# Inferential Analysis of the Supreme Court Citation Network

Christian S. Schmid* Ted Hsuan Yun Chen † Bruce A. Desmarais ‡
David R. Hunter §

April 9, 2018

**Abstract**

The significance and influence of US Supreme Court majority opinions derive in large part to opinions' roles as precedents for future opinions. A growing body of literature seeks to understand what drives the use of opinions as precedents through the study of Supreme Court case citation patterns. We raise two limitations of existing work on Supreme Court citations. First, dyadic citations are typically aggregated to the case level before they are analyzed. Second, citations are treated as if they arise independent of one another. We make the case that studying Supreme Court citations as a dynamic citation network overcomes both of these limitations. We present a methodology for studying citations between US Supreme Court opinions at the dyadic level, as a network. This methodology—the citation temporal exponential random graph model—enables researchers to account for the effects of case characteristics and complex forms of network dependence in citation formation. We apply this methodology to a network that includes all Supreme Court cases decided between 1937 and 2015. We find evidence for dependence processes, including reciprocity, transitivity, popularity, and activity. The dependence effects that we identify are as substantively and statistically significant as the effects of the exogenous covariates we include in the model. The summary result from this analysis is that theoretical models of Supreme Court citation formation should consider both the effects of case characteristics and the structure of past citations.[1]

*Department of Statistics, The Pennsylvania State University, schmid@psu.edu
†Department of Political Science, The Pennsylvania State University, thc126@psu.edu
‡Department of Political Science, The Pennsylvania State University, bdesmarais@psu.edu
§Department of Statistics, The Pennsylvania State University, dhunter@stata.psu.edu
[1] Prepared for presentation at the "The Enemy of My Enemy: Judicial Politics meets Network Analysis" conference within a conference at the 2018 Annual Conference of the Southern Political Science Association. This work was supported in part by NSF grants SES- 1558661, SES-1637089, SES-1619644, and CISE-1320219.

# 1 Introduction

Majority opinions written by the United States Supreme Court exercise their authority and influence, in part, through their roles as precedents in future Supreme Court decisions and opinions. The findings regarding the extent and exact nature of the influences of precedent have been mixed, but the balance of the literature finds that past decisions exert some form of influence on the justices' decision making (Knight and Epstein, 1996; Gillman, 2001; Richards and Kritzer, 2002; Hansford and Spriggs, 2006; Bailey and Maltzman, 2008, 2011). Despite a considerable body of research that focuses on the way in which relevant precedents shape decision making on the Court, relatively little work has focused on understanding which past opinions are cited by an opinion. Our focus in this paper is to provide what is, to our knowledge, the first comprehensive analysis of exactly which cases are cited by a case. We follow an emerging body of work on legal citations, and treat the system of citations as a network (e.g., Harris, 1982; Caldeira, 1988; Fowler et al., 2007; Fowler and Jeon, 2008; Bommarito II, Katz and Zelner, 2009; Lupu and Voeten, 2012; Pelc, 2014).

We are not the first to ask what predicts the citations in US Supreme Court Opinions. Indeed, a voluminous body of work has sought to explain how many times an opinion is cited (e.g., Cross, 2010; Benjamin and Desmarais, 2012), when in a lifecycle an opinion is cited (e.g., Black and Spriggs, 2013; Spriggs and Hansford, 2001), and how many cases are cited by an opinion (e.g., Lupu and Fowler, 2013)—all focused on the US Supreme Court. One common feature of the research design in all of these studies is that the observations are at the case or case-time level. The outcome variables in these analyses are defined as measures of the number of citations to a case over a period of time, the number of citations to a case at a particular time, or a measurement on the cases cited by a case. None of these studies treats citation in its micro-level form—as a relationship between two opinions, the citing opinion and the cited opinion. We are aware of one prior study, Clark and Lauderdale (2010), in which a statistical model is used to analyze dyadic citations between cases. However, Clark and Lauderdale (2010) use a dyadic latent variable model in order to estimate ideal points for Supreme Court Opinions, but does not use any explanatory variables to predict the formation of citation ties between opinions. We build upon this literature both methodologically and substantively. Methodologically, we develop a novel statistical framework for modeling directed dyadic citations. Second, we apply this methodology to a half-century of directed dyadic citations between US Supreme Court opinions.

There exist two broad reasons why empirical analyses of citations are best defined on the directed dyad level, not the case level. The first is that directed dyadic analyses can test both dyadic and case-level hypotheses. For example, case-level analyses can model whether opinions supported by a liberal majority coalition are more likely than those supported by a conservative majority coalition to be cited heavily in the future, but they cannot precisely model whether liberal cases will be cited more by liberal cases than by conservative cases. Thus, the first reason for analyzing citations at the dyadic level is to expand the set of hypotheses that can be represented in the model. The second reason for studying citations at the directed dyadic level is that, as articulated in the growing literature on legal citation networks, citations form complex networks in which a citation at one point in time may influence future citations. This phenomenon of complex dependence is very common in networks of many types, but processes specific to Supreme Court citations create interdependence in citations. For example, if opinion $i$ relies heavily on opinion $j$ as precedent, opinion $i$ is likely to discuss the legal basis for opinion $j$, and as a consequence cite some of the opinions cited by opinion $j$. Suppose opinion $k$ is cited by opinion $j$. Opinion $k$ is more likely to be cited by opinion $i$ because opinion $i$ relies heavily on $j$, and opinion $j$ cites $k$. This is a special case of a very common process on networks referred to as “triad closure”. Complex dependence is theoretically interesting on its own merits, but the effects of covariates cannot be reliably identified—either in terms of coefficient values or standard errors—without accounting for the interdependence inherent in networks (Desmarais and Cranmer, 2017).

In what follows we develop a theoretical case that citations on the US Supreme Court are characterized by forms of complex dependence that are common in networks. We then develop an extension of a model—the exponential random graph model—that can incorporate both exogenous covariates and complex forms of interdependence into a directed dyadic analysis of citations. Finally, we develop and estimate a specification of this model in an analysis of US Supreme Court citations between 1937 and 2001. We find robust support for the inherent complexity underpinning the formation of citation ties, and show that incorporating complex dependence into the model of citation formation significantly improves the model’s predictive performance.

*Comment from Benjamin Kassov: I think the discussion in the introduction of the paper focusing on a "case-based" versus network based analysis is a bit short. I was left slightly confused by what a "case-based" analysis of citation actually is or what it means. I can see several different conceptions of this, some being more problematic than others in a methodological sense, although none of them are ideal for analyzing networks and all likely suffer some amount of OV bias.*

## 2 Complex Interdependence in Supreme Court Citations

When it comes to the development and testing of theory, the defining feature of networks is that the fundamental element under study—the relationship between two units (i.e., the citation from one opinion to another) is a piece of a complex web of relations. The formation (or lack thereof) of that relationship cannot be fully understood without considering how the relationship fits into the complex web. Analytical designs that account only for covariates in explaining tie formation are incomplete theoretically, and, as a consequence, are subject to a form of omitted variable bias (Cranmer and Desmarais, 2016). Citations in legal opinions are unique in terms of the windows into network dependencies offered by the texts of the opinions. A number of common structural dependencies that are found in networks are likely to apply to citations in Supreme Court opinions. In this section we present these dependence forms, and document the mechanisms by which they arise through archetypal passages in example opinions.

**Transitivity:** In a network of directed relations (e.g., A cites B, but B doesn't cite A) transitivity refers to the tendency for A to send a tie to C if A sends a tie to B and B sends a tie to C (Holland and Leinhardt, 1971). In undirected networks, transitivity is simply the process by which friends of friends become friends (i.e., a friend of a friend is a friend). The term, "transitive closure" refers to a tie forming from A to C in response to extant ties from A to B and B to C. When writing opinions, Supreme Court justices present the legal bases for their rulings, which often involves discussing the most primary/relevant precedents underpinning these legal bases, but also the precedents and legal rules on which the primary precedents were based. This process of presenting several layers/levels of precedent in an opinion follows the structure of transitive closure exactly—opinion A cites opinion B as a primary precedent, and then cites opinion C because opinion B cites opinion C. The two examples presented below illustrate this process.

In the first example, a passage from Kansas v. Marsh (548 U.S. 163, 2006)—a case considering the constitutionality of a death sentence statute in Kansas. In this example, the case Stringer v. Black is cited by Kansas v. Marsh as a case that is quoted by Sochor v. Florida. The primary precedent under discussion in this passage of the opinion is Sochor v. Florida, but Stringer v. Black is cited as a result of its role in the Sochor v. Florida opinion.

