). A more structured approach, which still allows considerable automation, is the use of the citations between opinions as data

(e.g., Clark and Lauderdale 2010). The most structured approaches manually encode legally relevant information about the issues or rules encapsulated in opinions (e.g., Segal 1984; Maltzman, Spriggs, and Wahlbeck 2000; Kritzer and Richards 2002; Bartels 2009; Bailey and Maltzman 2011).

It is important to recognize the trade-offs between these data types. Using raw text as data demands a great deal of the model used to analyze the data: the information is all there in the text, but existing "bag of words" models for analyzing text use little of that information. Legal texts have many phrases of art, and much of the important action in the evolution of doctrine is in the creation of new, typically multi-word, terminology (e.g., the "Lemon test"). The move to analyzing citations instead of the raw text adds relational structure to the data, but also gives up a tremendous amount of information. The lost information is not only the arguments in the opinion texts, but also the substantive meaning of the citation. Some of the latter can be recovered by manual coding (Spriggs and Hansford 2000; Clark and Lauderdale 2010), but the former is lost entirely. Direct manual encoding of legally relevant information places almost all the demands on the careful specification and execution of the coding process, whereas subsequent modeling is relatively straightforward because the quantities of most interest are being directly assessed by the coders.

2.2 *Models of Doctrine*

Models of doctrine have differentially made use of these various sources of data about opinion content. Roughly speaking, those models fall into a small number of types: categorical models, spatial models, and classification models. *Categorical models* seek to divide opinions into substantively defined groups. Perhaps the most widely applied measures of judicial opinion content are the issue and issue area codes included in the Supreme Court database, which identify the one or two major legal issues with which each case deals and place the cases into discrete "bins." Beyond these and other expert codings, recent innovations in the analysis of texts have yielded tools that have proven (or could prove) useful for describing the global structure of Supreme Court doctrine. Among these tools are the discrete topic model (e.g., Quinn et al. 2010), and the closely related mixture topic model (e.g., Blei, Ng, and Jordan 2003). These models also seek to categorize opinions into "bins," though using a probability model of the words in the opinion rather than expert judgment. A related approach involves using network analysis tools to identify clusters of cases that are closely linked (through their citations)—so-called community detection models (e.g., Porter, Onnela, and Mucha 2009; Bommarito et al. 2010). However, these approaches are better suited to identifying the substantive and factual similarity among cases rather than how the doctrine within those issues has evolved through a particular sequence of cases.

Spatial models, by contrast, seek to capture variation in the ideological valence espoused in opinions, rather than the factual or substantive context of the case. While some research has had success in placing opinions relative to each other in terms of the ideological valence of their policy content (e.g., McGuire and Vanberg 2005; Clark and Lauderdale 2010), those models do not contain information about the nonspatial (i.e., logical, doctrinal) relationships among opinions. Moreover, spatial models do not leverage the sequential nature of judge-made law. As a consequence, those approaches are not suited to the study of doctrinal development.

Classification models seek to uncover the underlying doctrinal content by analyzing a set of case outcomes. A recent example is Kastellec (2010), who employs classification and regression trees (CARTs) to infer the doctrinal structure within search and seizure cases. His approach uses dichotomous, manually encoded, factual indicators from search and seizure cases along with the observed case outcome (did the Supreme Court allow the evidence to be admitted or not) to assess which facts are most relevant for the Supreme Court's decision. However, while the CART approach may be able to effectively describe the decision tree that best explains the pattern of outcomes from individual cases, it is not well suited to capture the doctrinal content of individual cases or the substantive relationships among cases. Nor can it be used to empirically evaluate theories about how the Court develops doctrine or will decide future cases. Kastellec's approach grows out of a long tradition in judicial politics research that involves the coding of facts from cases

and then estimating an empirical model on distinct subsets of cases to assess the predictive power of those facts and whether the "influence" of those facts on the case outcome changes over time (e.g., Kort 1957; Segal 1984; McGuire 1990; George and Epstein 1992; Songer and Haire 1992; Ignagni 1994; Benesh 2002; Richards and Kritzer 2002; Bartels 2009; Lax and Rader 2009). Finally, an important limitation on classification models is that they rely on manual coding of factual patterns in individual cases, and how one identifies the potentially relevant facts to be coded remains a serious problem.

2.3 *A Genealogical Model of Doctrine*

We provide a new tool for characterizing Supreme Court doctrine by combining some of the attractive theoretical features of the classification approach for thinking about legal rules with the tractability of the categorical and spatial approaches for analyzing data. We begin with a different conceptualization of the underlying processes that generate opinions. We see Supreme Court doctrine as a growing genealogy of court opinions that sequentially build from each other to shape and elaborate law. Fortunately, legal reasoning has a particular structure, which we can exploit to systematically study the patterns in Supreme Court doctrine.

Judges who write opinions, and Supreme Court justices in particular, use those opinions to communicate their interpretation of the law to others—lower court judges, lawyers, police officers, legal academics, the public at large, etc. (Maltz 2000; Bueno de Mesquita and Stephenson 2002). To do so, they follow the method of legal reasoning known as the *case method*, which involves identifying the most legally relevant precedents, explaining their relationship to the instant case, and interpreting precedent and applying it to the new case (Levi 1949). Indeed, by contrast with civil law systems, the case method is deeply embedded in the framework of American legal education (e.g., Patterson 1951) and serves as the bedrock of American legal argumentation. Because the judges are essentially engaged in argument by analogy, they engage the most legally relevant precedent the most, though they occasionally discuss other precedents that can help clarify where the new case fits into the line of doctrine, any implications it may have for those other cases, and how the new case is different (or not) from past ones. Critically, though, in order to achieve their primary goal of advancing the line of doctrine in which they are working, Supreme Court opinions draw primarily on the most directly relevant precedents. In other words, the opinions engage and discuss the most legally relevant precedents the most. Given the way that Supreme Court opinions are structured, we adopt a particular assumption on which our model rests. Specifically, it is our assumption that even mere citation counts reveal much of the structure of these genealogies. Opinions cite close relatives more frequently than more distantly related opinions.

An example helps illuminate this feature of doctrine. Abortion opinions in the 1970s and 1980s heavily cited the logic of *Roe v. Wade*, while partial-birth abortion cases in the past decade have more heavily cited *Planned Parenthood v. Casey* than *Roe*. Indeed, it is widely claimed that *Casey* replaced *Roe* as the new authority, and most central case, in abortion law. *Casey* eschewed the trimester framework that had dominated all abortion law jurisprudence for twenty years, replacing it with a new "undue burden" standard. Thus, whereas traditional analyses of Supreme Court citation networks find that *Roe* is a very "authoritative" opinion, most scholars of the law would agree that *Roe* has been supplanted by *Casey*. This can be seen in the fact that contemporary precedents give *Casey* a place of prominence and cite it more heavily within a given opinion than they do *Roe*, which is now cited less often within any given opinion (even though it is nearly always cited at least once). What this example illustrates, then, is that the significance of Supreme Court opinions as authoritative statements of the law changes over time. By only identifying the fact of citation between opinions, previous research has missed the more important issue of which cases are more authoritative, as measured by their prominence in the opinion itself. That prominence is correlated with how many times it is invoked in the opinion's reasoning.

In designing and implementing our genealogical model, we make two advances that facilitate future empirical analysis of legal doctrine and the evaluation of a range of theoretical models at the heart of judicial politics. First, we extract more substantively meaningful information from citation

patterns than has been done previously. Doing this requires connecting citation data to an underlying model of how legal doctrine evolves over time. In particular, we contend that the extent to which a new case relies on a past precedent, as measured by the extent to which the new case cites the past precedent, conveys information about the doctrinal connection between the two cases. Second, our model allows the researcher to overcome the limitation imposed on past research by the reliance on labor-intensive, and potentially problematic, manual coding of data from judicial opinions.

3 Estimating the Genealogy of Doctrine Using Citations

We propose modeling judicial doctrine as a genealogical structure. Doctrine evolves through a series of cases: new cases building upon, clarifying, and elaborating older cases (e.g., Patterson 1951; Kornhauser 1992a). Older cases cannot, of course, build upon newer cases. In this sense, a new case can be thought of as a "child" case of one or more "parent" cases. In this section, we develop a probability model that provides an estimate of which cases are most aptly described as parent cases to each new case. This allows us to provide simple descriptions of how a line of legal argumentation or doctrine develops and evolves through a series of linked cases.

3.1 *Data*

Previous studies of Supreme Court citation patterns—network and otherwise—make use of data consisting of every case-pair dyad for which at least one citation exists (e.g., Fowler et al. 2007; Clark and Lauderdale 2010). The collection of these data sets is in-and-of-itself an impressive feat. However, as has been emphasized already, there is a richer source of information on the citation network: how many times a given opinion references a previous opinion. There exists no previously assembled data set containing this information, so we assemble a data set consisting of every case-dyad pair among cases decided by the Supreme Court since its first session in 1790 through the October 2009 term. For each dyad, we identify the number of references from the first opinion ("the case") to the second opinion ("the precedent"). We code both full string citations and subsequent mentions of the case name within the text to be citations.[1]

To obtain these data, we use the full text of nearly all Supreme Court opinions, which is available from OpenJurist (http://www.openjurist.org). We then use a Perl script to search through the full text of each opinion to locate every reference to a Supreme Court case. This process involves a two-pass procedure, the first pass identifying only full citations and the second pass identifying mentions of the precedent employing an abbreviation of the citation (typically the first petitioner's name or the respondent's name if the USA is the petitioner).[2] The resulting data compose a citation count matrix, Y, with dimensions 18713×18713 (there are 18713 cases in our data[3]). In this matrix,

[1] All data and replication materials are available in Clark and Lauderdale (2012).

