2

SOME IDEAS ON HOW POLITICAL SCIENTISTS CAN DEVELOP REAL-WORLD IMPLICATIONS FROM THEIR RESEARCH (WITHOUT BECOMING POLICY WONKS OR LAW PROFESSORS)

Lee Epstein, Jack Knight, and Andrew D. Martin

The behaviorally minded student of politics is prepared to describe values as empirical data; but, qua “scientist” he seeks to avoid prescription or inquiry into the grounds on which judgments of value can properly be made.

—Robert A. Dahl (1961)

Political scientists and law professors will always harbor different methodological orientations—political scientists, frankly, have higher standards both for modeling and empirical testing, while law professors are more preoccupied with interpreting legal texts and providing normative recommendations.

—Eric Posner and Alan Sykes (2013b)

These quotes tell it like it is. A law review article without a prescriptive component is an oxymoron; a social science paper with a normative conclusion is just moronic.

How we got to this sorry state is not just a mystery to us; it is the height of irony. Implications for the real world—whether normative, policy, design, or otherwise—should carry more weight when we derive them through a process that uses models and data. That is because statistical inference enables us (and our readers) to gauge the uncertainty (or, if you prefer, the reliability) of our conclusions with some degree of precision. In most other kinds of research, including many articles in the law reviews, scholars base their prescriptions on little more than speculation, intuition, and sheer guesswork.

The question is how we can convince (quantitative) political scientists interested in law and legal institutions to be more attentive to the real world. The answer, we believe, does not lie in asking them to reorient their questions to address current policy concerns; most will not do it. The real career payoffs in political science
Developing Real-World Implications

15

come not from asking narrow questions (Did law X have a + or – effect? Did the
framers mean A or B or C? Was case 1 or 2 correctly decided?); they follow from
theoretical and methodological breakthroughs. If, and only if, political scientists
think they can make contributions of this sort will they undertake research with
the potential to influence law and public policy.

We see two paths here, neither of which is especially novel. One is for us to take
the extra step and develop the real-world implications of our research. Another is to
assess rigorously the speculation so rampant in the law reviews and policy reports.

In what follows we develop both. Along the way, we gesture toward the larger
project of developing a methodology for generating and evaluating policy impli-
cations. Much more than hand wave we cannot do in this short chapter. But we
do hope others will pursue this project in the not-so-distant future. It would repre-
sent a real advance relative to the easy fixes we explore here.

Developing the Real-World Implications
of Political Science Research

There is almost no study in political science lacking in real-world implications. If
you need convincing, turn to any law review article that makes use of our research
and you will find plenty. It is just that we political scientists have failed to develop
them ourselves.

We should—though we admit this is a debatable proposition. Law professors
have done a much better job than political scientists in communicating the scien-
tific literature on judging to relevant legal and policy actors. We would go so far as
to say that our research got zero real-world traction until legal academics got into
the game. Only now, and not 70 years ago when we started this enterprise called
“judicial behavior,” are members of Congress and advocacy groups referencing
our studies, are journalists writing about them, and do judges want to learn more
about them (if only to tell us we are wrong!).

We really should not whine too much. Dissemination of our work is a good
thing regardless of who is doing the disseminating. On the other hand, ceding to
law professors has its downsides. Sometimes legal academics do not characterize
our findings quite right (read: they adapt our results to suit their own purposes)—
no surprise since many are lawyers searching for material to bolster their arguments
and not social scientists trained to defeat their hypotheses. Then there is the matter
of credit. It does not help our discipline when we get little or no acclaim for impor-
tant breakthroughs. Just ask the political scientists at the National Science Foun-
dation (NSF). But do not bother with the seemingly clueless American Political
Science Association (APSA). A worse professional association is hard to imagine.

If you agree with us, then there is one simple fix. Instead of concluding our
articles with the (n)ever illuminating call for “more research,” how about a section
that contemplates the real-world implications of our findings? This recommenda-
tion should not be altogether unfamiliar to political scientists. For over a decade,
NSF has considered the potential “broader impacts” of proposals when deciding whether to fund them. 

Sometimes, perhaps often, the “impacts” or implications will center on institutional arrangements; others might focus on improving public policy and still others on doctrine. Whatever, developing them should not be hard because often it involves little more than thinking through the findings, although more systematic approaches are possible (and perhaps preferable) too. Here are a few examples from our field: judicial behavior.