> The statute thus addresses the risk of a morally unjustifiable death sentence, not by minimizing it as precedent unmistakably requires, but by guaranteeing that in equipoise cases the risk will be realized, by "placing a 'thumb [on] death's side of the scale,' " Sochor v. Florida, 504 U. S. 527, 532 (1992) (quoting Stringer v. Black, 503 U. S. 222, 232 (1992); alteration in original).

The second example is a passage from Seminole Tribe of Fla. v. Florida, (517 U.S. 44 1996)—a case addressing the rights of groups and citizens to sue states in federal court. In this example, Pennsylvania v. Union Gas Co (491 U.S. 1, 1989) is the primary precedent under discussion, and several cases are cited and discussed in terms of their roles as precedents in the Union Gas opinion.

> Never before the decision in Union Gas had we suggested that the bounds of Article

> III could be expanded by Congress operating pursuant to any constitutional provision other than the Fourteenth Amendment. Indeed, it had seemed fundamental that Congress could not expand the jurisdiction of the federal courts beyond the bounds of Article III. Marbury v. Madison, 1 Cranch 137 (1803). The plurality's citation of prior decisions for support was based upon what we believe to be a misreading of precedent. See Union Gas, 491 U. S., at 40-41 (SCALIA, J., dissenting). The plurality claimed support for its decision from a case holding the unremarkable, and completely unrelated, proposition that the States may waive their sovereign immunity, see id., at 14-15 (citing Parden v. Terminal Railway of Ala. Docks Dept., 377 U. S. 184 (1964)), and cited as precedent propositions that had been merely assumed for the sake of argument in earlier cases, see 491 U. S., at 15 (citing Welch v. Texas Dept. of Highways and Public Transp., 483 U. S., at 475-476, and n. 5, and County of Oneida v. Oneida Indian Nation of N. Y., 470 U. S., at 252).'

**Reciprocity:** Reciprocity (also referred to as mutuality) is the tendency for node B to send a tie to node A in response to or coordination with A sending a tie to B (Garlaschelli and Loffredo, 2004). It is typically not possible for reciprocated ties to form in legal opinions. Most citations reference past opinions that were issued before the citing opinion's case was even argued before the Court. However, opinions written within the same Supreme Court term are often drafted in tandem, and can cite each other reciprocally. Opinion A citing opinion B within the same term represents a signal that opinion B is relevant to the legal reasoning underpinning opinion A. Unlike the citations themselves, the applicability of legal rules or lines of reasoning across cases is not directed—if A is relevant to B, B is highly likely to be relevant to A. We expect opinions written within the same term to exhibit a high degree of reciprocity. Below we provide two example passages from opinions that illustrate the phenomenon of within-term reciprocity.

The first case in our example reciprocal dyad is a passage from Western Air Lines v. Criswell (472 U.S. 400, 1985)—a case considering mandatory retirement in the context of age discrimination laws. The second case in the dyad, Johnson v. Mayor of Baltimore (472 U.S. 353, 1985) is another case considering whether mandatory retirement violates the Age Discrimination in Employment Act. These cases addressed very similar legal questions, which increased the likelihood that they would inform each other, and the opinions were written within the same term, which made it possible for them to cite each other.

> *From Western Air Lines:* On a more specific level, Western argues that flight engineers must meet the same stringent qualifications as pilots, and that it was therefore quite logical to extend to flight engineers the FAA's age 60 retirement rule for pilots. Although the FAA's rule for pilots, adopted for safety reasons, is relevant evidence in the airline's BFOQ defense, it is not to be accorded conclusive weight. Johnson

> v. Mayor and City Council of Baltimore, ante at 472 U. S. 370-371. The extent to which the rule is probative varies with the weight of the evidence supporting its safety rationale and "the congruity between the . . . occupations at issue." Ante at 472 U. S. 371. In this case, the evidence clearly established that the FAA, Western, and other airlines all recognized that the qualifications for a flight engineer were less rigorous than those required for a pilot.
>
> *From Johnson:* The city, supported by several amici, argues for affirmance nonetheless. It asserts first that the federal civil service statute is not just a federal retirement provision unrelated to the ADEA, but in fact establishes age as a BFOQ for federal firefighters based on factors that properly go into that determination under the ADEA, see Western Air Lines, Inc. v. Criswell, post p. 472 U. S. 400. Second, the city asserts, a congressional finding that age is a BFOQ for a certain occupation is dispositive of that determination with respect to nonfederal employees in that occupation.

**Popularity** Popularity, also termed "preferential attachment" is the tendency for ties to be sent to nodes to which many ties have already been sent (Barabási and Albert, 1999). Citations to an opinion signal both the Court's awareness of the legal reasoning of the case and the Court's evaluation that the opinion is an authoritative precedent. The more citations, the stronger this signal. Landmark cases, or those that establish new legal rules, are particularly authoritative and accrue citations from most future opinions that follow the respective line of reasoning. The passage below, from Oregon v. Mitchell (400 U.S. 112, 1970)—a case on the legality of state age restrictions on voting in federal elections—illustrates this popularity dynamic. In this opinion passage Baker v. Carr is cited in reference to its role as a landmark precedent, and noted for the number of other cases by which it has been followed and for which an authoritative opinion is referenced, and even discussed in terms of the number of other cases by which it was followed.

> The first case in which this Court struck down a statute under the Equal Protection Clause of the Fourteenth Amendment was Strauder v. West Virginia, 100 U. S. 303, decided in the 1879 Term. [Footnote 2/1] In the 1961 Term, we squarely held that the manner of apportionment of members of a state legislature raised a justiciable question under the Equal Protection Clause, Baker v. Carr, 369 U. S. 186. That case was followed by numerous others, e.g.: that one person could not be given twice or 10 times the voting power of another person in a state-wide election merely because he lived in a rural area..."

**Activity** Activity, or sender activation, is the tendency for ties sent to beget more ties sent (Howard et al., 2016). Structurally, this means that there are few nodes that send a moderate number of ties—the number of ties sent is either small or fairly large. In terms of Supreme Court opinions, for each opinion discussed, that discussion is likely to raise other tangential issues on which the justices will want to draw upon past opinions. Furthermore, for each opinion that applies to and is cited by the current opinion, there is often a case/opinion that needs to be discussed in terms of why it does not apply. Justices often clarify not only those legal principles that apply, but often those that do not. The example passage that illustrates activity is from Kush v. Rutledge (460 U.S. 719, 1983)—a case addressing witness intimidation and due process protections. In the Kush opinion, Griffin v. Beckenridge (i.e., Griffin) is discussed at length in terms of its lack of applicability. It is common that opinions distinguish among a variety of legal rules in terms of their applicability to a case, which means that for each rule that is established in an applicable precedent, one or more need to be discussed in terms of why they do not apply.

> No allegations of racial or class-based invidiously discriminatory animus are required to establish a cause of action under the first part of 1985(2). The statutory provisions now codified at §1985 were originally enacted as §2 of the Civil Rights Act of 1871, and the substantive meaning of the 1871 Act has not been changed. The provisions relating to institutions and processes of the Federal Government (including the first part of §1985(2)) – unlike those encompassing activity that is usually of primary state concern (including the second part of §1985(2) and the part of ¶1985(3) involved in Griffin, supra – contain no language requiring that the conspirators act with intent to deprive their victims of the equal protection of the laws. Thus, the reasoning of Griffin is not applicable here, and, given the structure of §2 of the 1871 Act, it is clear"

## 3 Network Approaches to Studying Citations

The dependence processes discussed in the previous section represent new theoretical claims regarding the factors that account for citation formation among US Supreme Court opinions. Before we can test these claims, we consider methods for analyzing citation data as a network. Researchers from several fields have used network analysis to analyze citations—legal, patent, and scientific. Before describing our approach to modeling Supreme Court citations, we review the methods researchers have previously used to study citations. A citation is a directed link between two documents that indicates the citing document attributes the cited to be relevant to the evidentiary basis of the citing document. Collective patterns of citations within a certain domain make up citation networks. When raw citation data is formally represented as a network, the nodes are usually the documents themselves, and each directed arc is the existence of one or more cites from the sending to the receiving document. Given the permanent nature of documents in citations networks, these networks are acyclic (Leicht et al., 2007; Karrer and

Newman, 2009). Relatedly, citation networks grow as new documents and citations enter the network, but established arcs persist in all but exceptional cases.