[2] We only count citations, which omits references to cases by name only. For example, the plurality opinion in *Planned Parenthood v. Casey* repeatedly mentions *Roe v. Wade* with the word *Roe* only, rather than a parenthetical citation. Our coding procedure does not capture those references. This decision is the result of several considerations. Including such references is likely to overcount citations, as it can be difficult to distinguish case names from references to individuals or other actors that may be party to a past case. However, undercounting citations by our approach results in a measure that is still very highly correlated with the actual number of citations. As we find below, the *Casey* example is a nice case-in-point. Although we undercount the references to *Roe*, our model still estimates a very high probability that *Casey* is a direct doctrinal descendent of *Roe*.

[3] We note that the number of opinions in our data differs from other studies of Supreme Court citation networks (e.g. Fowler et al. 2007). There are several reasons for this difference. First, our sample is limited to full opinions delivered by the Court, in its appellate capacity (or under its constitutional original jurisdiction), which have citations in them to past cases. As a consequence, we do not include cases that the Court decided either without opinion or with a brief, cursory opinion with no citations in it. During the 19th century, especially during the early part of the 19th century, the Court decided many cases that met one of these two criteria. Second, our unit of analysis is the Court opinion, not the case citation. The Court sometimes consolidates multiple cases into a single opinion. As a consequence of these two selection criteria, there are fewer "cases" in our data than in studies that rely on *Shepard's* citations. However, after the early part of the 20th century (essentially, once the Court had virtually entirely discretionary jurisdiction), the number of opinions and number of case citations per *Shepard's* become much more comparable. As we discuss below, we focus our analysis in this article on those years.

20,372,788 cells take on the value zero (there were no observed citations for the dyad). The most frequent nonzero count is one (30,702 instances), and the maximum number of observed citations to a single precedent is thirty-one. However, 45,245 cells take on values greater than one, which is one source of information we take advantage of to estimate our model.

3.2 *Model*

Our data consist of an $n \times n$ matrix of citation counts Y, where element y_{ij} is the number of citations from the opinion represented in row i to the opinion represented in column j. Majority opinions are indexed chronologically, so the matrix of citation counts is strictly lower triangular (no opinion cites itself or a subsequent opinion).[4]

$$Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ y_{21} & 0 & 0 & 0 \\ y_{31} & y_{32} & 0 & 0 \\ y_{41} & y_{42} & y_{43} & 0 \end{bmatrix}$$

To model the probability of observing any particular distribution of citations to previous opinions, we use an approach similar to the *bag of words* approach to modeling textual data. The actual citations in a given opinion are assumed to come from a distribution over previous opinions, where each citation observed in the opinion is drawn from the bag of precedents with some vector of probabilities $\boldsymbol{\pi}$, which depend on the parameters of the model. In models of textual data, scholars typically use a categorical model for $\boldsymbol{\pi}$: There are a fixed number of topics, each with its own distribution of words. In those models, the observed distribution of words reveals information about which of the topics a document is most likely to come from (e.g., Blei, Ng, and Jordan 2003). For our problem, we need a different model: The body of precedent grows over time, and citations can only be to previous opinions.

To generate a probability model for citation counts, we conceptualize the structure of Supreme Court opinion as a *tree* or *genealogy*. Like a human genealogy or family tree, this structure consists of parent–child relationships. However, we are not constrained to have a specific number of parents for each child, so there are many kinds of genealogical structures that we could try to estimate. How do we decide the number of parents an opinion is allowed to have? A parent opinion is a precedent from which a new opinion directly builds. Although many previous opinions may be doctrinally or substantively relevant or connected, we seek to identify the past opinions a new opinion builds from most directly. A single-parent model is the simplest model we can consider: Each opinion is assumed to build directly from only one previous opinion. This is not to say that only one precedent is cited—or even that the parent is necessarily the most cited precedent—but as a simple way of describing the most important precedents for each new opinion, the single-parent model is elegant and relatively easy to estimate. Figure 1 shows an example of such a tree for modern reproductive rights law, where this structure works quite well. Alternatively, we can contemplate multiple-parent models, and we discus this option below.

Mathematically, we describe the structure of the tree using a "parent matrix," Ψ. This matrix is square—each opinion is represented by a row and a column, where the opinions are indexed chronologically. This matrix identifies the parent of opinion in row i with a 1 in column j, where the opinion represented by column j is the parent opinion to the one represented in row i. Thus, one can identify the parent opinion to any new opinion by finding the new opinion's row in Ψ and locating the column(s) that takes on the value 1. A corresponding "child matrix" is given by the transpose of the parent matrix, $\Psi^{\top}$. The matrix Ψ fully describes a tree of legal opinions. Thus, if one examines row i in $\Psi^{\top}$, each column where the row takes on the value 1 indicates a child case to

[4]The matrix, although strictly lower triangular in theory, is not quite so in practice. The reason is that sometimes two opinions are written essentially concurrently, delivered on the same day, and cross-cite each other. As a consequence, although one opinion is technically chronologically first, the two opinions do cite each other, leading to a few minor deviations from the strictly lower-triangular shape of the data matrix.

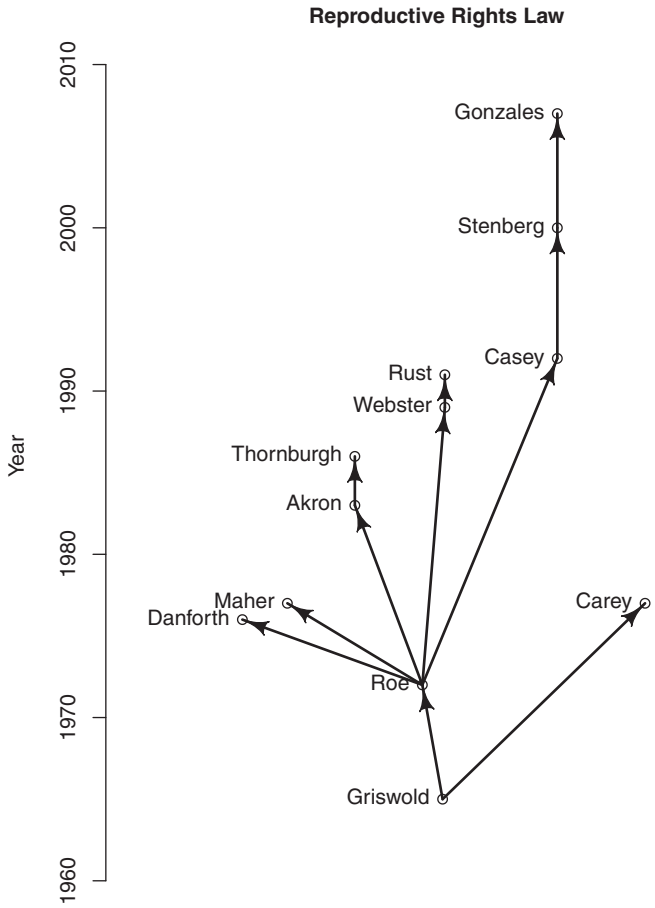

Fig. 1 Sample single-parent tree for modern reproductive rights cases. Tree estimated using a limited set of cases for illustrative purposes; below, we estimate the full doctrine of abortion law using all abortion cases.

opinion i. Consider a simple tree with three opinions, 1, 2, 3, in which opinions 2 and 3 are attached to opinion 1 (opinion 1 is the parent to opinions 2 and 3). In this example, Ψ is simply

$$\Psi = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \qquad \Psi^{\top} = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{1}$$

In the probability model described below, we assume that the relative probability of a new opinion i citing a precedent opinion j is a function of the distance between the two opinions in the tree. Distance is defined as the number of opinions through which one must go in order to get to another opinion. In Fig. 1, for example, *Rust* is one step away from *Webster*, two steps away from *Roe*, and four steps away from *Stenberg*. We denote as $D_{ij}(\Psi)$ the distance between opinion i and opinion j in the genealogy tree described by the matrix Ψ. It is our assumption that this distance is a function of citation patterns and yields information about the doctrinal connections among cases.

3.3 *Specification*

There are many ways to model the vector of probabilities that each opinion cites the other one conditional upon the distance between the two opinions. First, state our model and then explain our modeling choices and available alternatives. We assume that the log-odds of citation decrease linearly in the minimum distance from the new opinion $i \in 2, 3, \ldots, n$ to the precedent j along the

branches of the tree. For opinion i, the probability of a given citation being to precedent $j \in 1, 2, \ldots, i-1$ is $\pi_{ij}(\beta, \Psi)$. We assume that the probability distribution of the data, conditional on this vector of relative citation rates, is multivariate Polya (Dirichlet-Multinomial) distributed with parameters π_{ij} and ν. Our probability model is given by

$$\pi_{ij}(\beta, \Psi) = \frac{e\left(\beta \cdot D_{ij}(\Psi)\right)}{\sum_{j'=1}^{i-1} e\left(\beta \cdot D_{ij'}(\Psi)\right)} \tag{2}$$

$$p(y_i|\pi_i, \nu) = \frac{\Gamma\left(1 + \sum_{j=1}^{i-1} y_{ij}\right)}{\prod_{j=1}^{i-1} \Gamma\left(1 + y_{ij}\right)} \cdot \frac{\Gamma\left(\sum_{j=1}^{i-1} \nu\pi_{ij}\right)}{\Gamma\left(\left(\sum_{j=1}^{i-1} y_{ij}\right) + \left(\sum_{j=1}^{i-1} \nu\pi_{ij}\right)\right)} \cdot \prod_{j=1}^{i-1} \frac{\Gamma\left(y_{ij} + \nu\pi_{ij}\right)}{\Gamma\left(\nu\pi_{ij}\right)}. \tag{3}$$

Equations 2 and 3 describe our probability model for the citation counts, conditional on scalar parameters β and ν and a strictly lower-triangular binary matrix Ψ. The parameter β measures the effect of distance $D_{ij}(\Psi)$ on the propensity to cite a precedent. This functional form treats steps up the tree and down the tree as equivalent for the purposes of calculating relative probability of citation. We restrict β to be negative by prior assumption, enforcing the logic that the probability of citation is decreasing (rather than increasing) in the distance from opinion to precedent along the tree.

