# LOOKING OVER A CROWD— DO MORE INTERPRETIVE SOURCES MEAN MORE DISCRETION?

ADAM M. SAMAHA*

*Observers have suggested that adding sources of interpretation tends to increase interpreter discretion. The idea is embedded in a quip, attributed to Judge Harold Leventhal, that citing legislative history is like "looking over a crowd and picking out your friends." Participants in debates over interpretive method have applied the idea to the proliferation of other sources as well, including canons of construction and originalist history. But the logic of "more sources, more discretion" has escaped serious testing. And predicting the effect of source proliferation is not a matter of logic alone. The empirical study of how information loads affect behavior has grown dramatically in recent decades, though almost without notice in legal scholarship on interpretive method.*

*This Article tests the logic and evidence for "more sources, more discretion." The idea turns out to be incorrect, without more, as a matter of logic. Adding sources tends to reduce the chance of discretion using a simple model of interpretation. This starter model depicts judges as aggregators of source implications, and it draws on basic probability theory and computer simulations to illustrate. The analysis does change if we allow judges to "spin" or "cherry pick" sources, but without much hope for limiting discretion by limiting sources. Of course, judges will not always behave like machines executing instructions or otherwise follow the logic of these models. Thus the Article goes on to spotlight provocative empirical studies of information-load effects, develop working theories of interpreter behavior, and present new evidence.*

*After emphasizing that interpreters might ignore additional information at some point, the Article tests three other theories. First, an extended dataset casts doubt on an earlier study that linked a growing stock of precedents to increased judicial discretion. Adding to the pile of precedents seems to have no simple pattern of effect on discretion. Second, existing studies indicate that increasing information loads might prompt judges to promote the status quo, and new data suggest that this effect*

*depends on the type of information added. The number of sources cited in appel-lant briefs appears to have no effect on judges' willingness to affirm—in contrast with the number of words and issues presented, which may have opposing effects. Third, an expanded dataset supports an earlier finding that judges who face a large number of doctrinal factors might weight those factors in a quasi-legal fashion. This time-saving prioritization does not seem to follow conventional ideological lines.*

*With simple intuitions in doubt, thoughtful work remains to be done on the effects of source proliferation. Observers interested in judicial discretion have good reason to look beyond source proliferation to find it. And observers interested in institutional design have good reason to rethink the range of consequences when informa-tion is added to our judicial systems.*

## Introduction

Every decision is entangled with a choice, made thoughtfully or not, about how much information to consider. Recognition of this unavoidable choice has led to questions about what exactly happens when people are asked to process more information rather than less. In fact, data-saturated environments have inspired more than a gener-ation of researchers to study the processing of information flows when ordinary people, not machines, are the processors.[1] We have parallel

---

[1] *See infra* Section III.A (collecting sources).

work to do on legal systems in which heavy caseloads, broad missions, growing electronic databases, and other features can produce informational inundation. To advance that work, this Article takes up one curious, even nagging, idea about interpretation as a gateway into larger issues that surround the informational architecture of our legal system.

The idea is that, in statutory and constitutional cases, more sources of interpretation tend to yield more interpreter discretion. The idea is embedded in a biting quip attributed to Judge Harold Leventhal—that citing legislative history is like "looking over a crowd and picking out your friends."[2] Having a crowd to look over is part of the problem on this account.[3] A similar link between source proliferation and discretion appears in constitutional arguments. As Justice Stevens was retiring, Justice Scalia defended originalism as "less subjective because it depends upon a body of evidence susceptible of reasoned analysis rather than a variety of vague ethico-political First Principles whose combined conclusion can be found to point in any direction the judges favor."[4] The idea had surfaced earlier, along with the looking-over-a-crowd quip, in Justice Scalia's objection to social science as support for banning the death penalty for minors.[5] And the idea resurfaced later, during Chief Justice Roberts's confirmation hearings, when he criticized reliance on foreign law in domestic constitutional cases.[6]

---

[2] Patricia M. Wald, *Some Observations on the Use of Legislative History in the 1981 Supreme Court Term*, 68 Iowa L. Rev. 195, 214 (1983) (internal quotation marks omitted) (citing a conversation with the judge), *quoted in* Exxon Mobil Corp. v. Allapattah Servs., Inc., 545 U.S. 546, 568 (2005); *see also* Conroy v. Aniskoff, 507 U.S. 511, 519 (1993) (Scalia, J., concurring) (paraphrasing Leventhal similarly); Abner J. Mikva, *Statutory Interpretation: Getting the Law to Be Less Common*, 50 Ohio St. L.J. 979, 981–82 (1989) (similar).

[3] *See* John F. Manning & Matthew C. Stephenson, Legislation and Regulation 145 (2d ed. 2013) (describing new textualist critiques of legislative history); *id.* at 160–61 (summarizing the discretion debate, including Leventhal's notion).

[4] McDonald v. City of Chicago, 561 U.S. 742, 804 (2010) (Scalia, J., concurring) (responding to Justice Stevens's multifactor approach to substantive due process claims); *see also id.* at 805 ("What is more, his approach would not eliminate, but multiply, the hard questions courts must confront, since he would not *replace* history with moral philosophy, but would have courts consider *both*."); Gregory C. Sisk & Michael Heise, *Ideology "All the Way Down"? An Empirical Study of Establishment Clause Decisions in the Federal Courts*, 110 Mich. L. Rev. 1201, 1244–45 (2012) (quoting Justice Scalia in *McDonald*).

[5] *See* Roper v. Simmons, 543 U.S. 551, 617 (2005) (Scalia, J., dissenting).

[6] *Confirmation Hearing on the Nomination of John G. Roberts, Jr. to Be Chief Justice of the United States: Hearing Before the S. Comm. on the Judiciary*, 109th Cong. 201 (2005) (distinguishing foreign law from domestic precedent, which supposedly "can confine and shape the discretion of the judges"). Roberts contended:

> In foreign law you can find anything you want. If you don't find it in the
> decisions of France or Italy, it's in the decisions of Somalia or Japan or

At the same time, a "more sources, more discretion" idea is embraced by several critics of textualism, originalism, and formalism. Legal Realists might be the modern founders of the idea, with their suggestions that the proliferation of legal sources enhances the judge's ability to find convenient legal categories for post hoc rationalizations.[7] Karl Llewellyn's thrust-but-parry characterization of interpretive canons probably is the most famous version of the notion,[8] and at least a trace of the idea shows up in critical legal scholarship.[9] But "more sources, more discretion" is a mainstream counterpunch, too. Justice Stevens warned that a "rudderless, panoramic tour of American legal history" gives a judge "more than ample opportunity to 'look over the heads of the crowd and pick out [his] friends.'"[10] And "cherry-picking among interpretive sources is a problem for all methodologies," William Eskridge reminds us, which means that textualists along with others might enter "a crowded cocktail party and look[ ] over the heads of the guests for one's friends."[11]

Leventhal's quip is a meme.

The meme carries more than one element, however, which helps explain its attractiveness.[12] The idea hooks together the proliferation

---

Indonesia or wherever. As somebody said in another context, looking at foreign law for support is like looking out over a crowd and picking out your friends. . . . And that actually expands the discretion of the judge. It allows the judge to incorporate his or her own personal preferences . . . .

*Id.*

[7] *See* Mark Tushnet, The New Constitutional Order 120 (2003) ("For the realists, conclusions did not flow from principles: In a mature legal system whose doctrinal space was thickly populated, a judge given a principle articulated in some prior case could faithfully deploy that principle along with others equally available in the doctrinal universe to reach whatever result the judge thought socially desirable.").

[8] *See* Karl N. Llewellyn, *Remarks on the Theory of Appellate Decision and the Rules or Canons About How Statutes Are to Be Construed*, 3 Vand. L. Rev. 395, 401–06 (1950) (organizing canons into countervailing columns).

[9] *See* J.M. Balkin, *Taking Ideology Seriously: Ronald Dworkin and the CLS Critique*, 55 UMKC L. Rev. 392, 430 (1987) (summarizing a critique of Ronald Dworkin's work on interpretation—that judges are deluded from understanding that they impose ideological order upon legal sources, in part because "the materials of the law already contain justifications supporting every variety of liberal and conservative positions").

[10] McDonald v. City of Chicago, 561 U.S. 742, 908 (2010) (Stevens, J., dissenting) (alteration in original) (quoting *Roper*, 543 U.S. at 617 (Scalia, J., dissenting)).

[11] William N. Eskridge, Jr., *The New Textualism and Normative Canons*, 113 Colum. L. Rev. 531, 533–34 (2013) (reviewing Antonin Scalia & Bryan A. Garner, Reading Law: The Interpretation of Legal Texts (2012)); *see also* Paul W. Kahn & Kiel Brennan-Marquez, *Statutes and Democratic Self-Authorship*, 56 Wm. & Mary L. Rev. 115, 127 n.38 (2014) ("Of course, the critique can also be lodged in reverse.").

[12] The idea's attraction is not universal. As Adrian Vermeule has written, with appropriate restraint, "[i]n general, it is unclear whether judicial disagreement increases or diminishes as sources are added beyond the statutory provisions at hand. The only definite effect of adding further sources is to increase the costs of decisionmaking." Adrian Vermeule, Judging Under Uncertainty 189 (2006); *see also* Manning &

of sources with the power to pick preferred sources. This picking power is an important assumption, of course. Judge Leventhal was not singling out source proliferation alone. But the effect of source proliferation on decisions is important to understand, regardless. What independent role, if any, does source proliferation have in generating interpreter discretion? Does source proliferation increase discretion if we eliminate the power to cherry pick sources? If we reintroduce this power to pick, on what assumptions does source proliferation matter to discretion? If we develop reliable answers, what trade-offs are associated with increasing the number of sources? If we determine a roughly optimal number of sources to consider, can that choice be implemented in courts? This Article aims to make progress on each of these questions, using logic and evidence.

As a logical matter, the notion that discretion increases as sources increase is incorrect without more. Sometimes the opposite is true. Indeed adding sources tends to *drive down* the probability that all sources will cancel out into uncertainty and yield discretion, according to a simple model of interpretation. To show this, I construct a bare-bones model with clearly stated assumptions in which interpretive sources are equally likely to point in various directions and the implications of all sources are aggregated.[13] This starter model suggests extensions of the basic conclusion, as well as distinct patterns if we loosen the model's strictures.[14] The latter discussion introduces the possibilities of "ruleless" interpretive method and "willful" interpreters, along with interpreter ability to "spin" sources or "pick" among them. Even when we play with these looser models, however, the effect of source proliferation on discretion is modest and fragile. Logically and statistically, the power to make friends in a crowd by spinning sources overwhelms the importance of crowd size—and, crucially, the power to pick friends can generate plenty of discretion even in small crowds.[15]

---

STEPHENSON, *supra* note 3, at 145 (noting "a somewhat vaguer conception" among new textualists "that judges impermissibly acquire added policymaking discretion by relying on legislative history"); *cf.* David A. Strauss, *Common Law Constitutional Interpretation*, 63 U. CHI. L. REV. 877, 926–27 (1996) (contending that "[a] conscientious judge will find substantial guidance in a well developed body of precedent," and that "a judge who acknowledges only the text of the Constitution as a limit can, so to speak, go to town").

13   *See infra* Sections II.A–B.

14   *See infra* Sections II.C–D and Table 1.

15   The beginnings of this logical analysis were set out in a subsection of Adam M. Samaha, *On Law's Tiebreakers*, 77 U. CHI. L. REV. 1661, 1685–89 (2010) [hereinafter Samaha, *Tiebreakers*]. Here, the logical analysis is extended and revised, especially in the discussion of variations on and deviations from the starter model. *See infra* Sections II.C–D. The remainder of the analysis is new.

As an empirical matter, the effect of additional sources seems equally contingent. Useful studies of information-load effects are scattered and yet provocative when collected, organized, and extended. I foreground four working theories below. (1) *Ideology*: More sources might prompt judges to de-emphasize conventional sources of law and allow ideology greater influence, at least in ideologically charged fields of litigation.[16] However, neither theory nor evidence is strong on this score, as indicated by new data presented below. (2) *No effect*: Instead judges might economize on effort. Judges might ignore information after a certain point such that there is no effect on results, or treat additional information arbitrarily and make more mistakes.[17] New data on sources cited in appellate briefs suggest as much. (3) *Paper pushing*: Or judges facing heavier information loads might push the paperwork along by increasing deference to other decisionmakers or by attempting to reinforce a status quo.[18] Some findings on docket loads and case-specific information loads support this theory, though the results are mixed. (4) *Quasi-law*: Or judges might engage in a remarkable kind of quasi-legal weighting of valid sources.[19] According to this theory, formal law is honored imperfectly but decisions are not steered by ideology in the ordinary sense. Instead, judges attempt to distill the law's core priorities.

For support, the Article relies on a number of tools. Part of the analysis is heavy on logic and probabilities. Although the discussion avoids technical mathematics, the conclusions are supported by some standard probability theory, large permutation tables, and brute-force computer simulations to fill out logical implications. A second part of the analysis is empirical. In addition to spotlighting provocative studies and building working theories, the Article presents new data. On ideology, analyzing a new set of appellate decisions casts doubt on an earlier finding that increasing the stock of precedents ultimately increases judicial discretion.[20] On paper pushing, new data from appellate briefs reinforces an earlier finding that increasing brief length can reduce judicial preferences for affirmance; but the new data also indicate that more issues may increase affirmance rates, and the data fail to show a relationship between the number of sources cited and affirmance.[21] On quasi-law, an expanded set of district court decisions supports an earlier finding that a large number of doctrinal

---

[16] *See infra* Section III.B.
[17] *See infra* Section III.A.
[18] *See infra* Section III.C.
[19] *See infra* Section III.D.
[20] *See infra* text accompanying notes 162–65 and Table 2.
[21] *See infra* text accompanying notes 186–88 and Table 3.

factors may prompt judges to prioritize law's core factors, without evidence of conventional ideological influence.[22] But the data do not show that more spinning takes place when judges are asked to consider more factors. If anything, the opposite might be true.

All of this analysis aims to challenge simple intuitions about source proliferation while offering constructive paths forward. Controlling the number of interpretive sources does not seem likely to control judicial discretion, but setting out working theories for interpreter behavior will help develop a research agenda.[23] Moreover, lowering the number of sources for interpreters to consider might be justifiable anyway. Low-source methods might be cheaper and easier to implement than high-source methods, even if the quality of decisionmaking drops and the discretion level stays put.[24] In any event, the known trade-offs change once the various plausible effects of source proliferation are recognized.[25] Thinking practically about implementation and understanding the confluence of logical, cognitive, and institutional forces will help us predict the effects of adding interpretive sources—and build the informational architecture that works best for interpretation in the real world.

# I
## Beginnings

### A.  Trade-offs

A beautifully uncomplicated thought about information is that "more is better."[26] Assuming relevance, increasing the amount of information considered should increase the quality of decisionmaking. One version of the thought rests in the principle of total evidence for rational actors, which recommends the use of all available evidence when estimating the likelihood of various outcomes.[27] In 1967, the British mathematician and codebreaker Jack Good published a short formal proof demonstrating that, for those trying to maximize expected utility, "it pays to take into account further evidence,"

---

[22] *See infra* text accompanying notes 200–11 and Tables 4–5.

[23] *See infra* Section IV.A.

[24] *See infra* Section IV.B.

[25] *See infra* Section IV.C.

[26] *See* J. Edward Russo, *More Information Is Better: A Reevaluation of Jacoby, Speller and Kohn*, J. Consumer Res., Dec. 1974, at 68, 71 (undercutting the findings of an early study that had suggested that lower quality decisions followed from information overload).

[27] *See* I.J. Good, *On the Principle of Total Evidence*, 17 Brit. J. for Phil. Sci. 319, 319 (1967).

bracketing the cost of collecting and using the information.[28] Justice White was not a mathematician but his intuition was likewise in statutory cases. On the accumulation of legislative history, he wrote for the Court in 1991 that "common sense suggests that inquiry benefits from reviewing additional information rather than ignoring it."[29]

But White, Good, and the rest of us would not deny that gathering and processing information costs somebody something. People must devote time, mental effort, and other resources to processing information when they shoulder the duty to decide. It is true that technological progress allows us to offload evermore of the information-gathering and processing burden to machines. At the same time, increasing our reliance on databases, search engines, and computational software also means that resources must be devoted to developing, maintaining, and operating those technologies. As a first cut, then, adding information seems to increase decision quality but also decision costs, and thoughtful people will account for both consequences when they develop decision procedures.

Institutional design scholarship advertises just this trade-off,[30] with a catchy recommendation that institutional designers should try to minimize the sum of decision costs and error costs.[31] This formula is incomplete; the equation appears to have nothing on the benefit side of the ledger, for one thing. Still, the recommendation importantly suggests that driving down errors tends to drive up the cost of making decisions. Good system designers should tolerate expenditures of decisional resources, including information gathering and processing, to improve the probability of a good decision. But only to a point.

---

[28] *Id.* For deviations from this advice when risk aversion is added and the additional information might be misleading, see Lara Buchak, *Instrumental Rationality, Epistemic Rationality, and Evidence-Gathering*, 24 PHIL. PERSP. 85, 95–101 (2010).

[29] Wis. Pub. Intervenor v. Mortier, 501 U.S. 597, 611 n.4 (1991).

[30] *See* NEIL K. KOMESAR, IMPERFECT ALTERNATIVES 5 (1994) (emphasizing trade-offs across different institutions); David L. Weimer, *Institutional Design: Overview*, *in* INSTITUTIONAL DESIGN 1, 12 (David L. Weimer ed., 1995) (discussing decision costs and error costs in institutional design).

[31] *See, e.g.*, Cass R. Sunstein, *Foreword: Leaving Things Undecided*, 110 HARV. L. REV. 4, 16, 19 (1996) (suggesting that "good judges try to minimize the sum of decision costs and error costs," and that the idea is useful even though not all consequences are qualitatively similar or valued on the same metric); Frank H. Easterbrook, *The Limits of Antitrust*, 63 TEX. L. REV. 1, 16 (1984) [hereinafter Easterbrook, *Antitrust*] (using a similar formulation for antitrust law); *see also* Eric A. Posner, *Controlling Agencies with Cost-Benefit Analysis: A Positive Political Theory Perspective*, 68 U. CHI. L. REV. 1137, 1186 (2001) (referring to this formulation as a possible defense of cost-benefit analysis, but stressing normative goals, empirical uncertainty, and institutional setting); Adrian Vermeule, *Interpretive Choice*, 75 N.Y.U. L. REV. 74, 89 (2000) ("[A]s between two doctrines of equal accuracy, the one that imposes fewer decision costs would be selected by any plausible theory of authority.").

They should look for assurances that increases in estimated decision costs will purchase adequate increases in estimated decision quality, in light of the stakes of the decision.[32] The general implications for interpretive method in law are fairly straightforward: More interpretive sources will tend to increase the likelihood of a correct interpretation but also increase the cost of reaching decision. Information, in court as elsewhere, should be cut off at some stage. "We do not hold three-week trials about parking tickets."[33]

Behavioral research threatens this logic,[34] but legal scholarship on interpretation has developed in another direction and it warrants attention. A sizable part of this scholarship concentrates on discretion, particularly the extent to which various methods allow judges to influence law with policy preferences.[35] The "more sources, more discretion" idea is part of this debate. If true, the idea might well scramble standard trade-offs. At some point, adding information would not simply increase decision quality and decision costs, but also decision-maker discretion and ideological influence. Depending on how one feels about such discretion and ideology, the previously supposed trade-offs might disappear. Judicial discretion is not necessarily harmful, of course, and some amount of it seems unavoidable in human legal judgment. But suppose that an observer is concerned. Then source proliferation might be a major worry as well. Indeed if one takes the hardline position that judicial discretion and ideology always count as errors, then the argument for capping information becomes conclusive. After some number of interpretive sources, adding more could yield only losses, tallied in both decision costs and error costs. We might not hold three-week trials about anything.

In fact, interpreters always have simpler alternatives to tracking down every source and integrating every inference. Consider casebook staples such as *General Dynamics Land Systems, Inc. v. Cline*[36] and *Myers v. United States*.[37] In the former, the Supreme Court opinions showcased sophisticated interpretive moves, with Justices debating dictionary definitions, colloquialisms, canons, legislative

---

[32] *See, e.g.*, VERMEULE, *supra* note 12, at 166–67, 189 (noting that, whatever might be the other effects of adding interpretive sources, decision costs are likely to rise); Adam M. Samaha, *Undue Process*, 59 STAN. L. REV. 601, 616–20 (2006) (elaborating the idea of decision cost and associated trade-offs).

[33] Easterbrook, *Antitrust*, *supra* note 31, at 16.

[34] *See infra* Section III.A (noting that additional information might be ignored or increase mistakes); *infra* Section IV.C (returning to the connection between behavioral patterns and institutional design recommendations).

[35] *See supra* notes 2–7 (collecting sources).

[36] 540 U.S. 581 (2004).

[37] 272 U.S. 52 (1926).

history, and more before settling on a relatively narrow reading of the Age Discrimination in Employment Act.[38] The latter dispute produced nearly two-hundred pages of tree-killing opinions that struggled over the constitutional scope of the President's removal authority.[39]

Nearly lost to history are the mercifully concise lower court opinions in those cases. In *General Dynamics*, the district court briefly adverted to plain text, precedent, and statutory purpose before ordering the suit dismissed.[40] In *Myers*, the Court of Claims saw a gap in the constitutional text on removal authority, recapped a few Supreme Court discussions, and then held that uncertainty counseled deference to Congress.[41] Written opinions do not describe complete mental processes, of course. But perhaps the punchy analysis on display in these lower courts should be the high paragons of actual legal interpretation. Perhaps they represent not only judicial economy, but also prudence and restraint in the face of an alluring crowd of interpretive sources.

## B.   *Concepts*

To test the "more sources, more discretion" idea, however, we need to specify key concepts. Each concept—interpretation, source, discretion, and ideology—is a topic of concern in jurisprudence and beyond, but we may move quickly and simplify. First, we can use a relatively narrow conception of legal *interpretation* as referring to the process of identifying meaning in or assigning meaning to a given text. Broader scholarly conceptions are available, including those that reach the application of legal norms to case-specific facts or perhaps all of the central tasks of adjudication.[42] Much of the analysis below

---

38   *See General Dynamics*, 540 U.S. at 600 ("We see the text, structure, purpose, and history of the [Act], along with its relationship to other federal statutes, as showing that the statute does not mean to stop an employer from favoring an older employee over a younger one."); *infra* text accompanying notes 56–62 (summarizing the debate in *General Dynamics*).

39   *See Myers*, 272 U.S. at 106–77 (majority opinion); *id.* at 178–239 (McReynolds, J., dissenting); *id.* at 240–95 (Brandeis, J., dissenting); *infra* text accompanying notes 92–96 (summarizing history-based possibilities). Justice Holmes's dissent was a page long. *See Myers*, 272 U.S. at 177.

40   *See* Cline v. Gen. Dynamics Land Sys., Inc., 98 F. Supp. 2d 846, 848 (N.D. Ohio 2000), *rev'd*, 296 F.3d 466 (6th Cir. 2002), *rev'd*, 540 U.S. 581 (2004).

41   *See* Myers v. United States, 58 Ct. Cl. 199, 204–06 (1923), *aff'd*, 272 U.S. 52 (1926). The Court of Claims went on to hold that the suit was not timely, *see id.* at 206, making that court's constitutional analysis unnecessary even though concise.

42   *Compare, e.g.*, Kent Greenawalt, *Constitutional and Statutory Interpretation*, *in* The Oxford Handbook of Jurisprudence and Philosophy of Law 268, 268–70 (Jules Coleman & Scott Shapiro eds., 2002) (using the term to include text, original meaning, underlying rationale, application to particular cases, and stare decisis), *with* Steven Knapp & Walter Benn Michaels, *Intention, Identity, and the Constitution: A Response to David*

carries over to those broader notions, too, and the initial model below does connect legal interpretation with adjudicated case outcomes. But no serious problem should arise from using a simple and narrow conception of interpretation.

An interpretive *source* can be any item of information that is relevant and legally valid for interpreting a given text, such as a dictionary or a delegate's speech at a ratifying convention. This understanding may, but need not, include analytic resources for interpreting legal texts and handling close cases, such as presumptions, clear statement rules, and other canons. The analysis below will hold either way, with just a little willingness to divide information into subcategories.[43] Interpretive sources will then proliferate within every major methodological school. Statutory interpretation has competing textualist, purposivist, and dynamic approaches, among others,[44] and numerous subcategories of sources can be identified within each. Modern textualism, for instance, allows interpreters to rely on dictionaries, colloquialisms, terms of art, structural inferences, and canons within some notion of context.[45] We have not yet listed legislative history or judicial precedent, neither of which is a necessarily unitary source.[46] In constitutional interpretation, the labels differ but the basic

---

*Hoy*, *in* LEGAL HERMENEUTICS 187, 192–97 (Gregory Leyh ed., 1992) (identifying interpretation with authorial intent), *and* Gary Lawson, *On Reading Recipes . . . and Constitutions*, 85 GEO. L.J. 1823, 1834 (1997) (defining interpretation as the search for a text's original public meaning).

[43] That almost any proposed subcategory of information could be further subdivided will make my claims stronger. Multiplying the number of interpretive sources through subdivision means that no interpretive method can rely on only one "source," and that all interpretive methods face a similar question regarding the appropriate number of sources for decisionmaking.

[44] *See* Philip P. Frickey, *From the Big Sleep to the Big Heat: The Revival of Theory in Statutory Interpretation*, 77 MINN. L. REV. 241, 249–58 (1992) (surveying the development of competing interpretive approaches).

[45] *Compare* SCALIA & GARNER, *supra* note 11, at xi–xvi, 9 (discussing fifty-seven canons of statutory construction that the authors credit as valid), *with* Eskridge, *supra* note 11, at 536, 541–42 (counting 187 canons used by the Rehnquist and Roberts Courts, defined broadly enough to include precedent and legislative history). On some versions of textualism, the interpreter must consider a piece of statutory text "alongside the remainder of the *corpus juris*." Antonin Scalia, *Common-Law Courts in a Civil-Law System: The Role of United States Federal Courts in Interpreting the Constitution and Laws*, *in* A MATTER OF INTERPRETATION 3, 17 (Amy Gutmann ed., 1997).

[46] For legislative history, standard subdivisions include committee reports, sponsor statements, other floor debate, rejected amendments, and witness testimony. *See, e.g.*, Adrian Vermeule, *Legislative History and the Limits of Judicial Competence: The Untold Story of* Holy Trinity Church, 50 STAN. L. REV. 1833, 1879–80 (1998). For judicial precedent, the number of cases deemed relevant can range from one to dozens, with some of those cases receiving less weight than others—such as summary affirmances and perhaps less-deeply reasoned opinions. *See* District of Columbia v. Heller, 554 U.S. 570, 623–24 (2008) (discussing and limiting United States v. Miller, 307 U.S. 174 (1939)); *see also*

picture is the same. Several general methods persist: originalism, traditionalism, common-law constitutionalism, contemporary moral readings, deference, minimalism, and so on.[47] Even if some of these methods are ignored under a narrow definition of interpretation, the others are bound to encompass numerous sources. Originalism, for instance, is a family of ideas with many potentially relevant historical sources in a given case.[48]

For *discretion*, legal scholarship offers two popular versions. The first occurs when law, properly interpreted, authorizes the decisionmaker to take or withhold action within some outer boundary.[49] Law may grant the decisionmaker a range of permissible options and reasons on which to make a choice, but otherwise leaves that ultimate choice unconstrained. This conception is not prevalent in debates over interpretive method, where participants do not typically portray their own methods as assigning authority to choose; instead, discretion is sometimes said to be the bad effect of a competitor's method. The second version of discretion seems more fitting here. It occurs when law, properly interpreted, fails to resolve an issue that an official is obliged to resolve.[50] Law runs out in this version—presumably not by

---

Michael J. Gerhardt, *Super Precedent*, 90 Minn. L. Rev. 1204, 1205 (2006) (addressing variable precedential force and generative effect).

[47] *See* Adam M. Samaha, *Dead Hand Arguments and Constitutional Interpretation*, 108 Colum. L. Rev. 606, 639, 670 (2008).

[48] *See, e.g.*, Thomas B. Colby & Peter J. Smith, *Living Originalism*, 59 Duke L.J. 239, 247–62 (2009) (describing different strains of originalism); Vasan Kesavan & Michael Stokes Paulsen, *The Interpretive Force of the Constitution's Secret Drafting History*, 91 Geo. L.J. 1113, 1117–18, 1200–02 (2003) (contending that, even for those committed to original public meaning, secret drafting history may be given weight alongside contemporaneous dictionaries, *The Federalist Papers*, and other historical sources).

[49] *See, e.g.*, Henry M. Hart Jr. & Albert M. Sacks, The Legal Process: Basic Problems in the Making and Application of Law 144 (William N. Eskridge, Jr. & Philip P. Frickey eds., 1994) (defining discretion as "the power to choose between two or more courses of action each of which is thought of as permissible"); Steven J. Burton, *Particularism, Discretion, and the Rule of Law*, *in* 36 Nomos: The Rule of Law 178, 189, 193 (Ian Shapiro ed., 1994) (distinguishing situations in which law confers discretion on a decisionmaker from those in which law is indeterminate, although holding to the idea that law may include and exclude reasons for decision). Compare Ronald Dworkin, Taking Rights Seriously 32–33 (1977), which describes a "strong" form of discretion in which the decisionmaker is subject to some kind of criticism but is "simply not bound by standards set by the authority in question."

[50] *See, e.g.*, H.L.A. Hart, The Concept of Law 272 (2d ed. 1994) (depicting judges as "filling the gaps"); Todd E. Pettys, *Judicial Discretion in Constitutional Cases*, 26 J.L. & Pol. 123, 124 (2011) (describing indeterminacy or conflict of relevant legal texts, conventionally understood); Richard A. Posner, *Judicial Behavior and Performance: An Economic Approach*, 32 Fla. St. U. L. Rev. 1259, 1272 (2005) ("[T]here is a large open area . . . in which conventional legal materials do not dictate the outcome and the judge is forced to make a policy judgment . . . ."), *quoted in* Pauline T. Kim, *Lower Court Discretion*, 82 N.Y.U. L. Rev. 383, 408 (2007); *see also* Kent Greenawalt, *Discretion and Judicial Decision: The Elusive Quest for the Fetters that Bind Judges*, 75 Colum. L. Rev.

design—instead of affirmatively assigning to the decisionmaker a range of permissible options and reasons. The line between these versions of discretion might be smudged or practically unsustainable, and the general notion of legal discretion is slippery.[51] To kickstart the analysis, though, we can think about legal discretion as arising in situations where valid and relevant sources of law fail to dictate a unique outcome, leaving the interpreter to turn elsewhere.

One place an interpreter might turn after exhausting law's resources is *ideology*, which sometimes is suggested as an upshot alongside claims that more sources yield more discretion.[52] The notion of judicial ideology can be associated with several different influences. Judicial ideology can track partisan, liberal-conservative, left-right, or another salient divide over policy matters outside of the judiciary;[53] or it can be elaborated as the judge's worldview, perceptions, values, and preferences without reference to extra-judicial politics.[54] Any of these conceptions can signify the influence of judicial preferences beyond what existing law establishes. The task of pinning down the concept of ideology does reemerge in the development of strategies for measuring it. But those challenges will not impede our effort to specify claims about the effect of adding interpretive sources. For the moment, we can think about judicial ideology as a preference beyond existing law that an interpreter might use to reach a result.

---

359, 368 (1975) (suggesting discretion is present when "whatever external standards may be applicable either cannot be discovered by the decision-maker or do not yield clear answers to the questions that must be decided"). Compare DWORKIN, *supra* note 49, at 31–32, which describes "weak" forms of discretion in which the standards are vague or hard to apply, or the judgment is not reviewable.

[51] We might think that discretion is always constrained somehow, yet we also might think that adopting an interpretive method for oneself shows the exercise of discretion as much as its constraint.

[52] *See supra* notes 4–11 and accompanying text (collecting and quoting sources). If we assume that law has run out, we can avoid complications generated by the astute observation that "ideology" and "law" may partly collapse, insofar as either law authorizes the exercise of discretion or judges have an ideological preference for following law's sources. *See* Kim, *supra* note 50, at 404–08.

[53] *See, e.g.*, FRANK B. CROSS, DECISION MAKING IN THE U.S. COURTS OF APPEALS 13, 19–21 (2007) (asserting that, although partisan measures might be proxies, the empirical focus is personal judicial ideology on a liberal-conservative spectrum).

[54] *See, e.g.*, Bryan D. Lammon, *What We Talk About When We Talk About Ideology: Judicial Politics Scholarship and Naive Legal Realism*, 83 ST. JOHN'S L. REV. 231, 242, 249–50, 254–55 (2009) (defining ideology loosely to include beliefs, experiences, and perceptions about the world).

II

Logic

To gain leverage on the "more sources, more discretion" idea, we can add logic. First notice that, if there are *no* guidelines for interpretation, then a claim that more interpretive sources yield more interpreter discretion is nonsense. Unchannelled "interpretation" would be drenched with discretion regardless of the number of sources. We have to impose some structure on interpretation to make a source proliferation claim logically viable, although the amount and kind of structure is open to debate.

The analysis below begins by moving to the opposite extreme and imposing a vice-like structure on interpretation, but with room for residual uncertainty. This will allow us to assess the effect of adding sources on the probability of a decisive outcome within a rule-bound process, and then compare looser processes. As well, articulating the assumptions underlying the initial model will pinpoint plausible repositories of discretion—because, logically and all else equal, adding interpretive sources will not increase discretion. In this rule-bound model, more sources tend to yield less discretion.

### A. Assumptions

■ *Two options.* Suppose that a decisionmaker, such as a trial or appellate judge, must decide between two competing interpretations of a given statutory provision that will dictate the result in a particular case. The plaintiff's counsel or the prosecutor contends that the statute covers the defendant's (alleged) conduct, while the defendant's counsel contends otherwise. Any interpretive source—dictionaries, canons, legislative history, and so on—might or might not support one side's position over the other's. All of this is realistic enough, although later we will take up the possibility of more than two interpretive options.[55]

■ *Five source values.* Suppose further that each interpretive source might strongly or weakly support either the plaintiff's position or the defendant's position, or might be unclear. We can represent such interpretive sources as variables $\{s_1, s_2, \ldots s_n\}$ and their possible implications as one of five values $\{-2, -1, 0, +1, +2\}$, where the negative values indicate a strong or weak implication in the plaintiff's favor, zero indicates that the source is unclear, and the positive values indicate a weak or strong implication in the defendant's favor. This notion of sources and how they are viewed by interpreters is realistic

---

[55] *See infra* Section II.C. For simplicity, I refer to statutory cases here. The analysis holds for constitutional cases.

enough as well, although this model treats interpretive sources as if they have discontinuous strength levels instead of a smooth continuum of influence. Also, the model effectively assigns the same weight to all sources, even though the weight of a source might be distinguished from how strongly or clearly it appears to confirm or disconfirm a hypothesis. Such simplifications are useful to get a starter model running.

To illustrate differing source values, recall the careening discussion in *General Dynamics*.[56] There, the Supreme Court held that the Age Discrimination in Employment Act did not prohibit employers from favoring older workers over younger workers.[57] The case was well fought. A conscientious interpreter might first conclude that the dictionary meaning of "age" in the directive to not "discriminate against any individual . . . because of such individual's age"[58] points at least weakly toward the plaintiff's position that age means number of years since birth (-1).[59] The interpreter also might think that the colloquial meaning is ambiguous given various popular uses of "age" as both number of years and old age (0);[60] and that a presumption of consistent meaning with the bona fide occupational qualification provision of the statute, which uses "age" as number of years (-1), is canceled out by a presumption that the surrounding words in the antidiscrimination provision establish a special semantic context for "age" (+1).[61] But an interpreter might go on to conclude that the apparent purpose of the statute points strongly toward the defendant's position that "age" in the antidiscrimination provision means old age (+2).[62] With these numbers, the case is close but tilts toward the defendant (+1 total). These source groupings and numerical assignments are debatable, of course, and there was more to the case. But this much is adequate to illustrate the model.

▪ *Equal probability across values.* Suppose also that each interpretive source is equally likely to take each possible value. The implications of sources are independent and random in the sense of having equal probabilities. Applied to our case example, the notion is that,

---

[56] Gen. Dynamics Land Sys., Inc. v. Cline, 540 U.S. 581 (2004).

[57] *See id.* at 584.

[58] 29 U.S.C. § 623 (2012).

[59] *See General Dynamics*, 540 U.S. at 603 (Thomas, J., dissenting) (distinguishing primary and secondary dictionary meanings).

[60] *See id.* at 596 (majority opinion) ("[W]e understand the different choices of meaning that lie behind a sentence like 'Age can be shown by a driver's license,' and the statement, 'Age has left him a shut-in.'").

[61] *See id.* at 595 (referencing both presumptions, though deciding to follow the latter).

[62] *See id.* at 586–92 (recounting evidence of statutory purpose, including enacted findings and legislative history).

before the judges started interpreting the statute in *General Dynamics*, the only guess about probabilities was that each of the interpretive sources to be considered was equally likely to support either side strongly or weakly, or be unclear. The interpreter then assigns values to each source considered, without any presumptions or other priors that alter the equal probability constraint.

In real life, admittedly, the appropriate inference to draw from one source sometimes depends on the inference drawn from another source.[63] Additionally, we might make better predictions about the likely implications of subsets of sources; perhaps an identifiable subset usually lacks clarity or usually points in one direction. To be highly realistic, moreover, we should incorporate the possibility that an interpreter will spin a source's otherwise proper inference toward the interpreter's preferred direction. But we can explore these complications later. For now, assume a 20% chance across the board for each of the five possible source values, so that neither sources nor interpreters play favorites.

From a certain perspective, this setup tilts the universe of sources toward uncertainty. Not only might individual sources be unclear, but sources have no tendency to clump toward one side's position. One might instead believe or hope that the implications of interpretive sources usually will tend toward one case-dispositive direction, even if a fraction of the relevant sources is unclear or in tension. Our starter model rules out this pro-certainty position, and the choice is deliberate. The model accentuates the possibility of uncertainty, giving the "more sources, more discretion" notion a shot at success. The model also is intended as a nod to familiar dynamics in litigation. Easy cases often will not be disputed or will settle;[64] in cases that are fought out, lawyers from each side will collect and emphasize sources that they think favor their respective clients, thus building an initial foundation of sources for the court that is likely to be in tension.

---

[63] *See, e.g.*, Moseley v. V Secret Catalogue, Inc., 537 U.S. 418, 432 (2003) (referring to interpretive sources that "shed[ ] light" on each other); William N. Eskridge, Jr. & Philip P. Frickey, *Statutory Interpretation as Practical Reasoning*, 42 Stan. L. Rev. 321, 352 (1990) (contending that "a true dialogue with the text requires the interpreter to reconsider her preunderstandings as she considers the specific evidence in the case, and then to formulate a new understanding, which in turn is subject to reconsideration").

[64] *See* George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. Legal Stud. 1, 4–5 (1984) (modeling litigation selection effects, with symmetric stakes); Samaha, *Tiebreakers*, *supra* note 15, at 1731 ("Uncertainty among parties translates into higher probability of litigation to the end and probably correlates with difficulty among judges in reaching consensus. The outcome can be analogized to a coin flip, at least from the litigants' perspective.").

■ *Strict aggregation rule.* Suppose finally, and most controversially, that interpreters follow a strict rule for aggregating source implications. The formula is simply to add up the values of all sources and, if the total favors one side or the other in any amount, to decide the case in that direction. Plaintiff prevails on a negative total, defendant prevails on a positive total, and a zero total means that the correct result remains uncertain. If the result remains uncertain after every source's implication has been aggregated, then we can say law has run out and discretion must emerge to close the case. Thus, a zero total stands in for judicial discretion. In our initial rendering of *General Dynamics*, the numerical values for the cited sources totaled up to a close call in the defendant's favor (+1), but we might add a bit more legislative history in the form of a floor colloquy[65] and tip the total into overall uncertainty (0).[66] But all other totals count as decisive outcomes that interpreters will follow to resolve cases.

This modeling choice is debatable, too. One might instead liken the drawing of implications from a set of interpretive sources to the drawing of balls from an urn, as in a familiar probability illustration.[67] Then we should expect that drawing more sources will yield a distribution of results that better reflects the true distribution of all sources.[68]

---

[65] *See General Dynamics*, 540 U.S. at 598 (acknowledging a congressional floor statement indicating that younger workers should be able to sue when employers favor older workers). If this floor statement is persuasive enough, the total might be tipped back in favor of the plaintiffs.

[66] While contending that there is a single correct answer for each legal decisionmaker, even in hard cases, Ronald Dworkin once counted the conclusion that the case is a "tie" as a single correct answer. *See* DWORKIN, *supra* note 49, at 285. Of course, judges do not stop with uncertainty and then announce that their discretion takes over. *See* Samaha, *Tiebreakers*, *supra* note 15, at 1708. They find another source of law or presumption, such as deference to an administrative agency or the rule of lenity, to reach a decisive resolution when parties continue to demand one. But the model's allowance for uncertainty is necessary to make progress on understanding the relationship between source proliferation and discretion.

[67] *See, e.g.*, Dan M. Kahan et al., *"Ideology" or "Situation Sense"? An Experimental Investigation of Motivated Reasoning and Professional Judgment*, 164 U. PA. L. REV. 349, 400 (2016) (illustrating Bayesian updating of likelihood ratios after an additional draw from an urn—or bag). For an inquiry into panel effects along these sampling lines—in which groups of reviewing judges with a range of policy preferences and some sensitivity to facts make strategic decisions whether to conduct additional research—see Matthew Spitzer & Eric Talley, *Left, Right, and Center: Strategic Information Acquisition and Diversity in Judicial Panels*, 29 J.L. ECON. & ORG. 638, 645–60 (2013).

[68] In the text, I am using inexpert terms to refer to the (weak) law of large numbers and the central limit theorem(s). *See generally* E.L. LEHMANN, ELEMENTS OF LARGE-SAMPLE THEORY 47, 55, 72–73, 85 (1998) (indicating that a random sample's mean tends to converge with the overall distribution's mean as the sample size increases, and that the distribution of a large number of independent and identically distributed random variables with a mean and finite variance will be approximately normal); SIMON N. WOOD, CORE STATISTICS 14 (2015) (stating that the central limit theorem "justifies using the normal

If the universe of relevant sources truly is in equipoise, as our assumed source values and probabilities suggest, then the average (and sum) of all their values should be zero, and we should detect this number with more confidence as more sources are drawn.[69] Following this sampling logic would not do much good for the "more sources, more discretion" notion in general, however: More sources would straightforwardly tend to yield less discretion whenever the universe of sources actually does tilt in one direction.

Regardless, a good statistician might not make a decisive judgment based on a slight, perceived leaning in any large draw of sources, perhaps even on issues of law's meaning. Instead, she or he might only formulate hypotheses about law and wait for statistically significant differences from zero, whatever that means in the context of adjudication.

On the other hand, deciding cases is not exactly like taking samples. It is quite conventional to believe that judges should interpret law and decide cases based on all relevant and legally valid sources before them, reaching a decisive resolution based on the overall implications of those sources, however close this overall impression is to the applicable legal line—not sample fewer than all relevant sources, or treat interpretive results that are close to equipoise as within a margin of error. It is unorthodox, albeit intellectually productive, to think about judges as testing some party's hypothesis about law's meaning without necessarily announcing correct interpretations of law for deciding cases. Our model tries to represent the conventional adjudicative role by depicting the interpreters as collecting inferences from a number of interpretive sources that are identified according to law, then combining the inferences for the purpose of closing a case. This process of combining inferences to reach a decisive result can be thought of as simple addition, much like the metaphorical weighing of admitted evidence.[70]

We remain free to suspect that actual judges will not always behave as if they were following a strict formula. There is no such formula for judicial interpretation, anyway. But the goal here is to

---

[distribution] as an approximation in many situations where a random variable can be viewed as a sum of other random variables").

[69] *Cf.* D.H. Kaye, *Do We Need a Calculus of Weight to Understand Proof Beyond a Reasonable Doubt?*, 66 B.U. L. Rev. 657, 658 (1986) ("One must examine the completeness of a body of evidence and the circumstances under which the evidence was gathered if one is to assess its probative value.").

[70] *See, e.g.*, Gary Lawson, *Proving the Law*, 86 Nw. U. L. Rev. 859, 859–60, 868 (1992) [hereinafter Lawson, *Proving the Law*] (asserting the necessity of proof burdens for interpretation in adjudication, and likening such decision protocols to those for fact-oriented evidentiary issues).

investigate the independent effect of source proliferation on discretion. The investigation requires an effort to separate source proliferation from influences such as discretion over how individual sources are understood or how multiple sources are aggregated. Temporarily eliminating the possibility of discretion in those other areas is a way of isolating the effect of the number of sources. Loose formulas and willful judges will be reintroduced later.

## B.   *Probabilities*

With this set up, we can calculate probabilities that the decision process will yield uncertainty as interpretive sources pile up. These calculations can be done in the absence of a standard equation and without sophisticated mathematics. Probabilities can be obtained by brute computational force, through the construction of permutation tables for various numbers of sources alongside computer simulations to check and extend the emerging patterns.

The first permutation table is very simple: Each of the five possible values for a single source is equally likely, leaving a 20% chance of uncertainty and a 20% chance of each of the other possibilities—a relatively strong case for the plaintiff, a relatively weak case for the plaintiff, a relatively weak case for the defendant, and a relatively strong case for the defendant. The second table adds a second source and totals every permutation of values for both sources ($5^2 = 25$). Adding a second source neither increases nor decreases the chance of overall uncertainty. The chance of both sources canceling out is 20% again, although the remaining distribution of totals changes. The chance of a close case on either side (i.e., a total of -1 or +1) decreases from 20% to 16%, and the chance of a clearer case (i.e., a total < -1 or > +1) increases from 20% to 24% as the tails of the distribution get longer.

From that point on, adding another source decreases the chance of overall uncertainty and interpreter discretion, albeit more slowly with each additional source. With three sources, the chance of all sources canceling to zero declines to approximately 15%. With four sources, the chance is down to approximately 14%. By the time we reach seven sources and tens of thousands of permutations ($5^7 = 78,125$), the chance of overall uncertainty falls to about 10%. With each additional source, moreover, we extend the outer limits of the distribution. It is always possible, even though highly unlikely, that all or nearly all of these sources will point strongly in one direction. Furthermore, the chance of a close case continues to fall with the addition

of more sources. On this measure, cases become clearer with more sources. Figure 1 charts the results through seven interpretive sources.

Simulations help confirm and extend these results, as shown in Figure 2. A machine was coded to make equal-probability random draws of values {-2, -1, 0, +1, +2} for a given number of sources $\{s_1 \ldots s_{11}\}$, sum those values, and repeat the operation.[71] After one million iterations for each number of sources, the frequencies of the different sums were calculated. The result is a series of distribution points that are extremely similar to those displayed in Figure 1, along with an extension of the general pattern to eleven sources. The simulations again indicate that the chance of overall uncertainty continues to decline as sources are added.

FIGURE 1. PROBABILITIES FOR SUMS OF SOURCES WITH EQUALLY LIKELY VALUES {-2, -1, 0, +1, +2}



sum of sources added together

---
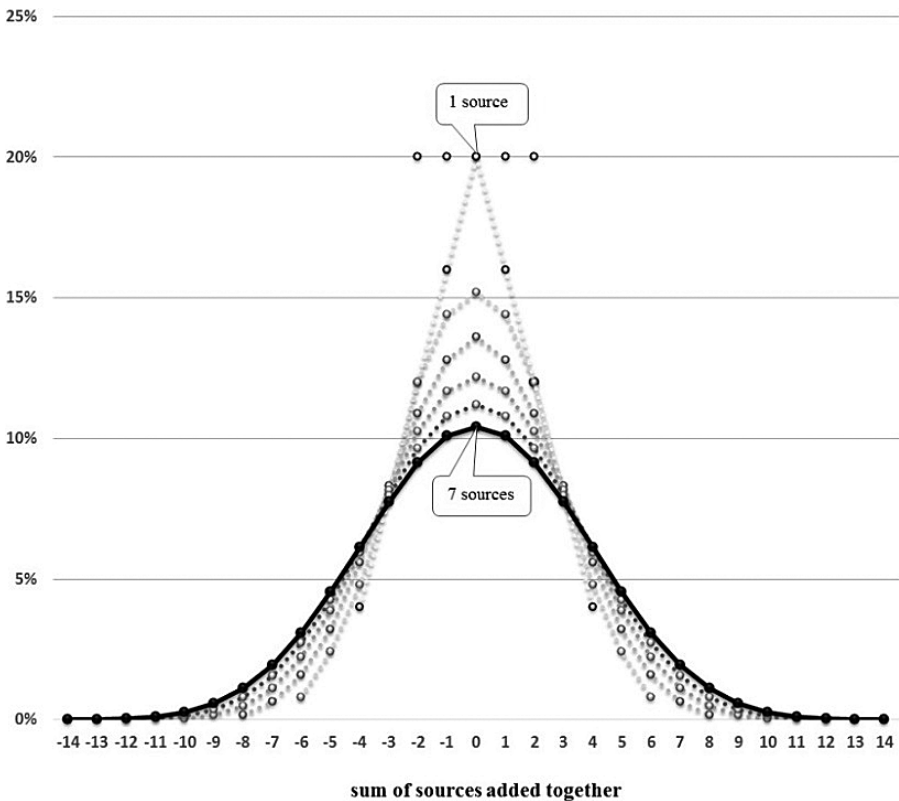
[71] I thank Nathan Noh for writing the machine code, which is available from the author upon request.

FIGURE 2. COMPUTER SIMULATED SUMS OF SOURCES WITH
EQUALLY LIKELY VALUES {-2, -1, 0, +1, +2}



sum of sources added together
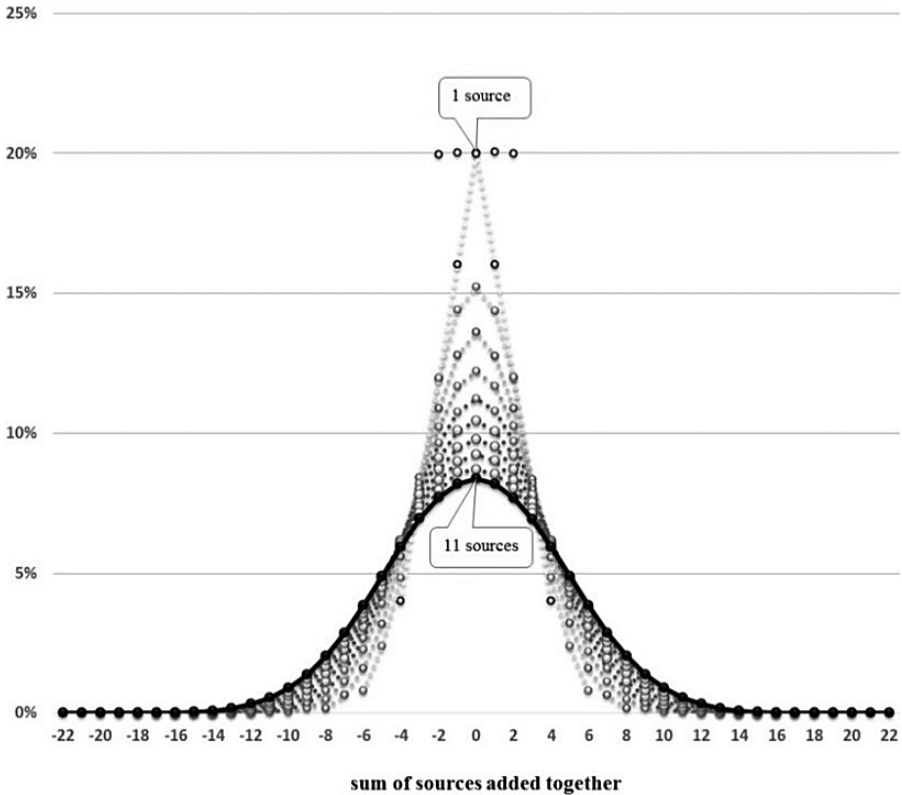
Comfortingly, a similar lesson about declining ties appears in
standard probability theory. A fairly tossed coin has a 50% chance of
landing heads (*H*) and a 50% chance of landing tails (*T*). An odd
number of coin tosses will never end in an equal number of heads and
tails, but the probability of heads and tails equaling out can be calcu-
lated for various even numbers of tosses. For example, with two tosses
there is a 50% chance of an equal number of heads and tails—two out
of four possible permutations (*H-T* and *T-H*, but not *H-H* or *T-T*).
Toss more coins, and this probability falls. The chance of equal num-
bers of heads and tails is 37.5% for four tosses, 31.25% for six tosses,
and then slowly declines toward zero as the number of tosses increases
toward infinity.[72]

---

[72] JOHN HAIGH, TAKING CHANCES 55 & tbl.4.1 (2003); *see also* WILLIAM FELLER, AN
INTRODUCTION TO PROBABILITY THEORY AND ITS APPLICATIONS 238–39 (1st ed. 1950)
(reporting the odds that an even number of heads and tails occurs for the first time at the
*n*th trial). Unlike the probabilities associated with the variables that I have constructed to

Intuitively, the chance of a bunch of coin tosses ending with a precisely equal number of heads and tails is far-fetched. Perhaps the same intuition applies to bunches of five-value variables in the model above. A similar idea has been applied in securities trading, to emphasize a low probability of exactly breaking even from various instances of securities fraud.[73] We should take care to note again that these situations involve the aggregation of source values; they differ from the accumulation of options, such as when an increasing number of criminal laws allow law enforcement officers more options to charge or not.[74] But this logical differentiation is part of the point.

Finally, note that our five-value sources do behave somewhat differently than binary coin tosses. The chance that the *proportion* of heads and tails will *approximately* even out does indeed increase with more coin tosses. The chance of a perfect tie falls with more tosses, but the proportion of heads and tails increasingly clumps toward that 50-50 split.[75] Our five-value source variables do not behave precisely this way. The chance of a (very) close case continues to fall along with the chance of a perfect tie across sources.

### C.   Variations

▪ *Always clear sources.* We have assumed that interpretive sources can themselves be unclear, but what if these sources are like coins with only two sides? Perhaps interpreters are so perceptive (or so filled with certitude) that they always find an implication in every source. Sources that always point in one of two directions alter the analysis, but they do not show the chance of uncertainty increasing as the number of those sources increase. For interpretive sources that can only take the values $\{-1, +1\}$, we get a zigzag pattern in the chance of everything canceling out. The chance of overall uncertainty with one such source is 0%, just as it is for any odd number of such sources. For even numbers of such sources, the chance of overall uncertainty peaks at 50% for two such sources and then declines thereafter toward 0%.[76] Hence the chance of uncertainty increases whenever the

---

model interpretation, there is a standard equation for the probability that the number of heads and tails will be even at the (even-numbered) *n*th toss. *See* HAIGH, *supra*, at 55, 323.

[73] *See* Alicia J. Davis, *Are Investors' Gains and Losses from Securities Fraud Equal over Time? Theory and Evidence* 3 (Univ. of Mich. L. Sch. Scholarship Repository, Law and Economics Working Paper No. 13, 2010), http://repository.law.umich.edu/law_econ_current/art13/.

[74] *See infra* Section II.D (discussing the effects of an accumulation of options, which may facilitate "cherry picking" when such behavior is unconstrained).

[75] *See* HAIGH, *supra* note 72, at 7–9, 54–56 (showing that the chance of falling within a fixed percentage on either side of 50% does increase with more tosses).

[76] *See* Samaha, *Tiebreakers*, *supra* note 15, at 1687–88.

interpreter moves from an odd to an even number of sources, although these increases diminish as sources increase, and the chance of uncertainty drops to 0% whenever the move is from an even to an odd number of sources.

We can give these never-unclear sources more nuance, allowing them to point strongly or weakly in either direction, by taking the values {-2, -1, +1, +2}. Now there is no simple pattern to report. The chance of overall uncertainty with only one such four-value source is again 0%. As more sources are added, this probability moves around from 25%, to approximately 9%, then back up to 14%, then down to just under 10%, then over 10%, and finally back down under 9% for seven such sources.[77] If it were possible to interpret law with only one source of information and if that source were never unclear, those concerned only with uncertainty and discretion could choose radical simplification in interpretation: look at $s_1$ and then stop. But even in this fantasy, we still have not established a pattern of more sources yielding more discretion. The opposite pattern often appears.

  ■ *Always unclear sources.* What if, instead, some source types are always unclear? Perhaps interpreters have so much difficulty understanding the implications of some sources (or interpreters are so filled with epistemic modesty) that zero is the proper value to assign. Justice Scalia's characterizations of legislative history and common-law constitutionalism came close to asserting systematic vagueness.[78] John Hart Ely's rejection of moral reason in constitutional decisionmaking had a similar quality—"[t]here simply does not exist *a* method of moral philosophy," he thought[79]—as do other commentators' doubts that originalist history offers determinate lessons for today's cases.[80]

Subsets of vague interpretive sources do not change the conclusion, however. Adding sources that always take the value zero adds nothing to the overall total. And if every interpretive source were unclear, uncertainty and discretion would not increase as more inference-free sources were added together. Instead interpreters would begin and end with no help from the sources examined. On our model's assumptions, zero-value sources have no influence on case results and so they cannot increase discretion. Yes, we should wonder

---

[77] These calculations and the relevant permutation tables are available from the author upon request.

[78] *See* McDonald v. City of Chicago, 561 U.S. 742, 804–05 (2010) (Scalia, J., concurring) (commenting on one form of common-law constitutionalism that includes moral reason); Scalia, *supra* note 45, at 32–34 (asserting that legislative history is basically unreliable and indeterminate as to collective intent).

[79] John Hart Ely, Democracy and Distrust 58 (1980).

[80] *See* David A. Strauss, The Living Constitution 17–18 (2010) (outlining common criticisms of originalism).

whether judges will manipulate unclear sources, consciously or not, such that those sources appear to support some result; and we should wonder whether judges will ignore or minimize the value of some sources. But not only do spinning and cherry picking violate our working assumptions, those possibilities have a complicated relationship with the number of sources considered—an observation to which we will return.[81]

People who invoke the "looking over a crowd" quip might picture a judge momentarily boxed in to a regrettable result, who then finds a friendly source in a new collection of sources. Of course this might happen, even in our crude model, but this scenario would occur regularly only with additional assumptions. We need reasons to believe that the first set of sources will tie the judge's hands while the second set will be friendly. Why would that be? True, interpreters might identify relatively clear subsets of sources and sequence the consideration of those sources first. But even with this sequencing rule, we could not show that more sources (as in a second subset of relatively less-clear sources) yield more discretion. As long as interpreters aggregate the implications of all sources, uncertainty and discretion would not increase when the second set of sources arrives. More sources yield no more guidance when those sources are unclear.

▪ *Negatively correlated sources.* A different concern about the foregoing logic is that certain subsets of sources systematically point in opposite directions when paired up. Perhaps one category of sources yields -1s predictably while a second category yields +1s predictably. This is a mathematical version of Llewellyn's portrayal of canons as opposing columns of "thrust but parry" moves in a game the outcome of which is determined by other forces.[82] Thickets of judicial precedent might be portrayed likewise.[83] There are narrower versions of these claims, too. On the statutory side, maybe a limited set of canon pairs are likely to be relevant and opposed. This might be so for the surrounding words canon and the consistent meaning canon, which were used against each other in *General Dynamics*.[84] On the constitutional side, maybe there is routine tension between ancient practices and contemporary moral reason.[85] A negative correlation in those

---

[81] *See infra* Section II.D.

[82] Llewellyn, *supra* note 8, at 401.

[83] *See infra* Section III.B.

[84] *See* Gen. Dynamics Land Sys., Inc. v. Cline, 540 U.S. 581, 595–96 (2004).

[85] *See, e.g.*, Adam M. Samaha, *Originalism's Expiration Date*, 30 CARDOZO L. REV. 1295, 1345–46 (2008) (examining arguments for deference to an earlier generation's normative judgments).

implications would not be terribly surprising, to the extent that behavior and values shift hard given time.

But negative correlations offer, at best, mixed logical support for the "more sources, more discretion" idea. Negatively correlated sources largely or entirely cancel out if aggregated, which reduces the number of useful sources from a particular baseline. If we think that losing a pair of negatively correlated sources "increases" uncertainty, from the baseline of no such negative correlation, it is because the interpreter now has *fewer* sources to use, not more. Assuming that there are decision costs for analyzing additional sources, we might want to throw out these couplets in the first place. Or we might find a reason to exclude one half of a negatively correlated pair. This move presumably would reduce decision costs while preserving sources with potentially useful implications for closing cases. But, as with always-unclear sources, the logic of "more sources, less discretion" still holds. We need other logic to believe that adding a source systematically increases the chance of adding a source that is negatively correlated with the implication of the preexisting set. That logic is elusive. A simple commitment to preserving one side of a conflicting canon pair, or favoring ancient history over modern values or vice versa, does nothing to show that more interpretive sources yield more uncertainty and discretion.

   ▪ *Lexically inferior tiebreakers.* What happens when an interpretive source is added in a lexically inferior tiebreaking position? That is, what is the effect on the chance of overall uncertainty when an additional source is considered if but only if the other sources yield uncertainty? Interpretive doctrine is sometimes lexically ordered in this fashion, such as when the Supreme Court indicates that legislative history may only be used to resolve residual statutory ambiguity;[86] that judges should defer to agency interpretations only when the statute is by some measure unclear;[87] or that the rule of lenity does not apply unless a number of other interpretive sources yield some measure of vagueness.[88]

The short answer is that, with a lexically inferior tiebreaker, more sources once again produce lower chances of overall uncertainty. The

---

[86] *See, e.g.*, Milner v. Dep't of the Navy, 562 U.S. 562, 574 (2011) ("Legislative history, for those who take it into account, is meant to clear up ambiguity, not create it.").

[87] *See, e.g.*, *General Dynamics*, 540 U.S. at 600 ("[D]eference to [the agency's] statutory interpretation is called for only when the devices of judicial construction have been tried and found to yield no clear sense of congressional intent.").

[88] *See, e.g.*, Moskal v. United States, 498 U.S. 103, 108 (1990) ("[W]e have always reserved lenity for those situations in which a reasonable doubt persists about a statute's intended scope even *after* resort to 'the language and structure, legislative history, and motivating policies' of the statute." (citations omitted)).

difference is that the decline is more rapid. Compared to adding another source to the mix, reserving that same source for a tiebreaker drives down the chance of overall uncertainty faster.[89] If the tiebreaker always points in one direction, such as lenity or agency deference, then the chance of overall uncertainty drops to zero.[90] There are many difficult trade-offs surrounding tiebreaking decision structures, and there are serious implementation challenges; these subjects deserve separate and extensive treatment. The point for now is that, for many variations, the basic logical analysis holds and indicates that more interpretive sources tend to reduce uncertainty.

■ *Lawmaking.* A more serious concern involves the assumption that interpreters face only two options. This assumption is in accord with a narrow dispute-resolution vision of the judicial role, which is only partly accurate. Courts regularly elaborate law's meaning into doctrine, which might take any number of specific forms. If judicial doctrine can influence behavior in or outside courts, the simplifying assumption of only two options seems problematic. Most important for present purposes, the number of options for specifying law might be related to the number of interpretive sources considered. The theory here is not airtight but perhaps consulting interpretive sources can be an inspiration for law-elaboration. William Eskridge maintains that legislative history has "hermeneutical value" in the sense of "open[ing] up the judicial mind to possibilities that might not have occurred to the judge."[91] The same thought generalizes to other interpretive sources, from dictionaries to precedent to founding era documents.

Take the marquee debate over the President's removal authority in *Myers*.[92] An interpreter might glance at Article II's vesting of "[t]he executive Power"[93] and suppose that the President retains exclusive authority to fire executive officers. Perhaps little creative thought is needed to suppose, instead, that the President's role in removing executive officers is an issue for Congress to decide through ordinary legislation.[94] After digging up some history, however, the interpreter might

---

[89] Samaha, *Tiebreakers*, *supra* note 15, at 1691–94.

[90] *Id.* at 1694.

[91] Eskridge, *supra* note 11, at 562 (emphasis omitted). I confess that I generally do not understand assertions that include the word "hermeneutic," but Eskridge's lucid discussion is a welcome exception.

[92] Myers v. United States, 272 U.S. 52, 176 (1926) (invalidating a statutory restriction on the President's power to remove first-class postmasters).

[93] U.S. Const. art. II, § 1.

[94] *See id.* art. I, § 8, cl. 18 (enumerating Congress's "necessary and proper" power); *Myers*, 272 U.S. at 183–84 (McReynolds, J., dissenting) (invoking this clause to support statutory removal restrictions).

suddenly see a plausible position in favor of the Senate having to consent to the removal of officers of the United States, on the logic that the process for removal should mirror the process for appointment in the absence of clear directions otherwise. Alexander Hamilton said so, before ratification.[95] Not much more historical research will reveal, perhaps shockingly, an alternative position under which impeachment is the sole lawful process for removing civil officers of the United States.[96] Consideration of numerous historical sources is not strictly necessary to develop all of these positions. But sustained research has a way of surprising open-minded readers.

More doctrinal possibilities do not necessarily yield more discretion, however. The question is how the interpreter will choose. For one more go, let us hold to our other assumptions. If we assume that sources are equally likely to offer any of a range of implications about a range of doctrinal specifications, and that all implications will be aggregated, we might find the same pattern of decreasing uncertainty as the number of sources increase. In fact, this is the pattern after making one change to our starter model.

Assume that interpretive sources take values on three dimensions instead of two, with each dimension representing a different specification for the law. To simplify, assume that each source takes the values {0, +1, +2} with equal likelihood for each of three possible specifications for law. The values for each dimension can be totaled across all sources and compared to see whether one specification emerges as the most supported, and the effect of adding sources on the probability of a tie for first place.[97] Doing so, we find a familiar pattern: The chance of no unique winner among the three possible specifications of law *drops* as we add sources—from approximately 44% with one three-dimensional source to approximately 32% with two such sources, then to approximately 27%, 24%, 20%, and 18% as we move up to seven such sources. At this point the number of permutations is enormously

---

[95] *See* THE FEDERALIST NO. 77, at 459 (Alexander Hamilton) (Clinton Rossiter ed., 1961) ("The consent of [the Senate] would be necessary to displace as well as to appoint.").

[96] *See* U.S. CONST. art. II, § 4 (stating conditions on which impeached and convicted officers "shall be removed"); 1 ANNALS OF CONG. 387 (1789) (Joseph Gales ed., 1834) (remarks of Rep. Smith) (expressing doubt about presidential removal and urging consideration of the impeachment-only position); *id.* at 477 (remarks of Rep. Huntington) (similar).

[97] For the suggestion that legal interpreters tend to select the best-supported alternative interpretation, along with a claim that this approach has been adopted without adequate care, see Lawson, *Proving the Law*, *supra* note 70, at 890–91.

large (over ten billion) and there is not much use in proceeding fur-
ther here.[98]

This is only a start. The ideal matrix of sources and possible
values might be more complex, and the foregoing three-dimensional
model is already complicated. Each source now has twenty-seven per-
mutations across the three possible specifications for law. Further-
more, any such model depends on a realistic probabilty that a source
will show support for more than one specification. What is the
number? Perhaps a source that shows support for one specification is
most likely to show no support for a second; and perhaps sources rele-
vant to the same dispute are most likely to offer implications for only
a few specifications. A useful model beyond two dimensions is plainly
difficult to build and confident generalizations about the relevant pat-
terns will have to await future work. In any event, a model for law-
making options will not resolve concrete cases. Adjudication requires
the application of law, and our three-dimensional model cannot
accomplish this. We probably must return to the earlier two-option
model, anyway, which is designed for choosing case results.[99]

## D.   *Deviations*

Strict interpretive rules do not exist for every contested case, of
course, and people do not always follow the interpretive norms that
do exist. Commentators have a range of views about the extent to
which different patches of interpretive method are constraining,[100] but

---

[98]  The proportion of zeros changes in this three-dimensional model: On each dimension
there is a 1/3 chance of zero on each draw, whereas our two-dimensional sources have a 1/5
chance of zero on each draw. A spreadsheet showing the one-source ($3^3 = 27$ permutations)
results and the two-source ($27^2 = 729$ permutations) results is on file with the author. Excel
cannot process five such sources ($27^5 = 14,348,907$ permutations). The results for five, six,
and seven such sources are based on two *C* programs written by Nathan Noh. These
algorithms produce results that match the results from permutation tables with one, two,
three, and four such sources. The machine code and the results are available from the
author upon request.

[99]  *See* Lewis A. Kornhauser, *Modeling Courts*, *in* Theoretical Foundations of Law
and Economics 1, 1–2 (Mark D. White ed., 2009) (distinguishing "case space" models,
which are composed of facts and in which judges might announce a policy but necessarily
adjudicate a particular case, from "policy space" models that are popular in political
science (emphasis omitted)); *see also* Scott Baker & Pauline T. Kim, *A Dynamic Model of
Doctrinal Choice*, 4 J. Legal Analysis 329, 332 (2012) (conceptualizing "legal rules as
guiding which cases (bundles of facts) should result in what outcomes").

[100]  *See, e.g.*, Jerome Frank, Courts on Trial: Myth and Reality in American
Justice 279 (1949) ("For precedential purposes, a case, then, means only what a judge in
any later case says it means."); James J. Brudney, *Canon Shortfalls and the Virtues of
Political Branch Interpretive Assets*, 98 Calif. L. Rev. 1199, 1231–32 (2010) ("[J]udges
who regularly rely on the canons have license to employ a systemic kind of discretion, in
contrast to judges who regularly invoke legislative history or agency deference."); Scalia,
*supra* note 45, at 35–36 ("Since there are no rules as to how much weight an element of

nobody should doubt that current interpretive norms leave a degree of discretion in a subset of cases. Similarly, commentators undoubtedly have a range of views about the extent to which judges try or are able to comply with prescribed interpretive norms, but nobody should doubt that interpreter compliance is imperfect.

A more realistic model of interpretation will allow a measure of *ruleless* method and *willful* interpreters. Adding another dimension of realism, we can separate two ways in which method can be ruleless and interpreters willful: *spinning* and *cherry picking*. Spinning involves an interpreter choosing to assign a preferred implication to a source. Picking involves an interpreter choosing a preferred source for emphasis while deemphasizing or ignoring other sources. Our starter model of interpretation disallowed both of those choices. It combined what amounted to a no-spinning rule (i.e., dictating equal probability across implications) with a no-picking rule (i.e., dictating that all source values be totaled). We can relax these strictures and see what happens to discretion as sources are added. The decisionmaking process might include spinning or picking or both; and spinning or picking might be the product of ruleless method or willful interpreters or both. Seeing these phenomena will complicate—and illuminate—the logical influence of source proliferation on discretion. One lesson will be fairly consistent, though: In generating interpreter discretion, the power to spin dominates the power to pick.

### 1.   Ruleless Spinning and Picking

If there were no rules for interpretation, the process would be saturated with discretion regardless of how many interpretive sources were available. Interpreter discretion would remain at a stable and high level whether the interpreter decided to consider one source or one hundred. There would be no interpretive rules to break, willfully or otherwise.

Now suppose that rule-following interpreters face a rule against cherry picking sources but no rule against spinning their implications. Interpreters will faithfully consider every relevant source but each source's implication is subject to the interpreter's control—such as the interpreter's preferred level of abstraction for the source,[101] or the interpreter's consideration of "context" in its most vacuous and elastic

---

legislative history is entitled to, it can usually be either relied upon or dismissed with equal plausibility.").

[101] *Cf.* SCALIA & GARNER, *supra* note 11, at 18–19 (characterizing purposivist statutory interpretation as subject to discretionary level-of-abstraction choices); Max Radin, *Statutory Interpretation*, 43 HARV. L. REV. 863, 876 (1930) (emphasizing that statutes have both immediate and ultimate purposes).

sense.[102] The same conclusion holds, actually: The number of sources will not affect interpreter discretion, which will remain stable and high. The unconstrained ability to spin sources, to and fro, as the interpreter desires means that discretion will enter the equation for every source and for every number of sources, even with no cherry picking. More sources would mean more sources to spin, and this might make the interpretive process more costly to conduct. But so far we have no reason to believe that this decision-cost increase would arrive with any increase in interpreter discretion.

In contrast, allowing cherry picking has a different and less powerful effect on discretion. If rule-following interpreters face a rule against spinning but no rule against cherry picking, the number of interpretive sources does indeed matter to ex-ante probabilities. And, unlike the fully rule-governed interpretive process in our starter model, increasing the number of interpretive sources *increases* the probability of interpreter discretion when cherry picking is allowed but spinning is prohibited. An interpreter who wants to have a choice over two possible results, for example, obviously must see at least two unspun sources to yield those two different implications.[103] An interpreter who already prefers one of two results might see the preferred result in only one unspun source, but increasing the number of unspun sources obviously increases the likelihood of seeing a friendly source in the growing crowd.

The trajectory of cherry-picking-based discretion can be estimated, this time with stock probability equations. The probability that an event with probability $p$ will occur at least once by the $n$th try is the complement of the probability that the event will not occur on any of $n$ tries—which is expressed as $1 - (1 - p)^n$ and which increases as $n$ increases.[104] So suppose that the interpreter prefers a plaintiff victory. On the assumptions from our earlier model, $p = 0.4$ (2/5 or 40%) that any one source either strongly supports (-2) or weakly supports (-1) the plaintiff's position. Hence the probability of one of these friendly sources not appearing on any given try is 0.6 (3/5 or 60%). The probability of finding at least one friendly source increases with each additional source examined, albeit more slowly with each successive

---

  [102] *Cf.* Lorillard Tobacco Co. v. Reilly, 533 U.S. 525, 542–46 (2001) (using "context" as a label for a variety of factors that might push against the ordinary understanding of statutory text). Readers need not be convinced that contextual factors or level of abstraction choices produce uninhibited interpreter discretion across all interpretive sources.

  [103] To simplify, I have left out the possibility of the first unspun source taking the value zero. The interpreter in this hypothetical is looking for two sources that point in two different directions, from which to pick.

  [104] *See* HAIGH, *supra* note 72, at 329–30.

FIGURE 3. PROBABILITY OF A FRIENDLY SOURCE IN THE CROWD,
WITH CHERRY PICKING BUT NO SPINNING AND EQUALLY
LIKELY SOURCE VALUES {-2, -1, 0, +1, +2}
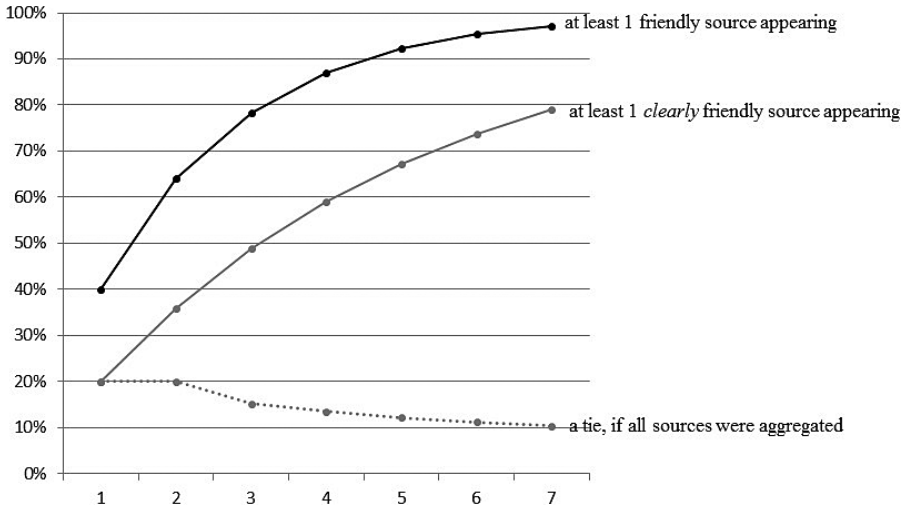


try. The chance of the first source being friendly is only 0.4. By the second try, the chance of at least one friendly source appearing jumps to $1 - 0.6^2 = 0.64$. By the third try, the number rises to $1 - 0.6^3 = 0.784$. By the fourth try, the number hits $1 - 0.6^4 = 0.8704$. And so on up.[105]

Charting the numbers through seven sources reveals stark contrasts (Figure 3). The top curve represents cherry picking but no spinning; the interpreter's chance of finding a friendly source increases rapidly as the number of sources increases. The middle curve represents a variation on cherry picking among unspun sources, where the interpreter waits for a clearly friendly source (-2). In contrast, the bottom curve represents our starter model of no cherry picking and no spinning, under which the chance of overall uncertainty and residual discretion decreases as the number of sources increases.

---

[105] A similar contrast holds for cherry picking where the interpreter waits to see the full set of possible implications among unspun sources, instead of looking for an already preferred source to appear. The math is more complicated, but the basic idea is that seeing $n$ implications requires at least $n$ sources, and that the probability of seeing all the implications increases as the number of sources increases. For instance, using our five-value variables {-2, -1, 0, +1, +2}, the probability of seeing at least one source favoring the plaintiff and also at least one source favoring the defendant is 0% with one source and about 58% with three sources. The relevant equation is for the chance of not seeing only non-positive or only non-negative sources, which in our case is $1 - \left(\frac{3}{5}\right)^n + \left(\frac{3}{5}\right)^n - \left(\frac{1}{5}\right)^n$, where the last term subtracts the chance of only zero values, which is already covered by the union of the preceding two terms.

The chart also helps show the stakes for developing an interpretive method that prevents cherry picking. Those stakes are low. If interpretive rules allow interpreters to spin sources, then rules against cherry picking sources mean little or nothing to the interpretive process and its results; discretion to spin will swamp lack of discretion to pick, and wash away the significance of the number of sources. Indeed we should wonder whether strict interpretive rules can be developed for drawing implications from every variety of interpretive source; and, just as important, if interpretive rules can stop spinning, why in the world can they not stop cherry picking, too. But if we ignore these complications and assume that the rules will constrain spinning and not picking, then minimizing the number of sources will reduce interpreter discretion. This slice of the analysis is consistent with Judge Leventhal's quip and the notion that more sources yield more discretion.

But even on these awkward assumptions, controlling the number of interpretive sources is not a powerful tool for controlling discretion. The number of unspun sources must be kept quite low to restrict the cherry-picking interpreter's options, using our assumptions. In the top-line numerical example shown in Figure 3, the chance of the interpreter finding a friend is already over 60% by the second unspun source and nearly 80% by the third. In the middle-line example, where the interpreter must find a clear friend, controlling the number of sources is more effective in controlling discretion but still loses its grip. By the third unspun source, the interpreter has a nearly 50% chance of finding a source clearly pointing in the preferred direction, and nearly 60% by the fourth source.

The lower bound on the number of interpretive sources in a real-world case is not immediately apparent and depends on how "sources" are individuated.[106] But no available interpretive method relies on only one source in any useful sense. Prevailing interpretive practices are open to multiple dictionaries and more than one kind of legislative history, for example.[107] For a contested and well-lawyered case, the number of individuated source citations will be far north of three. *General Dynamics* and *Myers* are just two high-profile examples among countless litigated cases. Reducing the number of sources might be effective for reducing interpreter discretion in busy lower courts facing relatively low stakes decisions (again assuming that spinning is controlled but cherry picking is not).[108] The same thought

---

[106] *See supra* note 43.

[107] *See supra* notes 44–48 and accompanying text.

[108] *Cf.* Chris Guthrie et al., *Blinking on the Bench: How Judges Decide Cases*, 93 CORNELL L. REV. 1, 5 (2007) (concentrating on trial judges and contending that they

applies when one implication is highly unlikely to appear in any source, in which case a relatively small number of source draws is unlikely to show a spread of options. But otherwise the prospects are poor for constraining discretion by limiting sources.

### 2. *Willful Spinning and Picking*

What if spinning and cherry picking are products of willful judging, instead of gaps in the rules for interpretation? Even with formally strict interpretive rules, an interpreter might manipulate the implications of sources whether by resort to a convenient level of generality or contextual factor or otherwise; and an interpreter might hunt for opportunities to support their preferences with inarguably supportive sources. Sympathetic observers and others might resist the thought that sitting judges engage in any spinning or picking consciously. But subconscious or unconscious mental processes, including motivated reasoning and coherence-forcing frameworks, would yield similar effects. I will use "willful" broadly enough to include all of these possibilities.

Like rulelessness, willfulness does reconnect source proliferation to discretion, but imperfectly. If willfulness includes both spinning and cherry picking, then, once again, the number of interpretive sources would be irrelevant to interpreter discretion. Discretion would remain high and stable regardless of the number of sources. This would be true regardless of which set of interpretive rules the legal system (unsuccessfully) imposed on willful interpreters. For the number of sources to matter, we must imagine interpreters who are not entirely willful. They must be constrained somehow. True, some people believe that interpreter constraint comes from emphasizing those interpretive sources that tend to have clear implications or more objective bases for evaluation. Textualists may promote ordinary understandings of statutory text, for example, while others may believe that legislative history ties the hands of interpreters. But those arguments are beside the point here. Those arguments pertain to how much discretion is allowed by different *kinds of sources*, not the influence of different *numbers of sources*.

Assuming that willful interpreters can be constrained in part, controlling their ability to spin is more significant than controlling their ability to pick. This point follows from the analysis above, where we

usually decide intuitively but sometimes override intuition with more careful deliberation); Dan Simon, *A Third View of the Black Box: Cognitive Coherence in Legal Decision Making*, 71 U. Chi. L. Rev. 511, 515–16 (2004) (discussing simplifying responses to complex decision tasks, but recognizing limits on the influence of coherence-based reasoning).

assumed that interpretive rules were followed: The ability to spin any given source is more powerful in producing interpreter discretion than the ability to pick among unspun sources. Willfulness would migrate from picking to spinning each source, and the number of sources would still be irrelevant to the degree of discretion. The key question for dealing with willfulness is whether interpretive rules can be developed and enforced to discipline spinning, or spinning plus cherry picking, when interpreters are not otherwise inclined to follow those rules.

A monitoring theory might be the best prospect for reconnecting source proliferation to discretion where interpreters are somewhat willful. The general idea is that some situations are easier for outsiders to monitor for compliance than others, and that monitoring more activity is more costly and more difficult than monitoring less activity.[109] Sometimes a given interpretive method is promoted along these lines, such as when textualism is offered as a more transparent method of judicial analysis that outsiders may double check with relative ease.[110] The pertinent version of this monitoring argument is not commonplace, but the idea would be that the proliferation of interpretive sources makes it more difficult for outsiders—lawyers, potential dissenters on shared panels, judges higher in the hierarchy, or any other observers with influential opinions—to check whether interpreters are drawing inferences appropriately and integrating inferences convincingly. More sources might mean more discretion because more sources mean less monitoring for compliance with interpretive rules, appropriately developed.[111]

---

[109] For one of many synopses in the law literature, see Elizabeth S. Scott & Robert E. Scott, *Parents as Fiduciaries*, 81 VA. L. REV. 2401 (1995). "Monitoring arrangements allow the principal to supervise the agent's actions so as to detect and sanction agent conduct in pursuit of selfish ends." *Id.* at 2421. A persuasive monitoring theory depends on which audiences judges care about. On appellate courts as monitors of trial judges, who might care about reputation or the burden of remands, see Christina L. Boyd, *The Hierarchical Influence of Courts of Appeals on District Courts*, 44 J. LEGAL STUD. 113, 115–16 (2015). On fellow panel members as monitors for manipulation of doctrine, when the panels are ideologically mixed, see Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155, 2156 (1998).

[110] *See* SCALIA & GARNER, *supra* note 11, at 16–17 (contending that textualist interpretation is more objective, more transparent, and easier to check for transgression into the imposition of personal preference).

[111] Outsiders are able to "monitor" interpreter behavior even if there are no interpretive rules. The monitoring would be conducted on some other measure of performance, such as socially beneficial consequences apart from compliance with interpretive norms within law in another sense. Here, I am assuming the value of an interpretive method within law and not defending or specifying much of its content, to explore the relationship to source proliferation and discretion (measured against the

The idea has logical traction. Limiting the number of interpretive sources would seem to decrease the burden for each monitor in detecting whether individual sources are being spun or over-weighted against a given baseline of interpretive rules. The monitors might be willful themselves, but a thriving monitoring process should produce useful information for many interested parties. Furthermore, if interpretive rules are especially hard to develop for drawing inferences from various sources, a relatively small number of sources might well ease the challenge of monitoring for improper spinning.

TABLE 1. LOGICAL RELATIONSHIPS BETWEEN SOURCE
PROLIFERATION AND INTERPRETER DISCRETION

|  | *Faithful Interpreter* | *Willful Interpreter* |
|---|---|---|
| *No Spinning, No Picking* | more sources, less discretion | *fixed monitoring*: more sources may have no effect on discretion; *variable monitoring*: more sources have an unclear effect |
| *No Spinning, Picking* | more sources, more discretion | *fixed monitoring*: more sources may have no effect on discretion; *variable monitoring*: more sources, more discretion |
| *Spinning, No Picking* | discretion regardless | discretion regardless |
| *Rulelessness* | discretion regardless | discretion regardless |

That said, the relationships among monitoring, source proliferation, and discretion can be complex (Table 1). Consider a legal regime with rules against *both* spinning and cherry picking (row 1), but with willful interpreters who might violate either rule unless they face effective monitoring (column 2). Even on the foregoing suppositions about better monitoring with fewer sources (variable monitoring, let us call it), the net effect of more sources on interpreter discretion remains decidedly unclear. Two forces would pull in different directions. On one side, monitoring would make these interpretive rules more effective when the number of sources is held down. On the other side, those same interpretive rules would be less able to reduce discretion when the number of sources is held down. This is the reverse lesson from the analysis in Section I.B above: As interpretive sources are reduced, discretion opens up when interpreters follow rules

against spinning and picking.[112] We cannot know which effect will dominate as the number of sources increases—the disciplining effect from interpretive rules or the liberating effect from laxer monitoring. For that we need additional theory, different assumptions, or behavioral evidence.

The power of monitoring also is subject to broader doubts, in addition to any insulating effects of salary and tenure protection in some courts and the fact that the risk of reversal and remand often is low and far in the future, for those judges who care. Another doubt involves the cause of deviations from interpretive norms. If willfulness is more subconscious than conscious, we may doubt the effectiveness of outsider monitoring. Moreover, we might consider the possibility that increasing the number of interpretive sources makes willful spinning harder to achieve, whatever the monitoring effects. Having to spin more sources is more laborious than having to spin fewer sources. This cost might suggest an internal damper on spinning, at least if cherry picking is constrained.

The analysis is rapidly entering the field of judicial preferences and incentives. There is more hard work to do in that field. What can be shown here already is a progression toward more realistic pictures of interpretation, more complications, and now a need for evidence.

## III

### EVIDENCE

The logic above might not match behavior in a world of limited resources, mistakes, and other influences. Because judges are a special kind of professional and because a given case might not trigger strong extralegal commitments, it is not soft-headed to think that methodological prescriptions will influence judicial decisions in part of the docket, even as the amount of information increases. Unfortunately we lack decisive evidence to mark these domains. Existing empirical studies have nothing like the precision or behavioral targets necessary to test the logical relationships compiled above. For the time being, we have to take small steps. One way forward is to pull together suggestive existing research involving a variety of informational sources, extend and reanalyze the available data, and build up testable working theories for what happens to results when interpreters face more rather than fewer interpretive sources.

---

[112] To repeat, dropping the no-spinning rule will yield discretion regardless of the number of sources. In that situation, monitoring cannot reduce interpreter discretion.

## A. *No Effect or Arbitrariness*

One jarring possibility, not to be forgotten, is that increasing the number of interpretive sources will have no influence on results. No effect on the amount of judicial discretion, no effect on the influence of ideology, no other patterned effect on case results.

First off, law might provide fairly tight rules that decisionmakers follow. In voting rights litigation, a field that is no stranger to ideological forces, relatively specific rules appear to guide judges in a patch of this field. Thus, Adam Cox and Tom Miles could not confirm an ideological influence on three arguably rule-like threshold factors under the test for vote dilution.[113] The difference in voting behavior between Democratic and Republican appointees hovered under 10% and was not statistically significant.[114] But the authors reported a significant and large difference between Democratic and Republican appointees on the totality-of-the-circumstances part of the doctrine. In this doctrinal area, the difference in voting behavior across judges ranged from 27% to 48% depending on the regression specification,[115] although the difference plunged and lost statistical significance in later years of the dataset.[116] Such studies support the mainstream view that rules constrain decisionmakers more than standards, whatever the other trade-offs.[117] When the judicial process has rules and basically

---

[113] *See* Adam B. Cox & Thomas J. Miles, *Judicial Ideology and the Transformation of Voting Rights Jurisprudence*, 75 U. CHI. L. REV. 1493, 1532–33 (2008). The three threshold factors, which can be converted into numerical tests, are (1) sufficient size and compactness of the minority group in a single-member district, (2) political cohesiveness, and (3) sufficient bloc voting among the majority to usually defeat the minority's preferred candidate. *See* Thornburg v. Gingles, 478 U.S. 30, 50–51 (1986).

[114] *See* Cox & Miles, *supra* note 113, at 1533.

[115] *See id.* at 1533–34. The nonthreshold factors under *Gingles* include, among others, "the history of voting-related discrimination in the State or political subdivision" and "the extent to which minority group members bear the effects of past discrimination . . . which hinder their ability to participate effectively in the political process." 478 U.S. at 44–45. The nonthreshold factors are more numerous than the threshold factors. It is possible that ideological influence increased with the *number* of factors faced by judges, independent of a *rules/standards* effect. But we cannot disentangle these influences with this dataset, to the extent that the threshold factors are rule-like and the nonthreshold factors are standard-like. In addition, because the threshold factors are supposed to be satisfied before the nonthreshold factors are considered, the authors note that perhaps "more easy than hard cases are screened out at the first step," or perhaps some judges vote insincerely at the first stage to avoid blocking the consideration of additional factors. Cox & Miles, *supra* note 113, at 1514–15. Testing the effects of source proliferation is cleaner when there is no lexically ordered decision procedure.

[116] *See* Cox & Miles, *supra* note 113, at 1534–35 (emphasizing changes in voting behavior after 1994).

[117] *See, e.g.*, VERMEULE, *supra* note 12, at 68 (itemizing trade-offs including the amount of information available for use at the time of application and the attendant risk of error, understood as a departure from the best all-things-considered result).

follows them, we should expect that more information means more work for judges—not more discretion.

Even when rule-driven adjudication is unrealistic, adding more sources still might not produce a pattern of influence. Information can be ignored when time is short.[118] However much information the law requires or allows judges to consider, and however much information attorneys and parties dump into our court systems, there is no guarantee that all of it will be considered. Judiciaries often do follow norms of addressing major arguments,[119] but these norms cannot ensure that all legally relevant information is actually taken into account. Written judicial opinions, even exhaustive opinions, are not a full corrective. Many judicial decisions are not subject to norms of written justification,[120] and a written opinion will not necessarily reflect the process of decision anyway.[121] Most judicial opinions, and perhaps especially the exhaustive opinions, are written by law clerks who might have no power over case results.[122] Regardless, imperfectly dutiful decisionmakers will stop reading and pondering after a certain point. Below we will see a potential example of more information having no demonstrable effect, all else equal: the number of sources cited in a sample of appellate briefs. The chance of reversal does not seem to vary with the number of sources cited in support of the appellant's position.[123]

---

[118] *See* Lisa D. Ordóñez et al., *Time-pressure Perception and Decision Making*, *in* 2 THE WILEY BLACKWELL HANDBOOK OF JUDGMENT AND DECISION MAKING 519, 522 (Gideon Keren & George Wu eds., 2015) (collecting studies indicating that decisionmakers sometimes cope with time constraints by filtering out information and limiting the option set); *cf.* Alan R. Dennis, *Information Processing in Group Decision Making: You Can Lead a Group to Information, but You Can't Make It Think*, 1993 ACAD. MGMT. BEST PAPERS PROC. 283, 285–86 (finding that a group connected electronically in a kind of wiki exchanged approximately 50% more information than a control group, but did not make better decisions and remembered less exchanged information).

[119] *See* Mathilde Cohen, *When Judges Have Reasons Not to Give Reasons: A Comparative Law Approach*, 72 WASH. & LEE L. REV. 483, 531 (2015) ("[F]ederal judges at all levels of the judicial hierarchy are prolific reason-givers.").

[120] *See id.* at 526–27 (observing that federal judges, unlike some state judges, are usually under no formal legal duty to explain their decisions); *see also* Taylor v. McKeithen, 407 U.S. 191, 194 n.4 (1972) ("[T]he courts of appeals should have wide latitude in their decisions of whether or how to write opinions.").

[121] *See* RICHARD A. WASSERSTROM, THE JUDICIAL DECISION: TOWARD A THEORY OF LEGAL JUSTIFICATION 25–27 (1961) (conceptually separating the process by which decisions are reached from the process by which they are justified).

[122] *See* G. Mitu Gulati & Richard A. Posner, *The Management of Staff by Federal Court of Appeals Judges* 4–7 (Univ. of Chi., Public Law Working Paper No. 531, 2015), http://ssrn.com/abstract=2590179 (reporting the results of interviews of approximately one quarter of U.S. Court of Appeals judges on their opinion-drafting practices).

[123] *See infra* Section III.C and Table 3.

Even when decisionmakers do consider additional information, the effect on decisions might lack rational patterns. Decisionmaking might become more arbitrary as the information load increases.[124] Arbitrariness can be operationalized in various ways, but one approach is to look for patterns that defy a preferred version of rational decisionmaking. For instance, in many decision situations, the sequence in which information is considered should, ideally, have no effect on results. This is true for totaling up numbers and for Bayesian updating.[125] Yet a tall stack of behavioral research shows order effects in various settings. Most relevant here, long lists of information can trigger a primacy effect in which earlier received information becomes more influential.[126] True, order effects may arise even for short lists of information,[127] and any effect can be highly sensitive to the decision environment.[128] Further, order effect studies have not yet been conducted on actual judges.[129] Existing behavioral evidence is just enough to raise the idea.

A related idea involves errors. A substantial body of research indicates that, when the informational load exceeds a given level, the decisionmaker's error rate may increase.[130] Much of this work is

---

[124] *See* Martin J. Eppler & Jeanne Mengis, *The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines*, 20 INFO. SOC'Y 325, 331, 333 & tbl.4 (2004) (listing, among the symptoms of information overload, "[a]rbitrary information analysis and organization" including ignoring information, losing control over information, superficial analysis, loss of differentiation, overestimation of peripheral cues, and misinterpretation); Barry Schwartz et al., *Maximizing Versus Satisficing: Happiness Is a Matter of Choice*, 83 J. PERSONALITY & SOC. PSYCHOL. 1178, 1179 (2002) ("To choose the best among 30 options is truly daunting. So rather than even try, people may disengage, choosing almost arbitrarily to complete the process.").

[125] *See* Adam M. Samaha, *Starting with the Text—On Sequencing Effects in Statutory Interpretation and Beyond*, 8 J. LEGAL ANALYSIS 439, 456–57 (2016).

[126] *See id.* at 463–64 (drawing from Robin M. Hogarth & Hillel J. Einhorn, *Order Effects in Belief Updating: The Belief-Adjustment Model*, 24 COGNITIVE PSYCHOL. 1, 6, 17–19, 38 (1992)). Declining attention is one of the mechanisms for triggering primacy. *See id.*

[127] *See id.* at 464–65.

[128] *See id.* at 465–69.

[129] *See id.* at 467. Sometimes judges do show decision patterns similar to those of nonprofessionals. *See* Chris Guthrie et al., *Inside the Judicial Mind*, 86 CORNELL L. REV. 777, 784 (2001) (presenting findings); Avani Mehta Sood, *Motivated Cognition in Legal Judgments—An Analytic Review*, 9 ANN. REV. L. & SOC. SCI. 307, 315–16, 319 (2013) (reviewing literature).

[130] *See, e.g.*, Melvin Aron Eisenberg, *The Limits of Cognition and the Limits of Contract*, 47 STAN. L. REV. 211, 214 (1995) ("[I]mperfections in human processing ability increase as decisions become more complex . . . ."); Ordóñez et al., *supra* note 118, at 522 (collecting studies indicating that time-pressured decisionmakers sometimes screen out information and exhibit more rigidity and less creativity, although suggesting that this effect is possibly ameliorated by training and experience).

experimental and involves settings far distant from the institutional environment of judges, who must interpret statutes or decide constitutional claims.[131] However, consider an early field study conducted by Charles O'Reilly in county welfare offices, which have a passing resemblance to busy trial courts. County employees were asked questions designed to measure their felt informational load with respect to welfare applications. Meanwhile, the decisions of many of these employees had been audited for errors in granting benefits to ineligible applicants or in granting excessive benefits.[132] Employee reports of heavier informational loads correlated in a statistically significant way with higher error rates detected in the audits.[133]

Of course, self-reported feelings like these do not necessarily reveal actual information use. Also, formulating objective tests for errors in judicial decisions is a major challenge, making it difficult to apply this research design to court systems. Nonetheless, existing research is sufficient to put increased errors on the list of plausible effects when interpretive sources increase.[134]

## B.  Ideology

Few of the studies noted above were aimed at courts, but other working theories can rely on concentrated efforts to study judicial behavior. One of these theories accords with Judge Leventhal's quip. We lack an easily quantified conception of judicial discretion, but we do have familiar, if crude, numerical proxies for judicial ideology. Increased ideological influence is one consequence of increased inter-

---

[131] Compare consumer choice studies, in which a choice is not required and where there is evidence that some consumers are deterred from purchasing when the option set increases. *See, e.g.*, Sheena S. Iyengar & Mark R. Lepper, *When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?*, 79 J. PERSONALITY & SOC. PSYCHOL. 995, 996–97 (2000) (reporting results of field studies involving, for instance, exotic jams). Also, for a study that separates the number of *options* from the number of option *attributes*, see Naresh K. Malhotra, *Information Load and Consumer Decision Making*, 8 J. CONSUMER RES. 419 (1982). That research found that increasing either options or attributes was independently associated with decreasing accuracy, measured by respondents' own preferences, but the effects were apparently triggered at different numerical levels. *See id.* at 424, 427–28. We also may put aside research on decisionmaker stress and regret, *see, e.g.*, Schwartz et al., *supra* note 124, at 1179, 1193 (distinguishing more-regretful maximizers from less-careful satisficers), which effects may count toward decision costs without necessarily affecting discretion or ideological influence.

[132] *See* Charles A. O'Reilly, III, *Individuals and Information Overload in Organizations: Is More Necessarily Better?*, 23 ACAD. MGMT. J. 684, 691–92 (1980) (reporting results of the welfare agent decisionmaking field study, entitled "Study 2").

[133] *See id.* at 692 (reporting a correlation of 0.20 at $p < 0.05$ between reported information load and audited error rate per case).

[134] *Cf.* John J. Gibbons, *Illuminating the Invisible Court of Appeals*, 19 SETON HALL L. REV. 484, 486–87 (1989) (questioning the quality of appellate review as docket load increases). The phenomenon includes honest nonideological mistakes.

preter discretion.[135] The question, then, is whether there is good empirical evidence of increasing ideological influence as interpretive sources increase. We have seen that Cox and Miles's examination of voting rights cases showed decreasing, not increasing, ideological influence in later years.[136] But their explanation for this trend had nothing to do with an increasing stock of precedent.[137]

There are competing predictions about the effects of precedent accumulation on judicial discretion. Alec Stone Sweet has written that the zone for judicial discretion and lawmaking shrinks as the number of precedents increases. "If the law is path dependent," he explained, "this zone will narrow over time, as the rule is adjudicated."[138] Frank Easterbrook has advanced the opposite prediction for the Supreme Court. He has alleged that Court precedent inevitably will become inconsistent, and "[t]he availability of inconsistent precedents allows the Justices to 'prove' anything they like, without fear of contradiction."[139] A broader and perhaps less sophisticated version of the prediction is that more precedents inspire judges to ignore any aggregated implications and to cherry pick. A more likely yet still broad version is that judges respect precedent to a degree, but that

---

[135] *See supra* Section I.B (discussing conceptions of discretion and ideology).

[136] *See* Cox & Miles, *supra* note 113, at 1534–35; *see also supra* notes 115–16 and accompanying text.

[137] *See* Cox & Miles, *supra* note 113, at 1495 (noting "[c]hanges over time in the types of suits brought and the political realities on the ground"); *see also* LEE EPSTEIN ET AL., THE BEHAVIOR OF FEDERAL JUDGES 114–16 (2013) (reporting an increase in ideological voting at the Supreme Court in cases decided from 1980–2009 compared to 1937–1979, and suggesting a relationship to "the growing homogeneity of the political parties," but also noting a potential case-selection effect toward new types of legal disputes).

[138] Alec Stone Sweet, *Path Dependence, Precedent, and Judicial Power*, in ON LAW, POLITICS AND JUDICIALIZATION 112, 130 (Martin Shapiro & Alec Stone Sweet eds., 2002); *see also* DWORKIN, *supra* note 49, at 286 (suggesting that a primitive legal system is more likely to generate true ties than an advanced legal system); Ronald Dworkin, *Law as Interpretation*, in THE POLITICS OF INTERPRETATION 249, 262 & n.4 (W.J.T. Mitchell ed., 1983) (suggesting that later authors in a chain novel exercise may have less freedom because fewer options will adequately fit what has been written already).

[139] Frank H. Easterbrook, *Ways of Criticizing the Court*, 95 HARV. L. REV. 802, 831 (1982) (applying Arrow's impossibility theorem); *see also* JEFFREY A. SEGAL & HAROLD J. SPAETH, THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED 77–78 (2002) ("[P]recedents lie on both sides of most every controversy, at least at the appellate level. . . . [M]erely consult . . . the litigants' briefs . . . ."). Compare Mark V. Tushnet, *Following the Rules Laid Down: A Critique of Interpretivism and Neutral Principles*, 96 HARV. L. REV. 781, 822 (1983), which indicates that there is (almost) always a mathematical function that can be constructed through which an earlier pattern of data can be made consistent with another data point, and that working with judicial precedent has a similar logical flexibility. *See also id.* at 824 ("[I]f we truly allow all reasonable arguments to be made and possibly accepted, we abandon the notion of rule-following entirely."). The suggestion seems not to indicate increasing or decreasing discretion as data points (i.e., cases or sources) accumulate, but rather sustained discretion in the absence of sociological constraint.

more precedents inspire more ideas about how law might be formulated and applied.[140]

The leading empirical effort does not cleanly support a "more sources, more ideology" claim, but the study is provocative.[141] Stefanie Lindquist and Frank Cross isolated cases deciding whether a defendant had acted "under color of" law for purposes of 42 U.S.C. § 1983, which the authors considered "cases of ideological import."[142] The authors collected cases in seven circuit courts during the thirty-year period after Section 1983 litigation was energized by *Monroe v. Pape*.[143] To measure the constraint of precedent, the authors used the inverse of the influence of judicial ideology.[144] Precedent and ideology are not the only influences on judicial decisions, but the force of precedent may be crowded out when ideological influence increases. To measure ideology, the authors scored appellate judges based on lawyers' perceptions.[145] The influence of ideology on votes was estimated and compared for each of three time periods,[146] controlling for a few other variables such as the tendency of respondents to prevail on appeal.[147]

The study found neither consistently increasing nor decreasing ideological influence over time. Instead the study suggests a curious U-shaped curve.[148] During a first stage of doctrinal development (1961–1975), ideology correlated with votes in a statistically significant

---

[140] *See supra* text accompanying note 91 (noting Eskridge's suggestion regarding legislative history's inspirational effect).

[141] *See* Stefanie A. Lindquist & Frank B. Cross, *Empirically Testing Dworkin's Chain Novel Theory: Studying the Path of Precedent*, 80 N.Y.U. L. Rev. 1156 (2005).

[142] *Id.* at 1193; *see also id.* at 1205 (warning that judges might become more tightly bound to precedent over time in "less ideologically charged areas").

[143] 365 U.S. 167, 173–74, 183–84 (1961) (concluding that "under color of" law is broad enough to include misuse of power in violation of state law, whether or not state law appears to offer relief, if the defendant was clothed with the authority of state law); *see* Erwin Chemerinsky, Federal Jurisdiction 490–93 (5th ed. 2007) (describing the doctrinal change in *Monroe*); Lindquist & Cross, *supra* note 141, at 1190–91 & n.161 (listing the Second, Third, Fourth, Fifth, Seventh, Ninth, and District of Columbia Circuits as within the dataset).

[144] *See* Lindquist & Cross, *supra* note 141, at 1177.

[145] *See id.* at 1181 (explaining that a five-point scale ranging from -2.0 to 2.0 was applied based on lawyer impressions reported in the *Almanac of the Federal Judiciary*, with the ideology of the appointing president used for a few judges who were not included in the *Almanac*).

[146] *See id.* at 1193–94. During certain time periods, when so-called conservative interests use section 1983 against so-called liberal state action, this research design can be thrown off.

[147] *See id.* at 1182, 1192–93.

[148] *See id.* at 1203–04.

way.[149] Perhaps not enough cases had been decided to firmly guide the post-*Monroe* analysis. During a second stage (1976–1982), the authors state that the coefficient for ideology shrunk and lost statistical significance,[150] although they also emphasize that the ideology coefficients for the first two stages were not significantly different from each other.[151] Perhaps judges were succumbing to the hemming influence of a growing thicket of precedent—or perhaps nothing very new was happening. During a third stage (1983–1990), the ideology coefficient peaked and regained statistical significance.[152] Perhaps the stock of precedent had finally increased enough to offer numerous options and induce cherry picking.[153]

There is no simple explanation for this pattern, and neither of the competing hypotheses identified in the study predicted it. The research design is open to debate, as well. Lindquist and Cross's choice of three time periods for the selected circuits is not fully defended.[154] Second, their ideology scores have downsides. Lawyers' anonymous impressions are useful sources, but these impressions were translated into numbers and the scoring is hard to evaluate or replicate.[155] Third, governing doctrine might have changed in unobserved ways. Relevant Supreme Court cases can have differing levels of doctrinal vagueness and tension over time, as the Court revamps tests or restarts battles.[156] Fourth, and difficult to fix, any ideological trend is

---

[149] *See id.* at 1194. Their Table 5 reports results for the first period, and then whether the influences of the variables were significantly different in the second and third periods *compared to the first period*. I am emphasizing a different comparison.

[150] *See id.* at 1203–04.

[151] *See id.* at 1194.

[152] *See id.* at 1194, 1203–04 (referring to unreported results in which the time periods were evaluated separately).

[153] *See id.* at 1204 (contending that, by the third period, "proliferation of available prior decisions . . . expand[ed] judges' discretion to decide cases in accordance with their attitudes," but declining to reach "firm conclusions from these initial results").

[154] *See id.* at 1193–94 (noting that the groupings were roughly equal in number).

[155] Also, these attorney observations might be based on the votes that the study was attempting to predict. This presents a circularity issue. *See* Lee Epstein et al., *The Judicial Common Space*, 23 J.L. ECON. & ORG. 303, 314–15 (2007) (suggesting that researchers remove the case category of interest from the ideology dataset). On a separate note, Lindquist and Cross do cautiously indicate that the ideological intensity of appointed judges might have changed over time in ways not captured by simple ideology scores. *See* Lindquist & Cross, *supra* note 141, at 1204.

[156] *See* Lugar v. Edmondson Oil Co., 457 U.S. 922, 935 (1982) (recognizing a connection between color of law issues and constitutional state action doctrine, in that doctrinal tests for the latter are sufficient for the former); *id.* at 939 ("Whether these different tests are actually different in operation or simply different ways of characterizing the necessarily fact-bound inquiry that confronts the Court in such a situation need not be resolved here."); Cox & Miles, *supra* note 113, at 1495 (relating ideological influence to doctrinal rules and standards); George L. Priest, *Measuring Legal Change*, 3 J.L. ECON. & ORG. 193, 201 (1987) (supposing a transitional increase in litigant uncertainty after legal change).

hard to confirm without accounting for litigation selection effects.[157] Accumulating precedents might dictate results in an ever-growing domain of situations even as we observe persistent, increasing, or unstable levels of discretion-driven ideological influence. The parties' litigation decisions might drift toward unsettled areas of law.[158] A statute that is "very broad in its terms"[159] offers running room for lawyers to select away from legally settled areas.[160]

Lindquist and Cross's data are not available for reanalysis, but an economical way of following up is to add a fourth period of cases. Tracking the design of their study, Roy Germano[161] and I coded and analyzed a post-1990 set of Section 1983 cases that addressed state action. The new dataset adds approximately ninety cases and nearly 290 judge votes within the same circuits examined by the earlier study, making the new set comparable in size to the first three case groupings in Lindquist and Cross's study. We also followed Linquist and Cross by including dummy variables for whether the Section 1983 plaintiff was the appellant (to account for affirmance bias), whether the plaintiff was pro se (to account for weak advocacy), and whether the appellate jurisdiction was in the South (the Fourth or Fifth Circuit here, on the theory that these jurisdictions might be less friendly to Section 1983 plaintiffs). Once again following suit, we calculated an aggregate measure of Supreme Court ideology for the date of each

---

[157] *See* Priest & Klein, *supra* note 64, at 4–5, 17 (modeling litigation selection effects); Frederick Schauer, *Judging in a Corner of the Law*, 61 S. Cal. L. Rev. 1717, 1722–27 (1988) (discussing appeals); *cf.* Richard E. Miller & Austin Sarat, *Grievances, Claims, and Disputes: Assessing the Adversary Culture*, 15 L. & Soc'y Rev. 525, 536–43 (1981) (showing a winnowing process from grievance, claim, dispute, and use of courts). A key assumption involves similar stakes for all sides to the dispute. Empirical challenges to strong selection-effect hypotheses include Theodore Eisenberg, *Testing the Selection Effect: A New Theoretical Framework with Empirical Tests*, 19 J. Legal Stud. 337 (1990).

[158] Thus, after a court holds that public defenders are usually not state actors when representing clients, *see* Polk County v. Dodson, 454 U.S. 312, 324–25 (1981), subsequent complaints against public defenders may include allegations of conspiracies with state officials, *see* Tower v. Glover, 467 U.S. 914, 920 (1984) ("Glover alleges that petitioners conspired with state officials, and his complaint, therefore, includes an adequate allegation of conduct 'under color of' state law.").

[159] Lindquist & Cross, *supra* note 141, at 1205.

[160] Much to their credit, Lindquist and Cross were alert to a possible "agenda effect" and they identified arguably "routine issues" that might be subject to greater precedential force, such as claims against private schools or attorneys. *Id.* at 1196. The study finds at best mixed support for reduced ideological influence in these areas. *See id.* at 1198–1200 & tbls. 7–9. Still, selection effects push cases toward hard questions regardless of the general subject matter, and the study's subcategories are not narrow enough to box out these effects.

[161] Research Scholar, New York University School of Law.

circuit court decision (to account for potential responsiveness to shifting Supreme Court composition).[162]

The circuit judge ideology scores constructed by Lindquist and Cross are difficult to replicate, so we used the now-familiar Judicial Common Space scores for our fourth time period.[163] Because Common Space scores are not the same as Lindquist and Cross's scores, we should be cautious about making simple comparisons of the magnitude of ideological influence across the four time periods. But using Common Space scores is a reasonable and efficient way of extending the investigation of ideological influence over time.

As it happens, the Common Space scores for judicial ideology are not statistically significant in our fourth time period (Table 2). The coefficient is in the expected direction, with more conservative scores possibly associated with judge votes against state action. But the standard errors are too large. This result holds for a bare-bones statistical model in which judicial ideology is the sole independent variable (Model 1), a model with variables for plaintiff appellants and pro se plaintiffs (Model 2), and a third model, which matches the Lindquist and Cross model, with variables for the South and Supreme Court ideology (Model 3). The coefficients for these additional variables are in the theoretically expected directions, but again the standard errors are mostly too large to reach statistical significance. The exception involves pro se plaintiffs. They appear to have more difficulty getting pro-state-action votes from judges in this fourth time period, although the correlation is only marginally significant ($p < 0.10$).

Now judicial ideology has no pattern that fits any stock theory. Our fourth period looks like Lindquist and Cross's second period, in which ideology loses statistical significance, while ideological influence during the first and third periods remains unchallenged. Probably no cogent hypothesis would have suggested a sine wave for ideological influence as precedent accumulates.[164] Possibly there is no demonstrable relationship between the accumulation of precedents

---

162 These Segal-Cover scores for judicial ideology range from 0.0 to 1.0 for each justice, with higher numbers representing more liberal ideology. *See* Jeffrey A. Segal, *Perceived Qualifications and Ideologies of Supreme Court Nominees, 1937–2012*, http://www.stony brook.edu/commcms/polisci/jsegal/QualTable.pdf (last visited Mar. 13, 2017). The mean score for all nine justices was used to represent Supreme Court ideology.

163 Judicial Common Space scores range from -1.0 to 1.0, with higher numbers representing *more conservative* ideology.

164 A different hypothesis, which is not testable without the original dataset, is that waves of new judicial appointments changed the probability of an ideologically mixed panel. Perhaps there were more mixed panels during periods with apparently less ideological influence. For an important and early contribution to the idea that ideologically mixed panels dampen ideological influence, see Cross & Tiller, *supra* note 109, at 2156. For another pioneering study of the same effect in the context of challenges to Environmental

and the influence of judicial ideology, at least without efforts to understand the content of individual precedents and selection effects. True, nothing presented here rules out a relationship. Additional investigation in the future seems worthwhile. But even apart from possible differences with other types of interpretive sources that do not really proliferate,[165] ideology-boosting theories of accumulating precedents lack adequate evidentiary support to be conventional wisdom.

TABLE 2. LOGIT MODELS OF VOTES AGAINST STATE ACTION IN
SECTION 1983 CASES, 1991–1998

| Independent Variables | Model 1 | Model 2 | Model 3 |
|---|---|---|---|
| Judge Ideology (Judicial Common Space scores) | 0.3386804 (0.4256161) | 0.4369414 (0.4349204) | 0.4709925 (0.4504678) |
| Plaintiff Appellant | | 0.1696803 (0.3160719) | 0.1564279 (0.3503046) |
| Plaintiff Pro Se | | 1.229526* (0.6359029) [0.67 → 0.87] | 1.186941* (0.6537856) [0.67 → 0.87] |
| South | | | -0.0662674 (0.3120929) |
| Supreme Court Ideology (Segal-Cover scores, mean) | | | -0.3449875 (1.098529) |
| Constant | 0.7309393*** (0.131453) | 0.5116616* (0.2873298) | 0.6269603 (0.4350511) |
| N | 289 | 289 | 289 |

\* $p < 0.10$, \*\* $p < 0.05$, \*\*\* $p < 0.01$. Standard errors are in parentheses. In brackets are predicted probabilities of a judge voting to find *no* state action as one variable changes [minimum value → maximum value] and all other variables are held at their means. Only some circuit courts are represented, following Lindquist and Cross. Note that higher Judicial Common Space scores are meant to represent more liberal judge ideology, while higher Segal-Cover scores are meant to represent more conservative judge ideology.

## C.   Paper Pushing

Other studies find information-load effects that do not implicate ideology, conventionally understood. Instead they suggest paper pushing. When workload increases, perhaps judges stop thinking as hard about individual cases and start finding ways to defer. Paul Carrington noted a pattern like this back in 1969: "Pressure of time may create a tendency to give greater deference to primary decision

---

Protection Agency actions, see Richard L. Revesz, *Environmental Regulation, Ideology, and the D.C. Circuit*, 83 VA. L. REV. 1717, 1721, 1764 (1997).

[165] Many issues of interpretation focus on a few or a few dozen interpretive sources—such as canons and different forms of legislative history—that do not grow over time.

makers."[166] In appellate courts, deference can be reflected in higher affirmance rates.[167] In trial courts, deference can be reflected in higher dismissal rates,[168] if we view dismissal as deference to private ordering and other nonjudicial systems. Alternatively, we might view affirmance and dismissal as bows to the status quo, defined as an upstream decision by a trial judge or nonjudicial system.[169] Thus, denying or delaying parole requests has been depicted as favoring the status quo of incarceration—and one study reports that a judge is more likely to stick with that status quo when the judge's mental resources are depleted.[170]

A creative effort in this line is Bert Huang's.[171] After the attacks of September 11, 2001, two federal circuit courts experienced a surge of immigration appeals from the Board of Immigration Appeals. Suddenly the Second and Ninth Circuits had nearly 50% more cases to close.[172] Huang studied what happened in a different part of the docket: appeals in civil cases outside the surging immigration area

---

[166] Paul D. Carrington, *Crowded Dockets and the Courts of Appeals: The Threat to the Function of Review and the National Law*, 82 HARV. L. REV. 542, 554 (1969). Carrington also indicated that time pressure might increase the influence of judges' "personal values," by crowding out thorough adversarial presentation, deliberation, and explanation. *See id.* at 555.

[167] *See* RICHARD A. POSNER, THE FEDERAL COURTS: CHALLENGE AND REFORM 176, 345 (1999) (asserting this relationship between docket load and affirmance rates); Chris Guthrie & Tracey E. George, *The Futility of Appeal: Disciplinary Insights into the "Affirmance Effect" on the United States Courts of Appeals*, 32 FLA. ST. U. L. REV. 357, 361 & n.15 (2005) (raising questions about the direction of causation, but finding that a one standard deviation increase of 9500 cases in the appellate docket was associated with an increase in the affirmance rate of about 6%).

[168] *See* POSNER, *supra* note 167, at 178–80 (suggesting silent redefinition of the tests for dismissal and summary judgment).

[169] *See* Guthrie & George, *supra* note 167, at 377–79 (using lower court judgments as the status quo, and noting a bias in which greater responsibility is associated with "action" over "inaction"). On competing status quos, especially the choice to focus on either upstream processes or downstream results, see Adam M. Samaha, *On the Problem of Legal Change*, 103 GEO. L.J. 97, 115–17 (2014).

[170] This is the horrifying result in Shai Danziger et al., *Extraneous Factors in Judicial Decisions*, 108 PROC. NAT'L ACAD. SCI. 6889, 6889 (2011), which studied several Israeli judges who processed a couple dozen felon parole cases per day with two breaks for food. The probability of a parole grant plummeted from about 60% to about 20% and lower, depending on whether the request was entertained toward the beginning or the end of each mini-session. *See id.* at 6890–92 & tbl.1 (noting that the results were not upset by controlling for some legally relevant variables, and that participants were unlikely to have strategically controlled the sequence of requests). The study lacked ideology scores. But the legally relevant variables of prior incarceration and availability of a rehabilitation program were significantly related to parole outcomes, while the applicants' sex and ethnicity were not. *See id.* at 6891.

[171] *See* Bert I. Huang, *Lightened Scrutiny*, 124 HARV. L. REV. 1109, 1115 (2011).

[172] *See id.* at 1113–14 & nn.13–14; *id.* at 1122–25 & figs.1, 2 & 3.

from 2001 through 2005.[173] In the non-surge appeals, affirmance rates increased in a statistically significant way—on the order of 4% higher from pre-surge baselines that were already about 85%—using similar cases in other circuits as a control group.[174] The surging caseload from one federal agency was unlikely to change the quality of the other civil cases appealed from the district courts, thereby increasing the chances that the increase in affirmance rates for those other cases was caused by an exogenous work surge.[175] Surging deference is one straightforward characterization, if not a deep explanation, of the resulting judicial behavior.

Huang's study is a terrific contribution but it was not designed to reveal mechanisms of decision. The data is aggregated at the circuit level without information on individual judges or their chambers-level decision structures.[176] Also, without assigning ideology scores to individual judges and coding case results, we cannot thoroughly test the effect of caseload on ideological influence. Still, Huang did offer reason to believe that ideology was not a major factor. He studied a marked drop and then a rise in the number of Second Circuit judges, with the rise from new Democratic appointees.[177] He found a rise and then a drop in affirmance rates, suggesting that ideology was not responsible. "[T]he change in 'politics' went in the same direction each time"—the ratio of Democratic appointees to Republican appointees increased both during and after the vacancy spike—"yet the changes in reversals went in opposite directions, first collapsing, then rebounding."[178]

More important for present purposes, Huang's study was not aimed at information loads *within* cases. Judges might deal with an increased stack of cases differently from an increased amount of information within a case. The parole study mentioned above offers a bit of support for this distinction. That study did find that the time elapsed

---

[173] *See id.* at 1114–15, 1127–28 (excluding criminal and habeas appeals for being legally unstable).

[174] *See id.* at 1130–35 (combining and separating straight reversals, partial reversals, and remands); *id.* at 1136 (claiming that "the smooth path of appeals filings suggests that would-be appellants have not reacted in large numbers to the outcome changes reported here"); *id.* at 1150, 1152 (reporting specific estimates, and finding the same patterns with pro se cases and U.S. Government cases excluded).

[175] *See id.* at 1122.

[176] *See id.* at 1116, 1129 (discussing possible ways in which chambers might deal with heavy caseloads).

[177] *See id.* at 1126–27.

[178] *Id.* at 1136. Furthermore, "the 'natural experiment' (or difference-in-differences) setup already accounts for 'political' change, to an extent, because such changes tend to be shared by all the circuits, seeing as their new judges are all appointed by the same President at any given time." *Id.* at 1136–37.

within each decisionmaking session was negatively correlated with the likelihood of a parole grant, but the independent significance of elapsed time fell away when the authors added the ordinal position of the parole application.[179] In other words, the order of a case within the decision sequence appeared to dominate the total minutes of work within a work session.[180] This result "hints" that judges' mental resources were depleted by "the act of making decisions rather than simply elapsed time."[181] If correct, this explanation counsels hesitation before we extend lessons from caseload increases to information increases within each of those decisions.

Existing work on "source surges" within cases is scarce but an enlightening recent effort deserves attention. Gregory Sisk and Michael Heise sampled lawyers' briefs in the Ninth Circuit's civil docket, checked word counts, and found that longer appellant briefs are associated with a *higher* chance of reversal.[182] This positive correlation is independent of attorney experience, whether the case was orally argued, the procedural posture, and several other variables.[183] Briefs that near the word limit appear to fall off in their seemingly positive effect,[184] but the pattern otherwise tends to indicate that more words mean a better chance of success for the appellant in this sample. Perhaps more information pressed on judges within a particular case increases the chances that an appellate judge will dig in and want to disrupt the result below—even if more total cases dumped on an appellate judge's docket reduces those same chances.

Sisk and Heise's study leaves open other possibilities, though. First, the study has word counts but not the number of issues presented. Lawyers sometimes argue in the alternative. If longer briefs correlate with a larger number of legally independent arguments for reversal, then the study might have picked up the influence of alternative arguments creating multiple chances to upset the result below. Either way, multiple issues present a distinctive kind of information load that is worth investigating. Second, Sisk and Heise's study did not measure information load in terms of law or legal sources. Perhaps longer briefs tend to cite more sources, or perhaps more

---

[179] *See* Danziger et al., *supra* note 170, at 6890.

[180] As another cautionary note, these two measures were highly correlated. *See id.*

[181] *Id.*

[182] *See* Gregory C. Sisk & Michael Heise, *"Too Many Notes"? An Empirical Study of Advocacy in Federal Appeals*, 12 J. Empirical Legal Stud. 578, 586, 590–93 (2015) [hereinafter Sisk & Heise, *Too Many Notes*].

[183] *See id.* at 583–84. The authors did not code judicial ideology and ideological direction of the judge votes, however, which are difficult tasks when facing a diverse set of cases. *See id.* at 583 n.23.

[184] *See id.* at 592 fig.2.

sources tend to overcomplicate appeals. Third, brief length might follow case complexity in some sense. District judges' struggles with complicated cases might generate more mistakes and opportunities to obtain at least partial reversal, as the authors explain.[185]

To learn more, Roy Germano and I replicated Sisk and Heise's results and then added two independent variables: the number of issues presented and the number of sources in the appellant's brief.[186] Not all briefs in the original dataset are publicly available for additional coding, unfortunately.[187] And we lack a reliable measure for case complexity, although the number of issues and sources might be proxies. Within these limits, we tested whether the number of issues presented or the number of sources cited by the appellant were correlated with judge votes to reverse, retaining all of Sisk and Heise's independent variables in the model.

Experienced lawyers may suspect that any association between reversals and the number of issues or citations is not strictly positive or negative. It would not be surprising to learn that increasing the number of issues that are independently adequate grounds for reversal increases the likelihood of reversal, but only up to a point when the appeal becomes excessively complicated.[188] The most plausible hypothesis for citations is less obvious. Perhaps rising from zero to some number of legal sources cited helps the appellant, but the effect levels off or turns negative at some point.

The first headline result is that more issues presented drag *down* the appellant success rate, and without undercutting the positive effect of more words (Table 3 and Figure 4). More issues correlate with a falling chance of a vote to reverse and at a high level of statistical significance ($p < 0.01$). The raw reversal rates do appear to bounce up

---

[185] *See id.* at 594.

[186] Two research assistants independently coded the number of issues presented from the relevant stand-alone sections of the appellants' briefs, including subissues that appeared to be discrete issues. The research assistants largely agreed in their coding, and the author reviewed a number of briefs with subissues presented. For citations, a research assistant and I divided up the appellants' briefs and manually counted the number of citations listed in the appellants' tables of authorities.

[187] Social Security case briefs are filed under seal. *See* Sisk & Heise, *Too Many Notes*, *supra* note 182, at 578 n.*, 583. For these cases, Sisk and Heise obtained word counts from the clerk's office, *see id.*, but we had to drop those cases from our revised data set. Running their model without these omitted briefs yields largely similar results, but the appellee lawyer experience variable loses significance. *See infra* Table 3 (reporting results for Model 2). Also note that the substantive results do not change if we exclude appellant briefs that the Court of Appeals allowed to exceed the ordinary word limit. Thanks to Bert Huang for suggesting this variation.

[188] If instead the issues are logically dependent such that the relevant party must prevail on each and every one, then the presence of multiple issues would tend to decrease the chance of success on appeal, logically and all else equal.

and down as issues are added, rising from one to two issues, for example, then dropping at three issues and at five. But our regression model indicates that the coefficient for number of issues is negative. Perhaps many of these issues are logically dependent, not alternative grounds for reversal. Or perhaps more issues often complicate appeals in ways that reduce persuasiveness, even if those issues are legally independent grounds for reversal. In terms of paper-pushing affirmances, adding issues to a case might be like adding cases to the judge's docket.

Nonetheless, appellant brief length remains correlated with reversal. This result suggests that a brief with an attractively low number of issues presented may nonetheless benefit from extensive argument on any given issue. And the negative effect of complicating an appeal with additional issues—perhaps with legal dependence, or an implausibly large number of alleged errors, or just a mentally challenging number of major logical transitions—might be counterbalanced with more extensive argument in support of however many issues are presented. In terms of paper-pushing tendencies, then, adding words does not seem to have the same effect on judges as adding issues or cases. Reading more words might be a light mental lift, as long as the judge is not shifting mental gears to confront a new issue.

Another headline is that the number of sources cited by the appellant seems to have *no effect*. The relationship between the number of unique citations in an appellant's brief and the probability of a vote to reverse looks positive at first glance, which could suggest that more sources might soften any usual preference to push paper by affirming. But the coefficient is not statistically significant at conventional levels given the standard error. So we cannot suggest an association. Perhaps those who consume legal sources in judiciaries, be they judges or law clerks or staff attorneys, are unmoved by large or small numbers. Future work will have to help determine more definitively whether and why an increase in legal sources cited has a different effect on paper pushing than increases in words and issues.

## Table 3. Logit Models of Votes to Reverse in Ninth Circuit Civil Appeals, 2010–2013

| Independent Variables | Model 1 Replication of Sisk & Heise | Model 2 Replication without briefs under seal | Model 3 Issues and citations variables added |
|---|---|---|---|
| Appellant Brief Length, measured in words | 0.000106*** (0.0000331) | 0.0000742** (0.0000378) | 0.000139*** (0.0000403) [0.14 → 0.59] |
| Appellee Brief Length, measured in words | 0.0000144 (0.0000320) | 0.0000350 (0.0000334) | 0.0000337 (0.0000338) |
| Orally Argued | 1.190*** (0.293) | 1.091*** (0.387) | 0.976** (0.379) |
| Published Opinion | 0.551** (0.259) | 0.654** (0.254) | 0.609** (0.258) |
| Appellant Lawyer Experience | -0.00142 (0.00191) | -0.00148 (0.00209) | -0.00220 (0.00224) |
| Appellee Lawyer Experience | -0.00489** (0.00234) | -0.00318 (0.00235) | -0.00238 (0.00230) |
| Issue Type: Civil Procedure | 1.602*** (0.349) | 1.443*** (0.358) | 1.417*** (0.370) |
| Party Type: Business Appellee | -0.953*** (0.267) | -1.324*** (0.299) | -1.180*** (0.338) |
| Appellant Issues | | | -0.212*** (0.0451) [0.44 → 0.05] |
| Appellant Citations | | | 0.00211 (0.00324) |
| Constant | -3.003*** (0.598) | -2.589*** (0.684) | -2.662*** (0.730) |
| N | 810 | 672 | 672 |

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors, clustered at the judge level, are in parentheses. The dependent variable is a judge's vote to "reverse," which includes reverse full stop, reverse in part, and vacate and remand in part. In brackets are predicted probabilities of a vote to reverse as one variable changes [minimum value → maximum value] and all other variables are held at their means. The reference point for Civil Procedure issue type is Constitutional issue type; the reference point for Business party type is Individual party type. To simplify, several variables that are part of Sisk and Heise's model and listed in their Table 2— i.e., four additional issue types, five additional appellant party types, six corresponding appellee party types, and five procedural stages—are not displayed in our Table but are included in our models. Most of those coefficients are not statistically significant.

FIGURE 4. PREDICTED RATES OF VOTES TO REVERSE BY
APPELLANT BRIEF LENGTH, ISSUES, AND CITATIONS
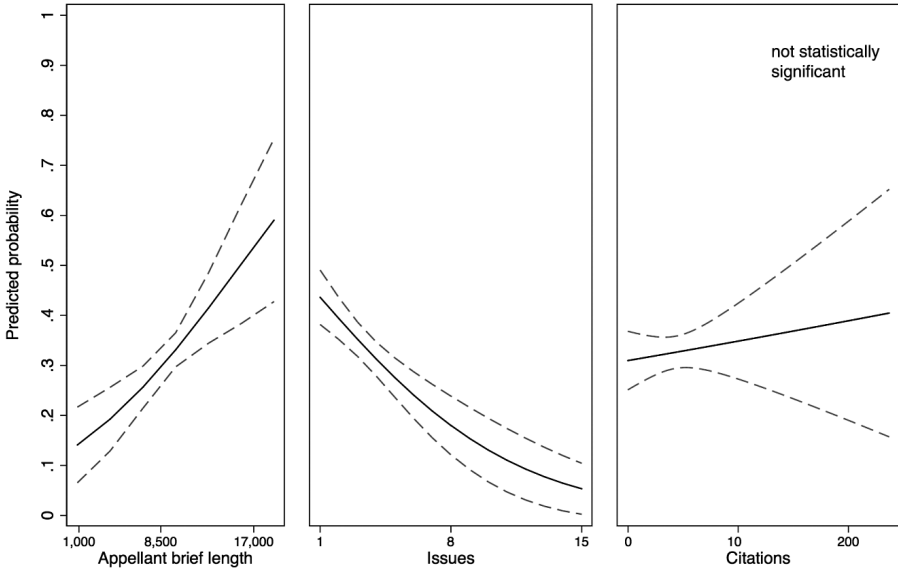


We should interpret these results cautiously. For the number of issues, we did not make the judgments necessary to distinguish arguments in the alternative from logically dependent arguments. The former are theoretically related to a higher probability of reversal, the latter are not. We might assume that lawyers for appellants usually present issues as alternative grounds for reversal or remand, while bundling together subissues on which they must prevail across the board. But we lack data on this. In addition, for the number of sources, we have citation counts for the appellants' briefs, not for the appellees'. Appellee citations might usually track the number of citations that the appellant chooses to open with, but will not always. Furthermore, we did not separate out different types of sources cited. Case citations tend to dominate in these briefs, but different types of sources might have different effects on judges and staff.

With those caveats, we have not found evidence that judges who face more as opposed to fewer sources are any more or less likely to follow a paper-pushing preference. Equally important, we now have evidence that information loads of different kinds have different effects within a particular case and across the docket. We see mixed effects from different types of information load within the same case, and this mixture was not easy to predict from what we have learned about docket loads. Converting our initial results into oversimplified slogans, if not memes: "more cases, more affirmances"—"more issues,

more affirmances"—"more words, more reversals"—"more sources, no difference."

## D.   Quasi-Law

We can build out the idea of cognitive-load reductions by adding an unorthodox strain of legal influence. Ordinary people often take mental shortcuts that produce triage-like patterns in results.[189] For instance, people may preemptively exclude options.[190] Ideally a person might consider several dimensions of value in comparing all available options but, when facing a large number of possibilities, the person might instead chop down the number of options eligible for thorough consideration by screening for thresholds on one or two core dimensions. Such mental shortcutting is believable for low stakes consumer decisions, where markets and technology lower search costs yet also increase the number of easily identified options. Although there is much less on-point evidence of this behavior among judges, the pattern is worth searching for.

The reintroduction of law at this stage might seem odd, given that we have been discussing discretion as a phenomenon in which law runs out. But we can entertain a "quasi-law" picture of judicial behavior, in which decisionmakers depart from some of law's formal demands while attempting to follow a principled prioritization within the law that did not run out. When facing heavy cognitive loads imposed by formal law, perhaps interpreters will not thoroughly examine all of the legally relevant sources but instead prioritize those sources that are consistent with what they believe are law's priorities. Quasi-law is a compromise in which judges are pushed by extralegal resource pressures but reintroduce law to help guide prioritization.[191]

Evidentiary support comes from trademark litigation. District judges may face an awfully large number of legally relevant factors in

---

189   *See* John W. Payne et al., The Adaptive Decision Maker 34, 37 (1993) (stating that people tend to take shortcuts away from aggregating all information when choice tasks become more complex).

190   *See* Richard H. Thaler et al., *Choice Architecture*, *in* The Behavioral Foundations of Public Policy 428, 435–36 (Eldar Shafir ed., 2013) (distinguishing a "compensatory" decision strategy in which all options and attributes are considered and traded off, and "elimination by aspects" for large and complex option sets in which important attributes are assigned cut-off levels that are used to eliminate certain options); *see also* Simon, *supra* note 108, at 513 ("[T]he mind shuns cognitively complex and difficult decision tasks by reconstructing them into easy ones, yielding strong, confident conclusions.").

191   *Cf.* Michael Asimow, *Public Participation in the Adoption of Interpretive Rules and Policy Statements*, 75 Mich. L. Rev. 520, 566 (1977) ("Having decided whether the interpretive regulation is valid, the court mentions all the factors it can muster that tend to support its decision.").

deciding whether there is a likelihood of consumer confusion. As Barton Beebe explains,[192] the number of doctrinal factors imposed by the circuit courts ranges from half a dozen to over a dozen[193]—even counting "[a]ny other established fact probative of the effect of use" as only one factor.[194] Not every factor seems to influence decisions, however. Only one or two factors seemed to matter much in Beebe's analysis of over 300 district court opinions from 2000 to 2004.[195] For instance, the judge's finding on similarity of marks was nearly dispositive, one way or the other, while proximity of goods appeared to operate as a necessary condition for a plaintiff win.[196] Those two factors are not supposed to be dispositive.[197] Judges nonetheless appeared to spin the other factors to agree with their findings on the key factors,[198] which Beebe suggests is a defensible heuristic strategy akin to "take the best, ignore the rest."[199] Perhaps the district courts' focus and spinning were instigated by the circuit courts' attempts to impose a large number of doctrinal considerations.

On the other hand, spinning might not increase after six factors. Roy Germano and I reanalyzed the data to check whether courts seemed more likely to spin doctrinal factors toward the final result as the number of doctrinal factors increased. Most of the circuits' doctrines crowded around eight factors, but there was some variation. We could not detect a statistically significant difference across circuits with six, seven, eight, or ten factors in terms of stampeding the factors toward a finding of confusion or a finding of no confusion.[200] In fact,

---

[192] *See* Barton Beebe, *An Empirical Study of the Multifactor Tests for Trademark Infringement*, 94 Calif. L. Rev. 1581 (2006) [hereinafter Beebe, *Multifactor Tests*].

[193] *Compare* SquirtCo v. Seven-Up Co., 628 F.2d 1086, 1091 (8th Cir. 1980) (listing six factors), *with* Stone Lion Capital Partners, L.P. v. Lion Capital LLP, 746 F.3d 1317, 1319–21 (Fed. Cir. 2014) (listing thirteen factors), *and In re* E. I. DuPont DeNemours & Co., 476 F.2d 1357, 1361 (C.C.P.A. 1973) (same).

[194] *DuPont*, 476 F.2d at 1361.

[195] *See* Beebe, *Multifactor Tests*, *supra* note 192, at 1584, 1649–50; *see also id.* at 1600 & n.8 (explaining the use of classification trees).

[196] *See id.* at 1608.

[197] *See, e.g.*, Gateway, Inc. v. Companion Prods., Inc., 384 F.3d 503, 509 (8th Cir. 2004) ("No single factor is dispositive . . . ."); Checkpoint Sys., Inc. v. Check Point Software Techs., Inc., 269 F.3d 270, 280 (3d Cir. 2001) (same). *But cf.* A&H Sportswear, Inc. v. Victoria's Secret Stores, Inc., 237 F.3d 198, 214, 216 (3d Cir. 2000) ("The single most important factor . . . is mark similarity.").

[198] *See* Beebe, *Multifactor Tests*, *supra* note 192, at 1587.

[199] *Id.* at 1604 (quoting Gerd Gigerenzer & Daniel G. Goldstein, *Reasoning the Fast and Frugal Way: Models of Bounded Rationality*, 103 Psych. Rev. 650, 653 (1996)).

[200] Beebe calculated a "stampede score" for each case by dividing the number of factors said to favor confusion by the number of factors said to not favor confusion, excluding factors that were not discussed or held irrelevant, and separating the cases finding a likelihood of confusion from those that did not. *See id.* at 1617. A score of 0.00 indicates an equal number of factors were said to point in each direction. In our reanalysis, the mean

any trend might be negative, with somewhat less spinning as the number of factors increases.

To further the inquiry, we chose two circuits at the extremes and coded additional cases through 2010. The Eighth Circuit's doctrine has the smallest number of factors (six),[201] while the Third Circuit's doctrine has the largest number (ten) among circuits with an adequate number of cases to analyze.[202] No major differences are apparent. The similarity of marks factor is slightly less outcome determinative in this case set, but about equally powerful in both circuits. A district court's ruling on the similarity factor lines up with the court's overall result on likelihood of confusion in 86% of the Eighth Circuit cases and 90% of the Third Circuit cases. In both circuits, a no-similarity finding is an even stronger predictor of an overall no-confusion result than a similarity finding is of an overall confusion result.

Moreover, the amount of spinning looks *higher* in the circuit with the *lower* number of factors (Table 4). In the expanded dataset, the Eighth Circuit's average stampede scores for pro-confusion decisions and no-confusion decisions (0.76 and -0.53) are somewhat more extreme than the Third Circuit's scores (0.65 and -0.30).[203] Although the number of observations is small and these differences are not statistically significant, they do not show more factors yielding more spinning. If we combine pro-confusion and no-confusion decisions and

---

stampede score appeared to trend a bit toward zero as doctrinal factors increased, although moving from seven to eight factors appeared to increase stampeding in cases finding no likelihood of confusion. Following Beebe, we restricted this reanalysis to bench trials and preliminary injunction rulings.

[201] *See* SquirtCo v. Seven-Up Co., 628 F.2d 1086, 1091 (8th Cir. 1980); *see also* Kemp v. Bumble Bee Seafoods, Inc., 398 F.3d 1049, 1053 (8th Cir. 2005) (characterizing *SquirtCo* as listing "(1) the strength of the owner's mark; (2) the similarity of the owner's mark and the alleged infringer's mark; (3) the degree to which the products compete with each other; (4) the alleged infringer's intent to 'pass off' its goods as those of the trademark owner; (5) incidents of actual confusion; and (6) the type of product, its costs and conditions of purchase").

[202] *See Checkpoint Systems*, 269 F.3d at 280 (listing "(1) [t]he degree of similarity between the owner's mark and the alleged infringing mark; (2) the strength of the owner's mark; (3) the price of the goods and other factors indicative of the care and attention expected of consumers when making a purchase; (4) the length of time the defendant has used the mark without evidence of actual confusion; (5) the intent of the defendant in adopting the mark; (6) the evidence of actual confusion; (7) whether the goods, though not competing, are marketed through the same channels of trade and advertised through the same media; (8) the extent to which the targets of the parties' sales efforts are the same; (9) the relationship of the goods in the minds of consumers because of the similarity of functions; and (10) other facts suggesting that the consuming public might expect the prior owner to manufacture a product in the defendant's market or that he is likely to expand into that market"). The Federal Circuit had embraced a thirteen-factor test at one point, but no sampled district court cases used that test. *See* Beebe, *Multifactor Tests*, *supra* note 192, at 1603 & n.97.

[203] On calculating stampede scores, see *supra* note 200.

convert the stampede scores to absolute values, the Eighth Circuit's average stampede score (0.66) remains somewhat higher than the Third Circuit's score (0.51). This difference is statistically significant, albeit weakly ($p < 0.10$). So the results are not strong evidence of more factors yielding less spinning, but the results certainly do not show the opposite. For all we know, six and ten factors instigate similar amounts of spinning; capping the number of factors at six might not make any difference.[204] To be fair, though, stampede scores represent ratios of factors pointing in one direction or another, not the absolute number of sources potentially getting spun. One might think that a similar stampede score in a circuit with a larger number of factors is more impressive evidence of spinning.[205]

TABLE 4. STAMPEDING FACTORS IN TRADEMARK CONFUSION
DECISIONS, 2000–2010

| Jurisdiction | Stampede score for pro-confusion results (mean) | Stampede score for no-confusion results (mean) | Stampede score for both results (mean absolute value) |
|---|---|---|---|
| Eighth Circuit (6 factors) | 0.76 ($n = 17$) | -0.53 ($n = 12$) | 0.66 ($n = 29$) |
| Third Circuit (10 factors) | 0.65 ($n = 19$) | -0.30 ($n = 20$) | 0.51 ($n = 39$) |
| Difference | 0.11 | -0.23 | 0.15* |

\* $p < 0.10$. Each district court decision is given a "stampede score" by dividing the number of doctrinal factors said to favor confusion by the number of factors said to not favor confusion, excluding factors not discussed or held irrelevant. Positive scores indicate stampeding toward confusion; negative scores indicate stampeding toward no confusion; a score of 0.00 indicates no stampeding. Following Beebe, only decisions on motion for preliminary injunction or after bench trial are included.

In any event, we might wonder whether ideology is playing a role. Judges present trademark confusion doctrine as a complex standard, not a strict mathematical equation. "[W]e do not count beans," as one court put it.[206] Intriguingly, Beebe reported that he tested for ideolog-

---

[204] *See supra* Section II.D.2 (disentangling the logical relationship between number of sources and spinning); *cf.* Barton Beebe, *An Empirical Study of U.S. Copyright Fair Use Opinions, 1978–2005*, 156 U. PA. L. REV. 549, 555 (2008) (finding no such evidence of stampeding factors in the four-factor fair use test in copyright doctrine).

[205] Again, spinning takes mental effort and might be somewhat self-limiting for those who draft opinions. *See supra* Section II.D.

[206] Dreamwerks Prod. Grp., Inc. v. SKG Studio, 142 F.3d 1127, 1129 (9th Cir. 1998) (adding that "[t]he factors should not be rigidly weighed"); *see also Kemp*, 398 F.3d at 1054 ("[T]he ultimate determination of whether confusion is likely is not to be mechanically determined through rigid application of the factors."); Scott Fetzer Co. v. House of Vacuums Inc., 381 F.3d 477, 485 (5th Cir. 2004) ("They do not apply mechanically to every

ical influence and could not confirm it.[207] We replicated Beebe's test, which used the ideology of the appointing president for each district judge and which included independent variables for judge gender, age, experience, and whether the judge found in favor of the trademark claimant on any of five core doctrinal factors.[208] We then substituted a more nuanced measure of judicial ideology developed by Christina Boyd and others, which integrates the ideology of the judge's home-state senators when they share the party of the appointing president.[209] Neither ideology score was related in a statistically significant way to whether judges found a likelihood of confusion (Table 5).[210]

Further testing produced similar results. We added an independent variable that interacts ideology with the number of doctrinal factors in the relevant circuit, to test whether the influence of ideology might be conditional on the number of factors in play. This variable did not approach statistical significance either. Finally, we returned to the doctrinal extremes of the Eighth and Third Circuits and used our expanded case set to enhance the analysis. We again included independent variables for the judge's experience, age, and gender. We

---

case and can serve only as guides, not as an exact calculus."); Thane Int'l, Inc. v. Trek Bicycle Corp., 305 F.3d 894, 901 (9th Cir. 2002) ("The list of factors is not a score-card . . . ."); A&H Sportswear, Inc. v. Victoria's Secret Stores, Inc., 237 F.3d 198, 215 (3d Cir. 2000) ("[T]he . . . test is a qualitative inquiry. Not all factors will be relevant in all cases; further, the different factors may properly be accorded different weights depending on the particular factual setting."). *But cf.* Arrow Fastener Co. v. Stanley Works, 59 F.3d 384, 400 (2d Cir. 1995) ("[I]t is incumbent upon the district judge to engage in a deliberate review of each factor, and, if a factor is inapplicable to a case, to explain why.").

[207] *See* Beebe, *Multifactor Tests*, *supra* note 192, at 1648 n.255.

[208] *See id.* (discussing Poole Common Space scores and other variables). Beebe's data, which he generously shared with us, make clear that he used the Common Space score for each judge's appointing president. We corrected two coding errors regarding the judges, one on gender and one on the judge's identity. Also, in Beebe's model, "bridge the gap" (Factor 4) was included but was not statistically significant. Beebe might have meant to include "intent" (Factor 6), which is emphasized in his article and which is statistically significant in our models. When we substituted intent for bridge the gap, intent and the other four (judge-determined) doctrinal factors were statistically significant independent variables. Judge ideology was not.

[209] These scores are available from Christina L. Boyd's *Federal District Court Judge Ideology Data*, http://clboyd.net/ideology.html (last visited Mar. 13, 2017). We shorthand the database for the 110th Congress, which we use here, as "Boyd110." These scores are based on contributions to the measurement of judicial ideology in Micheal W. Giles et al., *Picking Federal Judges: A Note on Policy and Partisan Selection Agendas*, 54 POL. RES. Q. 623 (2001), and Epstein et al., *supra* note 155.

[210] Table 5 reports results using Common Space scores and not including the judge-determined votes on individual doctrinal factors. Also note that Beebe coded rulings in multiple dimensions (confusion, no confusion) and with multiple gradations. *See* Beebe, *Multifactor Tests*, *supra* note 192, at 1609. In our later tests, reported in Table 5, we dropped cases in which the judge was coded as not finding confusion and also not finding *no* confusion. This combination can appear in the dataset when a judge denies a party's motion for summary judgment.

found no statistically significant relationship between ideology and confusion decisions in either circuit, and the standard errors for the ideology coefficient are large. Although the number of cases is not very large, examining the doctrinal extremes is a fairly direct test of the "more factors, more ideology" theory in this field of litigation.

We cannot rule out the influence of extralegal preferences, though. Quite possibly we lack a good proxy for judicial ideology in the trademark field, at least for lower court cases. The old left-right, unidimensional measure of ideology might be inapt.[211] Furthermore, unexpected results appear. Judge gender did correlate with trademark confusion decisions in the Third Circuit (with a larger number of factors), but not in the Eighth Circuit (with a smaller number of factors). Conversely, judge age correlates negatively and experience correlates positively with trademark confusion decisions in the Eighth Circuit, but this time the Third Circuit shows no such relationships. We might construct creative theories for why these judge traits are better related to trademark policy preferences than Common Space scores. For instance, maybe young Republican appointees were relatively sympathetic to intellectual property claims and grew more so as they gained experience adjudicating cases. But the coefficients are strangely large. More important for present purposes, extraordinary creativity is needed to explain why different judge traits would pop up as significant in different circuits—and then connect the differences to the different number of doctrinal factors.

---

[211] *Cf.* Barton Beebe, *Does Judicial Ideology Affect Copyright Fair Use Outcomes?: Evidence from the Fair Use Case Law*, 31 COLUM. J.L. & ARTS 517, 517–18, 521 (2008) (distinguishing empirical findings in "ideologically-charged areas," and reporting an inability to confirm ideological influence in four-factor fair use decisions in copyright). Insightful contrary indications—for Supreme Court Justices—can be found in Tonja Jacobi & Matthew Sag, *Taking the Measure of Ideology: Empirically Measuring Supreme Court Cases*, 98 GEO. L.J. 1, 18–24, 31–35 (2009) (reporting that the mean Martin-Quinn ideology score for the majority coalition of justices correlates with votes on intellectual property claims, and fits the voting data better than some other ideology proxies). On Martin-Quinn scores as a type of "agnostic coding," see Joshua B. Fischman & David S. Law, *What Is Judicial Ideology, and How Should We Measure It?*, 29 WASH. U. J.L. & POL'Y 133, 162–63 (2009) (explaining that, for instance, "a 'liberal' Justice is one who often voted with Brennan, or who often voted with Justices who often voted with Brennan").

TABLE 5. PROBIT MODELS OF TRADEMARK CONFUSION DECISIONS,
2000–2004 AND 2000–2010

| Independent Variables | All Circuits 2000–2004 (6–13 factors) | 8th Circuit 2000–2010 (6 factors) | 3d Circuit 2000–2010 (10 factors) |
|---|---|---|---|
| Judge Ideology (Boyd110) | 0.2953176 (1.772508) | -0.5287681 (0.7229969) | 0.2190972 (0.5894955) |
| Number of Factors | -0.0926901 (0.0877199) | | |
| Ideology x Number of Factors | -0.0474575 (0.2288821) | | |
| Judge Age | -0.0020679 (0.0126003) | -0.2706737** (0.1069577) [0.84 → 0.16] | 0.0452478 (0.0286765) |
| Judge Experience | 0.0130261 (0.0143229) | 0.3159829** (0.1490445) [0.26 → 0.78] | -0.0288283 (0.0391823) |
| Judge Gender Male | -0.1536432 (0.192948) | -0.5395515 (0.692386) | -0.9670206** (0.4729968) [0.71 → 0.36] |
| Time Trend | | 0.3075713** (0.1338319) | 0.0390438 (0.0603703) |
| Preliminary Injunction | | 1.48898** (0.6791173) | -0.1756016 (0.3967649) |
| Constant | 0.798497 (0.7938781) | -604.6371** (265.9287) | -80.07935 (120.9729) |
| N | 270 | 40 | 51 |

** $p < 0.05$. Standard errors are in parentheses. In brackets are predicted probabilities of finding confusion as one variable changes [minimum value → maximum value (Judge Gender), or one standard deviation above → one standard deviation below the mean (Judge Age, Judge Experience)] while all other variables are held at their means.

What seems clear is that ideology plays different roles in different parts of the docket.[212] In many voting rights cases, conventional ideological influence is plausible; less so for many trademark cases. In either situation, increasing the number of doctrinal factors need not change the influence of ideology significantly. Equally interesting, concentrating on one or two factors and spinning the remaining factors in the same direction is different from embracing the status quo or deferring to another decisionmaker. The evidence above makes way for a quasi-law working theory of judicial behavior. The quasi-law idea amounts to triage, but not guided by extralegal preferences. Instead the interpreter attempts to prioritize the factors that formal

---

[212] *See* Adam M. Samaha & Roy Germano, *Are Commercial Speech Cases Ideological? An Empirical Inquiry*, 25 WM. & MARY BILL RTS. J. 827, 833–34 n.25 (2017) (collecting sources); *id.* at 872 (reporting new evidence of ideological influence in some fields of constitutional litigation but not others).

law seems to value most, or what formal law ought to value most if it cannot be executed perfectly. In a second-best world of resource constraints or limits on interpreter motivation, the interpreter concentrates on a lawyerly understanding of law's core values.

Think about similarity of marks. The judge's finding on this factor predicts the vast majority of the trademark confusion results in the dataset, and the importance if not the magnitude of this factor is unsurprising. Similarity is part of every circuit's confusion doctrine; and similarity probably is indispensable to the notion of consumer confusion. True, the similarity factor is a hazy standard, with courts unhelpfully alluding to sights, sounds, and meaning.[213] But evaluating similarity need not depend on elaborate investigation or extended deliberation, to the extent that visual impressions and other snap judgments drive this factor.[214] Indeed accurately estimating the probability of consumer confusion can require a snap judgment, which often is how consumers actually formulate impressions and make purchasing decisions. Prioritizing the judge's impressions about the similarity of marks, therefore, tends toward the high values of trademark law at bargain basement prices. Decision costs are reduced without judges necessarily substituting their personal policy preferences for law's priorities, best understood, even if district judges are "short-circuit[ing] the multifactor test."[215]

Whether quasi-law adequately explains these decisions requires a deeper study of law. But Beebe's analysis is in accord. He contends that the ideal trademark confusion test would carefully survey the relevant class of consumers to estimate confusion.[216] Because the ideal usually is not realistic, judges might turn to a small set of cheap and reliable factors that are close enough to the ideal. Beebe maintains that the existing piles of doctrinal factors "often distract from their ultimate purpose: to estimate what is actually occurring or will occur in the marketplace" and that "considerations such as the comparative quality of the parties' goods or the inherent distinctiveness of the plaintiff's mark rarely aid in this inquiry."[217] District judge decisions in this area seem to reflect felt resource constraints combined with a

---

[213] *See, e.g.*, SquirtCo v. Seven-Up Co., 628 F.2d 1086, 1091 (8th Cir. 1980) ("Similarity is based on an examination of the marks as a whole, including visual impression and sound."); Beebe, *Multifactor Tests*, *supra* note 192, at 1624–25 (calling the test "nebulous").

[214] 4 J. THOMAS MCCARTHY, MCCARTHY ON TRADEMARKS AND UNFAIR COMPETITION § 23:25 (4th ed. 2016) ("Similarity of appearance between marks is really nothing more than a subjective 'eyeball' test.").

[215] Beebe, *Multifactor Tests*, *supra* note 192, at 1586.

[216] *See id.* at 1645.

[217] *Id.*

plausible understanding of law's priorities.[218] More sources may, at least within some range, yield more focus on law's core values.

IV

LESSONS

Now we can step back and consolidate lessons.

A.    *Conventional Wisdom*

We should resist any conventional wisdom that increasing the number of interpretive sources will increase interpreter discretion. As a logical matter, the claim is incomplete at best. If judges follow rules against spinning and cherry picking sources, then increasing the number of sources will reduce discretion.[219] If instead there are no hardline interpretive rules, or if the rules do not constrain spinning, then discretion will persist—it will flood the interpretive process regardless of the number of sources. If, oddly, spinning is constrained while cherry picking is not, then discretion indeed will increase as the number of sources increases. But even then discretion can bloom with only a few sources in hand. And although monitoring willful inter-preters can be easier with fewer sources, if the interpretive rules cover both spinning and cherry picking, then having fewer sources probably will simultaneously make monitoring more effective and make those rules less effective at reducing discretion. The net effect is unclear.

This leaves a "more sources, more discretion" claim in a precar-ious position. When the rules are tight and the interpreters are faithful, discretion will tend to fall as sources increase. But if the rules are lax or the interpreters are very willful, discretion will prevail regardless. Only in arguably odd decision situations, in which we can imagine that interpreters are unable to spin yet able to pick, should we predict that discretion will increase as sources increase. This is neither to assert that these situations do not exist, nor that the analysis here has blanketed all feasible interpretive approaches. We are working from simple models to make analytic progress on a hard problem. Still, our analysis has produced a good sketch of the complex logical interactions among source proliferation, interpretive rules, interpreter

---

[218] For an experimental study, not involving actual judges, in which time constraints appear to increase reliance on law's content and decrease the influence of ideology on case outcomes, see Brian Sheppard, *Judging Under Pressure: A Behavioral Examination of the Relationship Between Legal Decisionmaking and Time*, 39 FLA. ST. U. L. REV. 931 (2012). The author nonetheless observed that "[s]ubjects under time limitation, particularly those who followed the law, experienced a considerable drop in the strength of their belief that they were doing the right thing in deciding their cases as they did." *Id.* at 939.

[219] *See supra* Section II.B and Table 1.

type, and resulting levels of discretion. The sketch defies any simple conventional wisdom.

This logic will not always match behavior in the real world of judging, as I have emphasized. Using old and new data, however, we find little support for a "more sources, more discretion" pattern. As an empirical matter, four working theories emerge with some evidentiary support—not only (1) increased ideological influence, for which the key study of precedents is now in doubt, but also (2) no effect or arbitrariness, as interpreters ignore additional information or stumble into error, (3) paper pushing, as interpreters reduce cognitive load by deferring to other officials or by sticking to a status quo, and (4) quasi-law, as interpreters decline to weight all legally relevant information equally, yet focus on what law seems to prize.

The hard work going forward includes marking the domains in which each theory is most plausible, measuring the magnitude of effects, and specifying the mechanisms. A starting point is looking for fields of litigation in which contested issues seem ideologically charged or salient (like election law) and areas in which judges probably are less inclined to expend decisional resources (like trademark). Ideologically charged issues suggest ideological influence, obviously, while areas of low judicial interest are more likely to follow other patterns involving mental shortcuts.

Equally important, we need more research that measures the humdrum informational inputs, not just the headline decisional outputs, for cases processed by our court systems. Publically accessible lawyers' briefs, other filings, and hearings are crucial resources, along with sensitive research into the backroom decision architecture of judges' chambers. Studies such as Sisk and Heise's and this Article's extension are a start. Combined with knowledge of formal law, quantifying information inputs will help us test and expand our set of working theories.

### B.   *Implementation Issues*

The foregoing offers logically and empirically plausible effects of varying the number of interpretive sources, but not every number is equally realistic. Turning to hard-bitten issues of implementation, we can detect a hard-wired slant toward fewer sources. Decisionmakers are likely to lean toward processing less rather than more information, per day and per decision. This leaning will not always hold, but it is a sufficiently realistic assumption to inform prescriptions for judicial interpretation. Not every bit of information will be processed thought-

fully by judges, no matter how many sources parties are allowed to cite and discuss.

Courts already enforce some dull-bladed rules for keeping per-case information loads manageable, such as word limits for briefing and time limits for oral arguments.[220] If need be, these rules can be extended to limit the number if not the kinds of sources cited by lawyers. Even without rule enhancements (or without rules, full stop) many lawyers will think that simple arguments with fewer sources are most effective.[221] Source proliferation is self-limited.

Furthermore, prescribing the number of interpretive sources might be easier than formulating interpretive rules. The explanation is not so much that judges ignore interpretive rules, where they exist, but that trans-substantive interpretive rules are difficult for anyone to draft even if they will be followed. The rule of lenity always points in the same direction, but no one has yet specified when statutes are adequately vague with respect to given applications such that lenity is properly invoked. At least equally important, perhaps no one will ever be able to draft tight rules against spinning each of the multitude of valid interpretive sources.

Yes, judges have ambled toward a rough structure for statutory interpretation, if not constitutional adjudication, and additional steps are possible. But judges are not close to Felix Frankfurter's playful—or wistful—reference to an interpretive "table of logarithms."[222] The difficulty in designing handcuff-like interpretive rules makes the prospect of limiting interpretive sources more attractive. The alternative has the nearly self-implementing advantage of making decisions easier rather than more complicated for interpreters who feel the mental weight of their dockets.

These considerations might support a low-source design for legal interpretation but, to be clear, this strategy alone will *not* control interpreter discretion. A large barrier to using a low-source strategy against discretion is that, as we have seen, such discretion can erupt after only a few sources, even when spinning is constrained and everyone behaves logically. All else equal, the probability of a friendly

---

[220] *See, e.g.*, Sup. Ct. R. 33.1(g) (setting word limits for briefs in the U.S. Supreme Court); Fed. R. App. P. 32(a) (setting page or word limits for appellate briefs in federal circuit courts); 7th Cir. R. 34(b)(1) ("The amount of time allotted for oral argument will be set based on the nature of the case."); 2016 Cal. R. 8.520(c) (setting word limits for briefs in the California Supreme Court).

[221] *See, e.g.*, Jennifer F. Novak, *The Value of Simplifying Case Presentation*, L.A. Lawyer, Feb. 2005, at 12, 12 ("Keep it simple.").

[222] Felix Frankfurter, *Some Reflections on the Reading of Statutes*, 47 Colum. L. Rev. 527, 543 (1947) ("Unhappily, there is no table of logarithms for statutory construction.").

source appearing can increase rapidly as sources are added.[223] If spinning is not constrained, controlling the number of sources will be utterly ineffectual against discretion.

Of course there are perfectly good reasons to keep information loads light apart from opposing judicial discretion; and judicial discretion might be accepted instead of attacked where courts are superior to the other imperfect alternatives. But hopes should be low—very low—for calibrating judicial discretion through otherwise comfortable reductions in information flows. Fortunately or not, some of our energy still should be devoted to developing interpretive rules or guidelines, including better specification of weights, priorities, and sequences for valid interpretive sources. Indeed, we should direct attention at strategies for ensuring that interpreters consider each valid and relevant source, at all. The larger problem for interpretation might well be that sources are sometimes lost in the mental shuffle, not that some sources become unduly influential after all sources are scrutinized.

## C.   Institutional Design

The last point leads to a valuable reminder: However significant discretion might be, discretion is only one consideration in sound institutional design. Other factors are equally or more important to shaping well-functioning institutions, including judiciaries, and the foregoing analysis has implications for a general approach to institutional design. Sensible institutional design must account for the quantity, quality, and complexity of information to be processed, and the analysis above offers logic and evidence on the effects of source proliferation.

There is some good news for conventional wisdom on institutional design. Prior learning suggested that increased information tends to increase decision quality but also increase decision costs.[224] The trade-offs are not always easy to make well, but they are familiar. Part of the logic presented above reinforces the importance of these trade-offs, insofar as interpreter discretion should count as an "error cost" for society. A negative characterization of discretion ought to be controversial, but suppose we embrace the dark view. Then such discretion-based error costs would tend to decline, logically speaking, as more interpretive information is pressed through the system, albeit at some additional cost in making decisions.

---

[223] *See supra* Section II.D.1 and Figure 3.
[224] *See supra* Section I.A.

But we have sound reasons for reconsidering this conventional wisdom, too—which can produce even better news. The problematic trade-offs (quality, costs) depend on the interpretive rules being tight and the interpreters being faithful, which we know is not always realistic. If we think that discretion is a cost and then switch our assumptions such that interpreters cherry pick without spinning sources, then we are in the happy situation of being able to drive down decision costs by limiting sources while also driving out the error of discretion. These assumptions are, again, fragile and often implausible but worth noting. Less exotically, information overload also can induce errors quite apart from the exercise of discretion,[225] and the lesson is similar. So, too, for the admittedly sketchy evidence of ideological influence increasing as precedent accumulates. Under these conditions, however rare, pursuing a low-source strategy makes abundantly good sense. The vexing trade-offs fade away if more information yields only additional costs, even though the quality of decisions might be capped at an uncomfortably low level.

Turning to other working theories of behavior, the predicted effects are more mixed and less convenient. First, we have evidence suggesting that sometimes decisionmakers ignore additional information.[226] If so, adding sources will neither increase nor decrease decision quality. For the same reason, decision costs need not increase either, apart from any upstream costs to lawyers and others of collecting and presenting information that gets ignored. This pattern of decisionmaker behavior indicates that increasing information does not necessarily increase decision costs, or anything else, despite the general formulations of traditional institutional design. The idea that people may fail to process information presents a broad behavioral challenge to ordinary cost-benefit analysis as applied to institutional design, and this challenge is not special to issues of interpretation.

More curiously, even strangely, we also have evidence that increased information is sometimes processed in rule-defying yet potentially beneficial ways. Take the indications of paper pushing through increased deference or status quo preferences. In a judiciary, these moves tend to push power down from appellate tribunals to the trial level and, in turn, out to private ordering and ordinary politics. Those moves are not any improvement per se in the legal system, societal well-being, or other values. Still, there seems to be an inverse relationship between the amount of information processed by courts

---

[225] *See supra* text accompanying notes 130–34.

[226] *See supra* text accompanying notes 118–23 (discussing the general proposition); *supra* Table 3 (failing to find that the number of sources cited affected reversal rates); *cf. supra* Table 4 (failing to find that spinning increased as doctrinal factors increased).

and the amount of power centralized in the top tiers of the formal judicial hierarchy. This connection is worth further thought. In their most flattering light, information torrents in judiciaries can effectively disperse decisionmaking power to low-level decisionmakers who are aware of and rightly sensitive to the features of particular situations, along with other complex social systems such as markets, politics, and social norms. Source proliferation can yield decentralization.

Finally and most hopefully, flooding a legal institution with information might help crystallize law's priorities. The phenomenon of quasi-legal reason minimizes burdens on decisionmakers who are unwilling or unable to carry out deep investigations of every consideration made relevant by formal law. But the phenomenon also includes an effort to isolate what seems to matter most within law, on a principled lawyer's understanding of what law prizes. Of course this process of distillation can go wrong, it cannot be fully detached from the preferences of the decisionmakers, and there might be better ways of achieving the distillation than flooding judiciaries with cases, doctrinal factors, and evidentiary disputes. But if we are willing to consider another charitable view of a behavioral pattern, we might see quasi-law as a process in which a felt necessity driven by information burdens helps deliver legal insight without very much invention.[227]

Because the behavioral effects of source proliferation are varied and not yet fully specified for clearly identified domains in law, those responsible for developing these institutions are faced with tricky interim judgments. Choices about information loads are inescapable, yet our understanding of information-load effects is highly imperfect.[228] One option is to enumerate the plausible behavioral consequences (no effect, more mistakes, paper pushing, quasi-law, and so on), assign probabilities to each as best we can, and then take into account this weighted range of possibilities. The resulting estimates can be integrated with other considerations in designing judiciaries to process less or more information. Another option is to wait for more confident identification of domains in which each behavioral pattern is likely to emerge before making any significant move to increase or decrease information flows within courts. However, if it is best to wait, there is little sense in holding firm to any conventional wisdom about

---

[227] Any beneficial effects of source proliferation on human decisionmaking should influence our interest in automating law-related decisions and our investments in machine learning to facilitate such decisions.

[228] *Cf.* Vermeule, *supra* note 12, at 173–81 (offering decision strategies for making choices as to interpretive method while difficult empirical questions are investigated, including a strategy of concentrating on known effects over speculative effects).

what source proliferation does to judicial decisionmaking. At minimum, those old ideas should be abandoned.

CONCLUSION

As a logical matter, there is little reason to believe that adding interpretive sources will systematically increase interpreter discretion. This effect depends on the type of interpreter and the interpretive rules, which can just as easily reduce discretion as sources increase. As an empirical matter, the link between source proliferation and interpreter discretion is not well supported either.

Predicting the effects of information loads on law's decisionmakers remains challenging, though. For their part, judges are professionals of a particular stripe and the governing architecture of judicial decisionmaking is opaque. This Article and prior work are only enough to develop plausible working theories for the effects of source proliferation.

But there should be no real doubt that conventional wisdom in this field—however meme-like—is vulnerable to careful logic and sustained empirical investigation. Observers interested in judicial discretion have good reason, both logical and empirical, to look beyond source proliferation to find it. And observers interested in institutional design have good reason to rethink the range of consequences when judges look out over a crowd of sources.