Networks researchers have developed approaches to understanding the mechanisms that drive citations between documents. Work in this area generally proceeds by specifying a set of citation behaviors, then proposing a model to capture the combination of these behavioral rules (Simkin and Roychowdhury, 2007). Assessment of the model is based on how well it fits citation distributions. Researchers working in this area tend to focus on modeling the growth of the citation network as governed by a degree-based mechanism such as preferential attachment (Barabási and Albert, 1999; Vazquez, 2001), and the age of the paper (Jeong, Néda and Barabási, 2003; Eom and Fortunato, 2011; Wang, Song and Barabási, 2013). Regarding age, the previous standard approach was to treat the probability of citation as function of degree (i.e., the number of citations previously accrued by the document) and age (Hajra and Sen, 2005, 2006), then examine the distribution of citations. More recently, work has been done in relaxing the assumption that the effect of degree is static, instead allowing it to vary with time (Wang, Yu and Yu, 2008).

Another approach to understanding the generative process of citation networks is to examine the existence of network motifs, or subnetwork structures, that can be interpreted as measuring different generative mechanisms. For example, a triangle in which case A cites both cases B and C, and case B cites case C, is a type of motif—a triadic motif. If researchers find that a triangle is a highly prevalent motif in a network, that finding would suggest that triad closure is a prevalent process by which the network is generated. The analysis of network motifs is done by comparing the prevalence of the motifs in the observed network to the prevalence of the motifs in networks drawn from a relevant null distribution. This null model must also be characterized by features of citation networks discussed earlier, meaning that it must be a directed acyclic graph with unweighted arcs (Carstens, 2016). Karrer and Newman (2009) proposes the use of null models based on fixed-degree sequences. In these models, the nodes in the networks (i.e., opinions) have the same, or close to the same, numbers of ties. Conditional on the number of ties accrued by each node, ties are randomly re-wired such that there is no systematic pattern regarding the nodes to which any node connects.

From this brief review we see two broad components of past methods for analyzing citation networks. First, node features, age in particular, are used to model the rate at which a document is cited. Second, theoretical models are used to account for prevalent graph structures. In the next section we describe a methodology that can be used to integrate node attributes and graph structures/motifs into a comprehensive model of US Supreme Court citations.

# 4 The Citation Temporal Exponential Random Graph Model

We build upon existing methods for network analysis to define a model that fits the constraints to which Supreme Court citation networks are subject. The Supreme Court citation network is nearly a directed acyclic graph—a graph in which there can be no loops/cycles (e.g., ties cannot be reciprocated, which would constitute a two-node loop). If two cases are decided in different terms, the latter case can cite the earlier case, but not the reverse. However, as noted above, if two opinions are written in the same term, they can cite each other reciprocally. We implement a version of the temporal exponential random graph model (TERGM)—a model that is commonly used for longitudinal network data (e.g., Cranmer, Desmarais and Menninga, 2012; Desmarais and Cranmer, 2013; Clark and Caro, 2013; Block et al., 2018; Graif, Lungeanu and Yetter, 2017)—that is subject to the constraints specific to the Supreme Court citation network. The TERGM is a model that can simultaneously incorporate structural dependencies such as transitivity and reciprocity and exogenous covariates (e.g. the issue area of the citing case, the ideological distance between the opinion authors of the two cases in the dyad) to explain tie formation. If the ties do not depend upon each other, and the coefficients associated with the structural dependencies are either assumed or estimated to be zero, the TERGM reduces to a logistic regression in which the observation is a directed dyad (Cranmer and Desmarais, 2010). In this section we describe the citation TERGM (c-TERGM), and explain estimation methods.

Let $C_t$ be the case-to-case adjacency matrix at time $t$ such that $C_t^{ij} \in \{0,1\}$ is a binary indicator of whether case $i$ cites $j$. Furthermore, let

$$\mathcal{C}_t = \{C_{\leq t} \in \{0,1\}^{(N_t \cdot (N_t - 1))/2} : C_{\leq t}^{ij} \in \{0,1\}\}$$

be the set of all possible adjacency matrices among $N_t$ cases at time $t$. $N_t$ denotes the number of cases in the network at time $t$. Note that the cardinality of $\mathcal{C}_t$ increases exponentially for every newly added case. The probability of observing $C_t$ given past citations $C_{<t}$, where $C_{\leq t}$ is the network up to time $t$ is defined as

$$P(C_t | C_{<t}, \theta) = \frac{\exp\left(\theta^T \cdot h(C_{\leq t})\right)}{\sum_{Z_t^* \in \mathcal{C}_t} \exp\left(\theta^T \cdot h(Z_{\leq t}^*)\right)} \tag{1}$$

where $\theta \in \mathbb{R}^q$ is a $q$-dimensional vector of parameters, $h : \mathcal{C}_t \to \mathbb{R}^q$ is a q-dimensional vector of different statistics and $\kappa(\theta) := \sum_{Z_t^* \in \mathcal{C}_t} \exp\left(\theta^T \cdot h(Z_{\leq t}^*)\right)$ is a normalization constant that ensures that Equation (1) defines a probability function on $\mathcal{C}_t$.

The c-TERGM is specified through the decision regarding which network statistics $h(\cdot)$ to incorporate. We include the following network statistics for the Supreme Court citation network, which are derived from the interdependence hypotheses described in the previous section:

$$h_{edges} : \mathcal{C}_t \to \mathbb{R} \quad , \quad C_{\leq t} \to \sum_{ij} C^{ij}_{\leq t},$$

the number of citations. $h_{edges}$ performs the function of an intercept, and models the expected value of any given edge.

$$h_{outstar} : \mathcal{C}_t \to \mathbb{R} \quad , \quad C_{\leq t} \to \sum_{i=1} \binom{\sum_{j \neq i} C^{ij}_{\leq t}}{2}$$

the number of out-two-stars. An out-two-star is a configuration in which one node sends to two other nodes. The number of these configurations grows quadratically as the origins of ties concentrate on a few highly active/social senders. The out-two-stars configuration is commonly used to model activity (sender activation) in ERGMs.

$$h_{instar} : \mathcal{C}_t \to \mathbb{R} \quad , \quad C_{\leq t} \to \sum_{j=1} \binom{\sum_{i \neq j} C^{ij}_{\leq t}}{2}$$

the number of in-two-stars. An in-two-star is a configuration in which one node receives ties from two other nodes. The number of these configurations grows quadratically as the destinations of ties concentrate on a few highly popular recipients. The out-two-stars configuration is commonly used to model popularity in ERGMs.

$$\begin{aligned} h_{triangle} : \mathcal{C}_t \to \mathbb{R} \quad , \quad C_{\leq t} \to \sum_{j<i<k} & C^{ij} \cdot C^{jk} \cdot C^{ki} + C^{ji} \cdot C^{jk} \cdot C^{ki} + C^{ij} \cdot C^{kj} \cdot C^{ki} \\ & + C^{ji} \cdot C^{kj} \cdot C^{ki} + C^{ij} \cdot C^{jk} \cdot C^{ik} + C^{ji} \cdot C^{jk} \cdot C^{ik} \\ & + C^{ij} \cdot C^{kj} \cdot C^{ik} + C^{ji} \cdot C^{kj} \cdot C^{ik}, \end{aligned}$$

The number of triangles in the network. Triangles measure triad closure, as discussed above.

$$h_{reciprocity} : \mathcal{C}_t \to \mathbb{R} \quad , \quad C_{\leq t} \to \sum_{ij} C^{ij}_{\leq t} C^{ji}_{\leq t},$$

The number of reciprocal ties in the network. Even though the reciprocity of ties is not a meaningful network statistic for most citation networks, it may still play a role in the Supreme Court citation network. This statistic comes with the restriction that it only can appear among cases that entered the network at the same time.

$$h_{covariate} : \mathcal{C}_t \to \mathbb{R} \quad , \quad C_{\leq t} \to \sum_{ij} C^{ij}_{\leq t} X^{ij}_{\leq t},$$

the effect of an exogenous covariate ($X$). We include several exogenous covariates based on this standard statistic formulation. These covariates are discussed below.

The structure of this model is very similar to a conventional (T)ERGM setup. The main difference is that $\mathcal{C}_t$ excludes adjacency matrices in which there are loops that include edges sent at different time points. Loops can exist in the Supreme Court citation network, but only among cases decided in the same term. Furthermore, a case $i$ can only cite a case $j$ if $j$ entered the network before or at the same time as $i$ did.

The normalizing constant in the denominator of the probability function (i.e., Equation 1) is computationally intractable. This means that straightforward methods of maximum likelihood estimation (MLE) are infeasible with ERGM family models. Methods of Monte Carlo approximation are well established (Hunter and Handcock, 2006; Van Duijn, Gile and Handcock, 2009; Hummel, Hunter and Handcock, 2012), but with almost $10,000$ vertices, even the methods of Monte Carlo MLE are computationally expensive (Schmid and Desmarais, 2017). To estimate the parameters of the c-TERGM, we rely on two features of the model and data that facilitate the application of an approximate inference method that is faster than Monte Carlo approximation—maximum pseudolikelihood. First, our sample size is very large. Second, under the c-TERGM we assume that citations are formed in each time period conditional upon the citations that exist in the past. We estimate the c-TERGM via bootstrapped pseudolikelihood (Desmarais and Cranmer, 2012, 2010).