The parameter ν measures the degree of overdispersion of the multivariate Polya distribution by comparison to the multinomial distribution, with smaller values indicating greater overdispersion. We want to allow for the possibility that other factors besides the modeled genealogy of precedents play a role in determining how frequently an opinion cites each relevant precedent.[5] The multivariate Polya distribution assumes that there is an *unobserved* vector of probabilities characterizing relative citation rates for each individual opinion, which itself is drawn from a Dirichlet distribution that is a function of location in the tree. The multivariate Polya distribution (sometimes called the Dirichlet-multinomial) captures the inevitable fact that our model for $\boldsymbol{\pi_i}$ will not be exact for every opinion in the tree. Importantly, using this distribution protects us against being over confident in where we locate opinions on the tree due to model misspecification.

The estimated parent matrix Ψ has a single element that is equal to 1 in each row (except for the initial, parentless opinion). To gain intuition for how Ψ shapes the relative citation probabilities, it is useful to return to the example of the three-opinion tree described by Equation (1). Recall that the tree has two nodes both connected to an initial node. Denote the vector of relative probabilities of citation by each case i to each precedent opinion, conditional on case i's parent being precedent j as $\boldsymbol{\pi_i}(\Psi_{ij} = 1)$. Here, Ψ_{ij} refers to the tree where opinion i is a child opinion to opinion j. We can then express the vector of relative probabilities $\boldsymbol{\pi_4}$ for a new opinion $i = 4$ citing each of the three existing opinions, conditional on which existing opinion is its parent:

$$\begin{aligned}
\pi_4(\Psi_{41} = 1) &\propto \left(e^{\beta}, e^{2\beta}, e^{2\beta}\right) \\
\pi_4(\Psi_{42} = 1) &\propto \left(e^{2\beta}, e^{\beta}, e^{3\beta}\right) \\
\pi_4(\Psi_{43} = 1) &\propto \left(e^{2\beta}, e^{3\beta}, e^{\beta}\right).
\end{aligned}$$

Since $\beta < 0$, if opinion 4 is attached to opinion 1, it will cite opinion 1 at the highest rate, and opinions 2 and 3 at the same, lower rate. This is because, following our assumptions, opinion 4 is only one step away from opinion 1 but two steps away from both opinions 2 and 3. If it is attached to opinion 2, it will cite opinion 2 most, then opinion 1, and then opinion 3, as it is 1, 2, and 3 steps away from those opinions, respectively. Similarly, if it is attached to opinion 3, it will cite opinion 3 most, then opinion 1, and then opinion 2. These predicted probabilities are always different for all possible locations that a new opinion might attach to the existing tree, which guarantees that (conditional on our assumptions) the model will be statistically identified.

[5] For example, the author of the opinion, the quality of the precedents, among other things (Hansford and Spriggs 2006).

3.4 *Modeling Choices and Extensions*

Before presenting the results of our estimation strategy and results, we discuss four modeling choices that we have made, the rationale and implications behind our choices, and potential extensions of our model.

3.4.1 Functional form

One modeling choice concerns the functional form we have specified for our probability model. There are two pieces in particular that warrant justification. First, we have assumed that the probability that a precedent is cited can be represented as a logit function of the linear distance between a new case and the precedent in the tree. We considered several more heavily parameterized specifications, but found little improvement in fit. We considered models where the log odds were permitted to vary as a power of distance other than 1, but recovered estimates for the power that were very near 1. We considered models estimating different β parameters for upward and downward steps in the tree, and found that the resulting parameter estimates tended to be nearly identical. A single, linear penalty parameter β for distance in the tree works well in the cases we have considered. One might also consider changing the link function, perhaps representing the probability with a probit instead of logit link function, which would allow the probability of citation to drop off slightly more quickly in distance between cases.

3.4.2 Multiple parents

As noted, our model assumes that each case has a single-parent case. However, there are reasons one might prefer a more flexible model that allows each case to have multiple parents. That additional flexibility, of course, comes with a cost. It requires making assumptions about the relative frequency of two-parent nodes and how strong the relationship between two cases must be in order to assign a second case as a parent. One indicator that a single-parent model may not be correct would be if the posterior probability of "parenthood" for the most likely parent were very low—i.e., the model results in a fairly uncertain posterior about which precedent is each case's parent. As we discuss below, our single-parent model results in relatively high posterior probabilities of parenthood for each case's most likely parent. We have estimated two-parent models on the data we consider and find that the single-parent models fit the data nearly as well. Thus, we focus here on the single-parent model for its simplicity and parsimony.

3.4.3 Multiple founders

A related modeling issue concerns how many cases are assumed to be foundational cases in the area of law modeled. We have focused here on a model that assumes one founding case, from which all subsequent cases build (the case chronologically first in a given subset of cases); however, one could allow for multiple founding cases. Estimating this model requires making a decision about the number of parents a case may have, as well as how to model probability of citation for unconnected cases. Under such a structure, lines of doctrine building from each founding case will remain separate forever, whereas a multiple-parent model would allow those lines of doctrine to recombine—i.e., a new case could have parents that descend from the distinct foundational cases. That flexibility, though, requires making assumptions about the citation relationships between cases that are not connected in the tree, as well as precisely how many cases should be treated as foundational. Because we have selected substantively defined subsets of cases to which to apply our model, we believe a single-foundational case model is appropriate, and experimentation with the number of foundational cases suggests there is only negligible gain from adding additional foundational cases. By contrast, applying the model to more broadly defined sets of cases might imply that a multiple-foundational case model would be more appropriate. One possible avenue for future modeling extensions would be to estimate the number of foundational cases that best fits a given data set.

3.4.4 Selecting data to model

When estimating the model we focus on in this article, one must choose a coherent set of cases to consider. This is relatively easy if the subset has been chosen to be a particular area of law with a key founding opinion. For example, in reproductive law, *Griswold* is the obvious starting point both chronologically and in terms of the development of the law in that area. However, what if we had a data set that also included *Katz v. United States*, a major precedent on what constitutes an illegal search under the Fourth Amendment? *Katz* was decided two years after *Griswold*, but it cites *Griswold* in only a single footnote in the majority opinion because it is in a distant area of the law. In the model we focus on here, if *Griswold* was the first case and *Katz* was the second case in the data set, the estimator would be forced to attach *Katz* to *Griswold*, despite the two cases having little relationship. One way to solve this problem is to only look at subsets of cases that are closely related and to choose the initial parent. Another way to solve the problem is to move to a multiple parent and founder model. To consider the entire corpus of Supreme Court decisions at once, one would need to move to such a model.[6] In the application of our model here, we apply the probability model to discrete groups of substantively linked cases, as identified by the Supreme Court database. The first founding opinion is always the first opinion in the range of time we study. Of course, this choice creates a limitation in our ability to make direct comparisons across areas of the law. Our grouping of cases represents our substantive judgment about how best to group cases (note, we often combine multiple "issues"), balancing these competing interests. Particular applications in the future should consider carefully these competing interests and model the data in the way best suited to their particular needs.

For all these reasons, our model is far from the only way to estimate the kind of genealogical structure we are interested in. There are a variety of more complicated assumptions that one could make about relative citation rates, number of parents, and other components of the model. Our setup has the virtue of being fairly transparent in its assumptions about how the rate at which an opinion cites different precedents depends on its genealogical relationship to those precedents and is relatively easy to estimate (as we show in the next section). We believe it is an important point of departure for consideration of an entire class of models that distill citation count data into parsimonious genealogical descriptions.

3.5 *Estimation*

Using this probability model, we could estimate β, ν, and Ψ by either maximum likelihood (ML) or Bayesian MCMC techniques. However, maximizing the likelihood over the discrete tree structures is computationally infeasible for anything beyond very small trees. The number of possible trees grows as $(n-1)!$. Although a reasonable search algorithm need not consider every one of these trees, this is still a very difficult maximization problem. These problems are not entirely mitigated by using a Bayesian MCMC estimator: If there are multiple modes that describe very different trees, the sampler will still have a difficult time moving between the two modes. However, we do get the nontrivial benefit of an uncertainty estimate over possible trees, which is a substantial advantage over ML in discrete parameter spaces. Thus, we focus on the Bayesian MCMC estimator.

[6]In principle, one can imagine going all the way back to the U.S. Constitution—either taken monolithically or as clauses and amendments separately—as the initial founding precedent(s) and count when it is cited or mentioned in the opinions. This strategy creates its own problems, though, because citations to the Constitution are fundamentally different than to precedent, because it is the Constitution that the Court is interpreting. Large numbers of citations to the Constitution need not imply that an opinion is a direct child of the Constitution, so the relatively simple structure we describe for relative citation rates may be inappropriate if citations to the Constitution are included in the data. What is more, many of the Court's cases are not constitutional cases but arise from the Court's functions outside constitutional interpretation, such as statutory construction, common law, and judicial administration. Identifying the "foundational" authorities in such cases further complicates the matter.

We adopt a proper uniform prior over trees: For each new opinion, each preceding opinion has the same prior probability of being the parent, $p(\Psi_{ij}=1)=1/(i-1)$ iff $j<i$, otherwise $p(\Psi_{ij}=1)=0$. We also adopt uniform priors for $\beta \sim \mathcal{U}(-2, 0)$ and for $\nu \sim \mathcal{U}(0, 50)$.[7]

Our MCMC estimation consists of independence Metropolis steps for individual parameters, conditioning on the other parameters. We set start values $\beta=-1$, $\nu=10$, and Ψ such that all opinions are attached to the first chronological node (a star network topology). To sample the parent matrix Ψ, we step through the opinions (starting with $i=2$) in chronological order. For each opinion, we use a independence Metropolis step to propose—and then accept or reject—alternative connection points for node i. The proposal distribution for the Metropolis step for a given row of Ψ is proportional to the citation counts in the corresponding row of Y plus a small constant.[8] Having completed a cycle through each row of Ψ, we sample β and ν. The proposal distributions for the independence Metropolis steps for β and ν are the parameters' prior distributions.

