

Above the Rules: A Response to Epstein and King

Frank Cross†
Michael Heise††
Gregory C. Sisk†††

Much of what Professors Lee Epstein and Gary King say in *The Rules of Inference*¹ is true and good. Indeed, their discussion of research methods provides a very helpful guide for those who produce and consume empirical legal research, both quantitative and qualitative. As they suggest, greater attention to methodology will both inform and enhance a particular research genre—quantitative empirical legal scholarship—that is quickly coming into its own.² Unfortunately, to make their point, Epstein and King devote the bulk of their article to an unremitting and excessive attack on the current state of empirical legal research methodology. Although some of their attacks are well aimed, on too many occasions their shots miss the targets they seek. More important, however, is that their assault on legal scholarship violates many of their own rules of inference. These violations degrade their analyses and erode confidence in their conclusions and recommendations. In the end, despite its promising thesis, their article becomes an exemplar of how descriptive research should not be conducted.

Epstein and King's article might be considered a theoretical argument for certain rules for drawing inferences, but as they note, most works presenting a theory are grounded in descriptive claims. That is, such works make "at least some claims about the world based on observation or experience."³ This is manifestly true of their own article. They present a fairly elaborate proposal for change in the legal publishing process.⁴ They call for "substantial improvements in legal scholarship,"⁵ which they assert is currently "*deeply flawed*."⁶ They de-

† Herbert D. Kelleher Centennial Professor of Business Law, University of Texas School of Law; Professor, McCombs School of Business, University of Texas at Austin.

†† Professor of Law, Case Western Reserve University.

††† Richard M. & Anita Calkins Distinguished Professor of Law, Drake University.

¹ 69 U Chi L Rev 1 (2002).

² Increasing attention focuses on empirical legal scholarship as a developing research genre. See, for example, *Symposium on Empirical and Experimental Legal Scholarship*, 2002 U Ill L Rev (forthcoming).

³ Epstein and King, 69 U Chi L Rev at 3 (cited in note 1).

⁴ See *id* at 125–30.

⁵ *Id* at 133.

⁶ *Id* at 6.

vote a considerable portion of their article to establishing this descriptive claim about the shortcomings of legal research. While we are in substantial agreement with their rules of inference, given the structure and nature of their arguments it seems only fair to apply those rules to Epstein and King's descriptive claims about the current state of research. Under the scrutiny of their own standards, support for their assertions about the prevailing condition of empirical legal research is sorely wanting. We review their effort in terms of their selection and collection of data, revelation of the research process, definition and testing of hypotheses, and drawing inferences from the empirical tests conducted.⁷

Before turning to the rules of inference, it is important to acknowledge the authors' wise admonition that researchers conducting studies should approach their work with a proper scholarly attitude. Epstein and King observe that researchers should "seek out all evidence against their 'favored' theory"⁸ and objectively test the validity of that theory with data.⁹ Regrettably, their own article reveals little evidence of such an effort. While we cannot know the subjective intent of the authors, their article does not read like an evenhanded assessment in which they seek to examine all the vulnerabilities of their claims about legal research.¹⁰ Indeed, another prominent political

⁷ Given our space constraints, we have not followed the exact, lengthy outline of Epstein and King's article but have somewhat condensed their suggestions.

⁸ Epstein and King, 69 U Chi L Rev at 76 (cited in note 1).

⁹ Id at 76–80.

¹⁰ We will detail only two examples of problems with their treatment of work by others, by way of points of personal privilege. Epstein and King twice criticize the coding of variables in research authored in part by Professor Cross. Id at 87–89, 95–96 (criticizing Frank B. Cross and Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 Yale L J 2155 (1998)). They object that the research coded judges by the party of the appointing president as a proxy for the judges' ideological preferences, arguing that other coding approaches might have more precisely captured those preferences. Id. While the general point is worthy of consideration, their use of the point in criticism is misplaced. In fact, coding by the party of the appointing president was the virtually universal convention at the time, employed by appellate court research conducted by political scientists and published in peer-reviewed journals. See Daniel R. Pinello, *Linking Party to Judicial Ideology in American Courts: A Meta-analysis*, 20 Just Sys J 219, 222 (1999) (observing that scholars have traditionally used this coding convention). Pinello's meta-analysis used eighty-four studies relying on such coding, including twenty-six studies of federal circuit courts, most of which were published in peer-reviewed journals. Id at 225–26. The only alternative coding approach suggested by Epstein and King is a set of scores developed by Professor Giles that had not yet been published. Epstein and King, 69 U Chi L Rev at 95–96 (cited in note 1) (discussing Michael Giles, Virginia Hettlinger, and Todd Peppers, *An Alternative Measure of Preferences for Federal Judges*, 54 Polit Rsrch Q 623 (2001)). The criticism thus amounts to an objection that our 1998 article did not employ data that were first made available in 2001. It is also noteworthy that the Cross and Tiller article was selected as one of only three empirical articles on the process of judicial decisionmaking included in Professor Epstein's recently published anthology of political science research on courts. See Walter F. Murphy, C. Herman Pritchett, and Lee Epstein, *Courts, Judges, and Politics: An Introduction to the Judicial Process* 642–47 (McGraw-Hill 5th ed 2002). Yet in the Epstein

scientist (albeit from a different branch of the discipline) has recently characterized several of the works attacked by Epstein and King as “very good behavioralist work.”¹¹

I. COLLECTING DATA, SELECTING OBSERVATIONS, AND TARGETING POPULATIONS

A. Collecting Data

As Epstein and King understand well, a study can be no better than the data on which it is based, and that data need to be properly drawn and defined. Epstein and King set forth particular rules for how a random probability sample should be drawn.¹² Gathering data and constructing datasets are also critical to the important goal of replication. The ability of scholars to replicate each other’s work independently is a central component of the scholarly enterprise, and it is one that Epstein and King rightly emphasize.¹³

To illustrate problems they find in legal scholarship on these points, Epstein and King brand one study by Theodore Eisenberg and his colleagues¹⁴ as “problematic”¹⁵ and criticize the researchers for in-

and King article, our work is enumerated among the flawed pieces of empirical research found in law reviews. The fact that this coding matter was presented as a research flaw, rather than an issue to be considered, evidences the authors’ predetermined agenda.

Epstein and King also criticize research conducted in part by Professors Heise and Sisk. See Epstein and King, 69 U Chi L Rev at 56–57 (cited in note 1) (criticizing Gregory C. Sisk, Michael Heise, and Andrew P. Morriss, *Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning*, 73 NYU L Rev 1377 (1998)). They complain about a single sentence fragment in a 124-page article as inadequately acknowledging the full extent of prior research in a theoretical field. Epstein and King, 69 U Chi L Rev at 56–57 (cited in note 1). In context, the statement that no stable public choice theory has been established to explain judicial decision-making remains correct, in our view. See Sisk, Heise, and Morriss, 73 NYU L Rev at 1391–92. Nonetheless, we are gratified that Epstein and King immediately qualify this criticism by acknowledging that we have been “well integrated into the academic community” and that our article in other respects does include references to and analysis of prior research. Epstein and King, 69 U Chi L Rev at 57 (cited in note 1). Indeed, that particular article cites to and engages more than forty works found in the valuable library of academic journals existing beyond the law reviews. See Sisk, Heise, and Morriss, 73 NYU L Rev at 1385–96, 1451–98. While we certainly differ with Epstein and King on many points, we wholeheartedly agree that law professors should acknowledge and integrate their work with that of our colleagues elsewhere in the academy, fully appreciating the rich lode of research on law-related matters available in the journals of other disciplines. Epstein and King also criticize the three sources we cite to regarding the problem of multicollinearity. Epstein and King, 69 U Chi L Rev at 11 n 28 (cited in note 1). However, they assert error and provide no evidence for their claim. We remain confident that our caution in applying the most stringent standard to this common problem remains the better part of scholarly discretion.

¹¹ Howard Gillman, *What’s Law Got to Do With It?: Judicial Behaviorists Test the “Legal Model” of Judicial Decision Making*, 26 Law & Soc Inquiry 465, 467 (2001).

¹² Epstein and King, 69 U Chi L Rev at 108–12 (cited in note 1).

¹³ See id at 38–45.

¹⁴ Theodore Eisenberg, Stephen P. Garvey, and Martin T. Wells, *Jury Responsibility in Capital Sentencing: An Empirical Study*, 44 Buff L Rev 339 (1996).

adequately describing their sample selection for a study of jurors in capital cases, thereby impeding efforts to replicate their work by other scholars.¹⁶ Epstein and King go on to demand that Eisenberg and his colleagues provide in their article's text even more detail than they already supply about how they derived their random sample of jurors.¹⁷ Likewise, Epstein and King criticize Steven Engel's research¹⁸ because he may have considered only "key" events and therefore requires readers to rely on his subjective evaluation of data he implies are representative.¹⁹ Like Eisenberg and his coauthors, Engel's methods are taken to task by Epstein and King in part because of the burden they place on the reader²⁰ and because they impede replication.²¹ But Epstein and King's article evidences these same shortcomings to a degree far greater than the research they have criticized.

One first notices that Epstein and King's study of empirical legal research is obviously not replicable. Their failure to permit replication of their work is especially odd in light of their express admonition that "another researcher should be able to understand, evaluate, build on, and reproduce the research without any additional information from the author."²² An independent researcher seeking to replicate the claims made by Epstein and King might well retrace their "top law review" and "empirical in title" methods of data gathering. But telling us that they also "read widely" and reviewed articles voluntarily sent to them precludes other scholars from replicating their work.²³ They criticize one study because it was "impossible to specify, from [the authors'] description, the number in their original sample."²⁴ This is obviously the case with their article. Assuming that there was no overlap between the "top law review" and "empirical in title" methods (which

¹⁵ Epstein and King, 69 U Chi L Rev at 40 (cited in note 1).

¹⁶ Id at 40–41.

¹⁷ By noting the criticisms that Epstein and King level against the Eisenberg team as well as other researchers and articles, we do not mean to endorse Epstein and King's charges. In this instance, we find many of the complaints made by Epstein and King to be hypertechnical, that is, the Eisenberg team provides the information that Epstein and King seek, just not in the style and detail that pleases them. See Eisenberg, Garvey, and Wells, 44 Buff L Rev at 349–50 (cited in note 14). Moreover, the Eisenberg team goes on to caution readers about the limitations of its conclusions imposed by its methodology. Id at 350.

¹⁸ Steven A. Engel, *The McCulloch Theory of the Fourteenth Amendment: City of Boerne v. Flores and the Original Understanding of Section 5*, 109 Yale L J 115 (1999).

¹⁹ Epstein and King, U Chi L Rev at 42 (cited in note 1).

²⁰ Epstein and King correctly note that "the burden of proof in empirical research always remains with the researcher." Id at 41.

²¹ Id at 42.

²² Id at 38.

²³ See Part I.B (discussing Epstein and King's method of selecting observations).

²⁴ Epstein and King, 69 U Chi L Rev at 104 (cited in note 1).

they do not disclose), we can only assume that they considered some number of articles greater than 353.²⁵

Further restricting the ability of others to retrace Epstein and King's research steps and verify their results are the amount and type of discretion they used in gathering their data. They characterize their data as "something very roughly approximating a representative sample of all empirical research in the law reviews"²⁶ that, if anything, is skewed to include above-average legal research. Although we are not quite sure what they mean by "roughly approximating," we are quite sure that their level of methodological rigor falls far short of what they rightly expect from others. Even though their guesses could strike some as plausible, Epstein and King are generally not satisfied with the plausible, at least when it comes to other researchers. Indeed, they object to the fact that in many legal articles, "the researcher exercises complete discretion over what observations to include (sometimes called 'purposive sampling'),"²⁷ yet they appear to exercise considerable purposive discretion of their own in selecting articles for criticism.

B. Selecting Observations

Similarly, Epstein and King explain that among a researcher's most critical decisions—upon which a study's validity partially pivots and which bridges measurement and estimation—is the selection of data.²⁸ Interestingly, Epstein and King reveal that they selected only "articles published over the last five years in six top law reviews."²⁹ They identified 162 such articles.³⁰ They also read all 231 articles in any law review published from 1990 to 2000 that had the word "empirical" in the title³¹ and supplemented these sources by "reading widely through law reviews, following citations, and reading further."³² In addition, they considered some articles voluntarily sent to them by legal academics.³³ They give no further information on how many articles fell into their sample, let alone a list of the articles they selected for

²⁵ The 353 figure represents our best guess as to the number of articles that Epstein and King included in their sample. In their Article they reference 231 articles published in all U.S. law reviews between 1990–2000, see Epstein and King, 69 *U Chi L Rev* at 16 (cited in note 1), fifty top-cited articles written by legal academics and published in law reviews, *id.*, and seventy-two articles from the four leading peer-reviewed law journals, *id.* at 16 n 40.

²⁶ *Id.* at 16.

²⁷ *Id.* at 105.

²⁸ *Id.* at 99–114.

²⁹ *Id.* at 3 n 5.

³⁰ *Id.* at 48 & n 137.

³¹ *Id.* at 16.

³² *Id.*

³³ *Id.*

consideration. Indeed, they leave us with little more than our imaginations to piece together what they did after generating a thick pile of thousands of pages of law review articles. Their actions raise numerous problems.

Epstein and King's use of discretion in selecting observations is even more troublesome (and ironic) in light of their criticisms of a common choice made by many legal scholars studying judicial decisionmaking. Epstein and King correctly note that most such studies focus exclusively on published judicial decisions rather than the full universe of published and unpublished opinions.³⁴ The limitations of judicial decisionmaking studies that draw on only published opinions are well known and understood. Epstein and King identify a few of them in their critique of Kerr's study on how the *Chevron* doctrine has fared in the courts of appeals.³⁵

Practicality is one common reason that drives many legal scholars to invoke a selection rule that limits data to published judicial opinions. Indeed, Epstein and King call for only as much data as is "feasible."³⁶ Even if a complete set of unpublished decisions were available, issues and costs relating to accessibility could easily and quickly mount. In short, in some (perhaps many) instances, a researcher's goal of including the entire universe of unpublished judicial opinions might not, in Epstein and King's words, be "feasible." The impracticability of data access should not, we suggest, entirely close off a more limited and less ambitious line of scholarly inquiry. Instead, when decisions to exclude unpublished decisions are made, it is incumbent upon the responsible researcher to ensure that the study's conclusions account for the data limitation. Finally, in many studies a focus on published (and the exclusion of unpublished) decisions will not necessarily skew the results in any obvious direction.³⁷

Likewise, Epstein and King write, "when an opportunity exists to collect more data, we should generally take advantage of it."³⁸ Practicality is not comfortably available to them as an explanation for limiting their data selection. In contrast to those whom they criticize, Ep-

³⁴ Id at 106.

³⁵ Id at 106–08 (discussing Orin S. Kerr, *Shedding Light on Chevron: An Empirical Study of the Chevron Doctrine in the U.S. Courts of Appeals*, 15 Yale J Reg 1 (1998)).

³⁶ Epstein and King, 69 U Chi L Rev at 102 (cited in note 1).

³⁷ The criticism of reliance on published opinions has been made and addressed in previous research. One might well justify exclusive focus on published opinions on the grounds that these holdings are the ones with precedential value. In fact, with the expansion of Lexis and Westlaw, most unpublished opinions are now available for and employed by researchers. See, for example, Richard L. Revesz, *Ideology, Collegiality, and the D.C. Circuit: A Reply to Chief Judge Harry T. Edwards*, 85 Va L Rev 805, 809–10 (1999) (using unpublished opinions made available by Lexis and Westlaw).

³⁸ Epstein and King, 69 U Chi L Rev at 102 (cited in note 1).

stein and King invoke data selection decisions even though the larger universe is both readily available and accessible. Put slightly differently, it is entirely feasible to assume that they could have extended their sample beyond law reviews, especially in light of the breadth of their conclusions. Similarly, access to empirical legal scholarship not found in law reviews (at least to the published studies) cannot be used to justify their decision, as Epstein and King expressly urge law professors to mine the treasure trove of research journals “in a broad array of disciplines,” that is, beyond the law reviews available on Westlaw and Lexis (and elsewhere).³⁹ Finally, once data gathering and coding tasks are complete, Epstein and King point out the value of authors creating “storehouses for data sets they have created or deposit[ing] data and documentation in public archives.”⁴⁰ Indeed, as social scientists, Epstein and King are especially well positioned to access the growing amount of empirical legal scholarship published in scholarly journals other than law reviews and inaccessible through Westlaw or Lexis searches. Yet they have not storehoused their data from this study in such a manner.

Even though the far-ranging critiques that Epstein and King level and the inferences they wish readers to draw traverse well beyond the boundaries marked by law review publication (“*the current state of empirical legal scholarship is deeply flawed*”⁴¹), for their sample they limit themselves to, in their own words, “something very roughly approximating a representative sample of all empirical research in the law reviews.”⁴² They offer neither justification nor explanation for their decisions to draw the data selection lines where they did. Even more damaging is their failure to adequately rein in the breathtaking sweep of their conclusions in light of their restrictive data selection decisions.

C. Targeting Populations

Epstein and King further observe that a “critical step” in research “is to identify the target *population*.”⁴³ This identification is necessary in order to ensure that the data sample corresponds to the research target. They criticize a study by Nicole Veilleux examining all the Supreme Court stay decisions published over a fifteen-year period for failing to specify the period for which she seeks to draw inferences.⁴⁴ They question whether she can fairly draw inferences about decisions

³⁹ Id at 58.

⁴⁰ Id at 46.

⁴¹ Id at 6.

⁴² Id at 16.

⁴³ Id at 99.

⁴⁴ Id at 101.

outside the time frame of her data.⁴⁵ Yet Epstein and King examine only a subset of data from an even shorter time period (five or ten years) and yet seem to draw astonishingly broad conclusions about the entire universe of empirical legal research.

II. FULLY REVEALING THE RESEARCH PROCESS

Epstein and King stress that researchers need to reveal much more about how they undertook their research.⁴⁶ To promote greater accuracy and less uncertainty, they argue that researchers—within the text of their published works—should “reveal far more about the process by which they generated and observed their data—the whole process from the time the world produced the phenomena of interest to the moment when the data were in their possession and considered final.”⁴⁷ They correctly emphasize that the methods employed should be “transparent” in order for others to appreciate the significance of the research.⁴⁸ Epstein and King stress that researchers must identify variables and then “*measure* those variables and derive *estimates*.”⁴⁹ They also note that “[a] major source of unreliability in measurement is vagueness.”⁵⁰ Despite these cautions, Epstein and King themselves are imprecise in describing their process of data generation and utterly obscure how they observed and classified that data.

The authors provide us with almost no information in the text of their own article about the analytical processes they employ, other than a nonreplicable description of how they collected and read approximately 353 or more empirical studies⁵¹ published in various law reviews. We are told nothing about what they did with the raw data, how (or whether) they divided the studies into categories for comparable analysis, or how they managed to identify the many errors (that is, violations of their rules) that they tell us are present in each study. Nor do they explain how they took into account the different contexts, various pertinent methodologies, and unique characteristics of the hundreds of studies they analyzed, which span a wide range of legal subjects—at least some of which presumably are not fields in which Epstein and King possess substantive expertise.

Given their preceding prescriptions for conducting research, it is nothing short of remarkable that Epstein and King expressly concede

⁴⁵ Id.

⁴⁶ Id at 34.

⁴⁷ Id.

⁴⁸ Id at 43.

⁴⁹ Id at 80.

⁵⁰ Id at 83.

⁵¹ As previously discussed, we simply cannot discern from their article the precise number of articles they included in their sample. See note 25 and accompanying text.

that they “do not tabulate the particular types of mistakes” identified.⁵² Perhaps they deserve some credit for disclosing such a devastating admission, even if they do so in a footnote. They state that they tried to tabulate but found that “most articles make such a complicated and interrelated set of mistakes” that they “could not find a way to quantify errors.”⁵³ While Epstein and King question the reliability and validity of coding judgments in prior research, they themselves apparently made *no* coding decisions whatsoever (save, of course, for their decision not to code). One suspects that if they were reviewing their own “could-not-find-a-way-to-code” statement in a law review article written by someone else, they would have considered the declaration a cop-out at best. In any event, their inability to code and quantify the errors they so prominently dissect should have precluded them from advancing their strong qualitative conclusions about the state of empirical legal research.

The rigor prescribed by Epstein and King in their rules of inference is important because it forces researchers to examine their presumptions and present their data and inferences clearly. It is also important because of the possibility that a “researcher, however inadvertently, has biased a measure in favor of his or her pet hypothesis.”⁵⁴ Consequently, “human judgment should be removed as much as possible.”⁵⁵ This very possibility manifests itself in their own work.

Epstein and King confess to subjectivity in the manner by which they put their data to use. As they acknowledge, the individual studies they hold up for public critique in their article were selected because they illustrated a particular part of their argument against the quality of legal research.⁵⁶ Thus, they effectively admit that the examples adduced in their article were not selected in any fashion even remotely approaching random, but rather were chosen to further their advocacy purpose. And this violates yet another of their rules. In empirical research, they contend that theories and hypotheses should not be defended “as if they were clients in need of the best possible representation.”⁵⁷ Yet the legal studies converted into object lessons in their article were selected to promote their own theories and conclusions. Although they also insist that these examples of error are representative or typical (however that might be defined in this context) of the flaws found in other research published in the law reviews, they again pro-

⁵² Epstein and King, 69 U Chi L Rev at 15 n 36 (cited in note 1).

⁵³ *Id.*

⁵⁴ *Id.* at 83.

⁵⁵ *Id.* at 85.

⁵⁶ See *id.* at 15.

⁵⁷ *Id.* at 10.

vide no objective basis to evaluate this claim independently. Consequently, Epstein and King's study appears to fulfill its own prophecy.

In this regard, consider their treatment of a study conducted by Vicki Schultz and Stephen Petterson on the rate of success by plaintiffs in Title VII employment discrimination cases in which courts addressed the "lack of interest" defense.⁵⁸ Comparing the rate of success in such cases rendered during two time periods (1967–77 versus 1978–89), Schultz and Petterson found a "striking trend" in which the rate of success by plaintiffs declined from 86 to 40 percent.⁵⁹ Epstein and King critique these researchers' choice of those particular time segments for comparison, arguing that "we can radically alter the conclusion reached by Schultz and Petterson by (1) assuming that all twelve losses [by plaintiffs] in the 1978 to 1989 period actually occurred between 1978 and 1980 and (2) moving those losses into a reconfigured data category (1967–1980)."⁶⁰ Of course, we indeed can change the outcome by making that assumption, but the assumption is one that Epstein and King acknowledge is "contrived" and "extreme."⁶¹ Moreover, they level this criticism even though they concede "we [Epstein and King] do not know for sure."⁶² While it might have been preferable for Schultz and Petterson to present their data on an annualized basis, Epstein and King have at best nitpicked at the article's conclusions.

III. DEFINING AND TESTING HYPOTHESES

Epstein and King urge that the "critical first step in making a descriptive inference is to identify the target of the inference."⁶³ If the target is not identified "unambiguously," they argue that "a research project cannot be reasonably evaluated, and hence cannot be successful."⁶⁴ Yet they themselves are extremely obscure when it comes to identifying their own target. It is unclear whether Epstein and King are criticizing empirical legal research published (a) in law journals,⁶⁵ (b) by law professors,⁶⁶ or (c) by law professors who lack Ph.D.s.⁶⁷ They seem to make all three claims, but their research design is ill-equipped

⁵⁸ Vicki Schultz and Stephen Petterson, *Race, Gender, Work, and Choice: An Empirical Study of the Lack of Interest Defense in Title VII Cases Challenging Job Segregation*, 59 U Chi L Rev 1073 (1992).

⁵⁹ *Id.* at 1097–98.

⁶⁰ Epstein and King, 69 U Chi L Rev at 52 (cited in note 1).

⁶¹ *Id.*

⁶² *Id.* at 52.

⁶³ *Id.* at 30.

⁶⁴ *Id.* at 31.

⁶⁵ *Id.* at 11 n 28.

⁶⁶ See, for example, *id.* at 1, 114.

⁶⁷ *Id.* at 9–10.

to test at least the latter two claims. The research they expressly criticize includes studies that were not conducted by law professors⁶⁸ and studies conducted by law professors and others who hold Ph.D.s⁶⁹ — evidence that, if anything, contradicts their conclusions. Had they coded some dependent variable of research quality, which they either could not or would not do, they could have coded authors and tested their hypotheses with data.

Epstein and King present no hypotheses explicitly and clearly. Instead, they present a rather amorphous critique of empirical legal research and, as previously described, pull a handful of examples from the literature to support their position. Hence, it appears that they developed their “theory from the same evidence [they] used to evaluate it, [so] the theory was not vulnerable to being proven wrong.”⁷⁰ While Epstein and King urge that theories must produce observable implications,⁷¹ that researchers should delineate how implications can be observed,⁷² and that researchers should measure and estimate their data,⁷³ they openly defy all of these rules. Even if their theories are correct, such a conclusion cannot properly or responsibly flow from their study.

As Epstein and King aptly note, the provision of summary descriptive statistics is valuable.⁷⁴ Having previously indicated that they could not code variables,⁷⁵ they are by definition precluded from generating and presenting any helpful statistics. However, this absence of empirical evidence somehow did not deter them from advancing the conclusion that a “healthy portion of the work published in the law reviews” is “improperly conducted.”⁷⁶ The phrase “healthy portion” does not really qualify as a helpful finding, nor does it have a real basis in their limited evidence. Mysteriously (indeed, almost incredibly), they suggest that a tabulation of the sort required by rules of inference “would serve no material purpose.”⁷⁷

⁶⁸ For example, with respect to Cross and Tiller, 107 Yale L J 2155 (cited in note 10), neither Cross nor Tiller was formally a professor at a law school at the time that their article was published.

⁶⁹ For example, among the five researchers involved with the two studies we know best, Cross and Tiller, 107 Yale L J 2155 (cited in note 10), and Sisk, Heise, and Morriss, 73 NYU L Rev 1377 (cited in note 10), three of the researchers (Tiller, Heise and Morriss) hold Ph.D.s. Tiller, Heise, and Morriss also hold J.D. degrees.

⁷⁰ Epstein and King, 69 U Chi L Rev at 64 (cited in note 1).

⁷¹ Id at 65–70.

⁷² Id at 73–76.

⁷³ Id at 80–99.

⁷⁴ See id at 25.

⁷⁵ Id at 15 n 36.

⁷⁶ Id at 12 n 29.

⁷⁷ Id at 15 n 36.

The limitations of Epstein and King's claims are apparent from a simple effort to code their own article. While the precise number of articles in their data set is not known, it appears to be in the hundreds. They appear to criticize two articles for failure to consider rival hypotheses,⁷⁸ two articles for coding reliability problems,⁷⁹ and three articles for coding validity problems.⁸⁰ The proportion of articles in which shortcomings are identified is very small, even if one were to accept all the authors' characterizations of work by others.⁸¹ While some additional number of articles surely suffers similar shortcomings, Epstein and King provide no evidence that enables a measure of the magnitude of the problems they associate with legal research. The shortcomings of their article are well encapsulated by the passage in which they claim that the law reviews are "replete" with a particular error and back that claim with an "exemplary" reference to a single article.⁸² This, of course, is exactly the sort of "proof by anecdote" that their rules rightly reject. Indeed, "proof by anecdote" is precisely the type of problem that empirical legal scholarship seeks to address.

IV. DRAWING INFERENCES FROM TESTS

Epstein and King provide guidance for drawing conclusions from the data and hypotheses tested. They appropriately caution against drawing broad inferences from a limited data source. Epstein and King object that legal research often contains "stridently stated, but overly confident, conclusions"⁸³ that go well beyond what the research will support. They caution authors not to be "definitive,"⁸⁴ but to estimate the degree of uncertainty inherent in their conclusions.⁸⁵ Indeed, they use the phrase "all inference in research is uncertain" in a sub-heading in their article.⁸⁶

As they display not a whiff of uncertainty in their descriptive claims, Epstein and King violate this prescription with particular flagrancy. They claim that "many" law professors engaged in empirical research have "not" turned their interest "into productive research practices,"⁸⁷ that "few law professors conducting empirical research now seem to have much facility with the rules of inference,"⁸⁸ and that

⁷⁸ See *id.* at 76–80.

⁷⁹ See *id.* at 83–87.

⁸⁰ See *id.* at 87–97.

⁸¹ See note 9.

⁸² See Epstein and King, 69 *U Chi L Rev* at 111 (cited in note 1).

⁸³ *Id.* at 7.

⁸⁴ *Id.* at 51.

⁸⁵ *Id.* at 50.

⁸⁶ *Id.* at 49.

⁸⁷ *Id.* at 11.

⁸⁸ *Id.* at 114.

the deeply flawed empirical legal scholarship displays “little awareness of, much less compliance with, the rules of inference.”⁸⁹ They imply that 100 percent of the 231 articles with “empirical” in the title violate a rule of inference,⁹⁰ but do not cite all 231 articles, much less establish that each of them contains such a violation. Most remarkably, they contend that “the complete list of all law review articles devoted to improving, understanding, explicating, or adapting the rules of inference is as follows: none.”⁹¹ Epstein and King are no doubt well aware that the means to establish this extreme claim with confidence would involve the survey of every law review article ever written, which of course they have not performed.

Other self-deconstructing ironies in their article abound. Epstein and King question the value of the law review publishing process and the benefits of receiving peer review,⁹² but they obviously have chosen to publish their article in a law review, without the benefit of formal blind peer review. They point out the benefits of interdisciplinary co-authorship and observe that social scientists “would benefit from the substantive expertise law professors” offer,⁹³ but have chosen not to avail themselves of the full array of those benefits.

Epstein and King also put forth a fairly detailed proposal for reforming the law review publishing system used by legal academics, a proposal that they “cannot help but believe will lead to substantial improvements in legal scholarship.”⁹⁴ The basic thrust of this proposal is to make the system more like the convention in social science and many other disciplines. The necessary data to support such a recommendation, though, must be comparative. No amount of trashing legal scholarship for inferential errors can make the case for such a shift without a comparable evaluation of the social science scholarship that they prefer.⁹⁵ We are mindful of shortcomings in the current state of legal research and suspect that many of Epstein and King’s proposals on this point might well lead to an improvement. Although their suggestions are not entirely new,⁹⁶ they warrant debate and discussion. But they certainly have not established the necessary inferences to

⁸⁹ Id at 6.

⁹⁰ See id at 17 n 41.

⁹¹ Id at 11.

⁹² Id at 125–30.

⁹³ Id at 123.

⁹⁴ Id at 133.

⁹⁵ Again, Epstein and King expressly disclaim any suggestion “that empirical research appearing in law reviews is always, or even usually, worse than articles in the journals of other scholarly disciplines.” Id at 18. Yet exactly this conclusion would seem crucial to a proposal that legal research shift to an approach used by other disciplines.

⁹⁶ Many of Epstein and King’s suggestions and more are considered in a 1995 symposium issue of the Stanford Law Review focusing on the law review process. See *Law Review Conference*, 47 Stan L Rev 1117 (1995).

support their proposal, and they ignore the considerable literature criticizing aspects of the peer-review process. Indeed, the peer-review process itself looks insupportable by the very rules of inference established by Epstein and King.⁹⁷

Epstein and King's proposal for revamping legal publication is intriguing and has some appeal. However, it is important to note that peer-reviewed legal journals currently exist in most legal specialty areas.⁹⁸ This diverse publication marketplace frees academics to make their own judgments about the appropriate research outlets for their work and the value they will ascribe to those outlets when evaluating the research of others. The existence of student-edited journals need

⁹⁷ While the peer-review process sounds good, its reliability and validity is questionable. When tested, the peer-review process has shown itself to be wanting. The classic study is Douglas P. Peters and Stephen J. Ceci, *Peer-Review Practices of Psychological Journals: The Fate of Published Articles Submitted Again*, 5 *Beh & Brain Sci* 187 (1982). In this research, the authors took already-published articles and resubmitted them under less prestigious names. Only 8 percent of the reviewers detected the resubmissions and 89 percent of the reviewers rejected the publication of the articles on grounds such as poor statistical treatment or methodology. *Id.* at 189–91. A more recent study found that the level of agreement between peer reviewers for journals of neuroscience and evaluations of reports at clinical neuroscience conferences was only slightly greater than would be expected by random chance alone. See Peter M. Rothwell and Christopher N. Martyn, *Reproducibility of Peer Review in Clinical Neuroscience: Is Agreement Between Reviewers Any Greater than Would Be Expected by Chance Alone?*, 123 *Brain* 1964, 1966 (2000). See also Stephen Cole, Jonathan R. Cole, and Gary A. Simon, *Chance and Consensus in Peer Review*, 214 *Science* 881, 885 (1981) (reporting substantial inconsistencies in peer review of National Science Foundation grant applications). The poor reproducibility of peer review has also been reported in other fields. See Rothwell and Martyn, 123 *Brain* at 1967.

The peer-review process has other apparent shortcomings. There is reason to believe that the process screens out articles based on the reviewer's attitudinal predispositions rather than methodological validity. See, for example, Stephen J. Ceci, Douglas Peters, and Jonathan Plotkin, *Human Subjects Review, Personal Values, and the Regulation of Social Science Research*, 40 *Am Psych* 994, 1001 (Sept 1985) (reporting that reviewers invoked "seemingly objective scientific criteria" to criticize research proposals "whose real offense might be their social and political distastefulness"); David F. Horrobin, *The Philosophical Basis of Peer Review and the Suppression of Innovation*, 263 *JAMA* 1438 (1990) (arguing that the peer-review process inhibits the dissemination of new ideas). There is evidence that reviewers give article submissions relatively little scrutiny before judging them. See Effie J. Chan, Note, *The "Brave New World" of Daubert: True Peer Review, Editorial Peer Review, and Scientific Validity*, 70 *NYU L Rev* 100, 118–19 (1995) (reporting research to this effect in medical journals). The majority of scholars in numerous disciplines believe that the peer-review process contains a bias for established researchers from prestigious institutions. See Robert L. Jacobsen, *Scholars Fault Journals and College Libraries in Survey by Council of Learned Societies*, 42 *Chron of Higher Educ* 1, 1 (Aug 6, 1986). The assessments of reviewers are also influenced by "conflicts of interest due to friendship or competition and rivalry between the reviewer and the authors." Rothwell and Martyn, 123 *Brain* at 1968. Reviewer recommendations may even be influenced by the typeface used in submissions. See Gideon Koren, *A Simple Way to Improve the Chances for Acceptance of Your Scientific Paper*, 315 *New Eng J Med* 1298 (1986). When "deliberately flawed papers" have been sent out for peer review, the proportion of errors picked up by reviewers is low. Rothwell and Martyn, 123 *Brain* at 1968.

⁹⁸ Epstein and King are aware of at least four such journals. See Epstein and King, 69 *U Chi L Rev* at 16 (cited in note 1). These journals include the *Journal of Law & Economics*, the *Journal of Law, Economics, & Organization*, the *Journal of Legal Studies*, and the *Law & Society Review*.

not be exclusive of peer-reviewed journals but can offer a supplement not available in other disciplines. The high standing granted by law schools to publications in student-edited journals is testimony to the value of those journals. The low standing given to the journals by other disciplines may simply be testimony to their lack of familiarity with the process. Given the shortcomings of the peer-review process,⁹⁹ it is quite plausible that the combination of student-edited and peer-reviewed journals found in legal scholarship is a desirable outcome.

Moreover, whatever particular defects might limit student-edited law journals as a publication system, these defects do not necessarily attach to all the scholarly articles contained in those journals. To be sure, the quality of scholarship (empirical and nonempirical, theoretical and doctrinal) that appears in student-edited law reviews varies, sometimes tremendously. However, several studies published in law reviews—and subject to attack by Epstein and King—are written by researchers who also publish studies in peer-reviewed journals. Presumably, the knowledge and skills they apply to the studies passing muster in the peer-review process do not disappear whenever they select an alternative publication forum.¹⁰⁰ Accordingly, assessments of scholarly quality for any given article should pivot on the particular merits of the article itself rather than whatever one might feel about the procedures employed by the publication forum.

CONCLUSION

In their section on General Guidelines,¹⁰¹ Epstein and King repeat an admonition echoed in most graduate schools. We recall it well and feel it bears repeating as it captures the spirit behind our Response to Epstein and King:

When you publish your paper, out there somewhere will be a graduate student holding a yellow-lined pad of paper with your name scrawled at the top. Expect everything you do to be scrutinized, any rival explanation you have not explored to be tested, and every possible way you could be proven wrong to be dissected. Anticipate. Get to each of these areas before this gradu-

⁹⁹ See note 97. We need not establish that the peer-review process is worse than the current law review publication process to justify a position that the simultaneous existence of both distinct processes is preferable to either one by itself.

¹⁰⁰ The number of legal scholars that publishes in both student-edited law reviews and peer-reviewed journals is both large and growing. Any list, even just to buttress our point, would be hopelessly incomplete. However, the small pool of scholars to whom we cite in our Response provides ample evidence of and support for our point. They include Professors Cross, Eisenberg, Epstein, Heise, Morriss, Revesz, Tiller, and Wells. See notes 1, 10, 14, 37, and 105.

¹⁰¹ Epstein and King, 69 *U Chi L Rev* at 38–54 (cited in note 1).

ate student. When you do empirical research, follow the rules of inference.¹⁰²

Simply put, our main point about the Epstein and King article is that they should follow the very good advice they offer to others. Their purported goal of seeking to improve empirical legal scholarship by articulating rules of inference and applying those rules to the works of others is doomed in part by their own numerous violations of those very rules. We can only hope that their deeply flawed effort to achieve their goal will neither diminish nor deflect attention from a worthwhile and important thesis.

Of the rules of inference set out by Epstein and King, we endorse them all (even when we differ on their application) and feel that a few deserve particular attention. First, articles aspiring to establish a descriptive truth, whether about courts or the world, should be as objective and evenhanded as possible and not grounded in an agenda with a preordained result.¹⁰³ Second, articles should be fully transparent in their procedures and claims, which necessarily requires that researchers use more rigor in presenting their methodology and inferential claims.¹⁰⁴ Third, legal researchers should take advantage of and build upon the considerable body of research on law found in other disciplines, such as political science, economics, and the behavioral sciences.¹⁰⁵

We want to emphasize that embedded in our Response is a larger point that overshadows their article as well as our Response to it. The larger point, one that Epstein and King articulate and we heartily endorse, is that empirical legal scholarship (or, for that matter, all empirical scholarship) needs to adhere to—or at the very least aspire to—the rules of inference. Again, on this crucial point we have no quarrel with and, indeed, join Epstein and King. We take issue only with the way in which they develop their thesis. And attention to how they muster their arguments warrants focus, as the “how” is not with-

¹⁰² Id at 38.

¹⁰³ See id at 76–80.

¹⁰⁴ See id at 38–44.

¹⁰⁵ See id at 56–59. While legal researchers need to account for relevant research from other disciplines, this is not unique to the law, and political scientists could also benefit from more attention to legal research. See Frank B. Cross, *Political Science and the New Legal Realism: A Case of Unfortunate Interdisciplinary Ignorance*, 92 Nw U L Rev 251, 309–10 (1997). The situation in law reviews is improving, though. While Epstein and King support their criticism of legal research only with some casual comments made on a listserv, Epstein and King, 69 U Chi L Rev at 57 & nn 164–66 (cited in note 1), they also note that a more rigorous empirical analysis has found that law review citations to empirical research generally have increased dramatically. Id at 6 & n 17. Legal scholars would benefit from greater use of research from political science and other disciplines, though, including work by nonquantitative interpretivists such as Mark Graber, Howard Gillman, and Keith Whittington.

out consequence. In the life of the mind, scholarship, knowledge, and ideas are rarely advanced by “blood sport.”¹⁰⁶ Epstein and King’s article further illustrates why this is so.

Notwithstanding the shortcomings of *The Rules of Inference*, Lee Epstein and Gary King remain outstanding scholars.¹⁰⁷ We agree substantially with much in their directions on conducting empirical research and rules for drawing inferences. We would caution that their rules are aspirational, and very few studies in any field are clearly in full compliance with all of those rules.¹⁰⁸ Legal researchers should not be intimidated by these rules and the authors’ associated criticisms. Researchers should not be deterred from attempting empirical research but should simply strive to attain the article’s inferential standards insofar as practicable. We fear that Epstein and King were overcome by zeal in the intensity of their criticism of the current state of legal research. Such research warrants and would benefit from a well-conducted study and fair criticism, but Epstein and King’s polemic does not really tell us much about the true state of empirical legal research. Two of us serve as peer reviewers for social science journals and, had their article been submitted to us for review, we would have suggested that the authors revise and resubmit.

¹⁰⁶ Epstein and King, 69 U Chi L Rev at 15 (cited in note 1).

¹⁰⁷ Indeed, their standing obviously lends some credence to their conclusions. But as Epstein and King observe, such arguments from authority are not strong evidence. See, for example, *id.* at 15, 34.

¹⁰⁸ One obvious example is the fact that peer-reviewed political science and economic studies of judicial decisionmaking by courts below the Supreme Court have commonly used the coding decisions criticized by Epstein and King. See note 10.