## 5 Correlates of Legal Citations

Theorizing that legal citations arise through network interdependence processes represents a potential contribution to the literature. To demonstrate the value of our contribution we must also incorporate established non-network processes that have been found to affect legal citations. We incorporate these processes through the specification of $h_{covariate}$ terms to be included in the c-TERGM.

We incorporate a statistic to model the degree to which cases cite those that are similar in terms of the ideological positions of the justices who supported the decision. We account for this effect following Spriggs and Hansford (2001), who find that cases are more likely to be overruled when the Court is ideologically distant from the median justice in the majority coalition that decided the case. Clark and Lauderdale (2010) estimates a latent coordinate model of Supreme Court opinions based on the network of case-to-case citations. They find that the majority opinion falls at the ideal point of the median member of the majority coalition in the case. We include a covariate term in which $X_{ij}$ is the absolute difference between the Martin-Quinn scores (Martin and Quinn, 2002) of the median justices in the majority coalitions for cases $i$ and $j$. We expect this variable to have a negative effect, which would correspond to cases citing those to

| | | | |
|---|---|---|---|
| 1 | Criminal Procedure | 8 | Economic Activity |
| 2 | Civil Rights | 9 | Judicial Power |
| 3 | First Amendment | 10 | Federalism |
| 4 | Due Process | 11 | Interstate Relations |
| 5 | Privacy | 12 | Federal Taxation |
| 6 | Attorneys | 13 | Miscellaneous |
| 7 | Unions | 14 | Private Action |

Table 1: Assigned numbers for the variable *Issue Area*. This information originates from the Supreme Court Database code book.

which they are ideologically similar.

We include two sets of dummy variables that account for the issue areas of cases. In one set of dummy variable—sender intercepts—the variable $X^{ij}$ indicates the issue area of the sending case ($i$). In the set of receiver intercepts $X^{ij}$ indicates the issue of the receiver case. Issue area data comes from the Supreme Court Database (SCDB) (Spaeth et al., 2014). We include these variables because Cross (2010) finds that the number of citations to Supreme Court opinions depends heavily on the issue area of the case. Table 1 states the different issue areas.

We model the way in which citations to a case depend upon the age of a case. For this we use a second-order polynomial in which one covariate $X^{ij}$ is defined as the age of case $j$ at the time that case $i$ is decided, and another term in which $X^{ij}$ is the squared age of case $j$ at the time that case $i$ is decided. We include these covariates because Black and Spriggs (2013) find that the number of citations to a Supreme Court case over time depends significantly on the age of the case, characterized by a sharp drop off and leveling out with age.

Benjamin and Desmarais (2012) study the propensity for cases to be overruled and cited in other negative ways. They find that cases with majority coalitions that are large and ideologically broad are less likely to be cited negatively. In our data we do not differentiate between negative and positive citations, but since the overwhelming majority of citations are positive, we hypothesize that the effects they found will be reversed in our analysis. We include one covariate in which $X^{ij}$ is the number of justices in the majority coalition for case $j$. We also include another covariate in which $X^{ij}$ is the absolute difference between the maximum and minimum ideal points of the justices in the majority coalition for case $j$. We expect both of these covariates to have positive effects.

- ToDo: mention new variable: overruled cases

*Comment from Benjamin Kassow: In section 5, I think the discussion is mostly good, but could be organized a bit differently. I think the discussion of the random intercepts is a bit confusing early on (it is explained more effectively later).*

# 6 Empirical Analysis

Our three data sources for this study include the Supreme Court Database (SCDB) (Spaeth et al., 2014), Martin-Quinn scores (Martin and Quinn, 2002), and Supreme Court citation data provided by the CourtListener Free Law Project (Lissner and Carver, 2010). We limit the Supreme Court terms included in our analysis to those that are covered by all three of these data source (1937–2015)[2]. The Supreme Court citation network from $1937 - 2015$ consists of $9,945$ cases and the breakdown of the data by the Court's Chief Justice is presented in Table 2. In order to only account for positive citations, we exclude 315 citations that caused the cited case to be overruled. As a result, the network has a total of $111,986$ ties. The number of triangles in the network is $250,717$ and the number of mutual ties is 56. The in- and outdegree distributions (i.e., the distributions of the number of citations sent and received by cases, respectively), are visualized in Figure 1. The maximum indegree is 230 and the maximum outdegree is 162.

| | Terms | Total Number Cases | Cases/Term |
|---|---|---|---|
| CE Hughes* | 1937 - 1941 | 628 | 125.6 |
| HF Stone | 1942 - 1946 | 756 | 151.2 |
| FM Vinson | 1946 - 1953 | 789 | 98.63 |
| E Warren | 1954 - 1969 | 2149 | 126.41 |
| WE Burger | 1970 - 1986 | 2805 | 155.83 |
| W Rehnquist | 1987 - 2005 | 2022 | 106.42 |
| J Roberts ** | 2006 - 2015 | 796 | 79.6 |

Table 2: For the time range of interest (1937 - 2015) this table displays the chief justices, the time range they served as chief justice, the number of cases in their time range as well as the average number of cases per year.
* CE Hughes served as chief justice from 1930 - 1941.
** J Roberts still serves as chief justice (retrieved 3/2018).

The degree distributions indicate that there is a long tail to both the number of citations sent and received. These long-tailed distributions provide preliminary evidence of both popularity and activity dynamics. Figure 2 displays the complete network. We see here that the densest rates of tie formation tend to be between consecutive courts (e.g., the Stone Court is much more tightly tied to the Hughes Court than the Rehnquist Court is to the Hughes Court). This pattern lends preliminary/descriptive support to the hypothesis that the rate of citations to a case decrease over time.

- update Figure 2 to time period 1937 - 2015
- add data source for overruling citations

[2]There were 145 cases that were listed in the SCDB but could not be matched to a case in the CourtListener data. We decided to exclude these 145 cases from our model.

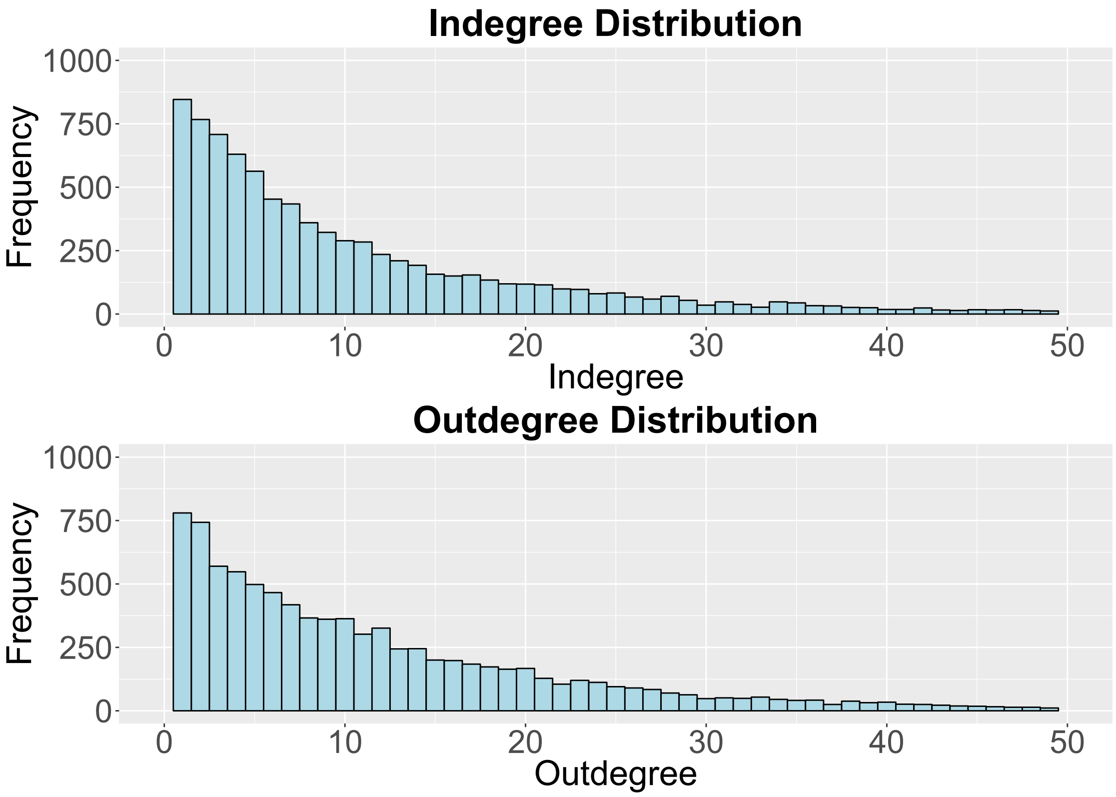


Figure 1: The in- and outdegree distribution of the Supreme Court Citation Network from 1937 - 2001. There are cases with an indegree >50, but they are not captured in this figure.