This sampling procedure has two notable features. First, although the independence Metropolis steps for the parameters β and ν are somewhat inefficient, independence Metropolis is a relatively efficient way to sample Ψ. One could compute the exact posterior probability for each possible location $1, 2, \ldots, i-1$ that opinion i might attach and perform a direct sample over that conditional distribution; however, this is less computationally efficient because it involves computing the likelihood function $i-1$ rather than 1 time for opinion i at each iteration.[9] The direct approach is somewhat more efficient at exploring the posterior per MCMC iteration, but the number of iterations that the Metropolis simulation can complete in the same time as the direct sampling approach grows approximately linearly as a function of the number of opinions in the data set. Our motivation for using independence Metropolis rather than a Metropolis or Metropolis-Hastings step that conditions on the existing parameter value when making proposals is the difficulty of tuning these procedures in a discrete parameter space. Proposal acceptance probabilities do not provide a useful diagnostic for Metropolis steps on discrete parameter spaces, because if the posterior distribution is heavily concentrated on a particular value, very few proposed moves should be accepted.

Second, because of the way the parent matrix specifies the tree structure, proposals to change the parent of case i from precedent j ($\Psi_{ij}=1$) to precedent j' ($\Psi_{ij'}=1$) involve moving the entire subtree of node i and its offspring. Heuristically, one can think of this as a proposal to chop off a part of the tree and graft it to a new location on the trunk. This is a good sampling procedure because it does not disturb the subtree of cases attached to case i. The contributions to the likelihood made by these subtrees of cases are being stochastically optimized in the MCMC, and so conserving their structure in proposals will tend to yield proposals with higher probabilities of acceptance. Each proposal does change the likelihood with respect to the citations from i and its offspring to all other cases in the data, and it is on the basis of this change that the decision to accept or reject the proposal is made.[10]

[7]We employ uninformative priors because we want to avoid using the information to which we apply our model twice. In particular, most sources of information that we might use to construct an informative prior—oral arguments, briefs from litigants, internal memoranda among the justices—contain the same information as the citation data we model, simply filtered through a different model. Thus, while informative priors might prove useful, we adopt our priors to be careful not to use our limited information twice. It also bears mentioning that when applying the model to a limited subset of cases (as we do below), we are in effect employing an informative prior, by placing 0 prior probability on cases outside the subset we examine.

[8]Adding a small constant to the observed citation counts ensures that uncited cases can still be proposed. We use $5/(i-1)$ as the constant: Scaling by $(i-1)^{-1}$ ensures that for large i most of the proposal distribution's weight is still on the cases that are actually cited.

[9]In this problem, by far the largest computational expense is computing the likelihood function. We find that using a Metropolis step for each row of Ψ yields a computational time that grows as $\approx n^{5/2}$, whereas sampling from the true distribution yields a simulation whose computational time grows as $\approx n^{7/2}$.

[10]The results we report below are based on a 2000-iteration simulation, with a 200-iteration burn-in period. Although these sample chains are short, diagnostics suggest that the model converges very quickly and that 2000 iterations are sufficient for a well-converged chain. Because the parameter space is discrete (other than for the β and ν parameters), visual inspection of traceplots is not a particularly useful means of evaluating convergence. We have instead compared the estimated parent across multiple chains of several subsets of cases. We have found that the two chains result in the same estimated parent for 97% of the cases. The differences in the most likely parents across the two chains are due to a

Although the parameter space of the model is large, it does not grow as fast as the data as we increase by the number of cases n. We are estimating two continuous variables for the entire model (β and ν), plus a single categorical variable over $i-1$ alternatives for each case $i \in 1, \ldots, n$. Each additional case provides information about citations to *all* of the previous cases. Many of these citation counts are zeros, but those zeros are very informative about the likely location of the new case within the tree. As the number of potential attachment points for case i grows, so does the information we gain from the precedents that it does and does not cite.

3.6 *Results*

We apply our estimator individually to twenty distinct subsets of cases decided since 1953: abortion, attorneys' fees, civil rights liability, desegregation, double jeopardy, establishment clause, federal utilities regulation, free exercise, juries, interstate relations, libel, national supremacy, obscenity, patents, preemption of state legislation, privacy, search and seizure, self incrimination, standing, and voting rights cases. We identify the cases falling into each group by reference to the "issue" code assigned to the case in the Supreme Court database (Spaeth et al. 2010).

3.6.1 Evaluating model fit

To assess how well our model fits our data, we first compare the actual number of citations from each opinion to each precedent with the predicted number of citations using our posterior estimates of the model parameters. This kind of posterior predictive check verifies that our functional form assumptions are not in conflict with the data. Conditioning on the total number of citations made in each case, we next use the parameter values at each iteration of the MCMC to calculate predicted counts. We then take the mean prediction across all iterations of the MCMC as the mean posterior predication.

Figure 2 shows the comparison between posterior predicted and actual citation counts for each of the twenty subsets of cases to which we apply our model. The y-axes measure the number of actual citations (the data), and the x-axes measure the predicted number of citations. We plot points for each case dyad consisting of a case i and a precedent $j < i$. In each plot, we include a line showing the diagonal and a smooth spline fit to the data. (To ease visual interpretation, we have plotted the points on the square-root scale, though the smooth spline was fitted to the untransformed data.) As these plots show, our model fits the data well. Our model tends to slightly over predict the citation counts for cases that are distant in the tree (where expected citation counts are low) and to slightly under predict the citation counts for cases that are nearby in the tree (where expected citation counts are high). This suggests that we have not found exactly the optimal functional form describing how the relative citation rates call off as a function of distance in the tree. However, the magnitude of these deviations is substantively very small, and our model fits the data well across all the subsets of cases.

A second way to evaluate model fit is to check how confident we can be about the parent–child relationships among these data. Figure 3 shows the distribution of the mean posterior probability for the highest probability parent node among all cases in each of the eleven subsets of cases we study. Among most sets of cases, the citation data and probability model provide an average probability of parenthood greater than 60% for the most likely parent. Critically, in nearly all circumstances, the distribution is skewed away from 0, suggesting that rather than identifying many possible parents, each with low posterior probability, we are identifying a handful of parents with relatively high likelihoods of being the parent and usually assigning greater than 60% probability to the most likely parent. One exception to this pattern is the set of interstate relations cases, where the median posterior probability of parenthood for the most likely parent is below 40%.

handful of cases that are very near posterior indifference between two parents. We have also compared the estimated posterior probability of parenthood for the most likely parent across two chains, and found this correlation to also be 97%.

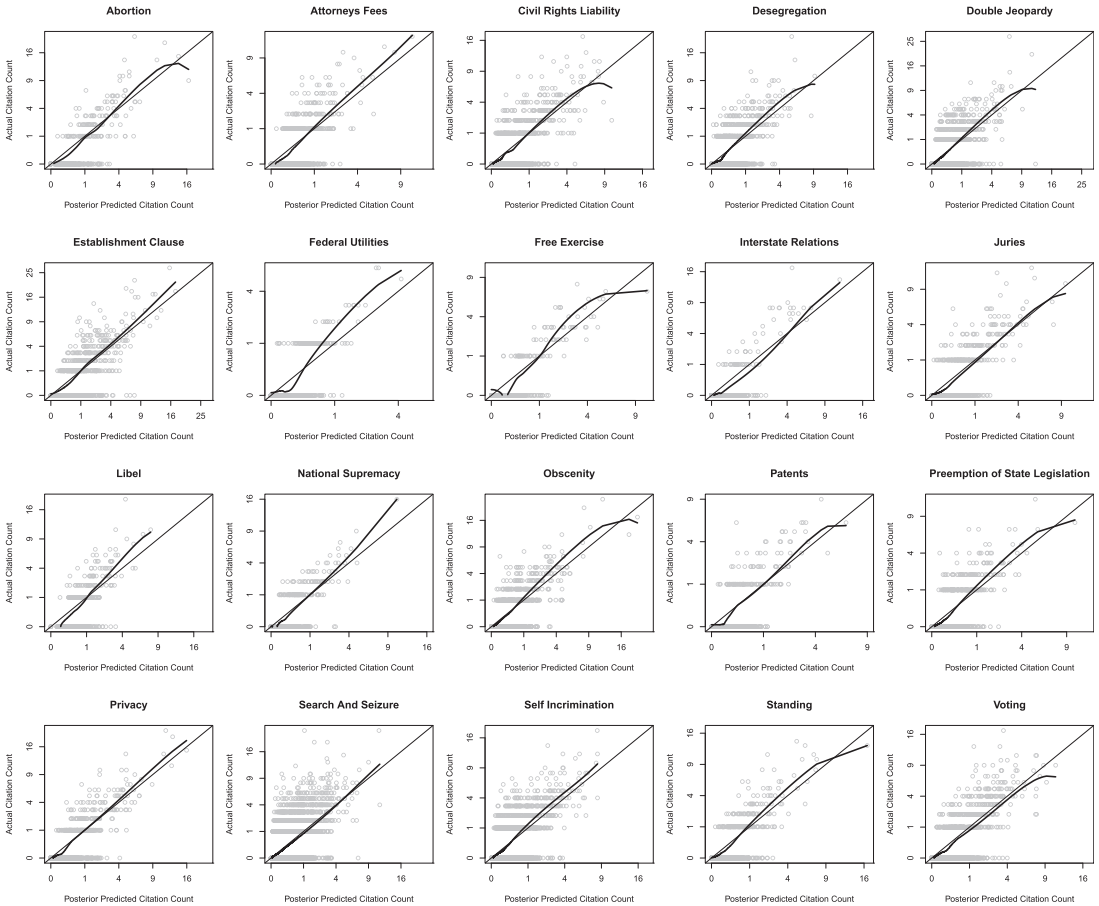

Fig. 2 Posterior predictive check. For each case subset, the *x*-axis shows the mean posterior predicted citation count and the *y*-axis shows the actual citation count for a case dyad. The loess closely approximates the diagonal in most case subsets, indicating that the model fits the data well.