**Gender**

Scores of studies have analyzed whether male and female judges decide cases distinctly—“individual effects”—and whether serving with a female judge causes male judges to behave differently—“panel” or “peer” effects. In our work, we found consistent gender effects in only one of the 13 areas we considered: sex discrimination in employment under Title VII (Boyd, Epstein, and Martin 2010).

Following standard operating procedures in political science, we conclude this article with—what else?!—a call for more research: “While we hope our study goes some distance toward answering important questions in the literature, we also think that the very questions we addressed here continue to deserve a prominent place on the scholarly agenda” (Boyd, Epstein, and Martin 2010, 407).

But imagine a different approach:

Our research finds that male judges are more likely to vote for the plaintiff in Title VII gender-based employment cases when they sit with a female judge. Why? We are not sure. Mechanisms consistent with our results include (1) learning/updating (the males continue to be more pro-plaintiff after serving on a mixed-gender panel), (2) suppressing (the males are only pro-plaintiff when sitting on a mixed panel), and (3) framing/deliberating (the males are only pro-plaintiff when sitting on a mixed panel and the case is orally argued so that the female judge can push the arguments her way). The failure to identify which is doing the work is a hole in our study because without a mechanism, we can offer no silver bullet but only some possibilities for institutional change. To see the problem suppose we desired more plaintiff-friendly outcomes. The mechanisms would point to different recommendations:

- **Learning/Updating**→Occasionally abandon random assignment; all males must sit on a mixed-gender panel at least once in Title VII cases.
- **Suppressing**→Appoint more female justices to maximize the likelihood of a mixed panel in Title VII cases.
- **Framing/Deliberating**→Mandate oral arguments in all Title VII cases when the panel is mixed.
Although many accounts of judicial behavior can accommodate different or even multiple motivations (e.g., Ferejohn and Weingast 1992; Knight and Epstein 1996), the conventional (political science) party line is that maximizing policy is of paramount, even exclusive, concern. Two of us (Epstein and Knight 2013) have written on the huge mistake of neglecting other motives. That is because research conducted by scholars (mostly outside of political science, but see Baum 1997, 2006) has demonstrated that the policy goal is hardly the only motivation; it may not even be dominant for many judges. The evidence is now so strong that it poses a serious challenge to the extremely (un)realist(ic) conception of judicial behavior that has dominated the study of law and legal institutions for generations.

Exploring these other motivations would lead not only to theoretical breakthroughs in the form of new conceptualizations of judicial behavior (a very good thing!) but also to potentially interesting real-world implications. Consider promotion. This would seem to be an important factor influencing the personal utility that judges gain from their work. It could be coincident with policy preferences: the higher judges sit in the hierarchy, the more important the cases they hear, and the greater the opportunity to influence the law. But we also know that promotion tends to increase job satisfaction, prestige and reputation, and salary.

And yet, many scholars tend to dismiss promotion as a driver for life-tenured judges. They say it is such a “random” (Cooter 1983, 1029) or low-probability event that it could not possibly influence judicial behavior.

We agree that promotion is not a relevant motivation for most U.S. Supreme Court justices. What of lower court judges? Though promotion rates are low, the odds for some judges are substantially higher because they possess certain markers (e.g., their age, pedigree, and even district/circuit) that place them in the genuine promotion pool, and they know it (Epstein, Landes, and Posner 2013a).

Perhaps this explains the growing number of empirical studies showing that these judges, the “auditioners,” seem motivated by the possibility of promotion to higher status (and paying) jobs, and will make choices based on this motivation. Cohen (1991) demonstrated, and Sisk, Heise, and Morriss (1998) confirm, that district court judges with a reasonable chance of promotion were more likely to uphold the politically popular federal sentencing guidelines. Along similar lines, Epstein, Landes, and Posner (2013a) show that auditioners, not wanting to appear soft on crime, impose harsher sentences on criminal defendants. Gaille (1997) finds that the number of articles published by court of appeals judges standing some chance of elevation to the Supreme Court dropped “precipitously” after the Bork hearings. “Rather than the selection process constituting a one-sided choice by the president
and Congress,” he concluded, “it appears that prospective nominees themselves make important behind-the-scenes decisions that may influence the president’s choice and, ultimately, determine the composition of the Court” (Gaille 1997, 376).