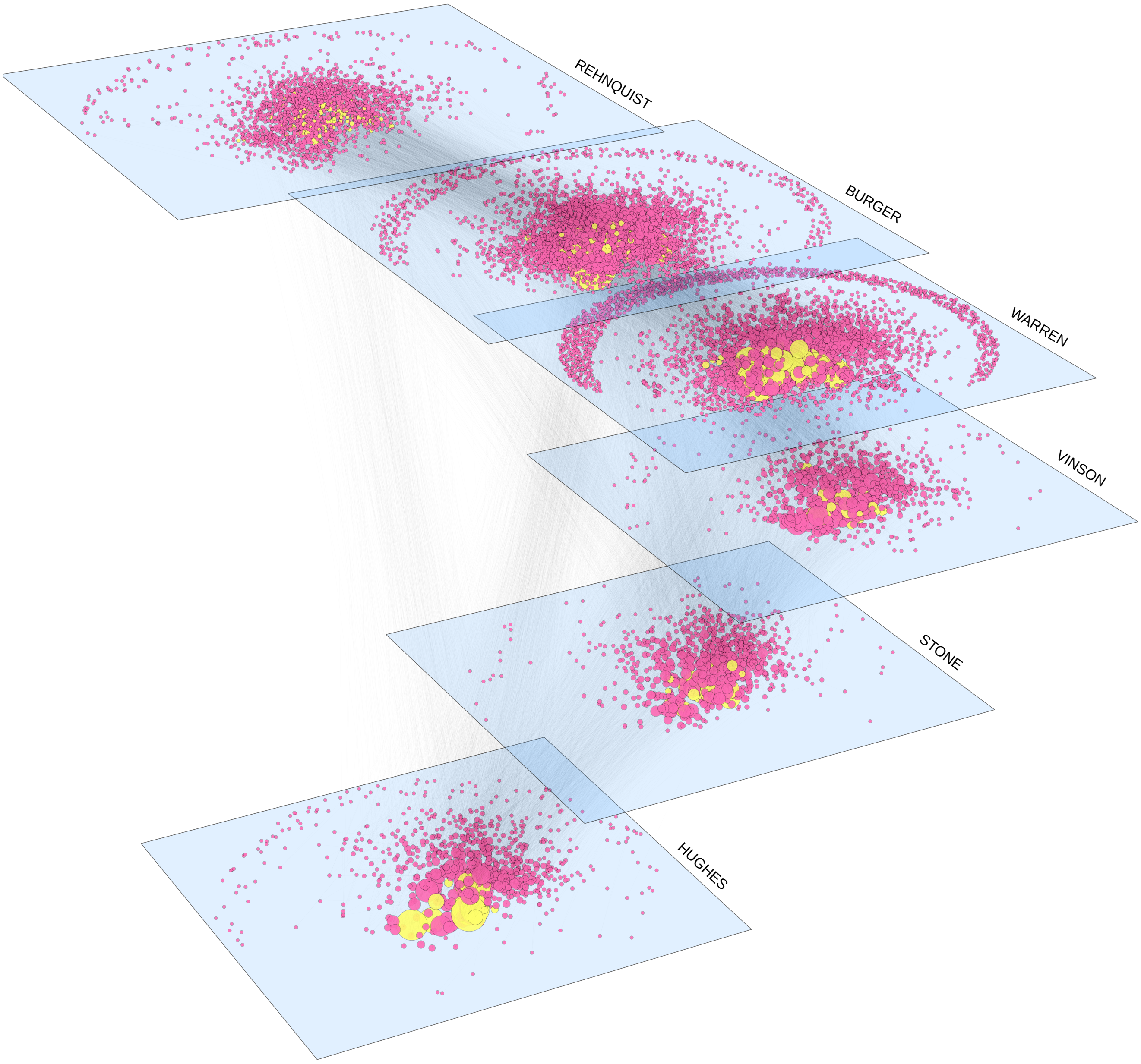


Figure 2: Supreme Court Citation Network, 1937-2001. Nodes are Supreme Court cases, with size based on incoming citations. Salient cases (*Oxford Guide*) are in yellow. Each network layer contains cases decided under a different chief justice. Because our data is temporally bound, cases for the Hughes and Rehnquist courts are not complete.

## 6.1 c-TERGM Results

The results from the c-TERGM are presented in Table 3 and Figures 3 and 4. Table 3 gives the coefficient estimates and 95% bootstrap confidence intervals. In Figures, 3 and 4, we depict coefficient estimates and confidence intervals for all of the effects for which we estimated trends (excluding the intercept—edge count—which is typically not interpreted substantively). The coefficients in an ERGM family model can be interpreted in the same way as logistic regression coefficients—they give the change in the log odds of a tie from $i$ to $j$ given a one unit increase in the respective variable. We first discuss the dependence effect estimates (panels (a)–(d) in Figure 3). The reciprocity effect is substantial, positive, and statistically significant for most of the period under study. Conditional upon cases $i$ and $j$ being considered within the same term, the log of the odds that opinion $i$ cites opinion $j$ increases by approximately 2 when opinion $j$ cites opinion $i$. The naive probability that any one case cites another case decided at the same time or previously in the Supreme Court Citation network is approximately 0.0019. This probability increases eight-fold, to approximately 0.015 if the citation would be reciprocal. The transitivity (i.e., triangle closure) effect is also substantial, positive, and statistically significant over the entire period studied. The probability that case $i$ cites case $j$ increases approximately five-fold if the citation would close a triangle involving cases $i$ and $j$. The activity and popularity effects are both estimated to be positive, but trend downward and are more modest in scale (and in the case of activity, not statistically significant later in the time series). Interpreting the popularity effect—the log odds that $i$ cites $j$ increases by 0.018–0.035 with every additional citation to case $j$. Based on the in-degree distribution, it appears that the popularity of cases can be classified by multiples of 10. If case $j$ increases from ten to twenty citations from other cases, the probability that case $i$ cites case $j$ increases from approximately 0.0030 to 0.0040. Overall, these results represent evidence of complex dependence in Supreme Court citation formation.

The effects of the exogenous covariates included in the c-TERGM (panels (e)–(k) in Figure 4) exhibit some surprising dynamics. The effects of absolute difference in Martin-Quinn score, and majority size of the recipient case all flip signs throughout the time period under study. This may be an artifact of the simple linear trends we have estimated, but is interesting nonetheless. The attitudinal model holds that ideological preferences drive decision-making on the US Supreme Court. Ideological selection of precedents would be evidenced by a negative effect of the absolute difference in MQ scores in the c-TERGM. We find this effect is negative and statistically significant only after 1975. The absolute difference in MQ scores variable is on the scale of approximately 0—3. At its most ideological point (i.e. 2001), the probability that case $i$ cites case $j$ decreases from approximately 0.024 to 0.018—by one third—if the ideological distance between the majority coalitions in the two cases increases by 3. The effects of the majority size and ideological breadth of case $j$ are similar in both scale and trend. It is not until later in the time series that they both exhibit their expected positive effects. The effect of same issue is positive and statistically significant, and substantial for the time period under study. The probability that $i$ cites $j$ increases by approximately three-fold over the naive probability of citation if $i$ and $j$ have the same legal issue area.

The results from the c-TERGM provide evidence that citations between US Supreme Court cases are driven both by exogenous case attributes and complex dependence processes. The scales of covariate and dependence effects are comparable.