3.6.2 An example recovered tree: abortion law

Setting aside the uncertainty in our estimates, we turn to the underlying doctrinal structure our model recovers. The most instructive way to summarize our results is visual. (We could alternatively express the estimated genealogy using the parent matrix representation depicted in Equation (1).) Figure 4 shows the posterior genealogy tree for the set of cases in the "abortion" issue. In this figure, the cases are aligned vertically according to the date they were decided; the text showing each case's name is scaled proportionally to the number of child cases we estimate for the case. The gray lines show parent–child links between cases. As noted above, it is important not to draw too strong conclusions about the first case in chronological order. We have assumed the first case in our data is the "founder" of the doctrine, which may not be true. Nevertheless, the early cases here reveal an intuitive and striking pattern. *Roe v. Wade* (bottom of Fig. 4) is the child case of *Eisenstadt v. Baird*. *Roe*, in turn, has nine direct child cases.

One notable child case is *Bellotti v. Baird*. This seminal case involved a challenge by a Massachusetts law that required minors to obtain parental consent before obtaining an abortion. The Court held that parental consent laws were constitutional as long as they provide an opportunity for a court to waive that requirement—a so-called "judicial bypass." *Bellotti*, in turn, has four child cases—*H.L. v. Matheson*, *Planned Parenthood v. Ashcroft*, *Hodgson v. Minnesota*, and *Ohio v. Akron*—each of which dealt with filling in further details about parental consent requirements. *Matheson* dealt with whether a state may require a doctor to notify a teenager's parents before performing an abortion, even if consent is not required (the Court said yes).

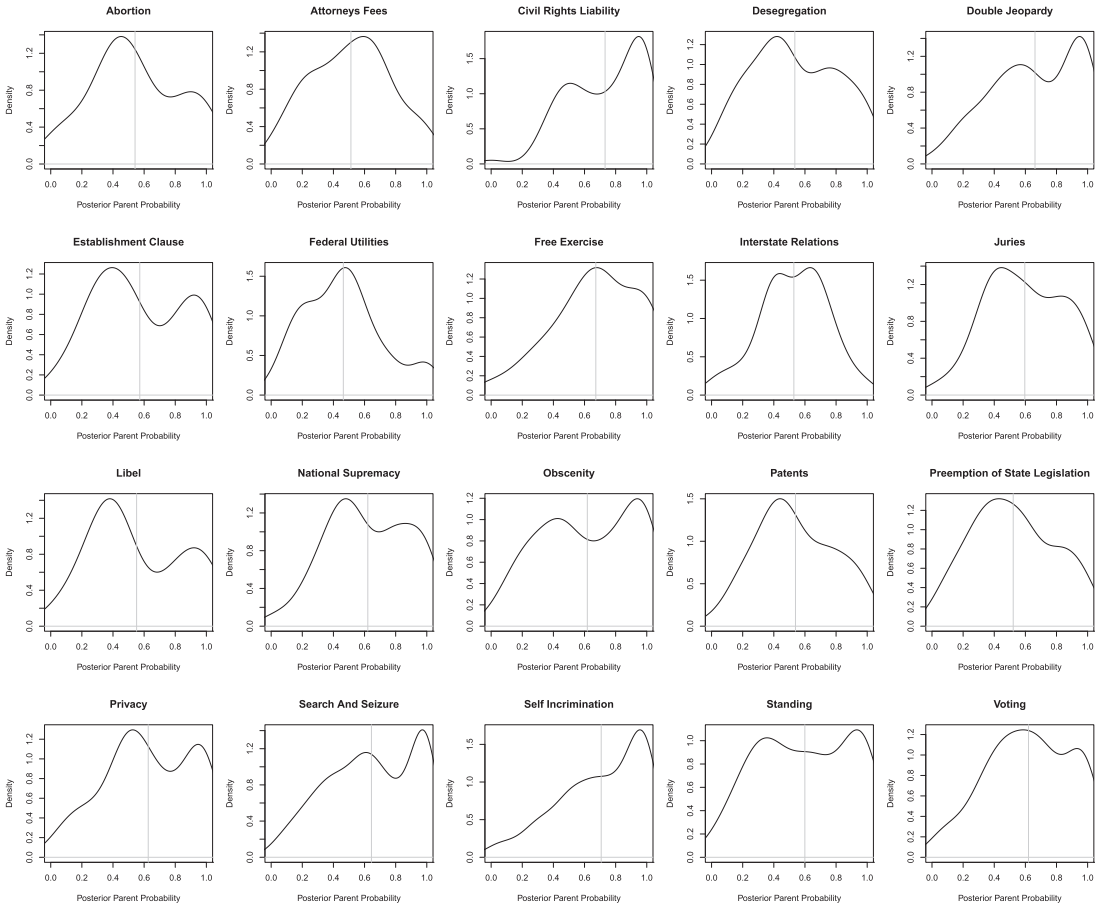

Fig. 3 Distribution of posterior probabilities of parent–child relationship for most likely parent. The *x*-axis shows the mean posterior probability of parenthood for the most likely parent for each case; in each panel, the vertical gray line shows the mean of the distribution.

Ashcroft similarly dealt with the conditions for granting a judicial bypass, identifying maturity as one of the primary criteria for the decision. *Hodgson* answered whether a parental consent law could require that *both* parents be informed. Finally, *Ohio v. Akron* dealt with whether the Constitution requires the availability of a judicial bypass to a parental notice law, as opposed to a parental consent law.

A distinct line of descendants from *Roe* are *Planned Parenthood v. Danforth*, *Colautti v. Franklin*, and *Anders v. Floyd*. These three cases dealt with the types of restrictions that can be placed on abortion availability under the "viability" standard. *Roe* identified viability as the point at which the state has a compelling interest in protecting a fetus and therefore may more heavily regulate or even fully proscribe abortion. However, the courts thereafter had to deal with exactly how to identify viability, and these cases represent that line of doctrine. The viability cases, in our model, emerge as a line of doctrine descending from *Roe* separately from the parental consent cases. This example illustrates that the legal, doctrinal connections among cases are being recovered by our model of citation patterns. We find comparable results across the various substantive areas we examine.

Clearly, the genealogy model is identifying substantive links among cases. What is more, it is identifying how cases sequentially build from one another. It is important to underscore here that unlike other approaches (e.g., Segal 1984; Kritzer and Richards 2002; Kastellec 2010; Bailey and Maltzman 2011), our model of legal evolution does not require the coding of case facts but rather simply of the legal authorities referenced in an opinion. This represents a considerable advantage over subjective manual coding of an opinion's substance and relevant factual content. Simply using

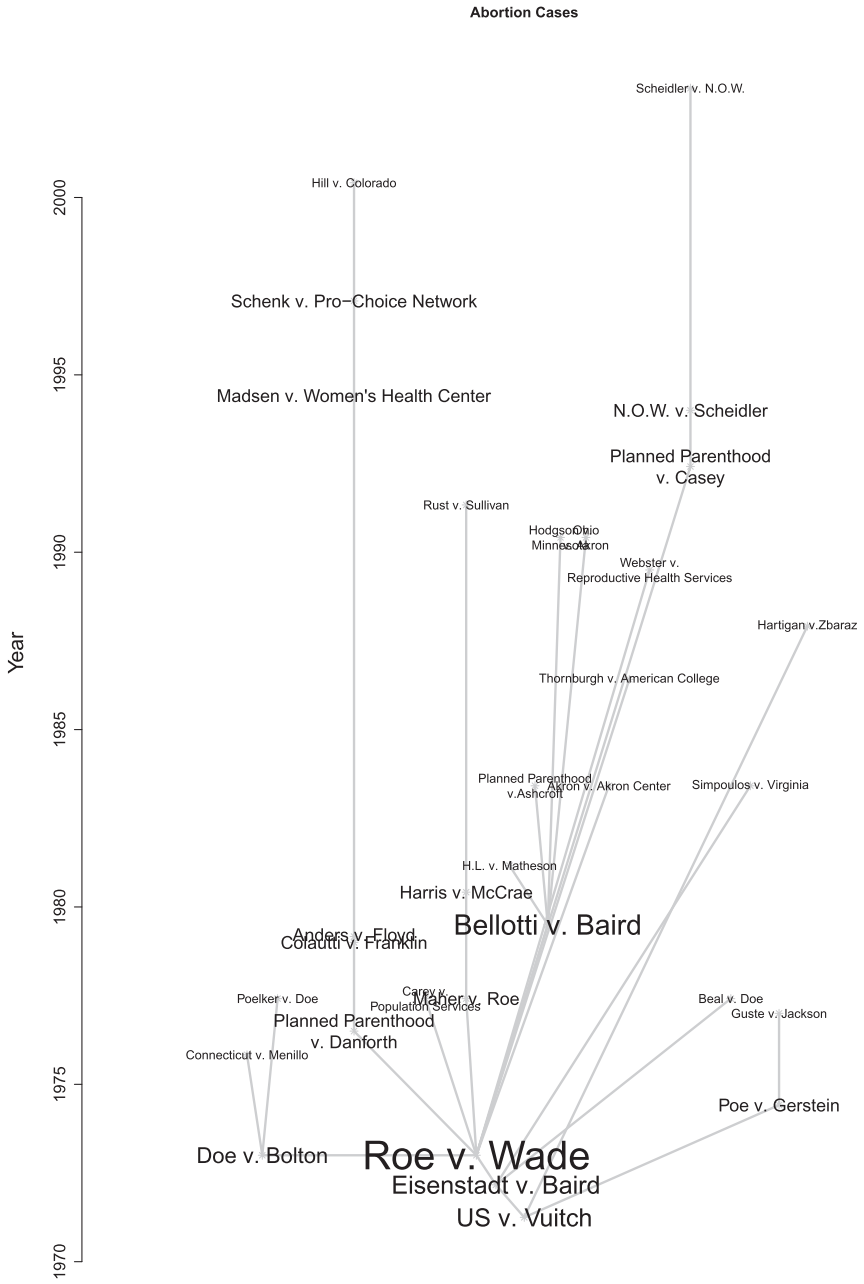

Fig. 4 Posterior genealogies of abortion cases. This figure shows estimated links among cases coded as abortion cases in the Supreme Court database. Estimates based on a 2500-iteration MCMC simulation; the horizontal dimension has no meaning and is used only to make visualization possible; cases have been assigned horizontal space in proportion to the number of child cases they have. Case names are scaled proportional to the number of child cases.

the sources of legal authority that are employed in an opinion reveals a striking and convincing portrait of the nuanced substantive and legal relationships among opinions.