These empirical studies lend (some) support for promotion as a motivating force;17 much more work needs to be done (and would it not be nice if political scientists contributed). But what should we make of promotion from a normative standpoint? Some have argued that banning promotions would foster judicial independence. “Judges who know that they are unlikely to be promoted lack the incentive to curry favor with those who could appoint them to more powerful, more prestigious, or more lucrative posts” (Klerman 1999, 455). Others disagree, claiming that promotion provides judges with a crucial incentive “to work hard and judge wisely” (Klerman 1999, 456). This is a debate political scientists could and should want to enter.

Independent Courts

There is a flourishing literature on the role of courts in promoting economic growth. The central argument is that courts independent from the government can act as a check on the government’s efforts to interfere with contracts and property rights.18

You would think political scientists would be front and center in testing these and related hypotheses. You would be wrong. With only limited exceptions,19 the chief empirical players are economists (e.g., LaPorta et al. 2004; Glaeser and Shleifer 2002; LaPorta, de Salines, and Shleifer 2008) with a sprinkling of law professors (e.g., Klerman and Mahoney 2005; Mahoney 2001).

This is unfortunate in more ways than one. First, much of the work suffers from precisely the same flaw as the early separation-of-powers studies (done primarily by legislative specialists): an unsophisticated understanding of courts (and perhaps politics). Rios-Figueroa and Staton (2013) demonstrate as much in their interrogation of the various measures of judicial independence, but there are other problems too (see Harvey 2011). Second, this is a subject that not only fits with our substantive interests but is also an area where real theoretical and methodological breakthroughs are possible and the implications for the real world abundant. The World Bank thinks so: “The rule of law is a principle of fundamental importance to the World Bank. It lies at the heart of what the Bank is, what it does, and what it aspires to accomplish.”20

Election Effects

Judicial elections is one of the very few areas where political scientists have developed real-world implications from their work. But the implications diverge depending on the study. Work focusing on the elections themselves or on public
opinion comes out in favor of electing judges. In their study of more than 500 elections and retentions for seats on state high courts, Bonneau and Hall (2009) “throw empirical grenades” at the claim that voters are uninterested in judicial elections or unqualified to vote because they know so little about the candidates. Neither is true—especially not in “expensive and contentious” races (Bonneau and Hall 2009, 47). From the results of experiments embedded in surveys, Gibson argues that, on balance, elections benefit (or at least do not harm) judges because they increase the legitimacy of their courts relative to other systems of selection (Gibson 2013).

Studies on the judging end shore up the potential costs of elections. One strand shows that judges respond to their constituents by voting in ways that reflect their constituents’, not their own, preferences. Of special concern is research demonstrating that judges who face the electorate to retain their jobs are especially tough on criminal defendants, in part because the public prefers law-and-order types to softies (e.g., Gordon and Huber 2007; Huber and Gordon 2004; Hall 1987, 1992; Berdejo and Yuchtman 2013).

The second strand is the literature we just mentioned on the relationship between judicial independence and economic freedom or growth. La Porta et al. (2004) find that courts with substantial independence (measured in part by the length of tenure) and constitutional review power are more likely to exhibit higher levels of economic freedom (the security of property rights, the number of legal procedures to start a business, and the level of worker protection and government ownership of banks). Klerman and Mahoney (2005) show that 18th-century laws in England providing greater job security to judges increased the value of financial assets in England.

We draw attention to the literatures on elections/public opinion versus those on judging with the hope that both sides will engage each other more seriously than they now do. Work in one area sometimes does not even bother citing work in the other. That is not an ideal way to contribute to knowledge or public policy.

The two literatures also raise more general—and very difficult—questions about developing real-world implications when the evidence points in different directions. How might we proceed in this not atypical situation? We see two ways. First, we should drill down to the research design or methodological differences driving the contradictory results. So doing would allow policy makers and other researchers to decide which result is more plausible based on their judgment about which design or methodological assumptions are closer to their interests. By way of example: in studies of judicial selection, how the research classifies Ohio and Michigan drives conclusions about political participation (Nelson, Caufield, and Martin 2013). Second, sometimes contradictory results are about different phenomena—in our example, money spent in judicial elections increases participation (a good thing), but electing judges slants judicial decisions against criminal defendants (a bad thing). It is important to quantify “how much” increase or
decrease in the various outcomes under study so that policy makers can trade them off in a more precise manner.

Assessing the Implications in Law Studies

A second approach to producing research with real-world implications is to use our approaches and methods to test the guesses (call them hypotheses) in law review articles, think-tank studies, and judicial opinions. Examples in the last category come from participants at a conference called “Testing the Constitution.”