- adjust results/numbers in text

| | Estimate | Lower Bound | Upper Bound | Significance |
|---|---|---|---|---|
| Edges | -6.249 | -6.412 | -6.110 | * |
| Instar(2) | 0.035 | 0.032 | 0.038 | * |
| Outstar(2) | 0.021 | 0.019 | 0.028 | * |
| Mutual | 2.071 | -4.157 | 4.072 | |
| Triangle | 1.507 | 1.431 | 1.559 | * |
| Martin Quinn Score | 0.079 | 0.029 | 0.130 | * |
| Same Issue Area Homophily | 1.447 | 1.389 | 1.515 | * |
| Year Difference | -0.068 | -0.079 | -0.058 | * |
| (Year Differnce)$^2$ | 0.0019 | 0.0016 | 0.0022 | * |
| Receiver Abs Diff MQ Score in Majority | 0.047 | 0.029 | 0.063 | * |
| Receiver Number Justices in Majority | -0.075 | -0.095 | -0.055 | * |
| Receiver Sender Year | 0.0039 | 0.0036 | 0.0042 | * |
| Overruled Cases | -1.043 | -1.624 | -0.564 | * |
| Sender Same Issue Area 2 | 0.142 | 0.100 | 0.185 | * |
| Sender Same Issue Area 3 | -0.337 | -0.403 | -0.270 | * |
| Sender Same Issue Area 4 | 0.498 | 0.438 | 0.561 | * |
| Sender Same Issue Area 5 | 0.182 | 0.031 | 0.305 | * |
| Sender Same Issue Area 6 | 0.536 | 0.422 | 0.648 | * |
| Sender Same Issue Area 7 | 0.370 | 0.316 | 0.426 | * |
| Sender Same Issue Area 8 | 0.107 | 0.059 | 0.156 | * |
| Sender Same Issue Area 9 | 0.278 | 0.238 | 0.329 | * |
| Sender Same Issue Area 10 | 0.371 | 0.317 | 0.433 | * |
| Sender Same Issue Area 11 | -0.030 | -0.151 | 0.164 | |
| Sender Same Issue Area 12 | 0.064 | -0.014 | 0.130 | |
| Sender Same Issue Area 13 | 0.608 | 0.423 | 0.792 | * |
| Sender Same Issue Area 14 | 0.232 | 0.075 | 0.393 | * |
| Receiver Same Issue Area 2 | 0.252 | 0.199 | 0.310 | * |
| Receiver Same Issue Area 3 | -0.092 | -0.194 | 0.014 | |
| Receiver Same Issue Area 4 | 0.481 | 0.397 | 0.570 | * |
| Receiver Same Issue Area 5 | 0.374 | 0.224 | 0.519 | * |
| Receiver Same Issue Area 6 | 0.521 | 0.346 | 0.652 | * |
| Receiver Same Issue Area 7 | 0.412 | 0.339 | 0.487 | * |
| Receiver Same Issue Area 8 | 0.174 | 0.134 | 0.232 | * |
| Receiver Same Issue Area 9 | 0.311 | 0.240 | 0.361 | * |
| Receiver Same Issue Area 10 | 0.506 | 0.429 | 0.577 | * |
| Receiver Same Issue Area 11 | 0.299 | 0.088 | 0.432 | * |
| Receiver Same Issue Area 12 | 0.350 | 0.276 | 0.445 | * |
| Receiver Same Issue Area 13 | 0.972 | 0.780 | 1.135 | * |
| Receiver Same Issue Area 14 | 0.517 | 0.200 | 0.740 | * |
| Instar(2) × Sender Year | -0.00026 | -0.00031 | -0.00021 | * |
| Outstar(2) × Sender Year | -0.00031 | -0.00045 | -0.00027 | * |
| Mutual × Sender Year | 0.110 | 0.010 | 0.524 | * |
| Triangle × Sender Year | -0.00022 | -0.0016 | 0.0013 | |
| Martin Quinn Score × Sender Year | -0.0019 | -0.0030 | -0.0009 | * |
| Same Issue Area × Sender Year | -0.0068 | -0.0081 | -0.0055 | * |
| Year Difference × Sender Year | 0.00046 | 0.00028 | 0.00065 | * |
| Year Difference$^2$× Sender Year | -0.000024 | -0.000028 | -0.000021 | * |
| MQ Score in Majority × Sender Year | -0.00053 | -0.00093 | -0.0018 | * |
| Justices in Majority × Sender Year | 0.0017 | 0.0013 | 0.0021 | * |
| Overruled Cases × Sender Year | 0.012 | 0.002 | 0.023 | * |

Table 3: Bootstrapped MPLE Results for the time period 1937 – 2015. A '*' indicates that the 2.5th and 97.5th quantile of the variable does not include '0' and as a result is statistically significant.

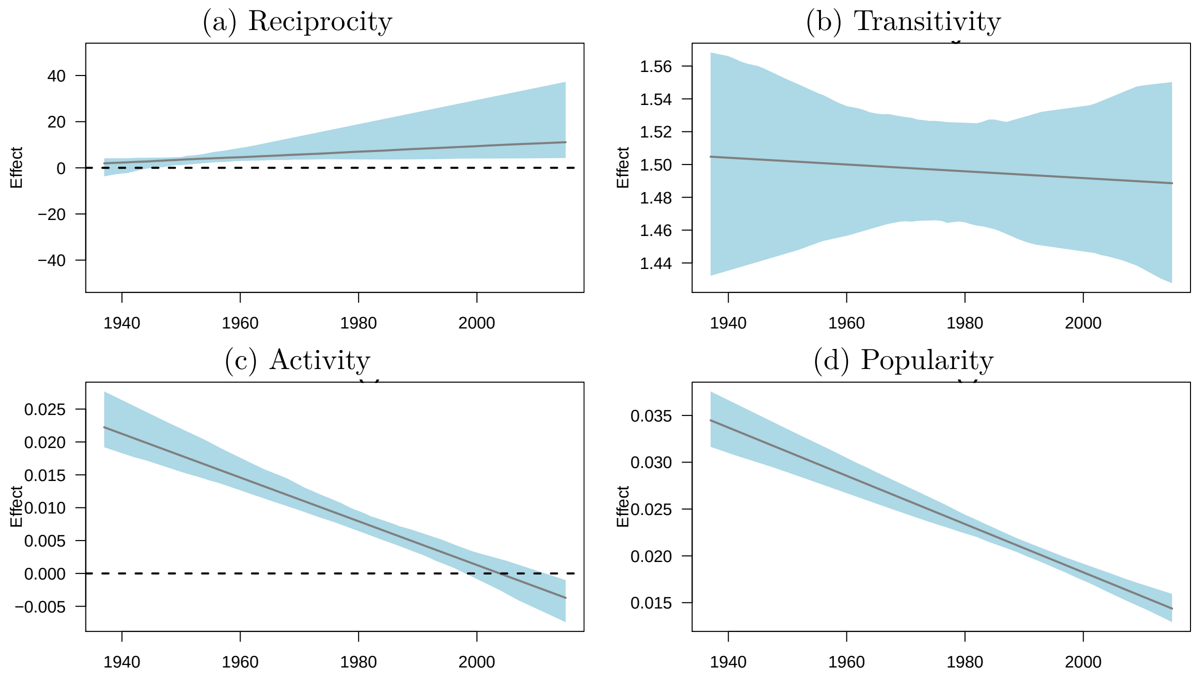


Figure 3: Trending effects for dependence effect estimates from the c-TERGM. The solid lines plot the point estimates, and the shaded areas span 95% confidence intervals.