4 Analysis: A Measure of Legal Significance

The genealogy of law model allows us to quantify a number of theoretically and substantively compelling quantities of interest. Perhaps chief among these are features of Supreme Court cases relating to their legal significance or impact. Previous attempts to quantify the importance of

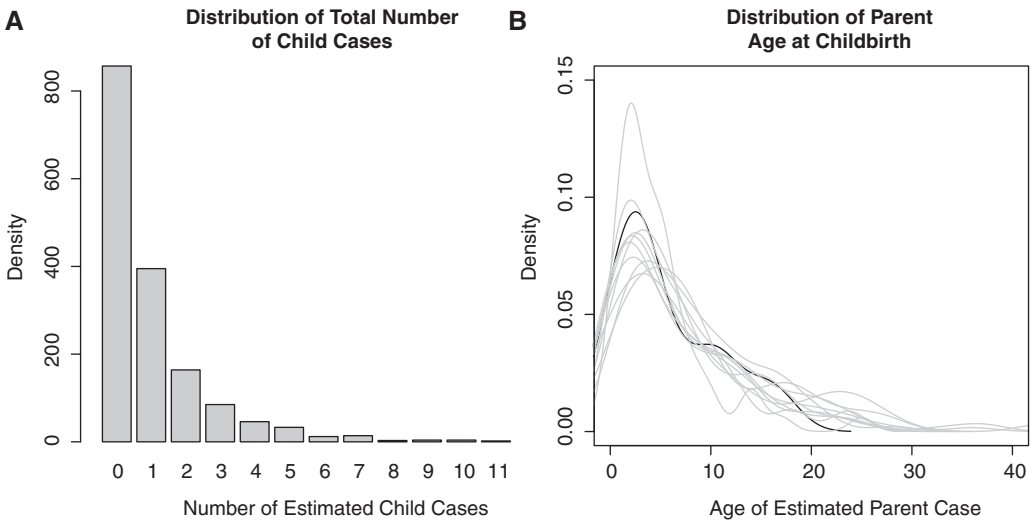

Fig. 5 Estimated fertility of Supreme Court precedents. (**A**) shows distribution of total number of child cases per precedent (all cases). (**B**) shows distribution of age of precedent at childbirth; abortion cases shown in black; all other areas in gray.

Supreme Court precedents have relied on troublesome approaches. For example, one common approach is to rely on post hoc subjective evaluations of which cases constitute "landmark" precedents (e.g., Biskupic and Witt 1997; Hall 1999). Another approach has been to rely on proxies, such as coverage by the *New York Times* (e.g., Staudt, Friedman, and Epstein 2007). Perhaps the closest approach to the one we advocate is the use of network analysis tools to identify the "centrality" of each case to the body of Supreme Court precedent (Fowler et al. 2007).

With our model, we can measure legal significance by the position a case holds in the emerging genealogy of law. There are several definitions of "influence" that one might choose to adopt that are compatible with our model (or its multi-parent extensions). For example, is a more consequential case one that spurs many new lines of inquiry (i.e., has many children)? Perhaps one wants to define consequence not in terms of direct children, but in terms of long, sustained lines of doctrine. Alternatively, one could imagine focusing instead on cases that tie together multiple lines of argumentation (i.e., those that have many parents). For the subsequent analysis, we focus on a measure of significance that is based on the number of "child cases" a precedent has.

4.1 *Fertility in Supreme Court Cases*

To measure legal significance, we calculate the total number of children each case has.[11] Figure 5 shows the distribution of the number of child cases a precedent has as well as the distribution of the age of precedents at "childbirth." In Fig. 5A, we aggregate all cases together and show the total number of child cases per precedent. As the figure makes clear, most cases have no direct descendants. Among precedents with child cases, the modal number of child cases is 1, and the frequency continues to decline steadily (though not strictly monotonically) over the range of estimated child cases per precedent.

Figure 5B shows the distribution of parental age at "childbirth." The black line shows the distribution of parent ages among abortion cases; the gray lines show the other nineteen groups of cases. The mean age of a parent case at childbirth is about seven years, with a standard deviation of about six years; the maximum age of an estimated parent case is forty-five years. These distributions reveal that any given case is most directly influenced by cases about seven years old,

[11]Because our model assigns a probability to each potential parent, we choose the precedent with the highest probability of being the parent as the parent for each case.

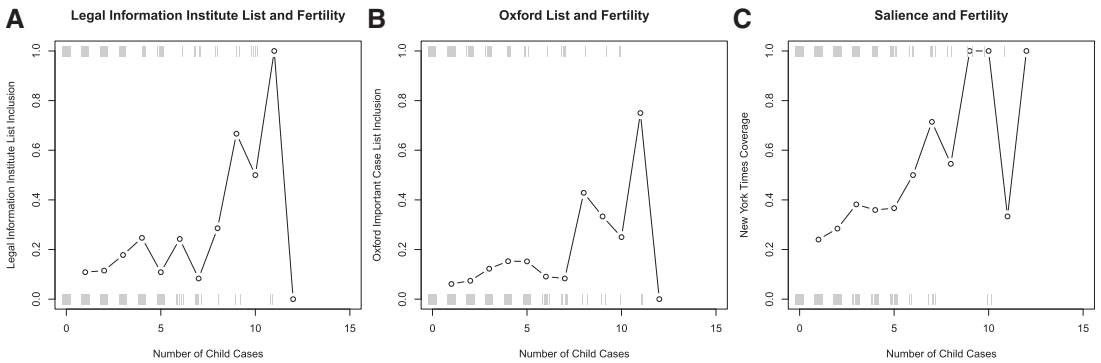

Fig. 6 Comparison of case fertility with expert measures of legal significance. (**A**) shows comparison of number of child cases with inclusion on the Oxford list of important cases; (**B**) shows comparison of number of child cases with inclusion on the Legal Information Institute list of important cases; (**C**) shows comparison of number of child cases with front-page coverage by the *New York Times*. Points are averages for each integer value of the number of child cases; data exclude outliers with more than twenty-five child cases.

whereas a few cases remain directly influential beyond ten to fifteen years, and only a small proportion of cases have much *direct* influence on the law more than two decades into their life.

4.2 *Estimates of Legal Significance*

Our estimates of legal fertility measure an aspect of legal significance that is related to, but not identical to, those captured by existing measures. As noted above, the significance of Supreme Court cases is usually measured with expert lists of "important" cases (Fowler and Jeon 2008) or media attention to cases (Staudt, Friedman, and Epstein 2007). We compare our estimates of case fertility with those expert lists in Fig. 6. More recently, scholars have leveraged *hub* and *authority* scores from network analyses of Supreme Court citations as proxies for legal importance (Fowler and Jeon 2008). We compare our estimates of legal significance to the network measures in Fig. 7.

Turning first to the comparison of our measures with the expert lists (Fig. 6), we find in each instance a positive relationship—the more children a case has, the more likely it is to be included on any of these indicators of case "importance." In Fig. 6A and B, we find that cases at the low end of the child case distribution—i.e., cases with fewer than five children—are highly unlikely to appear on either the Legal Information Institute or Oxford Guide lists of important cases. However, as one increases the number of child cases, we find that a case is more likely to be included on those lists. (Of course, as we approach the upper levels of estimated child cases, where there are fewer cases, the averages become more variable.) That is, the cases on those lists appear to be those that create or outline new areas of law that will be subsequently filled in or further refined by many subsequent cases. More derivative cases—those that do the filling in of questions left unanswered in major decisions—are, by comparison, unlikely to be included on one of the lists of important cases. We find a similar pattern when we look at the set of cases whose decisions are covered by the *New York Times*. As a measure of a case's legal significance, our measure of legal fertility correlates with subjective expert judgments about a case's importance, but not strongly.