- Commandeering. In the Printz v. United States (521 U.S. 898, 1997) line of commandeering cases, the majority operates under empirical assumptions about accountability—notably that if commandeering occurs citizens are unable to know whether to hold the states or the feds accountable for the government action. True?
- Election Spending. In Caperton v. Massey Coal Co., (556 U.S. 868, 2009), which requires elected judges to recuse themselves from cases in which interested parties spent a lot of money to elect them, the Court wrote, “Our decision today addresses an extraordinary situation where the Constitution requires recusal. Massey and its amici predict that various adverse consequences will follow from recognizing a constitutional violation here—ranging from a flood of recusal motions to unnecessary interference with judicial elections.” Has this happened?
- Exclusionary Rule. In Fourth Amendment cases the Supreme Court’s decisions are full of claims about how excluding evidence will or will not deter police misconduct. Which side has the better case?
- Abortion. In the partial-birth abortion case Gonzales v. Carhart (550 U.S. 124, 2007), the Court wrote, “While we find no reliable data to measure the phenomenon, it seems unexceptionable to conclude some women come to regret their choice to abort the infant life they once created and sustained.” Really?

These are quite specific examples; others are more general. Here is one from our own research, also centering on constitutional law: whether or not political scientists are aware of it, constitutional law scholars are less inclined or interested in debating legalism/realism then they are in the scores of legalistic “theories”—intellectual systems or methodologies really—supposedly adopted on neutral grounds to generate objective decisions. The seemingly scores of these systems go by such names as “originalism,” “textualism,” “the Constitution in exile,” “the Constitution as common law,” “the living Constitution,” “active liberty,” etc. The list goes on and on.

Invariably, adherents claim that their own approach is best (or at least better than the alternatives). But what does “best” mean? This is where a dose of political
Developing Real-World Implications

Science would be informative, as well as have some real-world impact because we can put some flesh on “best.” There are yardsticks here. For example,

- **Endurance.** All else being equal, are subsequent courts or even Congress less likely to overturn certain kinds of decisions—originalist, textualist, pragmatic, and so on?
- **Efficacy.** Which types of decisions are courts, other institutions, and the people more likely to follow?
- **Impact.** Are some methods more successful at generating real social, political, and economic change?
- **Neutrality.** Proponents of the different methods argue that theirs produces more neutral outcomes. True?

We could go on, but you get the drift. Words like “best” are thrown around all the time in white papers and the law journals—the best system for deciding cases, the best method for selecting judges, the best legal rule, the best procedure for assigning panels. You name it; there is a best. But rarely are the bests assessed rigorously against the alternatives.

Projects of the sort we propose not only would use political science constructs but they could also lead to methodological advances. In our example of assessing systems of interpretation, we could take advantage of the enormous progress in processing and categorizing text and in modeling networks.

To “take advantage,” however, requires more than deploying canned solutions; it requires collaboration with the technically oriented in the (social) sciences. That is because methodological advances in text or network analysis will be motivated by unique problems that are not amenable to off-the-shelf methods. Read: progress requires both technical skills and the kind of unique and specialized knowledge we possess about law and legal institutions. Fruitful collaborations of this sort already have developed in other pockets of political science,26 and there is every reason to suspect they will happen in ours. After all, tools for categorizing text and modeling networks should be catnip to us law and court scholars—scholars whose life’s work centers on reading texts and understanding how they interrelate.

**Notes**

We are grateful to the Guggenheim Foundation, National Science Foundation, the University of Southern California, and Washington University in St. Louis for supporting our research on law and legal institutions. We also thank Brandon Bartels and Chris Bonneau for very helpful comments.

1 The quote above is from Posner and Sykes (2013b). See also, e.g., Posner and Sykes (2013a, n. 4): “Political scientists (like economists) are oriented to producing descriptive hypothesis . . . and testing them using statistical methods . . . Lawyers are oriented to normative argument.”) and Smits 2009 (“[W]hat is the core of legal scholarship? in
my view . . . [it] tries to answer the normative question what ought to be. It is this ‘pre-
scriptive’ motive that makes legal science distinct from other disciplines.”

2 Dahl’s quote appears in Dahl (1961). See also Gerring and Yensowitz (2006): “Tradi-
tionally, the scientific study of politics has been associated with a value neutral approach
to politics.”

3 Epstein and King (2002) and Epstein and Martin (2013) make the same point.

4 We trace the origins to Pritchett (1941, 1948). For some disciplinary history, see
Epstein, Landes, and Posner (2013a); Segal and Spaeth (2002); and Maveety (2003).

5 We are sure there are lots of cites here. But we know that since we have become law
professors (or collaborate with law professors) our work has gotten a lot more play. To
provide one example, an earlier version of Epstein, Landes, and Posner (2013b)—Epstein,
Landes, and Posner (2010)—was referenced in Congress (Barriers to Justice and Account-
ability: How the Supreme Court’s Recent Rulings Will Affect Corporate Behavior: Hearing Before
the S. Judiciary Comm., 112th Cong. 1–2 2011); covered in the New York Times (“Justice
Offer Receptive Ear to Business Interests,” Dec. 19, 2010, A1) and in the Economist
(“Corporations and the Court,” June 25, 2011, 75); and refuted by the Federalist Society
(www.fed-soc.org/doclib/20111216_NewhouseEngage12.3.pdf). We also understand
that at least one Supreme Court justice read it and agreed with the conclusions.

6 You will have to trust us here. One of us has already slammed the legal academy enough
for one lifetime! See Epstein and King (2002)—characterized recently as “excoriat[ing]
law reviews for publishing so much empirical work that violated basic rules of infer-
ence” (Ho and Kramer 2013, 1198).

7 The NSF Grant Proposal Guide suggests that “[t]he Broader Impacts criterion encom-
passes the potential to benefit society and contribute to the achievement of specific,

8 With some emphasis on our own work so that we embarrass only ourselves.

9 For citations, see the web appendix to our study at http://epstein.usc.edu/research/
genderjudging.html. There is also a large literature on the effect of race on judging.
See, e.g., Cox and Miles (2008); Farhang and Wawro (2004); and Welch, Combs, and

10 This follows from the lack of deliberations in conference. It is one thing to call the
lawyer a fool at oral argument and quite another to call your colleague one (see Epstein,
Landes, and Posner 2013a). It is also possible that the attorneys orient their arguments
toward the female judge.

11 The recommendations would be the opposite if we desired more pro-defendant
outcomes.

12 We adapted some of this material in this section from Epstein and Knight (2013) and

13 As well as by perpetuating the pedantic, tiresome, and ultimately phony debate over law
(legalism) v. politics (realism).

14 See, e.g., Schauer (2000) and Segal and Spaeth (2002). Arthur Goldberg, in 1965, was
the last to leave for “higher” office—and his departure was not voluntary. Of course,
some associates might hope to become Chief Justice (Frank 1970), but that is truly a
rare event.

15 Eleven percent of all district court and 4 percent of all court of appeals judges have been
elevated.

16 In a separate study, Cohen (1992) found that judges with promotion potential gave
higher antitrust fines. See also Taha (2004).
The studies we mention focus on the U.S. federal courts. Promotion (and retention) has also figured into work on courts abroad. Salzberger and Fenn (1999) show that the lower the reversal rate, the higher the judge’s prestige and thus the higher the likelihood of promotion from a court of appeals to the House of Lords in England. Ramseyer and Rasmusen (1997, 2001) have written several papers on the relationship between judicial careers and judicial decisions in Japan. The upshot is that that Japanese judges who are reversed receive less prestigious responsibilities. For a study on German judges, see Schneider (2005).

Some of the literature relates judicial independence to common law systems; some of it studies independence through proxies, such as judicial tenure. Harvey (2011) and in part Rios-Figueroa and Staton (2013). Also North and Weingast (1989) were among the first to theorize on this connection (see also Hayek 1960). Weingast, though an economist by training, is a political scientist in practice.

We adopt some of the material in this section from Epstein (2013).


Held in October 2014 at the University of Chicago Law School. The idea was to bring together doctrinal constitutional law scholars and methodologists with the goal of testing empirical assertions in the Supreme Court’s constitutional law decisions. The examples below come from Barry Friedman, Adam Liptak, and Geoffrey Stone.

In this case, over $3,000,000.

Adapted from Epstein, Landes, and Posner (2013a, 2).

For example, Quinn et al. (2010) develop a dynamic topic model to understand political agendas using data from over 100,000 speeches in Congress. The coupling of substantive political scientists with methodologists and computer scientists resulted in the creation of a new method to trace the evolution of topics over time.

References


Developing Real-World Implications


PART II
Law and Decisionmaking