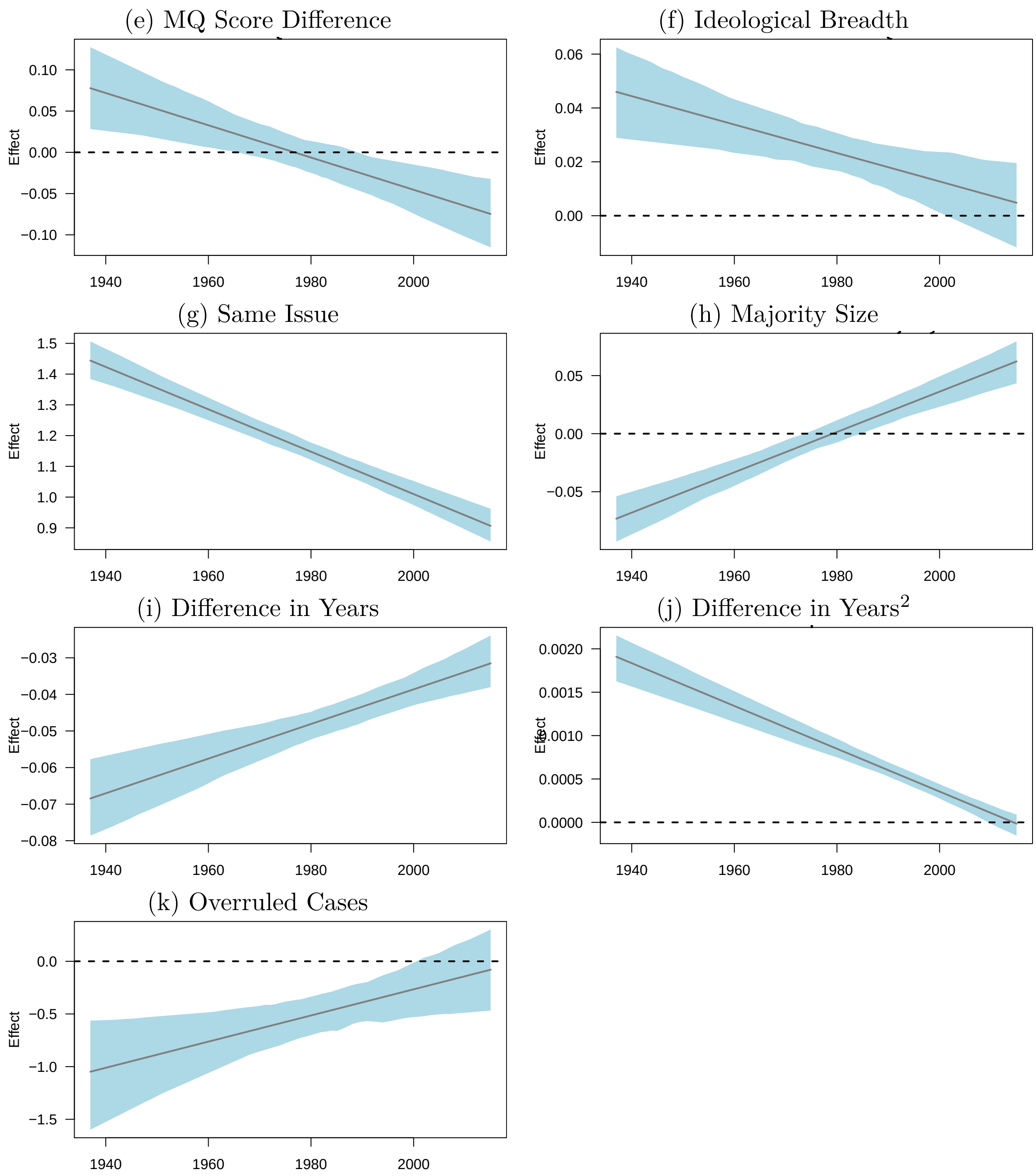


Figure 4: Trending effects for exogenous variables from the c-TERGM. The solid lines plot the point estimates, and the shaded areas span 95% confidence intervals.

## 6.2 Predictive Performance

Our case for studying legal citations at the directed dyadic level hinges upon the contribution to modeling offered by incorporating network dependence. To quantify this contribution, we use out-of-sample prediction. Predicting out-of-sample offers an unbiased and general purpose way to evaluate the contribution, in terms of model fit, of one or more terms/parameters in a model (Jensen and Cohen, 2000; Ward, Greenhill and Bakke, 2010). Unlike in-sample measures of model fit, out-of-sample methods are highly robust in avoiding overfitting, and work when we cannot accurately calculate the value of the likelihood function, as in the current case. Out-of-sample prediction is a common way to evaluate methods for modeling ties in networks, and has recently been applied to TERGMs in particular (Desmarais and Cranmer, 2013; Cranmer and Desmarais, 2017).

In our prediction experiment we randomly split the directed dyads into an 80% training set and a 20% test set. The parameters of the model are estimated using the directed dyads in the training set, and the parameters are used to form the conditional probability of a tie for all of the directed dyads in the test set. Directed dyads for which the conditional probability of a tie exceeds 0.5 are predicted to be citations. The experiment is run with the full model, and with a model that excludes all of the dependence terms (i.e., all terms involving reciprocity, in and out stars, and triangles)—the independent dyads model. We run this experiment for 10 iterations. Predictive performance is evaluated with three common and related measures—precision (i.e., the proportion of predicted citations that are actually citations), recall (the proportion of actual citations that are predicted to be citations), and the F1 score (the harmonic mean of precision and recall) (see, e.g., Makhoul et al., 1999, for discussion of these measures). All three measures are bounded between 0 and 1, with higher scores indicating better performance.

| | Independent Model | | Full Model | |
|---|---|---|---|---|
| | mean | range | mean | range |
| precision | 0.5589 | (0.5508, 0.5757) | 0.8677 | (0.865, 0.8718) |
| recall | 0.098 | (0.0943, 0.103) | 0.5943 | (0.5887, 0.6032) |
| F1 score | 0.1667 | (0.1618, 0.1748) | 0.7054 | (0.7013, 0.7119) |

Table 4: The predictive performance of the directed dyadic models with linear trends, over ten 80/20 train/test splits.

We see from Table 4 that, based on all three measures, the predictive performance of the model improves dramatically from adding the network dependence terms. The recall of the full model is particularly impressive, indicating that it recovers over half of the actual citations in the test set. This provides clear evidence that the full model, which includes covariates and network dependence terms, represents a more accurate and complete model of the process of citation formation in US Supreme Court opinions.

| | Independent Model | | Full Model | |
|---|---|---|---|---|
| | mean | range | mean | range |
| precision | 0.5872 | (0.5781, 0.5962) | 0.8622 | (0.8593, 0.8668) |
| recall | 0.0565 | (0.0547, 0.0587) | 0.5839 | (0.5777, 0.592) |
| F1 score | 0.103 | (0.1, 0.1069) | 0.6963 | (0.6915, 0.7027) |

Table 5: The predictive performance of the directed dyadic models with no trends, over ten 80/20 train/test splits.

| | Independent Model | | Full Model | |
|---|---|---|---|---|
| | mean | range | mean | range |
| precision | 0.6441 | (0.6302, 0.6536) | 0.8671 | (0.8626, 0.873) |
| recall | 0.0852 | (0.0832, 0.0863) | 0.594 | (0.5874, 0.6012) |
| F1 score | 0.1504 | (0.1473, 0.1525) | 0.705 | (0.7008, 0.712) |

Table 6: The predictive performance of the directed dyadic models with quadratic trends, over ten 80/20 train/test splits.

We also conduct a predictive experiment for the model that does not consider any trends and for the model that does consider quadratic instead of linear trends. The predictive performance results are given in tables 5 and 6. By comparing the full model results in tables 5 and 6 with the results of the full linear trend model in table 4 we can conclude all three models perform similarly well.

*Comment from Benjamin Kassow: In section 6.2 (predictive performance), I like what is here, but was hoping to see a bit more discussion about the predictive performance of the model, with some explanation of what the predictive performance means. I understand it (I think), but depending on where this gets sent to, I am not sure if all reviewers will. I also think tying this into the overall model explanation or discussion of model performance overall will be helpful for better integrating this section into the rest of the manuscript.*

# 7 Conclusion

We present a methodology for studying citations between US Supreme Court opinions at the dyadic level, as a network. This methodology—the citation-TERGM—enables researchers to include both exogenous covariates such as the ideological predisposition and age of a case, and dependence terms, such as transitivity and reciprocity, as explanations for citation formation.

We apply this methodology to a network that includes all Supreme Court cases decided between 1937 and 2015. We find, somewhat counterintuitively, that Supreme Court citations are highly reciprocal. We also find that citations are driven by dependencies such as triad closure, popularity, and activity. The dependence effects that we identify are as substantively and statistically significant as the effects of the exogenous covariates we include in the model. The summary result from this analysis is that theoretical models of Supreme Court citation formation should consider both the effects of case characteristics and the structure of past citations.

We recognize several limitations of the current analysis, which should be addressed in future iterations of this work or in future studies. First, our theoretical claims and analyses treat all citations—positive and negative—as equivalent. Both data and methodological limitations will need to be overcome to consider these two citation types. Second, there is nothing in the c-TERGM that governs the generation of new cases (i.e., cases are assumed to arise exogenously, unrelated to the history of the citation network). As new cases are often taken up by the Supreme Court in order to test, clarify, and/or challenge existing precedents, it is likely inappropriate to assume that new cases arise exogenous to the existing citation network. Overcoming this limitation would require methodological innovation in the c-TERGM as well as theoretical development regarding exactly how citation structure might predict case emergence.

## Bibliography

Bailey, Michael A and Forrest Maltzman. 2008. "Does legal doctrine matter? Unpacking law and policy preferences on the US Supreme Court." *American Political Science Review* 102(3):369–384.

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

Barabási, Albert-László and Réka Albert. 1999. "Emergence of scaling in random networks." *science* 286(5439):509–512.

Benjamin, Stuart Minor and Bruce A Desmarais. 2012. "Standing the test of time: The breadth of majority coalitions and the fate of us supreme court precedents." *Journal of Legal Analysis* 4(2):445–469.

Black, Ryan C and James F Spriggs. 2013. "The Citation and Depreciation of US Supreme Court Precedent." *Journal of Empirical Legal Studies* 10(2):325–358.

Block, Per, Johan Koskinen, James Hollway, Christian Steglich and Christoph Stadtfeld. 2018. "Change we can believe in: Comparing longitudinal network models on consistency, interpretability and predictive power." *Social Networks* 52:180–191.

Bommarito II, Michael J, Daniel Katz and Jon Zelner. 2009. Law as a seamless web?: comparison of various network representations of the united states supreme court corpus (1791-2005). In *Proceedings of the 12th international conference on artificial intelligence and law.* ACM pp. 234–235.

Caldeira, Gregory A. 1988. "Legal precedent: Structures of communication between state supreme courts." *Social networks* 10(1):29–55.

Carstens, J. 2016. "Topology of complex networks: models and analysis.".

Clark, Jennifer Hayes and Veronica Caro. 2013. "Multimember districts and the substantive representation of women: An analysis of legislative cosponsorship networks." *Politics & Gender* 9(1):1–30.

Clark, Tom S and Benjamin Lauderdale. 2010. "Locating Supreme Court opinions in doctrine space." *American Journal of Political Science* 54(4):871–890.

Cranmer, Skyler J and Bruce A Desmarais. 2010. "Inferential network analysis with exponential random graph models." *Political Analysis* 19(1):66–86.

Cranmer, Skyler J and Bruce A Desmarais. 2016. "A critique of dyadic design." *International Studies Quarterly* 60(2):355–362.

Cranmer, Skyler J and Bruce A Desmarais. 2017. "What Can We Learn from Predictive Modeling?" *Political Analysis* pp. 1–22.

Cranmer, Skyler J, Bruce A Desmarais and Elizabeth J Menninga. 2012. "Complex dependencies in the alliance network." *Conflict Management and Peace Science* 29(3):279–313.

Cross, Frank B. 2010. "Determinants of citations to Supreme Court opinions (and the remarkable influence of Justice Scalia)." *Supreme Court Economic Review* 18(1):177–202.

Desmarais, Bruce A and Skyler J Cranmer. 2010. Consistent confidence intervals for maximum pseudolikelihood estimators. In *Proceedings of the Neural Information Processing Systems 2010 Workshop on Computational Social Science and the Wisdom of Crowds.*

Desmarais, Bruce A and Skyler J Cranmer. 2012. "Statistical mechanics of networks: Estimation and uncertainty." *Physica A: Statistical Mechanics and its Applications* 391(4):1865–1876.

Desmarais, Bruce A and Skyler J Cranmer. 2013. "Forecasting the locational dynamics of transnational terrorism: A network analytic approach." *Security Informatics* 2(1):8.

Desmarais, Bruce A and Skyler J Cranmer. 2017. Statistical Inference in Political Networks Research. In *The Oxford Handbook of Political Networks.*

Eom, Young-Ho and Santo Fortunato. 2011. "Characterizing and modeling citation dynamics." *PloS one* 6(9):e24926.

Fowler, James H and Sangick Jeon. 2008. "The authority of Supreme Court precedent." *Social networks* 30(1):16–30.

Fowler, James H, Timothy R Johnson, James F Spriggs, Sangick Jeon and Paul J Wahlbeck. 2007. "Network analysis and the law: Measuring the legal importance of precedents at the US Supreme Court." *Political Analysis* 15(3):324–346.

Garlaschelli, Diego and Maria I Loffredo. 2004. "Patterns of link reciprocity in directed networks." *Physical Review Letters* 93(26):268701.

Gillman, Howard. 2001. "What's law got to do with it? Judicial behavioralists test the ?legal model? of judicial decision making." *Law & Social Inquiry* 26(2):465–504.

Graif, Corina, Alina Lungeanu and Alyssa M Yetter. 2017. "Neighborhood isolation in Chicago: Violent crime effects on structural isolation and homophily in inter-neighborhood commuting networks." *Social Networks* .

Hajra, Kamalika Basu and Parongama Sen. 2005. "Aging in citation networks." *Physica A: Statistical Mechanics and its Applications* 346(1):44–48.

Hajra, Kamalika Basu and Parongama Sen. 2006. "Modelling aging characteristics in citation networks." *Physica A: Statistical Mechanics and its Applications* 368(2):575–582.

Hansford, Thomas G and James F Spriggs. 2006. *The politics of precedent on the US Supreme Court*. Princeton University Press.

Harris, Peter. 1982. "Structural change in the communication of precedent among state supreme courts, 1870–1970." *Social Networks* 4(3):201–212.

Holland, Paul W and Samuel Leinhardt. 1971. "Transitivity in Structural Models of Small Groups." *Small Group Research* 2(2):107–124.

Howard, Michael, Emily Cox Pahnke, Warren Boeker et al. 2016. "Understanding network formation in strategy research: Exponential random graph models." *Strategic management journal* 37(1):22–44.

Hummel, Ruth M, David R Hunter and Mark S Handcock. 2012. "Improving simulation-based algorithms for fitting ERGMs." *Journal of Computational and Graphical Statistics* 21(4):920–939.

Hunter, David R and Mark S Handcock. 2006. "Inference in curved exponential family models for networks." *Journal of Computational and Graphical Statistics* 15(3):565–583.

Jensen, David D and Paul R Cohen. 2000. "Multiple comparisons in induction algorithms." *Machine Learning* 38(3):309–338.

Jeong, Hawoong, Zoltan Néda and Albert-László Barabási. 2003. "Measuring preferential attachment in evolving networks." *EPL (Europhysics Letters)* 61(4):567.

Karrer, Brian and Mark EJ Newman. 2009. "Random graph models for directed acyclic networks." *Physical Review E* 80(4):046110.

Knight, Jack and Lee Epstein. 1996. "The norm of stare decisis." *American Journal of Political Science* pp. 1018–1035.

Leicht, Elizabeth A, Gavin Clarkson, Kerby Shedden and Mark EJ Newman. 2007. "Large-scale structure of time evolving citation networks." *The European Physical Journal B-Condensed Matter and Complex Systems* 59(1):75–83.

Lissner, Michael and Brian W. Carver. 2010. "CourtListener.com: A platform for researching and staying abreast of the latest law.".
**URL:** *http://www.courtlistener.com*

Lupu, Yonatan and Erik Voeten. 2012. "Precedent in international courts: A network analysis of case citations by the European court of human rights." *British Journal of Political Science* 42(2):413–439.

Lupu, Yonatan and James H Fowler. 2013. "Strategic citations to precedent on the US Supreme Court." *The Journal of Legal Studies* 42(1):151–186.

Makhoul, John, Francis Kubala, Richard Schwartz, Ralph Weischedel et al. 1999. Performance measures for information extraction. In *Proceedings of DARPA broadcast news workshop*. pp. 249–252.

Martin, Andrew D and Kevin M Quinn. 2002. "Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999." *Political Analysis* 10(2):134–153.

Pelc, Krzysztof J. 2014. "The politics of precedent in international law: A social network application." *American Political Science Review* 108(3):547–564.

Richards, Mark J and Herbert M Kritzer. 2002. "Jurisprudential regimes in Supreme Court decision making." *American Political Science Review* 96(2):305–320.

Schmid, Christian S and Bruce A Desmarais. 2017. "Exponential Random Graph Models with Big Networks: Maximum Pseudolikelihood Estimation and the Parametric Bootstrap." *arXiv preprint arXiv:1708.02598* .

Simkin, Mikhail V and Vwani P Roychowdhury. 2007. "A mathematical theory of citing." *Journal of the Association for Information Science and Technology* 58(11):1661–1673.

Spaeth, Harold, Lee Epstein, Ted Ruger, Keith Whittington, Jeffrey Segal and Andrew D Martin. 2014. "Supreme Court Database Code Book.".
**URL:** *http://scdb.wustl.edu*

Spriggs, James F and Thomas G Hansford. 2001. "Explaining the overruling of US Supreme Court precedent." *The Journal of Politics* 63(4):1091–1111.

Van Duijn, Marijtje AJ, Krista J Gile and Mark S Handcock. 2009. "A framework for the comparison of maximum pseudo-likelihood and maximum likelihood estimation of exponential family random graph models." *Social Networks* 31(1):52–62.

Vazquez, Alexei. 2001. "Statistics of citation networks." *arXiv preprint cond-mat/0105031* .

Wang, Dashun, Chaoming Song and Albert-László Barabási. 2013. "Quantifying long-term scientific impact." *Science* 342(6154):127–132.

Wang, Mingyang, Guang Yu and Daren Yu. 2008. "Measuring the preferential attachment mechanism in citation networks." *Physica A: Statistical Mechanics and its Applications* 387(18):4692–4698.

Ward, Michael D, Brian D Greenhill and Kristin M Bakke. 2010. "The perils of policy by p-value: Predicting civil conflicts." *Journal of Peace Research* 47(4):363–375.