A similar story emerges when we consider the measures based on more standard network models: The measures we derive are correlated with the most substantively similar existing measures, but far from perfectly so. Figure 7 compares our estimates of the number of child cases an opinion has with Fowler and Jeon's measures of case importance. In particular, the Fowler and Jeon (2008) study imports two important concepts from network analysis to the study of Supreme Court opinions: *hub* and *authority* scores. As Fowler and Jeon (2008, 20) note: "A *hub* is a case that cites many other decisions, helping to define which legally relevant decisions are pertinent to a given precedent, while an *authority* is a case that is widely cited by other decisions." For ease of comparison, we normalize each set of estimates by centering on zero and

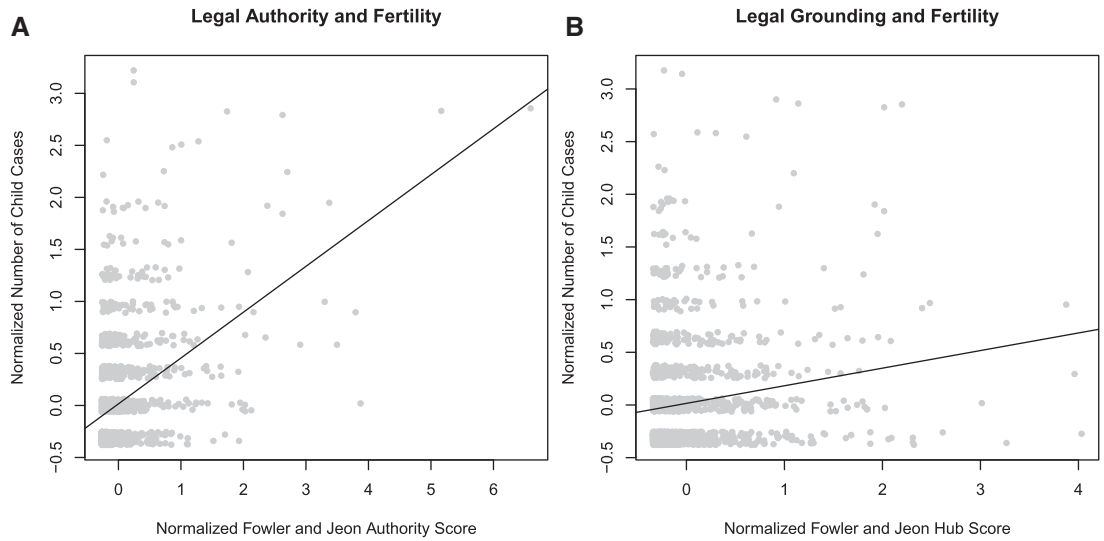

Fig. 7 Comparison of case fertility with authority and hub measures. (**A**) panel shows the correlation between the number of child cases and the Fowler and Jeon authority score. (**B**) panel shows the correlation between the number of child cases and the Fowler and Jeon hub score. Solid lines are linear-best-fit lines. All estimates derived from single-parent model assuming the parent is the case with the highest posterior probability of a direct connection.

dividing by two standard deviations. Unsurprisingly, our estimates are much more strongly associated with the Fowler and Jeon authority scores than the hub scores. Opinions with many children are necessarily highly cited opinions, but such opinions need not cite previous cases widely in the way that would lead to a high hub score. The key distinction between the measures is that our measure of legal fertility is sensitive to the number of subsequent cases that *primarily* cite a given precedent, whereas the authority scores are sensitive to the number of subsequence cases that cite a given precedent at all. Depending on the application being studied, this different emphasis may make one measure or the other more appropriate.

5 Application: Ideological Homogeneity and Legal Significance

In this section, we provide an application of our estimates to a question that has received recent scholarly attention. Staudt, Friedman, and Epstein (2007) examine the relationship between the ideological homogeneity of a decision's majority coalition and the significance of the Court's decision. We do not reconsider the underlying theoretical argument or empirical specification they present; we simply consider the empirical relationship they assert and the empirical operationalization they employ. Staudt et al. argue that more ideologically homogeneous decision coalitions should be more likely to make "consequential" law. Staudt, Friedman, and Epstein (2007, 364) note that they seek "to capture one central insight: the Justices' decisions will have a greater impact on the law when those deciding the cases are of a like mind." Thus, the critical concept in need of empirical operationalization is the impact each case has had *on the law*.

To measure legal significance, Staudt et al. rely on the measure of "salience" developed in Epstein and Segal (2000) as a proxy for legal significance. In particular, Epstein and Segal code whether each decision is reported the day after the decision on the front page of the *New York Times*. There are good reasons to suspect this measure is a tenuous proxy for their concept of interest. In particular, the measure has no direct, necessary connection to the underlying quantity—the impact a case has *on the law*. In fact, the measure was initially introduced as an index of the *political* salience of a Court decision. Moreover, if we suspect that the *New York Times* has an incentive to cover ideologically fractious decisions, then this variable will have measurement error with respect to the legal significance that is correlated with the primary variable of interest: the

homogeneity of the majority coalition. Our model provides an alternative measure to more directly capture the central concept of their theoretical argument: the *legal* rather than *political* consequence of a given decision. Using our measure of a case's subsequent legal influence, we can evaluate the robustness of their analysis. We replicate the analysis from Staudt et al. below using as the dependent variable (1) the total number of child cases each case has and (2) the number of cases normalized by the number of estimated child cases per case within the relevant subset of cases.

The key independent variable in the Staudt et al. analysis is the ideological homogeneity of the majority coalition. To measure this concept, the authors rely on the Martin and Quinn (2002) measures of judicial ideology. They use the standard deviation of the estimated ideal points of the majority coalition as a measure of the ideological heterogeneity of the majority. When an ideologically cohesive block of justices vote together in the majority, this measure decreases; when a majority coalition is made up of a wide spread of justices, this measure increases.

In addition to the key explanatory variable, Staudt et al. include in their analyses a host of control variables—the relative conservatism of the majority coalition, the size of the majority coalition, the total number of cases decided in the particular term, whether the case involves a formal alteration of Supreme Court precedent, and whether the case declares a federal law unconstitutional. Staudt et al. estimate the relationship between their key independent variable and these controls and the *New York Times* coverage using a probit specification, and their results are replicated in Table 1, which also shows the results of two linear regressions using our estimates of legal significance as the dependent variable.

There are points of both agreement and disagreement between our two analysis that bear discussion. Consider first the key independent variable in the Staudt et al. analysis. They find a statistically significant relationship between ideological diversity and legal significance, whereas we find no such statistically significant relationship. More important than statistical significance, the magnitude of the estimated relationship is substantively small in our analysis. A two-standard-deviation change in the value of ideological diversity is associated with a change of about 0.04 standard deviations in the point prediction of the number of child cases. A change from the minimum observed value to the maximum observed value is associated with only about a 0.1-standard-deviation change in the number of child cases. Not only is this relationship not statistically significant, but it is substantively negligible. To the extent our measure is more directly connected to the substantive quantity of interest—the legal influence of a court decision—this finding casts doubt on the conclusions that Staudt et al. draw.

Consider next the relationship between majority size and the significance of the decision. Although Staudt et al. find that the larger majorities are *less* likely to make significant law, as measured by the *New York Times* coverage, we essentially find no relationship. We believe this discrepancy is likely due to the fact that the operationalization of legal significance that Staudt et al. adopt actually captures underlying political salience and newsworthiness, which are lower when the decision itself is less contentious, regardless of its legal consequence.

Like Staudt et al., we find that the cases that overturn a precedent have higher legal impact. However, as just noted, our measure seems to be capturing a qualitatively different concept than Staudt et al. measure. Their measure is likely picking up the fact that overturning a precedent is newsworthy; our measure is likely to be picking up the fact that overturning a precedent often raises a range of subsidiary issues that the Court revisits in subsequent cases.

In brief, using our measure of legal significance rather than the Staudt et al. measure, we cannot replicate the key finding that greater ideological diversity is associated with lesser legal significance. We believe our measure is a better operationalization of the underlying quantity of *legal* significance. The discrepancies between our analysis and their analysis can be easily understood as deriving from the fact that their measure is contaminated by political salience and newsworthiness.

6 Conclusion

The genealogy model of Supreme Court doctrine that we have introduced offers a powerful new tool for the systematic study of features of judge-made law that interest a broad array of scholars. By conceptualizing the law as a genealogical structure, we have directly modeled the central feature

Table 1 Replication of Staudt, Freidman, and Epstein (2007) analysis

	Staudt et al.	*Case fertility*	*Case fertility (normalized)*
Ideological diversity of the majority	−0.43	−0.18	−0.06
	(0.12)	(0.13)	(0.04)
Number of justices in the majority	−0.16	0.02	<0.01
	(0.04)	(0.04)	(0.01)
Ideology of the majority (conservatism)	−0.29	−0.32	−0.10
	(0.06)	(0.07)	(0.02)
Number of cases	−0.01	<0.01	<0.01
	(0.002)	(<0.01)	(<0.01)
Overturn precedent	1.11	0.96	0.29
	(0.21)	(0.25)	(0.07)
Overturn federal law	1.55	0.36	0.10
	(0.26)	(0.35)	(0.11)
Constant	1.26	1.24	0.07
	(0.32)	(0.38)	(0.12)
N	2573	950	950

Notes. The first column of results reports the results from the original Staudt et al. analysis (probit coefficients and standard errors). The second and third columns of results replicate their analysis using our measure of legal significance as the dependent variable rather than their original variable (OLS coefficients and standard errors). All independent variables are as coded in Staudt, Friedman, and Epstein (2007); to ease comparison, we employ the same variable names in our table as do Staudt et al.

of judge-made law—the idea that cases build sequentially and systematically from each other to develop a line of doctrine. Our model thus formalizes one qualitative aspect of law at the heart of traditional approaches to studying law. In this way, we hope to further bridge the gulf between the legal academy's approach to studying judge-made law and the political science of law and courts. In the words of Friedman (2006), our model "takes law seriously."

Our approach not only provides a systematic representation of law, it allows us and future scholars to derive axiomatically defined measurements of any number of features of judge-made law. As an example, we have highlighted the ability of our model to yield systematic measures of legal significance. Numerous other possibilities present themselves, such as the ability to construct theoretically oriented measures of a case's legal impact or the extent to which any given case alters the nature and content of a line of doctrine. Critically, those measures require ex ante definition by the researcher of the theoretical concepts of interest and do not rely on ex post or proxy measures of those concepts. Our replication of the Staudt, Friedman, and Epstein (2007) analysis of the relationship between ideological homogeneity and legal significance reveals how our theoretically derived measures can yield substantively different conclusions than one finds using proxies that are poorly connected to the underlying quantities of interest.

