Norman Schofield · Gonzalo Caballero · Daniel Kselman *Editors*

Advances in Political Economy

Institutions, Modelling and Empirical Analysis

This book presents latest research in the field of Political Economy, dealing with the integration of economics and politics and the way institutions affect social decisions. The focus is on innovative topics such as an institutional analysis based on case studies; the influence of activists on political decisions; new techniques for analyzing elections, involving game theory and empirical methods.

Social Sciences / Political Science

ISBN 978-3-642-35238-6

springer.com

Schofield · Caballero
Kselman *Eds.*
Advances in Political Economy

Norman Schofield
Gonzalo Caballero
Daniel Kselman *Editors*

Advances in Political Economy

Institutions, Modelling and Empirical Analysis

Springer

too long to develop here (the detailed calculations are reported in previous versions of this paper). With the appropriate algebra we find that

$$P(v_R = V|primary) = \begin{cases} \frac{1}{2} & \text{if } \pi_{RI} \in (0, \underline{\pi}] \\ \pi_{RI}q^2 + q - \frac{1}{2}q^2 - \pi_{RI}q + \frac{1}{2}\pi_{RI} & \text{if } \pi_{RI} \in (\underline{\pi}, \frac{1}{2}) \\ \frac{1}{2}q + \frac{1}{4} & \text{if } \pi_{RI} = \frac{1}{2} \\ \pi_{RI}q - \pi_{RI}q^2 + \frac{1}{2}q^2 + \frac{1}{2}\pi_{RI} & \text{if } \pi_{RI} \in (\frac{1}{2}, \overline{\pi}) \\ \pi_{RI} & \text{if } \pi_{RI} \in [\overline{\pi}, 1) \end{cases}$$

I can now calculate the value of interest, S. The values above are used to calculate $S \equiv P(v_R = V|primary) - P(v_R = V|leadership)$, remembering that $P(v_R = V|leadership) = \pi_{RI}$. With some algebra and noting the continuity of S at $\pi_{RI} = \underline{\pi}$, $\pi_{RI} = \frac{1}{2}$ and $\pi_{RI} = \overline{\pi}$, we find that

$$S = \begin{cases} \frac{1}{2} - \pi_{RI} & \text{for } \pi_{RI} \in (0, \underline{\pi}] \\ \pi_{RI}q^2 - \pi_{RI}q - \frac{1}{2}q^2 - \frac{1}{2}\pi_{RI} + q & \text{for } \pi_{RI} \in [\underline{\pi}, \frac{1}{2}] \\ -\pi_{RI}q^2 + \pi_{RI}q + \frac{1}{2}q^2 - \frac{1}{2}\pi_{RI} & \text{for } \pi_{RI} \in [\frac{1}{2}, \overline{\pi}] \\ 0 & \text{for } \pi_{RI} \in [\overline{\pi}, 1) \end{cases}$$

which are the values we were looking for.

Now we need to analyze the sign of S. If $\pi_{RI} \in (0, \underline{\pi}]$ we have that $S = \frac{1}{2} - \pi_{RI} > 0 \Leftrightarrow \pi_{RI} < \frac{1}{2}$, but that is satisfied because $\pi_{RI} \leq \underline{\pi}$ and I have already noted that $\underline{\pi} < \frac{1}{2}$. If $\pi_{RI} \in [\underline{\pi}, \frac{1}{2}]$ we have that $S = \pi_{RI}q^2 - \pi_{RI}q - \frac{1}{2}q^2 - \frac{1}{2}\pi_{RI} + q > 0 \Leftrightarrow \pi_{RI} < \frac{2q-q^2}{1+2q-2q^2}$ (noting that $1 + 2q - 2q^2 > 0$) which is satisfied because $\frac{1}{2} < \frac{2q-q^2}{1+2q-2q^2}$. If $\pi_{RI} \in [\frac{1}{2}, \overline{\pi})$ we have that $S = -\pi_{RI}q^2 + \pi_{RI}q + \frac{1}{2}q^2 - \frac{1}{2}\pi_{RI} > 0 \Leftrightarrow \pi_{RI} < \frac{q^2}{1-2q+2q^2}$ which is satisfied because $\overline{\pi} = \frac{q^2}{1-2q+2q^2}$. And finally if $\pi_{RI} \in [\overline{\pi}, 1)$ we have $S = 0$. So we have indeed $S > 0$ for $\pi_{RI} \in (0, \underline{\pi}] \cup [\underline{\pi}, \frac{1}{2}] \cup [\frac{1}{2}, \overline{\pi})$ and $S = 0$ for $\pi_{RI} \in [\overline{\pi}, 1)$, as the lemma claims.

A.5 Proof of Lemma 3

I calculate the differential of S with respect to π_{RI} and check its sign. If $\pi_{RI} \in (0, \underline{\pi})$, $\frac{\partial S}{\partial \pi_{RI}} = -1$ which is strictly negative. If $\pi_{RI} \in (\underline{\pi}, \frac{1}{2})$, $\frac{\partial S}{\partial \pi_{RI}} = q^2 - q - \frac{1}{2}$ which is strictly negative for $q \in (\frac{1}{2}, 1)$. If $\pi_{RI} \in (\frac{1}{2}, \overline{\pi})$, $\frac{\partial S}{\partial \pi_{RI}} = -q^2 + 2q - 1$ which is strictly negative for $q \in (\frac{1}{2}, 1)$. So S is decreasing with π_{RI} in all those intervals. S is non-differentiable at $\pi_{RI} = \underline{\pi}$ and $\pi_{RI} = \frac{1}{2}$, but is continuous at both points, and is therefore decreasing just like their neighboring points. Hence S decreases with π_{RI} when $\pi_{RI} \in (0, \underline{\pi}) \cup \{\underline{\pi}\} \cup (\underline{\pi}, \frac{1}{2}) \cup \{\frac{1}{2}\} \cup (\frac{1}{2}, \overline{\pi})$.

If $\pi_{RI} \in [\overline{\pi}, 1)$, S is constant for all values of π_{RI} (and equal to zero), so an increase in π_{RI} will not affect it.

A.6 Proof of Lemma 4

I calculate the differential of S with respect to q and check its sign, remembering that the values of $\underline{\pi}$ and $\overline{\pi}$ are $\underline{\pi} = \frac{(1-q)^2}{1-2q+2q^2}$ and $\overline{\pi} = \frac{q^2}{1-2q+2q^2}$. According to the values of S in Theorem 1, if $\pi \in (0, \underline{\pi})$, $\frac{\partial S}{\partial q} = 0$; similarly if $\pi \in (\overline{\pi}, 1)$, $\frac{\partial S}{\partial q} = 0$. So in those intervals, S is unresponsive to marginal changes in q.

However, if $\pi \in (\underline{\pi}, \frac{1}{2})$, $\frac{\partial S}{\partial q} = 2\pi q - \pi + 1 - q$ which is strictly positive; if $\pi = \frac{1}{2}$, $\frac{\partial S}{\partial q} = \frac{1}{2}$ which is strictly positive; if $\pi \in (\frac{1}{2}, \overline{\pi})$, $\frac{\partial S}{\partial q} = -2\pi q + \pi + q$ which is strictly positive. So in those intervals, S is strictly increasing with marginal increases in q.

To analyze the cases where $\pi = \underline{\pi}$ and $\pi = \overline{\pi}$, note that $\frac{\partial}{\partial q}(\frac{(1-q)^2}{1-2q+2q^2}) < 0$, so with a marginal increase in q, $\underline{\pi}$ remains in the interval $[\frac{(1-q)^2}{1-2q+2q^2}, \frac{1}{2}]$ where I just proved that S is increasing with q. Similarly note that $\frac{\partial}{\partial q}(\frac{q^2}{1-2q+2q^2}) > 0$, so with a marginal increase in q, $\overline{\pi}$ remains in the interval $[\frac{1}{2}, \frac{q^2}{1-2q+2q^2}]$ where I just proved that S is increasing with q.

To summarize, S is unresponsive to marginal changes in q for $\pi \in (0, \underline{\pi}) \cup (\overline{\pi}, 1)$, and is strictly increasing with q for $\pi \in \{\underline{\pi}\} \cup (\underline{\pi}, \frac{1}{2}) \cup \{\frac{1}{2}\} \cup (\frac{1}{2}, \overline{\pi}) \cup \{\overline{\pi}\}$.

A.7 Proof of Lemma 5

See the proof of Lemma 1 in Serra (2011).

A.8 Proof of Theorem 3

See the proof of Theorem 2 in Serra (2011).

A.9 Proof of Theorem 4

For points 1, 2, 5, 6 and 7, see the proof of points 1, 2, 6, 7 and 8 of Theorem 4 in Serra (2011), respectively.

To study the effect of q (point 3 in the theorem), we note that it only has an indirect effect on T through its effect on S. I proved in Lemma 5 that q has a strictly positive effect on S whenever for $\pi_{RI} \in [\underline{\pi}, \overline{\pi}]$. And I have proved (in point 2 of the theorem) that S has a strictly positive effect on T. Therefore, combining both partial derivatives, I prove that q has a strictly positive effect on T whenever for $\pi_{RI} \in [\underline{\pi}, \overline{\pi}]$.

To study the effect of π_{RI} we must note that it has two effects on T: a direct effect, and an indirect effect through its effect on S. In total, we have that $\frac{dT}{d\pi_{RI}} = \frac{\partial T}{\partial \pi_{RI}} + \frac{\partial T}{\partial S}\frac{\partial S}{\partial \pi_{RI}}$. It is easy to calculate that $\frac{\partial T}{\partial \pi_{RI}} = \frac{-S[X_{RE}(1-\pi_L)-X_L\pi_L]}{(1-\pi_L)(\pi_{RI}+S)^2}$ which is strictly negative. On the other hand I just calculated that $\frac{\partial T}{\partial S}$ is strictly positive, and we know from Lemma 4 that $\frac{\partial S}{\partial \pi_{RI}}$ is non-positive. We therefore have that $\frac{dT}{d\pi_{RI}} < 0$ and T is strictly decreasing with π_{RI}.

A.10 *Proof of Theorem 5*

Note from Lemma 2 that $S = 0$ when $\pi_{RI} \in [\overline{\pi}, 1)$. And remember that $T \equiv \frac{S[X_{RE}(1-\pi_L)-X_L\pi_L]}{(1-\pi_L)(\pi_{RI}+S)}$. Hence, when $\pi_{RI} \in [\overline{\pi}, 1)$ we have that $T = 0$ for any value of the other parameters.

References

Adams J, Merrill S III (2008) Candidate and party strategies in two-stage elections beginning with a primary. Am J Polit Sci 52(2):344–359

Adams J, Merrill S III, Stone WJ, Simas EN (2009) When candidates value good government: a spatial model with applications to congressional elections. Manuscript

Ashworth S, de Mesquita EB (2009) Elections with platform and valence competition. Games Econ Behav 67(1):191–216

Austen-Smith D, Banks JS (1996) Information aggregation, rationality, and the Condorcet jury theorem. Am Polit Sci Rev 90(1):34–45. 1996

Caillaud B, Tirole J (2002) Parties as political intermediaries. Q J Econ 117(4):1453–1489

Callander S (2008) Political motivations. Rev Econ Stud 75:671–697

Calvert RL (1985) Robustness of the multidimensional voting model: candidate motivations, uncertainty, and convergence. Am J Polit Sci 29:69–95

Carrillo J, Castanheira M (2008) Information and strategic political polarisation. Econ J 118(530):845–874

Castanheira M, Crutzen BSY, Sahuguet N (2010) Party organization and electoral competition. J Law Econ Organ 26(2):212–242

Cho S-J, Kang I (2008) Open primaries and crossover voting. Paper presented at the annual national conference of the Midwest Political Science Association

Condorcet M de [1785](1994) Essai sur l'application de l'analyse a la probabilite des decisions rendues a la pluralite des voix. McLean I and Hewitt F (trans) Paris

Ezra M (2001) The benefits and burdens of congressional primary elections. In: Galderisi PF, Ezra M, Lyons M (eds) Congressional primaries and the politics of representation. Rowman & Littlefield, New York

Feddersen T, Pesendorfer W (1998) Convicting the innocent: the inferiority of unanimous jury verdicts under strategic voting. Am Polit Sci Rev 92(1):23–35. 1998

Jackson MO, Mathevet L, Mattes K (2007) Nomination processes and policy outcomes. Q J Polit Sci 2(1):67–94

Kselman D (2012) Median-activists or median-voters? A spatial model with intra-party competition. Manuscript, Juan March Institute

Meirowitz A (2005) Informational party primaries and strategic ambiguity. J Theor Polit 17(1):107–136

Meirowitz A (2008) Electoral contests, incumbency advantages, and campaign finance. J Polit 70(03):681–699

Owen G, Grofman B (2006) Two-stage electoral competition in two-party contests: persistent divergence of party positions. Soc Choice Welf 26(3):547. Heidelberg

Roemer JE (2001) Political competition: theory and applications. Harvard University Press, Cambridge

Schofield N (2007) The mean voter theorem: necessary and sufficient conditions for convergent equilibrium. Rev Econ Stud 74:965–980

Schofield N, Claassen C, Ozdemir U, Schnidman E, Zakharov AV (2008) Party positioning under proportional representation and plurality rule. Center in Political Economy, Washington University. Typescript

Schofield N, Sened I (2005) Multiparty competition in Israel, 1988–1996. Br J Polit Sci 35:635–663

Serra G (2010) Polarization of what? A model of elections with endogenous valence. J Polit 72(2):426–437

Serra G (2011) Why primaries? The party's trade-off between policy and valence. J Theor Polit 23(1):21–51

Snyder JM Jr, Ting MM (2011) Electoral selection with parties and primaries. Am J Polit Sci 55(4):781–795

Wittman D (1973) Parties as utility maximizers. Am Polit Sci Rev 67:490–498

Measuring the Latent Quality of Precedent: Scoring Vertices in a Network

John W. Patty, Elizabeth Maggie Penn, and Keith E. Schnakenberg

Examples of network data in political science are ubiquitous, and include records of legislative co-sponsorship, alliances between countries, social relationships, and judicial citations.[1] Numerical estimates of the *influence* of each node (*e.g.* legislator, country, citizen, opinion), defined in terms of its propensity to form a relationship with another node, are often of interest to an analyst in each of these examples. In this chapter we present a new approach to solving a common problem in the social sciences—that of estimating the influence of vertices in a network. Our approach assumes that observed levels of influence relate to an underlying latent "quality" of the vertices.[2] Although common methods for measuring influence in networks assume that each vertex has the potential to influence every other vertex, many networks reflect temporal, spatial, or other practical constraints that make this assumption implausible. We present a scoring method that is appropriate for measuring influ-

[1] The networks literature in political science is large and growing. Recent comprehensive reviews include Lazer (2011) and Ward et al. (2011). In addition, Fowler et al. (2011) summarize and discuss methodological issues with inference of causality in networks.

[2] The word "quality" is simply a placeholder, though one that is roughly descriptive (at least in common parlance) of the characteristic that our method is estimating. While one might be precise and use a term such as "citability," we note the traditional issues of scope and space constraints and, setting this larger issue to the side, default to the use of a real word to refer to the latent construct our method is attempting to detect and estimate.

This research was supported by NIH Grant # 1RC4LM010958-01.

J.W. Patty (✉) · E.M. Penn · K.E. Schnakenberg
Center in Political Economy, Washington University in Saint Louis, 1 Brookings Drive, Saint Louis, MO 63130, USA
e-mail: jpatty@wustl.edu

E.M. Penn
e-mail: penn@wustl.edu

K.E. Schnakenberg
e-mail: keith.schnakenberg@gmail.com

N. Schofield et al. (eds.), *Advances in Political Economy*,
DOI 10.1007/978-3-642-35239-3_12,

ence in networks where (1) some vertices cannot form an edge with certain vertices for reasons that are unrelated to their underlying "quality" and (2) each vertex may be influenced by a different number of other vertices, so that some edges reveal different amounts of information about the latent "quality" of the influencing vertices.

As an example, we rate the "quality" of Supreme Court decisions, which we define as the likelihood that the decision will be cited in a future decision. These decisions are readily analyzed by our method due to their connectedness—the Supreme Court's explicit usage of previous decisions as precedent for current and future decisions generates a network structure. The network data enable us to assess some instances when a given decision "succeeded" (*i.e.*, was cited in a later opinion) or "failed" (*i.e.*, was not cited in a later opinion). However, because later decisions cannot be cited by earlier opinions, the data do not allow us to observe whether a given opinion *would have been cited* by an earlier opinion. Our network structure is necessarily incomplete.

The method we describe and employ in this chapter is intended to deal explicitly with this problem of incompleteness. The method, developed and explored in more detail by Schnakenberg and Penn (2012), is founded on a simple (axiomatic) theoretical model that identifies each opinion's latent quality in an (unobserved) world in which every object has the potential to succeed or fail. The theoretical model identifies the relative quality of the objects under consideration by presuming that the observed successes are generated in accordance with the independence of irrelevant alternatives (IIA) choice axiom as described by Luce (1958). In a nutshell, the power of this axiom for our purposes is the ability to generate scores for alternatives that are not directly compared in the data. Substantively, these scores locate all opinions on a common scale.

1 Inferring Quality from Network Data

We conceive of our data as a network in this chapter. Accordingly we first lay out some preliminaries and then discuss how one applies the method to general network data. We represent the observed network data by a graph denoted by $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is a set of n vertices and E is a set of directed edges, where for any $v, w \in V$, $(v, w) \in E$ indicates that there is an edge *from v to w*.[3] We define a *community* to be a subset of vertices, $C \subseteq V$, with a *community structure* $\mathcal{C} = (C_1, \ldots, C_n)$ being a set of subsets of V, and C_i being the community of vertex i.

Underlying our model is an assumption that each vertex j in a community C_i has the potential to influence vertex i. To define this formally, let $\tilde{E}$ be a set of *potential interactions*, with $E \subseteq \tilde{E}$. If $(i, j) \in E$ then we know that i and j interacted with j influencing i, and so it is known that they had the *potential* to interact: it is known that $j \in C_i$. On the other hand, of course, $(i, k) \notin E$ need not imply that i could not

[3] In general network settings, we interpret a connection from v to w as implying that w "influences" or "is greater than" v. What is key for our purposes is that the notion of influence be conceptually tied to the notion of quality, as we have discussed earlier.

have been connected to k. Rather, it may be the case that opinion that i could have been connected to k, but the link was not created for some reason (possibly because k was not of high enough quality to influence i, possibly because k and i never had an opportunity to interact, or for some other independent factor(s)). Our community structure is designed to accommodate this fact, and in particular we assume that $k \in C_i$ implies that $(i,k) \in \tilde{E}$. Thus, k being in community C_i implies that k had the potential to influence i (*i.e.*, i had the opportunity to link to k), regardless of whether k may or may not have succeeded (*i.e.*, regardless of whether an edge between i and k is observed).

The second assumption of our model is that each vertex can be placed on a common scale representing the vertex's quality. We assume that vertices with higher latent qualities are more likely to have had successful (*i.e.*, influential) interactions with vertices that they had the potential to interact with. Thus, the higher latent quality of vertex i, the more likely that, for any given vertex $j \in V$, $(j,i) \in \tilde{E}$ implies that $(j,i) \in E$.

Our goal is to estimate each vertex's "latent quality" score subject to a network G and an observed or estimated community structure, $\mathcal{C}$. We conceive of our network and community structure as generating a collection of "contests" in which some vertices were influential, some had the potential to be influential but were not, and others had no potential to influence. These contests are represented by the set $\mathcal{S} = \{s \in V : (s,v) \in E \text{ for some } v \in V\}$. Thus, every vertex that was influenced represents the outcome of a contest.

Let $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$ represent each vertex's latent quality. Then for each $i \in \mathcal{S}$ we let the expected influence of vertex k in contest i (*i.e.*, probability of i connecting to k), which we denote by $E(i,k)$, equal 0 if $(i,k) \notin \tilde{E}$. Thus, k's expected influence in contest i is zero because in this opinion we assume that $k \notin C_i$, and thus k had no potential to influence i (*i.e.*, there is no chance that i will connect to k). Otherwise,

$$E(i,k) = \frac{x_k}{\sum_{j \in C_i} x_j}.$$

In words, the expected share of influence of k in a contest in which k has the potential to influence i is k's share of latent influence relative to the total latent influence of the vertices that can potentially influence i.

Similarly, we can calculate the share of *actual* influence of k in i, or $A(i,k)$, by looking at the total set of vertices that actually influenced i in the network described by G. This set is $W_i = \{w : (i,w) \in E\} \subseteq C_i$, and (without any additional information such as edge weights), k's share is $\frac{1}{|W_i|}$ if $k \in W_i$ and 0 otherwise. We can now utilize our network and community structure to estimate x subject to an unbiasedness constraint that is conditional on the community structure. The constraint is that

$$\sum_{s \in \mathcal{S}} E(s,i) = \sum_{s \in \mathcal{S}} A(s,i) \qquad \text{for all } i,$$

or that each vertex's total actual score equals their total expected score. Satisfaction of this constraint implies, given a correct community structure, that no vertex is estimated to be more or less influential than it actually was. Schnakenberg and Penn (2012) prove that, subject to a minimal connectedness condition, there exists a vector $x^* = (x_1^*, \dots, x_n^*)$ that solves the above system of equations and that is unique up to scalar multiplication.[4] Viewed substantively, this vector represents the relative qualities/influences of the different nodes. In particular, as x^* is uniquely identified up to scalar multiplication, the ratio of any two nodes' qualities,

$$\rho_j^i \equiv \frac{x_i}{x_j},$$

is uniquely identified. This ratio ρ_j^i represents the hypothetical relative frequency of selection/influence by node i versus that by node j in a future contest in which both nodes i and j compete (*i.e.*, for any future node that both i and j have the ability to exert influence on).

2 Measuring the Quality of Precedent

The use of judicial precedent by Supreme Court Justices—and, in particular, a focus on citations as an indication of this usage—has attracted sustained attention from legal and political science scholars for over 60 years.[5] Unsurprisingly, given the breadth of the topic, scholars have adopted various approaches to the study of precedent, but most have focused on the determinants of citation: in a nutshell, what factor or factors of an opinion augur revisitation of the opinion in future opinions?

Because our model imputes unobserved relationships between objects, it is particularly well-suited to analyzing networks in which certain links are impossible to observe. These types of networks could, for example, arise in situations in which vertices are indexed by time and a later vertex is incapable of influencing a vertex that preceded it.

We utilize a data set consisting of the collection of citations by United States Supreme Court majority opinions to Supreme Court majority opinions from 1791 to 2002. Thus, viewed in the theoretical framework presented above in Sect. 1, the vertices of our network are Supreme Court majority opinions, and if majority opinion i cites majority opinion j, we include the edge $(i, j) \in E$.

Before moving on, it is important to note what we are explicitly abstracting from in our operationalization of the judicial citation/precedent network. Most importantly, we omit consideration of all opinions other than the majority opinion. Both

[4] For reasons of space, we refer the interested reader to Schnakenberg and Penn (2012) for more details on the method.

[5] Seminal offerings include Merryman (1954) and Landes and Posner (1976), while more recent, book-length analyses include Hansford and Spriggs II (2006) and Gerhardt (2008). Other relevant citations are provided where appropriate in our discussion.

dissenting and concurring opinions are relevant for understanding both the bargaining processes at work in constructing the majority opinion and inferring the role and quality of precedent (*e.g.*, Carrubba et al. (2011)).[6] In addition, our approach ignores the citing opinion's treatment of the cited opinion (*e.g.*, favorable, critical, or distinguishing).[7,8] We leave each of these for future work.

Differentiating Cases: Community Structure As discussed earlier, the method we employ allows us to compare/score objects that have not been directly compared. Accordingly, it offers an analyst the freedom to "break up" the data in the sense of estimating (or, perhaps, observing) communities of objects that are less likely to be directly compared with one another. For the purposes of this chapter, we take into account only the temporal bias discussed earlier—later opinions cannot be cited by earlier opinions—and presume that each opinion is eligible (*i.e.*, "in competition") for citation by every subsequently rendered opinion.[9]

Thus we construct the community C_i for a given opinion i as follows. Letting $\text{Year}(i)$ be the year in which opinion i was heard, we assume that for any pair of vertices (*i.e.*, majority opinions), i, j,

$$\text{Year}(i) > \text{Year}(j) \quad \Leftrightarrow \quad j \in C_i.$$

In words, an opinion can be influenced by any and only opinions that strictly predate it.

Data We apply our method to Fowler and Jeon's Supreme Court majority opinion citation data (Fowler et al. (2007), Fowler and Jeon (2008)). There are a number of ways one might approach this data when considering the question of the quality or influence of each opinion. The most straightforward approach would rank all of the opinions that have been cited at least once (any opinion that is not cited by any other opinion in the database cannot be ranked). In this approach, every opinion is a *contest*, and each opinion that is cited at least once is a *contestant*.

Practical constraints prohibit us from ranking all of the opinions. Fortunately, our approach implies that we can examine any subset of the data and recover relative rankings that are (in theory) identical to the rankings that would be estimated from

---

[6] In addition, there are many interesting theoretical and empirical questions regarding how one should conceive of the relationship between opinions and opinions (*e.g.*, Bommarito et al. (2009)) that the data we employ here do not allow us to explore more fully.

[7] Practically speaking, there are a number of ways that scholars have developed and employed to consider this aspect of how Justices cite earlier opinions. For recent examples, see Clark and Lauderdale (2010), Spriggs II et al. (2011).

[8] We are not aware of any recent work that has differentiated citations by the number of times the citation occurs in the citing opinion.

[9] Note that, for simplicity, we approximate this "later than" relation in the sense that we presume (unrealistically) that, in any year, the Court cannot cite one opinion that is decided in that year in another opinion that is decided in that same year. Given the number of years that we consider, this approximation affects a *very* small proportion of the number of potential citations we consider.

the entire data set. Accordingly, we restrict our attention to the 100 most frequently cited opinions between 1946 and 2002. In graph theoretic terms, we examine the smallest subgraph containing all edges beginning or ending (or both) with an opinion whose *in degree* (number of times cited) ranks among the top 100 among the opinions rendered between 1946 and 2002. This graph contains many more than 100 opinions (3674, to be exact). After these opinions, and their incident edges, are selected, they are then used for our community detection algorithm, which we now describe.

Using the years of the opinions to create the communities as described earlier, we then solve for the influence scores of the opinions (*i.e.*, contestants) as follows. First, we choose the contestants in turn and, for each majority opinion (*i.e.*, contest) that was subsequent to an opinion and cited at least one member of the contestant's community, we count the contestant as having been participant (*i.e.*, available for citation) in that majority opinion/contest. If the contestant was cited in (*i.e.*, won) that contest, the contestant is awarded $1/|W|$ points, where W is the set of opinions (contestants) cited in that majority opinion (contest). Otherwise, the contestant is awarded 0 points in that contest. With this vector of scores for each contestant in each contest, it is then possible to directly apply the method developed by Schnakenberg and Penn (2012) to generate the latent influence scores of each majority opinion, $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_n)$.

These latent influence scores represent, in essence, the appeal of each majority opinion as a potential citation in any subsequent majority opinion. What this appeal represents in substantive terms is not unambiguous, of course. It might proxy for the degree to which the opinion is easily understood, the degree to which its conclusions are broadly applicable,[10] or perhaps the likelihood that the policy implications of the opinion support policies that are supported by a majority of justices in a typical opinion. Obviously, further study is necessary before offering a conclusion on the micro-level foundations of these scores. Such research will require inclusion of observed and estimated covariates distinguishing the various opinions and majority opinions.

3 Results

We now present the results of three related analyses. We first present our results for the 100 most-cited opinions rendered between 1946 and 2002.[11] Following that, we present the results for the 100 most-cited opinions since 1800.[12] Finally, we consider the 204 most-cited opinions since 1800 with an eye toward comparing the ranking

[10] Note that this is true *despite* the presumption that an opinion might have been feasible only in a subset of observed and subsequent majority opinions.

[11] This time period includes all cases in the Fowler and Jeon data for which Spaeth's rich descriptive data (Spaeth 2012) are also available.

[12] This time period includes all cases in the Fowler and Jeon data.

Table 1 Descriptive correlations with scores. Sample: Top 100 most-cited cases since 1946

Correlation (Age, Score):	−0.461
Correlation (# Cites, Score):	0.496
Correlation (# Cite/Year, Score):	0.787

of the 100 most-cited opinions since 1946 with the ranking of those cases when all opinions that have been cited at least as many times as these 100 are considered.

3.1 *Top 100 Opinions Since 1946*

Table 2 presents the opinions with the top 36 estimated latent quality scores for this period. This is the set of opinions for which the estimated quality score is greater than 1, which is by construction the average estimated quality score for the 100 cases.

This ranking is interesting in a number of ways. The top two majority opinions score significantly higher than all of the others.[13] The top-scoring opinion, *Chevron*, is a well-known case in administrative law with broad implications for the judicial review of bureaucratic decision-making. The second-ranked opinion, *Gregg*, clarified the constitutionality of the death penalty in the United States. Of course, the third highest scoring opinion is the famous *Miranda* decision in which the Court clarified the procedural rights of detained individuals.

Space prevents us from a full-throated treatment of the scores, but a few simple correlations are of interest. Table 1 presents three Pearson correlation coefficients relating the opinions' scores with, respectively, the age of the opinion, the number of subsequent opinions citing the opinion, and the number of subsequent opinions citing the opinion divided by the age of the opinion.

The negative correlation between the age of an opinion and its score is broadly in line with previous work on the depreciation of the precedential value (or, at least, usage) of judicial opinions.[14] It is important to note, however, that this effect is *potentially* at odds with the IIA axiom on which the scoring algorithm is based. We partially return to this question below when we expand the sample of opinions.

That the correlation between the opinions' scores and the number of times each opinion has been cited by a subsequent Supreme Court majority opinion is positive is not surprising: the score of an opinion is obviously positively responsive to the number of times that an opinion has been cited, *ceteris paribus*. Accordingly, the interesting aspect of the correlation is not that it is positive but, rather, that it is not closer to 1. Indeed, inspection of Table 2 indicates, *a fortiori*, that the rank-

[13] Note that the estimated scores for the top 100 opinions sum to 100, so these two opinions account for over 1/8th of the sum of the estimated scores. In other words, any opinion that cites exactly one of these 100 cases is predicted to cite either *Chevron* or *Gregg* almost 13 % of the time.

[14] See, for example, Black and Spriggs II (2010).