What is more, our model provides a never-before-developed ability to quantify a number of descriptive features of the law in which scholars are interested. For example, we have shown that our method can be used to assess the effect of overruling a precedent—does that end its "influence" or merely change the nature of the line of doctrine? Our model also yields insight into the "fertility" of Supreme Court doctrine. For how long does a case's direct influence last? That is, for how long does a case continue giving rise to new lines of argumentation? How many direct descendants do cases typically have?

The underlying approach we use to estimate legal genealogies is applicable to other texts that use citations, most obviously scholarly research. For example, in this article, the most frequently cited papers are a set of papers that use citation data to explore the relationships between Supreme Court cases (Fowler et al. 2007; Fowler and Jeon 2008; Clark and Lauderdale 2010). In a single-parent genealogical model, this article would naturally attach to one of these papers, which have multiple mutual citation links among themselves. This article is indeed part of an emerging methodological

literature on Supreme Court citation data.[12] The genealogical model using citation count data yields an accurate portrayal of the scholarly lineage of this article. More generally, estimates from a genealogical model might be genuinely useful to scholars who are interested in quickly distinguishing the small set of papers that centrally engage with a particular line of inquiry from the sometimes large set of papers that merely offer incidental citations.

There are, to be sure, limitations of both the model we have developed and the analysis we have presented here. One limitation is that the model is specifically tailored to model doctrine in a particular kind of legal setting—a common law system in which case law is at the center of the legal method. Thus, our model may be appropriately applied to places like the USA, Canada, or the United Kingdom.[13] However, the model would not be an appropriate representation of doctrine in a civil law system in places such as France, Germany, or Spain. Perhaps most importantly, our model does not incorporate information about the nature of the doctrinal content. Doctrine consists of more than just the sequence of cases that together construct the line of argumentation—the content of the cases is often just as important as the sequence. Our model captures how the cases logically build doctrine in sequence but do not capture the political or ideological direction in which the sequence of cases takes the law. We expect that future research can and should bring together those two components of doctrine to present a fuller quantitative representation of the law.[14] However, for now, it bears focusing on the structure of doctrine and exploiting the myriad features of judge-made law previously not subject to systematic, quantified examination. Finally, future research might seek to incorporate lower court opinions into the genealogy of law. Some current research examines how lower courts shape or influence Supreme Court doctrine and how the entire body of case law fits together (Carrubba and Clark, forthcoming), and we view this as a particularly promising application of our model and its extensions.

References

Bailey, Michael A., and Forrest Maltzman. 2011. *The constrained court: Law, politics, and the decisions justices make*. Princeton, NJ: Princeton University Press.

Bartels, Brandon L. 2009. The constraining capacity of legal doctrine on the U.S. Supreme Court. *American Political Science Review* 103(3):474–95.

Benesh, Sara C. 2002. *The U.S. Court of Appeals and the law of confessions: Perspectives on the hierarchy of justice*. New York: LFB Scholarly Publishing.

Biskupic, Joan, and Elder Witt. 1997. *Congressional Quarterly's guide to the U.S. Supreme Court*, 3rd ed. Washington, DC: Congressional Quarterly Press.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research* 2003(3):993–1022.

Bommarito, Michael J., Daniel Martin Katz, Jon Zelner, and James H. Fowler. 2010. Distance measures for dynamic citation models. *Physica* 389(19):4201–8.

Bueno de Mesquita, Ethan, and Matthew Stephenson. 2002. Informative precedent and intrajudicial communication. *American Political Science Review* 96(4):1–12.

Carrubba, Clifford J., and Tom S. Clark. Forthcoming. Rule Creation in a Political Hierarchy. *American Political Science Review*.

Clark, Tom S., and Benjamin Lauderdale. 2010. Locating Supreme Court opinions in doctrine space. *American Journal of Political Science* 54(4):871–90.

Clark, Tom S., and Benjamin Lauderdale. 2012. *Replication data for: The Genealogy of Law*. http://hdl.handle.net/1902.1/18176.

Clark, Tom S., and Clifford J. Carrubba. 2012. A Theory of Opinion Writing in the Judicial Hierarchy. *Journal of Politics* 74(2):584–603.

[12] A multiple-parent genealogical model might add a second link to Staudt, Friedman, and Epstein (2007), reflecting the fact that this article also contributes to a line of substantive research about the legal significance of Supreme Court opinions that was previously distant from the methodological literature about Supreme Court citations.

[13] Even within the USA and Canada, there are civil law jurisdictions where our model would be less appropriate. In particular, the State of Louisiana and the Province of Québec are both civil law jurisdictions whose courts operate under a hybrid of the national common law system and local civil law system.

[14] Clark and Lauderdale (2010) show that policy-relevant information can be gleaned from the nature of citations between opinions. A natural extension to our model of the genealogy of law is to modify the model to account for a policy/ideological dimension. The *x*-axis in Fig. 4, for example, could be modeled as a policy dimension (it currently has no substantive meaning at all).

Epstein, Lee, and Jeffrey A. Segal. 2000. Measuring issue salience. *American Journal of Political Science* 44(1):66–83.
Fowler, James H., and Sangick Jeon. 2008. The authority of Supreme Court precedent: A network analysis. *Social Networks* 30(1):16–30.
Fowler, James H., Timothy R. Johnston, James F. Spriggs II, Sangick Jeon, and Paul J. Wahlbeck. 2007. Network analysis and the law: Measuring the legal importance of precedents at the U.S. Supreme Court. *Political Analysis* 15(3):324–46.
Friedman, Barry. 2006. Taking law seriously. *Perspectives on Politics* 4(2):261–76.
Gennaioli, Nicola, and Andrei Shleifer. 2007. The evolution of common law. *Journal of Political Economy* 115(1):43–68.
George, Tracey E., and Lee Epstein. 1992. On the nature of Supreme Court decision making. *American Political Science Review* 86:323–37.
Hall, Kermit L. 1999. *The Oxford guide to United States Supreme Court decisions*. New York: Oxford University Press.
Hansford, Thomas G., and James F. Spriggs II. 2006. *The politics of precedent on the U.S. Supreme Court*. Princeton, NJ: Princeton University Press.
Ignagni, Joseph A. 1994. Explaining and predicting Supreme Court decision making: The Burger Court's establishment clause decisions. *Journal of Church and State* 36(2):301–21.
Kastellec, Jonathan P. 2010. The statistical analysis of judicial decisions and legal rules with classification trees. *Journal of Empirical Legal Studies* 7(2):202–30.
Kornhauser, Lewis A. 1992a. Modeling collegial courts II: Legal doctrine. *Journal of Law, Economics & Organization* 8:441–70.
———. 1992b. Modeling collegial courts I: Path dependence. *International Review of Law and Economics* 12:169–85.
Kort, Fred. 1957. Predicting Supreme Court decisions mathematically: A quantitative analysis of the "right to counsel" cases. *American Political Science Review* 51(1):1–12.
Kritzer, Herbert M., and Mark J. Richards. 2002. *Deciding the Supreme Court's administrative law cases: Does chevron matter?* Paper prepared for the Annual Meeting of the American Political Science Association.
Lax, Jeffrey R., and Kelly T. Rader. 2009. Legal constraints on Supreme Court decision making: Do jurisprudential regimes exist? *Journal of Politics* 72(2):273–84.
Lax, Jeffrey R. 2007. Constructing legal rules on appellate courts. *American Political Science Review* 101(3):591–604.
———. 2011. The new judicial politics of legal doctrine. *Annual Review of Political Science* 14:131–157.
Levi, Edward. 1949. *An introduction to legal reasoning*. Chicago: University of Chicago Press.
Maltzman, Forrest, James F. Spriggs II, and Paul J. Wahlbeck. 2000. *Crafting law on the Supreme Court: The collegial game*. New York: Cambridge University Press.
Maltz, Earl M. 2000. The function of Supreme Court opinions. *Houston Law Review* 37:1395–420.
Martin, Andrew D., and Kevin M. Quinn. 2002. Dynamic ideal point estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999. *Political Analysis* 10(2):134–53.
McGuire, Kevin T., and Georg Vanberg. 2005. *Mapping the policies of the U.S. Supreme Court: Data, opinions, and constitutional law*. September 1–5.
McGuire, Kevin T. 1990. Obscenity, libertarian values, and decision making in the Supreme Court. *American Politics Quarterly* 18(1):47–67.
Patterson, Edwin W. 1951. The case method in American legal education: Its origins and objectives. *Journal of Legal Education* 4(1):1–24.
Porter, Mason A., Jukka-Pekka Onnela, and Peter J. Mucha. 2009. Communities in networks. *Notices of the American Mathematical Society* 56(9):1082–97.
Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev. 2010. How to analyze political attention with minimal assumptions and costs. *American Journal of Political Science* 54(1):209–28.
Richards, Mark J., and Herbert M. Kritzer. 2002. Jurisprudential regimes in Supreme Court decision making. *American Political Science Review* 96(2):305–20.
Segal, Jeffrey A. 1984. Predicting Supreme Court cases probabilistically: The search and seizure cases, 1962–1981. *American Political Science Review* 78(4):891–900.
Songer, Donald R., and Susan B. Haire. 1992. Integrating alternative approaches to the study of judicial voting: Obscentity cases in the U.S. Courts of Appeals. *American Journal of Political Science* 36:963–82.
Spaeth, Harold J., Lee Epstein, Theodore W. Ruger, Keith E. Whittington, Jeffrey A. Segal, and Andrew D. Martin. 2010. *The Supreme Court database*. http://supremecourtdatabase.org.
Spriggs, James F. II, and Thomas Hansford. 2000. Measuring legal change: The reliability and validity of Shepard's citations. *Political Research Quarterly* 53(2):327–41.
Staudt, Nancy, Barry Friedman, and Lee Epstein. 2007. On the role of ideological homogeneity in generating consequential constitutional decisions. *North Carolina Law Review* 86(5):1299–332.
Sulam, Ian. 2011. Policy, precedent, indeterminacy: Using doctrine space to bridge across circuits. Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL.