Fall 2016

The Future of Empirical Legal Scholarship: Where Might We Go From Here?

Kathryn Zeiler
Boston University School of Law

THE FUTURE OF EMPIRICAL LEGAL SCHOLARSHIP: WHERE MIGHT WE GO FROM HERE?

Boston University School of Law
Public Law & Legal Theory Paper No. 16-22

Journal of Legal Education, Volume 66, Number 1
(Autumn 2016)

Kathryn Zeiler
Boston University School of Law

The Future of Empirical Legal Scholarship: Where Might We Go from Here?

Kathryn Zeiler

Introduction

For some time now, scholarly journals have been editing and publishing empirical legal studies. This includes both peer-reviewed journals, in fields such as law, economics, and psychology, and non-peer-reviewed, student-edited law journals. The number of empirical legal studies that show up in the pages of journals has been on the rise as empirical methods improve and researchers gain easy access to a growing number of data sets. This addition to legal scholarship is welcome after decades of theory's dominance. Prior to the entrance of empirical work onto the scene, we were drowning in a sea of predictions derived from unvalidated theories. Journals were filled with guesses, and we blindly applied theory in policymaking, if at all, in the absence of evidence that the relevant theoretical insights were accurate and robust. The situation has improved in recent years as we have sought to verify theories relevant to law and policy using data from a rapidly expanding number of sources.

This is the good news. The troubling news is that empirical scholarship in most fields is under fire. Several studies published during the past decade have revealed the tenuous nature of a shockingly high percentage of some of the most widely cited empirical results. Empiricists are scrambling to find effective ways to restore confidence in published empirical studies. The situation in the field of law is doubly worse. First, the average quality of empirical studies published in student-edited law reviews is undoubtedly lower than those published in peer-reviewed journals. This is partly because law review editorial boards, usually comprising solely law students, do not systematically require expert review of submitted work. Student editors are eager to publish empirical work, but too often lack the expertise to ensure that they publish only high-quality, replicable studies. This system encourages submissions by

Kathryn Zeiler is the Nancy Barton Scholar and Professor of Law, Boston University School of Law. kzeiler@bu.edu. Thanks to Jennifer Arlen, Dawn Chutkow, Cosette Creamer, Michael Frakes, Paul Gugliuzza, Michael Heise, Mat McCubbins, Gregory Mitchell, Michael Meurer and Kate O'Neill for comments and suggestions. Thanks also to my library liaison Jenna Fegreus and research assistant Megan Saumier-Smith (BU Law class of 2017).

authors who are interested in testing theory but lack sufficient expertise in empirical methods. As a result, empirical studies published in law reviews are often plagued by basic problems such as incorrect definitions of technical terms, application of incorrect tests, and misinterpretations of results. Second, although many experts have identified problems with studies published by law reviews, little is being done to remedy them.

This failure to improve is not for lack of ideas. Several methods experts have cataloged the problems and offered numerous suggestions for turning things around. Some have attempted to clarify the rules of inference via article-length tutorials. Others have suggested that student editors be required to take courses in empirical methods taught by experts, that experts be added to law review editorial boards, that nonexperts interested in producing empirical work either enroll in short methods courses or find methodologists willing to co-author, and that law journals design and implement data- and procedure-disclosure policies to facilitate easier replication of published studies. The list goes on. Only one of the ideas has taken hold. A number of universities offer short methods courses designed for law professors interested in producing empirical legal studies. While these courses certainly hold promise, they might unintentionally worsen the situation. The law professors gain knowledge sufficient to use fancy statistical software to employ technical empirical models to produce estimates that get neatly organized into impressive looking tables, but at least some of the short courses provide insufficient training to ensure that the professors can employ the methods at a level that would satisfy expert peer reviewers.

This essay argues that most ideas for improving law review editorial practices have failed because they do not capitalize on the interests of inexpert authors and student editors. To increase the chance of improving quality, the infrastructure of the publication system must be altered to generate incentives for authors to produce high-quality work, and some combination of carrots and sticks must be effectively employed to encourage student editors to ensure high-quality publications. Both authors and students have incentives to protect their reputations. Authors who submit empirical studies to law reviews might take steps to increase the quality of their work if critical reviews of it were posted publicly. Student editors might do more to ensure publication of only high-quality empirical studies if journals were publicly graded on the quality of published studies and their policies related to pre-publication peer review and data availability. The purpose of this essay is to argue that change will occur only if those with the power to effect change have sufficiently strong incentives to improve.

Who would expend the effort to take such steps? The proposal in the following pages relies on an assumption that a sufficient number of methods experts will pitch in to do the work required to generate strong incentives for change. These experts do a far bit now to promote high-quality empirical legal scholarship. Experts expend effort, for example, as directors and members of the Society for Empirical Legal Studies, an academic association

designed to encourage empirical work on legal issues, to stimulate interaction among scholars in law and other academic fields and to organize an annual conference on empirical legal studies.[1] This essay proposes the addition of a fourth mission: to join the efforts of empiricists in other fields who are working to reinvigorate confidence in published empirical work.

The essay is organized as follows. I first make the easy argument for the value of empirical legal studies. I then describe the ongoing quality crisis, not only in the law review literature but also in empirical literature across virtually all fields. I argue that the legal academy should join in the efforts in other disciplines to bolster confidence in published empirical work. The legal academy should, at the same time, take steps to remedy much more basic problems that tend not to plague peer-reviewed empirical scholarship, such as the use of objectively incorrect statistical tests and empirical specifications and the drawing of invalid inferences from empirical results. After laying out a set of possible causes of these basic problems, the essay offers a number of proposed solutions designed to capitalize on the interests of inexpert authors and inexpert editors. The solutions call for instituting incentives for authors and editors to increase the quality of published empirical work. I suggest that existing academic societies already engaged in building infrastructures to foster empirical work are best-suited to institute such incentives.

Value of Empirical Legal Scholarship

Few doubt the value of empirical legal scholarship.[2] It contributes both to the development of "a mature legal science,"[3] which aids in our endeavor to accurately describe and explain what we observe, and to informed policymaking. In both the sciences and social sciences, valid empirical data analysis is essential to verifying predictions based on theories and thus to developing confidence in the theories themselves. Researchers verify by deriving precise hypotheses from the theory under investigation, collecting data—from the field or from the laboratory, for example—and analyzing the

1. The Society for Empirical Legal Studies, http://www.lawschool.cornell.edu/SELS/about.cfm (last visited July 30, 2016).

2. *See generally* Michael Heise, *The Importance of Being Empirical*, 26 Pepp. L. Rev. 807 (1999); Michael Heise, *The Past, Present, and Future of Empirical Legal Scholarship: Judicial Decision Making and the New Empiricism*, 2002 U. Ill. L. Rev. 819 [hereinafter Heise, *Judicial Decision Making*] (comprising just one paper of a number of papers in a symposium issue dedicated to arguing for the benefits of empirical legal scholarship and encouraging its production). I adopt Michael Heise's definition of "empirical legal scholarship" as "the subset of empirical legal scholarship that uses statistical techniques and analyses. By statistical techniques and analyses I mean studies that employ data . . . that facilitate descriptions of or inferences to a larger sample or population as well as replication by other scholars." Heise, *Judicial Decision Making*, *supra*, at 821.

3. Thomas S. Ulen, *A Nobel Prize in Legal Science: Theory, Empirical Work, and the Scientific Method in the Study of Law*, 2002 U. Ill. L. Rev. 875, 900 (2002) ("[E]mpirical work is an absolutely vital part of the development of a mature legal science.").

data to confirm or disprove the hypotheses. Empirical verification is critical in our endeavor to develop models with strong predictive value.

The stronger a theory's predictive value, the more confidently we can apply the theory in prospective policymaking. Empirical validation helps us move from a conversation in the pages of academic literature to the application of science in the form of evidence-based policy. Theory is useful. Theory backed by a single, methodologically sound empirical study is better. We get closer to the ideal, however, when the theory we wish to apply is supported by a collection of methodologically sound empirical studies from a variety of contexts using a number of different methods and different data samples drawn from relevant populations. While robust empirical verification cannot guarantee that some adopted policy will work as intended, it can help guide us toward policies with the best chance of fulfilling their promise.

The usefulness of empirical legal research, however, depends heavily on the methods employed to produce it and on the validity of the inferences drawn from reported results. Its value diminishes if researchers fail to employ best practices and publishers of such studies do not maintain effective checks on the quality of published work.

Continuing Crisis of Low Quality

The quality of scholarship published in student-edited law reviews has been under attack for some time.[4] In more recent challenges, critics have railed against the poor quality of empirical legal studies published by the reviews.[5] My ongoing conversations with colleagues suggest that many in the legal academy and beyond share this negative appraisal. Likely few would dispute the claim that the average quality of empirical work published in law reviews lies somewhere below that of work published in peer-reviewed journals.[6]

4. *See, e.g.*, Roger C. Cramton, "*The Most Remarkable Institution": The American Law Review*, 36 J. Legal Educ. 1 (1986); Richard A. Posner, *The Future of the Student-Edited Law Review*, 47 Stan. L. Rev. 1131 (1995); Fred Rodell, *Goodbye to Law Reviews*, 23 Va. L. Rev. 38 (1936-1937); Fred Rodell, *Goodbye to Law Reviews—Revisited*, 48 Va. L. Rev. 279 (1962). Undoubtedly the legal academy continues to lament the low quality of all types of work published in student-edited law journals. The problems with quality certainly are not limited to empirical studies. The larger quality issues are beyond the scope of this narrow essay. The essay's bottom line, that quality improvements require capitalizing on the interests of the actors with control over quality, can be generalized to other types of scholarship published in law reviews.

5. *See*, *e.g.*, Elizabeth Chambliss, *When Do Facts Persuade? Some Thoughts on the Market for "Empirical Legal Studies,"* 71 L. & Contemp. Probs. 17 (2008); Lee Epstein & Gary King, *The Rules of Inference*, 69 U. Chi. L. Rev. 1 (2002). *See also* Frank Cross, Michael Heise & Gregory C. Sisk, *Above the Rules: A Response to Epstein and King*, 69 U. Chi. L. Rev. 135, 149 (2002) (forwarding a milder critique than Epstein and King's, but conceding that "[t]o be sure, the quality of scholarship (empirical and nonempirical, theoretical and doctrinal) that appears in student-edited law reviews varies, sometimes tremendously.").

6. *But see* Theodore Eisenberg, *The Origins, Nature, and Promise of Empirical Legal Studies and a Response to Concerns* (Cornell Law Faculty Publ'ns Paper 974, 2011), http://scholarship.law.cornell.edu/facpub/974.

To be sure, peer review is hardly a panacea when it comes to ensuring quality.[7] In fact, the social sciences academy is in the midst of its own quality crisis.[8] Recently methodologists have attempted to replicate results published in top social sciences journals with mixed success.[9] Others have produced evidence that journals suffer from publication bias, the tendency of editors to publish statistically significant results over null results.[10] This bias leads to an imbalance in published studies in the sense that studies that fail to identify predicted effects are less likely to see the light of day relative to those that find, perhaps just by chance, statistically significant effects. Still others have been working on ways to detect tricks researchers use to produce statistically significant results with the hopes of easing the road to publication.[11] Such findings suggest that the current peer-review system used by academic journals is insufficient to ensure high quality.

The legal academy and researchers in the social sciences have reacted differently to these concerns. Social science researchers have taken steps to advance solutions and reestablish credibility. Following the firestorm against psychology studies in 2012, Nobel Laureate Daniel Kahneman wrote an open letter to academic psychologists imploring them to "do something . . . and

7. *See* Gregory Mitchell, *Empirical Legal Scholarship as Scientific Dialogue*, 83 N.C. L. Rev. 167, 175 (2004) (reviewing critiques of peer review, offering a defense of the system and ultimately suggesting that student-edited law journals "compel[] the disclosure of important information about empirical research using a common methodological language so that the research may be subjected to critical scrutiny.").

8. *See*, *e.g.*, John Ioannidis & Chris Doucouliagos, *What's to Know About the Credibility of Empirical Economics?*, 27 J. Econ. Survs. 997 (2013). Similar problems have been identified in medical research and the hard sciences more generally. *See*, *e.g.*, John P.A. Ioannidis, *Why Most Published Research Findings Are False*, 2 PLoS Med. 0696 (2005). I focus here on the identified problems and reactions to them in the social sciences.

9. Andrew C. Chang & Phillip Li, *Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not"*, Fin. & Econ. Discussion Series 2015-083 (2015), https://www.federalreserve.gov/econresdata/feds/2015/files/2015083pap.pdf; Colin F. Camerer et al., *Evaluating Replicability of Laboratory Experiments in Economics*, 351 Science 1433 (2016); Stéphane Doyen et al., *Behavioral Priming: It's All in the Mind, but Whose Mind?*, 7 PLoS One e29081 (2012); Stuart J. Ritchie et al., *Failing the Future: Three Unsuccessful Attempts to Replicate Bem's 'Retroactive Facilitation of Recall' Effect*, 7 PLoS One e33423 (2012); Open Science Collaboration, *Estimating the Reproducibility of Psychological Science*, 349 Science 943 (2015). A healthy battle over the soundness of the replication methodology has ensued. *See*, *e.g.*, Christopher J. Anderson et al., *Response to Comment on "Estimating the Reproducibility of Psychological Science,"* 351 Science 1037 (2016); Daniel T. Gilbert et al., *Comment on "Estimating the Reproducibility of Psychological Science,"* 351 Science 1037 (2016).

10. *See*, *e.g.*, Gregory Francis, *Publication Bias and the Failure of Replication in Experimental Psychology*, 19 Psychonomic Bull. & Rev. 975 (2012); John P.A. Ioannidis et al., *The Power of Bias in Economics Research* (Deakin Univ. Working Paper No. 1, 2016), https://www.deakin.edu.au/___data/assets/pdf_file/0007/477763/2016_1.pdf.

11. *See, e.g.*, Joseph P. Simmons et al., *False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant*, 22 Psychol. Sci. 1359 (2011).

. . . do it collectively."[12] In response, a set of thirty-six research groups from around the globe formed the Many Labs Replication Project.[13] The goal of the project was to replicate thirteen effects reported in widely cited psychological studies.[14] They published all data sets and procedures through the Open Science Framework.[15] The Berkeley Initiative for Transparency in the Social Sciences was established in 2012 to "strengthen the quality of social science research and evidence used for policy-making."[16] A number of psychologists are now heading up projects to replicate findings and promote disclosure of materials sufficient to perform such replications.[17] Several private foundations recently donated $10 million to the Center for Open Science, founded in 2013 and directed by members of the academy to "increase openness, integrity, and reproducibility of scientific research."[18]

The field of political science has also taken steps toward increasing research transparency, including access to data. In 2010, the American Political Science Association (APSA) formed an ad hoc committee to formulate a plan.[19] The committee grew into an organization called Data Access & Research Transparency (DA-RT). Members of DA-RT drafted an ethics guide that was eventually adopted by APSA. In addition, a consortium of twenty-seven political science journal editors signed a joint statement to commit individually to implement a set of transparency policies by January 2016.[20] Two scholars kicked off the Political Science Replication Initiative to encourage and facilitate replication of empirical studies in political science.[21]

A similar movement is occurring in economics, but progress seems slower. A growing (but still small) number of economics journals have instituted data-availability policies with the goal of ensuring that those who wish to

12. Open Letter from Daniel Kahneman, Nature (Sept. 26, 2012), http://www.nature.com/polopoly_fs/7.6716.1349271308!/suppinfoFile/Kahneman%20Letter.pdf (suggesting that social psychologists who run experiments set up a "daisy chain" of replications).

13. Many Lab, https://osf.io/89vqh/ (last visited July 26, 2016).

14. *Id.* Ten of the results were successfully replicated.

15. Richard A. Klein et al., *Investigating Variation in Replicability: A "Many Labs" Replication Project*, Open Sci. Framework, https://osf.io/wx7ck/ (last visited July 30, 2016).

16. *Mission and Objectives*, Berkeley Initiative for Transparency Soc. Sci., http://www.bitss.org/about/ (last visited July 30, 2016).

17. Amy Novotney, *Reproducing Results*, 45 Monitor on Psychol. 32 (2014), http://www.apa.org/monitor/2014/09/results.aspx. *See also*, *Replication in Economics,* ReplicationWiki, http://replication.uni-goettingen.de/wiki/index.php/Main_Page (last visited July 30, 2016).

18. *COS Mission*, Ctr. for Open Sci., https://cos.io/about_mission/ (last visited July 30, 2016).

19. *About DART,* Data Access & Res. Transparency, http://www.dartstatement.org/#!about/c24vq (last visited July 30, 2016).

20. *The Journal Editors' Transparency Statement (JETS)*, Data Access & Res. Transparency, http://www.dartstatement.org/#!blank/c22sl (last visited July 30, 2016).

21. Pol. Sci. Replication Initiative, http://projects.iq.harvard.edu/psreplication/home (last visited July 30, 2016).

attempt replication of published results have the necessary information.[22] Researchers from the U.K. and New Zealand have formed the Replication Network, a website designed to "serve[ ] as a channel of communication to (i) update scholars about the state of replications in economics, and (ii) establish a network for the sharing of information and ideas."[23] Economists have offered ideas on how to create demand for replication.[24] Others have nudged us toward methods that physicists have adopted to avoid bias caused by "the desire to support one's theory, to refute one's competitors, to be first to report a phenomenon, or simply to avoid publishing 'odd' results."[25] The Allied Social Science Association held a workshop on replication in economics during its 2016 annual meeting.[26] Researchers at the Federal Reserve Bank are calling for the creation of a journal dedicated solely to replication.[27] The newly formed Critical Finance Review recently announced its plan to publish issues "dedicated to replicating the most influential empirical papers in financial economics."[28]

While empirical legal studies published in student-edited law reviews are experiencing the same problems currently plaguing the social sciences, they

22. Sven Vlaeminck, *Research Data Management in Economic Journals*, Open Econ., http://openeconomics.net/resources/data-policies-of-economic-journals/ [https://perma.cc/FXG7-X8TT] (last visited July 26, 2016) (reporting that 21% of a sample of 141 high-ranking economics journals had a data availability policy in place). A recent study reports, however, that very few of the journals with such policies consistently enforce them. Sven Vlaeminck & Lisa-Kristin Herrmann, *Data Policies and Data Archives: A New Paradigm for Academic Publishing in Economic Sciences?*, *in* New Avenues for Electronic Publishing in the Age of Infinite Collections and Citizen Science: Scale, Openness and Trust 145 (Birgit Schmidt & Milena Dobreva eds., 2015). Seventy-six percent of journals from a sample of economics journals (n = 346) with data availability policies (n = 71) require authors to submit data and procedures with the initial submission or before publication. *Id.* Twenty percent with data availability policies failed to make available the required information for any article in four issues investigated. *Id.* Only one journal with a policy made available data and procedures for every article in four issues. *Id.*

23. Replication Network, https://replicationnetwork.com/ (last visited July 26, 2016).

24. *See* John Cochrane, *Secret Data*, The Grumpy Economist (Dec. 28, 2015), http://johnhcochrane.blogspot.com/2015/12/secret-data.html (suggesting ways to create demand for replication including refusing to referee a paper if the author does not provide the data and procedures used to produce results, refusing to cite to articles if the data and procedures are not accessible, disclosing in tenure letters whether the candidate makes all data and procedures available, refusing to discuss conference papers unless the data and procedures are accessible, and refusing to invite to conferences those who do not disclose their data and procedures).

25. Robert MacCoun & Saul Perlmutter, *Blind Analysis: Hide Results to Seek the Truth*, 526 Nature 187, 187-88 (2015).

26. *Replication and Transparency in Economic Research*, Inst. for New Econ. Thinking, http://ineteconomics.org/community/events/replication-and-transparency-in-economic-research (last visited July 26, 2016).

27. Christian Zimmermann, *On the Need for a Replication Journal* (Fed. Res. Bank of St. Louis, Working Paper, No. 2015-016-A, 2015), http://research.stlouisfed.org/wp/2015/2015-016.pdf.

28. *RFP*, Critical Fin. Rev., http://cfr.ivo-welch.info/RFP.html (last updated Jan. 2016).

are grappling with much more basic problems, most of which likely are wrung out of studies published in peer-reviewed journals. Epstein and King list many of the problems caused by a lack of adherence to the standard rules of inference.[29] From my perspective, the problems are even more basic than their list suggests. Objective errors—errors that methodologists would uniformly agree are mistakes—seem common. A handful of examples include:

1. incorrectly defining statistical terms such as "p-value" and "statistical significance";
2. misinterpreting regression results (e.g., interpreting results as supporting claims of causation when they support only claims of correlation; misinterpreting a null result as proof of no causal relationship);
3. incorrectly applying statistical tests (e.g., performing a simple t-test on samples not drawn from mutually exclusive populations or when the population variances of two mutually exclusive groups are known to be unequal); and
4. failing to account for and verify assumptions necessary to produce valid results (e.g., the commonly employed ordinary least squares regression model assumes independent observations, linearity, the absence of observations that exert undue influence on the estimates, and normally distributed residuals, to name a few).

Given the lack of controversy around whether these types of problems constitute errors, we might ask whether we should expend any effort to reduce the likelihood that studies plagued by these sorts of basic errors get published. It could be that readers easily detect such errors and place appropriately low weight on invalid results. Even if this is the case, problematic studies can slow down scientific progress by shifting time and attention away from forward momentum and toward error correction.[30] On the policy front, concern might be quelled by claims that policymakers don't read anything published in law journals, let alone empirical work published in them.[31] Others assert, however, that at least some legislators, judges and others who craft law and make policy do refer to law journal publications[32] and that the articles influence (or at least justify) policy choices.[33] A (likely very) noisy estimate of how often

29. Epstein & King, *supra* note 5.

30. Michele Landis Dauber, *The Big Muddy*, 57 Stan. L. Rev. 1899, 1912 (2005) (arguing that pre-publication peer review is crucial because if low-quality work is published, disputes over methods must be played out in public). Others worry that "sensational sound bites" from problematic studies can become "political tools." Elizabeth Chambliss, *supra* note 5, at 29.

31. *See e.g.*, Alex Kozinski, *Who Gives a Hoot About Legal Scholarship?*, 37 Hous. L. Rev. 295 (2000).

32. *See e.g.*, David L. Schwartz & Lee Petherbridge, *The Use of Legal Scholarship by the Federal Courts of Appeals: An Empirical Study*, 96 Cornell L. Rev. 1345 (2011).

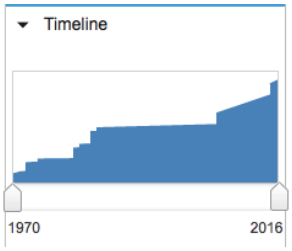
33. *See* Epstein & King, *supra* note 5, at 2, 4-6 ("[R]esearch that offers claims or makes inferences based on observations about the real world—on topics ranging from the imposition of the death penalty to the effect of court decisions on administrative agencies to the causes of fraud in the bankruptcy system to the use of various alternative dispute resolution mechanisms—

courts mention empirical studies published in law reviews suggests that these studies get at least some play in litigation.[34] Epstein and King note that "[t]he staying power of flawed and discredited legal studies can be extraordinary."[35] While studies with basic flaws certainly are not totally devoid of value,[36] they potentially cause more harm than good. At a minimum, the fact that our publication system does not weed them out raises concerns about the general quality of empirical legal studies published in law reviews, which potentially distracts us from important findings.

Proposed solutions with the most promise to increase the quality of empirical work in law reviews must account for the factors that allow poor quality studies to make it into and to slip through the editing process. The next section offers specifics about drivers that lead to low quality publications.

'can play an important role in public discourse . . . and can affect our political system's handling' of many issues.") (citing Ronald J. Tabak, *How Empirical Studies Can Affect Positively the Politics of the Death Penalty*, 83 Cornell L. Rev. 1431, 1431 (1998). *See also* Tracey L. Meares, *Three Objections to the Use of Empiricism in Criminal Law and Procedures—And Three Answers*, 2002 Univ. Ill. L. Rev. 851, 853-57 (discussing the fact that judges are not trained to evaluate the quality of empirical research and pointing to examples of checks on the system that might help, including training judges and relying on the adversarial nature of our adjudicatory process).

34. I searched all state and federal cases in Lexis Advance using the following query: **empirical w/s ("lj." or "l. rev." or "l.rev." or "j.l." or "law review")**, limiting the search to cases published from Jan. 1, 1970, to present (July 30, 2016). Lexis Advance Res., https://advance.lexis.com (last visited July 30, 2016). The Cases database includes both reported and unreported cases in federal and state jurisdictions. The search returned 1084 opinions split roughly evenly between federal court opinions (566) and state court opinions (518). A histogram of frequency by year reported by Lexis suggests that the number of citations has increased over time:



Unfortunately, Lexis does not provide counts on the graph's vertical axis. The search is both underinclusive and overinclusive. Some empirical studies published in law reviews do not include the word "empirical" in the title, and articles with the word "empirical" in the title are not necessarily empirical studies. Note also that mentions are not always citations to support a court's claim. For example, they might refer to studies cited in briefs submitted to the court. The point of the back-of-the-envelope count is to provide some rough evidence for the claim that empirical studies published in law reviews do make their way into briefs and opinions.

35. Epstein & King, *supra* note 5, at 17 n.42.

36. Faulty studies often raise interesting and important questions and can prompt others to employ sound methods to search for answers. *See* Theodore Eisenberg, *The Origins, Nature, and Promise of Empirical Legal Studies and a Response to Concerns*, 2011 U. Ill. L. Rev. 1713, 1721-22, 1730-31.

Potential Causes

How do basic errors survive the law review publication process? Both the authors of the studies and journals' editorial boards play a role. On the author side, lack of training, inadequate training and incentives to publish empirical research drive low quality. First, some authors of low-quality studies have little methodological training, if any at all.[37] Law reviews generally do not require authors of empirical studies to have any sort of empirical-methods training. To my knowledge, law review editors do not systematically assess authors' credentials when selecting articles for publication.

Second, and potentially more harmful, many obtain insufficient training. Law professors now have access to a number of short methods courses. The Center for Empirical Research in the Law at Washington University (St. Louis), for example, ran the fourteenth annual Conducting Empirical Legal Scholarship Workshop in June 2015.[38] The workshop is designed for "law school faculty, political science faculty, and graduate students interested in learning about empirical research and how to evaluate empirical work."[39] The workshop promises to provide "the formal training necessary to design, conduct, and assess empirical studies, and to use statistical software . . . to analyze and manage data."[40] Topics include (but are not limited to) research design, data sampling methods, statistical inference, hypothesis testing, linear regression analysis, discrete choice modeling, graphic display of data, and how to use a popular statistical software package.[41] The program is taught by two of the best methodologists in the field of empirical legal studies.[42] In theory, this sort of course might go a long way to increase the production of empirical legal scholarship. This course, however, and others like it are problematic in that they attempt to teach material generally covered over at least three semesterlong courses[43] for law professors, some of whom likely have no background or knowledge of statistics,[44] in *just three days*. Even if we

37. Epstein & King, *supra* note 5, at 9.

38. My intent here is not to single out this program. It's similar to a number of programs that law schools and other organizations run in an attempt to teach empirical methods to legal scholars.

39. *Conducting Empirical Legal Scholarship Workshop 2015,* Ctr. for Empirical Res. L., http://cerl.wustl.edu/training/cels15.php (last visited July 26, 2016) [hereinafter *Scholarship Workshop*].

40. *Id.*

41. *Id.*

42. *Id.*

43. The topics listed generally are covered in a course on statistics, a course on regression analysis and a course on discrete choice models, all semester-long courses.

44. "Participants need no background or knowledge of statistics to enroll in the workshop." *Scholarship Workshop*, *supra* note 39. George Mason's Law & Economics Center runs a similar training workshop that runs for five days and requires no background in empirics. *LEC Workshop on Empirical Methods for Law Professors*, L. & Econ Ctr., Geo. Mason U. Sch. L., http://masonlec.org/events/event/226-lec-workshop-empirical-methods-law-professors (last visited July 27, 2016).

assume that the average law professor can get up to speed with substantially less instruction than formally trained methodologists get, it certainly requires more than three days. Perhaps the most damaging aspect of these types of courses is that they provide basic training on how to use statistical software packages, allowing for the production of results that are often invalid but seem sophisticated to the untrained, student-editor eye.[45]

45. The University of Michigan's Inter-University Consortium for Political and Social Research (ICPSR) offers a number of summer courses ranging in length from two days to four weeks. *Summer Program in Quantitative Methods of Social Research*, ICPSR, http://www.icpsr.umich.edu/icpsrweb/sumprog/index.jsp (last visited July 27, 2016). Each course assumes different levels of knowledge. For example, the four-day course on multivariate modeling with Stata and R assumes "only familiarity with introductory statistics." *Multivariate Modeling with Stata and R*, ICSPR, http://www.icpsr.umich.edu/icpsrweb/sumprog/courses/0175 (last visited July 27, 2016). The course packs quite a bit into four days:

> We begin with a review/crash course in the linear multiple regression model, and then move on to other important multivariate models including binomial logit and probit models, the multinomial logit model and the mixed logit model. The workshop then considers the interpretation of interaction effects in linear and nonlinear models as well as multilevel models that analyze how socio-economic and political contexts influence individual behavior. Time series and pooled time series methods that investigate how factors such as policy interventions and socio-economic conditions affect dynamic outcomes also are considered. Additional topics, such as the analysis of spatial statistical models will be covered, based on student interests and time availability.

Id. Four days likely are insufficient to cover even one of these topics in the depth required to allow for the production of high-quality empirical work.

Northwestern University offers a five-day workshop on causal inference designed for "[q]uantitative empirical researchers (faculty and graduate students) in social science . . . ." *2016 Main Causal Inference Workshop*, Nw. Sch. L., http://www.law.northwestern.edu/research-faculty/conferences/causalinference/frequentist/ (last visited July 27, 2016). Organizers assume

> knowledge, at the level of an upper-level college econometrics or similar course, of multivariate regression, including OLS, logit, and probit; basic probability and statistics including conditional and compound probabilities, confidence intervals, t-statistics, and standard errors; and some understanding of instrumental variables. Despite its modest prerequisites, this course should be suitable for most researchers with PhD level training and for empirical legal scholars with reasonable but more limited training.

Id. Temporary statistical software licenses are provided; thus, it's possible that those with the elusive "reasonable but more limited" training are able to learn just enough to use the software to produce results but not enough to produce valid results.

Others have forwarded similar claims that use of sophisticated tools by unsophisticated researchers might reduce overall quality of empirical legal scholarship. *See* John F. Pfaff, A Plea for More Aggregation: The Looming Threat to Empirical Legal Scholarship 3-4 (July 16, 2010) (unpublished manuscript), http://ssrn.com/abstract=1641435 ("The rise of the powerful desktop computer and user-friendly statistics software has been a mixed blessing for empirical work. On the one hand, the high-quality work, relying on large datasets and powerful statistical and computational techniques, has never been better. On the other hand, the low-quality work has likely never been worse, as a growing pool of unsophisticated researchers try their hands at empirical analysis. Overall average quality may even be declining.").

Third, the legal academy is encouraging the production of empirical legal studies by all who are willing to produce them.[46] Some law schools have started to fund empirical legal scholarship.[47] Law school libraries have created webpages devoted to empirical legal research.[48] While these efforts certainly increase the production of high-quality empirical work, they also lower the barriers to entry for those without sufficient credentials, who are more likely to produce low-quality work. Requirements for tenure often include positive reviews of scholarship by scholars writing in the same field as the candidate, but law school tenure committees are not required to request reviews from internal and external colleagues with adequate methods training, and committees might have an incentive to bypass such reviewers to increase the likelihood of obtaining positive letters. Of course, once tenure is secured, incentives to produce high-quality empirical work are weakened by the ability to publish low-quality work in law reviews and the low probability that the published work will be challenged given the lack of incentives for empiricists to publicly criticize the methods of flawed studies.[49] This developing infrastructure fosters the production of empirical legal research, but it does not guarantee quality at even the most basic levels.

The problems on the supply side are exacerbated on the demand side. Law students are the sole members of the vast majority of law journal editorial boards. They have complete discretion over which articles to publish and requirements for final versions that appear in the pages of law reviews. Many in the academy have lamented the fact that law students are not qualified to

46. Tracey E. George, *An Empirical Study of Empirical Legal Scholarship: The Top Law Schools*, 81 Ind. L. J. 141, 141-42 (2006) ("Empirical legal scholarship (ELS) is arguably the next big thing in legal intellectual thought. . . . ELS recently and dramatically has expanded in law reviews, at conferences, and among leading law faculties.").

47. For example, see Shana Jackson, *Library Tip: Empirical Legal Research Support,* Yale L. Sch. (Nov. 9, 2015, 11:22 AM), http://library.law.yale.edu/news/library-tip-empirical-legal-research-support (offering support for Yale law students to perform empirical legal studies, which they sometimes use as job market papers on the entry-level law school job market).

48. *See*, *e.g.*, *Empirical Legal Research*, Robert Crown L. Library, Stan. L. Sch., https://law.stanford.edu/robert-crown-law-library/research-resources/empirical-legal-research/ (last visited July 27, 2016); *Empirical Legal Research*, Lillian Goldman L. Library, Yale L. Sch., http://library.law.yale.edu/tags/empirical-legal-research (last visited July 27, 2016).

49. Exceptions exist. For example, the Stanford Law Review (in volume 57, issue 6, May 2005) published several critiques of Richard Sander's empirical study of the impacts of affirmative action policies in law school admissions. See e.g., Ian Ayres & Richard Brooks, *Does Affirmative Action Reduce the Number of Black Lawyers?,* 57 Stan. L. Rev. 1807 (2005); David L. Chambers et al., *The Real Impact of Eliminating Affirmative Action in American Law Schools: An Empirical Critique of Richard Sander's Study*, 57 Stan. L. Rev. 1855 (2005). In addition, Ian Ayres and John Donohue famously challenged the work of John Lott and David Mustard. *See* Ian Ayres & John J. Donohue III, *Shooting Down the "More Guns, Less Crime" Hypothesis*, 55 Stan. L. Rev. 1193 (2003) (critiquing John R. Lott, Jr. & David B. Mustard, *Crime, Deterrence, and Right-to-Carry Concealed Handguns*, 26 J. Legal Stud. 1 (1997)).

select and edit legal studies, including empirical legal studies.[50] Law review editors seem intrigued by empirical studies related to legal topics, despite the fact that they are difficult to assess. For years, law reviews have been dominated by doctrinal and theoretical work. Empirical studies are relatively new to the scene, and their novelty and apparent usefulness make them especially appealing. Very few law students, however, are sufficiently trained in empirical methods to assess the quality of the work researchers submit. In many cases, the articles are published without any technical assessment. In some cases, law review editors request a review by a trained member of their law school's faculty. Occasionally, law review editors will seek out a methodologist outside their home school's faculty with knowledge of the study's substantive field. Peer review, though, is neither systematically sought nor required. Further, reputation effects are minimal at best. By the time published work is publicly criticized, the students responsible for publication have graduated. The rapid turnover of boards (once per year) makes it difficult for boards to implement and sustain policies to improve quality. In short, law journal editorial boards do not have the necessary training to identify high-quality empirical research nor the appropriate incentives to ensure that only high-quality empirical work is published. This, in turn, creates an incentive for scholars to submit low-quality empirical work that they believe might capture the attention of editors. And so it goes.

Many have recognized the quality problem and have offered suggestions to address it, but few solutions have taken hold, and some might have worsened the problem. The next section catalogs some of the suggestions and argues that they have failed to generate incentives both for scholars to produce high-quality work and for editors to publish only high-quality work. The section goes on to suggest ways to capitalize on the interests of scholars and editors to generate incentives for the production and publication of high-quality empirical legal studies.

Generating Incentives for High Quality

During the past couple decades, scholars have suggested ways to increase the quality of empirical scholarship published in law reviews. Epstein and King, for example, offer a set of suggestions, some of which are geared toward law student editors and others toward law school faculties.[51] They suggest law schools offer courses in empirical research taught by trained methodologists and argue that law review editors be required to take a course with a focus on empirical research design.[52] They recommend that law schools encourage law faculty interested in producing empirical research to develop the necessary methodological skills by taking methods courses, by attending short summer

50. *See*, *e.g.*, Epstein & King, *supra* note 5 (reviewing complaints that have been published in the literature); Posner, *supra* note 4, at 1132-35.

51. Epstein & King, *supra* note 5, at 115. *See also* Lee Epstein & Gary King, *Building an Infrastructure for Empirical Research in the Law*, 53 J. Legal Educ. 311 (2003).

52. Epstein & King, *supra* note 5, at 116.

training programs, and by collaborating with trained methodologists.[53] They advocate ways law schools can supply the resources necessary to support empirically minded faculty, such as providing necessary software and hardware, providing generous research assistance, offering funding for data sets and other necessary ingredients, and providing the infrastructure required to obtain external funding. They advise adding faculty members to law review editorial boards, requiring at least one blind peer review of each empirical article by a reviewer who has some expertise in the methods employed and the substantive area of law addressed by the work, and requiring board approval of all empirical articles before publication.[54] They also push for law review editors to establish data archiving policies that allow for easy replication of published results.[55]

Despite the myriad suggestions for ways to improve, little has changed in the way that law reviews select and edit empirical articles. Highly credentialed faculty are teaching empirical methods courses in a handful of law schools,

53. *Id*. at 119-20. Others have suggested that the legal academy has already moved a long way in this direction by hiring researchers trained in both law and empirical methods, by fostering collaboration between law scholars and methodologists, and by funding research workshops that bring together researchers from different disciplines. *See*, *e.g.*, Richard L. Revesz, *A Defense of Empirical Legal Scholarship*, 69 U. Chi. L. Rev. 169 (2002). David Van Zandt recommends the Northwestern Law School approach of hiring dually trained faculty who are capable of producing high-quality empirical legal research, supporting new and existing faculty who wish to pursue graduate degrees, insisting on greater productivity and quality in tenure and promotions, and placing greater weight on peer-reviewed scholarship in scholarship reviews. David E. Van Zandt, *Discipline-Based Faculty*, 53 J. Legal Educ. 332 (2003).

54. Epstein & King, *supra* note 5, at 127-28. *See also* Dauber, *supra* note 30, (echoing a call for peer review and citing to others who have done the same). Others are more pessimistic. Jack Goldsmith & Adrian Vermeule warn that the peer-review process might stifle innovation in scholarship. *See* Jack Goldsmith & Adrian Vermeule, *Empirical Methodology and Legal Scholarship*, 69 U. Chi. L. Rev. 153, 162 (2002) ("Peer review systems also raise the concern that senior scholars favor like-minded scholarship and choke off the channels of intellectual change and development."). Matthew Spitzer has voiced skepticism, predicting that authors facing peer review by both student-edited law reviews and peer-reviewed journals would opt to submit only to peer-reviewed journals to enjoy the benefits of working with professionals rather than students during the editing process. Matthew Spitzer, *Evaluating Valuing Empiricism (at Law Schools)*, 53 J. Legal. Educ. 328 (2003). This assumes, however, that a sufficient number of slots exist in peer-review journals to fill the demand for publication of empirical legal studies. In any event, Spitzer convincingly argues that moving to peer review would be costly in terms of the faculty time required, and that it would be politically difficult to wrest full control from student editors.

55. Epstein & King, *supra* note 5, at 130-33. Epstein and King also encourage empiricists to make their data available and to advertise their availability on their CVs, and call for law schools to reward researchers for doing so. Gregory Mitchell elaborates on the benefits of a push for journals to require a set of disclosures, including information about the methods employed to produce results, data analysis procedures and raw data, suggesting that it might act as a primary way to improve the "scientific dialogue." Mitchell, *supra* note 7. Michele Landis Dauber suggests that law reviews refuse to publish studies unless the author makes the data available for replication purposes. Dauber, *supra* note 30, at 1907-08. Matthew Spitzer agreed with the idea of documentation requirements and wrote that he planned to take the idea to a meeting of law school deans. *See* Spitzer, *supra* note 54.

but no school to my knowledge requires journal editors to take any such course.[56] Some dually trained scholars are publishing high-quality empirical studies in law reviews, but likely a larger number of insufficiently trained law faculty are able to place low-quality work in the reviews. No editorial boards, to my knowledge, include faculty members, and only two journals mention a peer-review process in their submission instructions.[57] Some are reaching out to methodologists to ask for help reviewing submitted work, but most journals have no system in place to check quality. Finally, just a small handful of top journals currently require disclosure of data and/or procedures for replication purposes.[58] If we have any hopes of effectuating change, we need to diagnose

56. At least one law school (Cornell) offers an optional course on empirical legal studies specifically designed for law review editors. E-mail from Dawn Chutkow, Visiting Prof. of Law & Exec. Ed. of *Journal of Empirical Legal Scholarship*, to author (May 19, 2015). The course was developed after student editors requested methods training. *Id*.

57. As of June 2016, websites for the main law reviews of law schools included in the top twenty of either the 2017 *U.S. News & World Report* law school ranking or Washington and Lee University's law journal ranking using combined score and limited to student-edited journals included the following policies on peer review of submissions. *See 2017 Best Law Schools*, U.S. News & World Rep., http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-law-schools (last visited July 30, 2016); *Law Journals: Submissions and Ranking*, Wash. & Lee U. Sch. L. L. Library, http://lawlib.wlu.edu/LJ/ (last visited July 30, 2016). Stanford conducts peer review of each article submission. *Article Submissions*, Stan. L. Rev., https://www.stanfordlawreview.org/submissions/issue-article/ (last visited July 30, 2016) [hereinafter *Stanford Submissions*]. Harvard subjects "many pieces" to faculty peer review. *Submissions,* Harv. L. Rev., http://harvardlawreview.org/submissions/ (last visited July 30, 2016); Columbia "strongly prefers" to subject submissions to peer review, but whether that happens depends on selection time frames and other extenuating circumstances. *Submissions Instructions*, Colum. L. Rev., http://columbialawreview.org/submissions-instructions/ (last visited July 30, 2016). University of Chicago "occasionally" conducts peer review. *Submissions,* U. Chi. L. Rev., https://lawreview.uchicago.edu/page/submissions (last visited July 30, 2016). Yale and NYU have an "extensive review process," without mention of peer review. *Volume 126 Submission Guidelines*, Yale L.J., http://www.yalelawjournal.org/files/V126SubmissionsGuidelines_fnm1a455.pdf (last visited July 30, 2016) [hereinafter *Yale Submissions*]; *Submissions*, N.Y.U. L. Rev., http://www.nyulawreview.org/submissions (last visited July 30, 2016) [hereinafter *NYU Submissions*]. The other top schools (Penn, UC Berkeley, Michigan, UVA, Duke, Northwestern, Cornell, Georgetown, Texas (Austin), Vanderbilt, Washington University (St. Louis), USC, Boston University, Iowa, Minnesota, Fordham, Notre Dame, and UCLA) make no mention of peer review in their online submission guide.

58. As of July 2016, websites for the top twenty-four main law reviews (as defined in footnote 57, *supra*) included the following policies: Yale recommends submission of the data. *Yale Submissions*, *supra* note 57. Stanford, NYU and Northwestern require publication of data on the journal's website, with narrow exceptions (NYU and Northwestern do not require submission of the data until after the article is accepted). *Stanford Submissions*, *supra* note 57; NYU *Submissions, supra* note 57; *Print Submissions*, Nw. L. Rev., http://www.northwesternlawreview.org/submissions (last visited July 30, 2016). UVA requires submission of data within seven days of accepting an offer to publish, encourages but does not require submission of data at the time of submission, and states a strong preference that the data be publicly available or disclosed upon request unless waived in extraordinary circumstances. *Submissions*, Va. L. Rev., http://www.virginialawreview.org/submissions/articles-essays (last visited July 30, 2016). Vanderbilt requires submission of procedures, methodology, or robustness checks

why law students and faculty have not been inspired to change the methods used to edit empirical studies in line with suggestions put forward over the past decade.

One possibility might be that past proposals do not sufficiently leverage the interests of the actors with power to effect change.[59] Law students have little interest in ensuring high-quality publications given that they spend at most one year on the editorial board and their reputations do not depend directly on the quality of the work they print in the issues published during their short time at the helm. In the unlikely event that the quality of a published study is challenged, the students can point to the author of the study as the culpable party. This suggests that the authors of empirical studies might have an incentive to avoid reputation hits, but the pressure is not sufficiently strong. While some studies on hot-button issues like the impacts of gun laws on crime[60] and the effects of affirmative action programs in law school admissions on the success of black lawyers[61] are challenged publicly following publication, researchers seem to have weak incentives, if any, to publicly criticize published studies that contain basic errors.[62] Thus, both authors who submit low-quality empirical legal studies to law reviews and law review editors who select and edit articles lack strong incentives to increase the quality of published work.

How might we better capitalize on the interests of authors and editors? Both have an interest in protecting their reputations. The key is to put into place mechanisms that will generate direct reputational hits for authors who submit low-quality empirical studies and for editors who publish low-quality studies that are impossible to replicate.

First, consider levers that might compel authors to take better care. Authors likely would focus more on quality if they expected reviews of their studies conducted by experts to be posted publicly. Certainly tenure-track law school faculty members would have a strong incentive to avoid reviews that publicly expose poor quality. Tenured faculty members' incentives would not be as strong, but the reputation effects imposed by public reviews might be sufficient

not included in the body of the article to be included as an appendix to the article and requests that the author prepare the data set to be sent upon request. *Submissions*, Vand. L. Rev., https://www.vanderbiltlawreview.org/about/submissions/ (last visited July 30, 2016). The other top schools (Harvard, Columbia, University of Chicago, Penn, UC Berkeley, Michigan, Duke, Cornell, Georgetown, Texas (Austin), Washington (St. Louis), USC, Boston University, Iowa, Minnesota, Fordham, Notre Dame and UCLA) have no posted data disclosure policy.

59. Matthew Spitzer recognized this early on. He lamented that change likely is impossible from the inside due to lack of incentives for reform. Spitzer, *supra* note 54.

60. *See* Ayres & Donohue, *supra* note 49.

61. *See* 57 Stan. L. Rev. 1745, 1745-2170 (2005) (issue 6, published in May 2005).

62. This claim is not based on any systematic empirical analysis. It is an assessment based on my own informal review of tables of contents of law journals, economics journals, and law and economics journals over the past decade.

to compel at least some to take steps to avoid negative reviews, especially if they are widely publicized.

How about incentives for student editors? Students might institute changes if their journals were graded on the quality of published empirical studies. Currently, the Law Journal Rankings Project, hosted by Washington and Lee University's School of Law, ranks law reviews annually based on a combination of impact and citation counts.[63] Some journals tout their ranking on their websites.[64] A Harvard Law School Library guide on deciding where to publish refers to Washington and Lee's ranking.[65] Rankings are discussed in popular online media outlets that cover law schools.[66] Presumably law student editors care about their ranking, and they take steps to publish articles that will contribute maximally to the journal's impact and citation scores. Student editors compete for articles submitted by heavily cited authors. Some evidence suggests students are motivated by journal prestige.[67]

Assuming these interests in fact motivate authors and students, we need an infrastructure that creates incentives based on these interests. Experts are needed to review published studies and journal policies and to compile and maintain a journal grading system.[68] Academic societies are natural collectors of such experts. Further, they are well-positioned to take the lead in a serious and sustained effort to generate incentives for authors and editors. Society boards comprise, at least in part, expert empirical methodologists.[69] These experts gain prestige through serving on the board, and the workload of

63. *Law Journals: Submissions and Ranking Explained*, Wash. & Lee U. Sch. L. L. Library, http://lawlib.wlu.edu/LJ/method.asp (last visited July 27, 2016). Eigenfactor.org also publishes a ranking of law-related journals using a method similar to the one Google uses to rank websites. *About the Eigenfactor Project*, Eigenfactor, http://52.6.43.8/about.php (last visited July 27, 2016); *Journal Ranking for LAW*, Eigenfactor, http://52.6.43.8/projects/journalRank/rankings.php?search=OM&year=2014&searchby=isicat&orderby=Eigenfactor (last visited July 27, 2016).

64. The Google search "law journal rankings project" links to instances of such advertisements.

65. June Casey, *Publishing in Law Reviews and Journals*, Harv. L. Sch. Library, http://guides.library.harvard.edu/c.php?g=309907&p=2070141 (last updated Dec. 8, 2015).

66. *See*, *e.g.*, David Lat, *America's Next Top Law Review: New Rankings!*, Above the L. (Feb. 24, 2014, 4:30 PM), http://abovethelaw.com/2014/02/americas-next-top-law-review-new-rankings/; Alicia Albertson, *Best Law Reviews: Stanford Tops List*, Nat'l Jurist (Mar. 7, 2014), http://www.nationaljurist.com/content/best-law-reviews-stanford-tops-list.

67. Jason P. Nance & Dylan J. Steinberg, *The Law Review Article Selection Process: Results from a National Study*, 71 Alb. L. Rev. 565, 585 (2008).

68. A grading system seems preferable to a system that ranks journals. Under a grading system, each journal would earn a grade (e.g., excellent, good, fair, poor) that corresponds to a transparent set of requirements.

69. The boards of the American Law and Economics Association (ALEA), http://www.amlecon.org/alea-officers.html (last visited July 27, 2016) and the Society for Empirical Legal Studies (SELS), http://www.lawschool.cornell.edu/SELS/about.cfm (last visited July 27, 2016), for example, always include several members with formal training in empirical methods who have published empirical studies in peer-reviewed journals.

nonofficers is light.[70] Serving as a board member likely would remain attractive if, during each year of service, members were required to review some number of empirical studies published in law reviews.[71] To maintain continuity, the boards could create standing committees, the chairs of which could work together to coordinate and edit online publications of published article reviews and to determine and widely publicize annual journal grades.[72]

Post-publication reviews would serve at least two purposes. First, they would act as a quality check on published work. Academics, law students, lawyers, judges, and other policymakers would have access to the entire set of reviews and the study author's response to the reviews, if any.[73] The reviews would assist readers who are unable to assess the quality of studies on their own. Second, the reviews would act as a teaching tool for both student editors and those aiming to sharpen their understanding of empirical methods, including authors of future studies. The reviews could be used to generate a catalog of the

70. I have served as a nonofficer on both the ALEA and SELS boards, and this has been my experience.

71. We might worry about the willingness of law professors to engage in this time-consuming process. The *South Carolina Law Review* experimented with peer review and reported that most law professors who were asked to review submissions were willing to complete timely reviews. *See* John P. Zimmer & Jason P. Luther, *Peer Review as an Aid to Article Selection in Student-Edited Legal Journals*, 60 S.C. L. Rev. 959, 969-972 (2009). In addition, most if not all methods experts are involved in the review systems of peer-reviewed journals. Peer review will not be new to them, although public posting of reviews will be.

    Even if a substantial number of volunteers step forward, we might worry about capacity given the number of empirical studies published in student-edited journals. To roughly estimate the number of empirical studies published by student-edited law journals, I searched Lexis Advance's Law Reviews and Journals database using the terms **"Table 1" and "statistical significance" and publication ("law journal" or "law review")** for publication between Jan. 1, 2006 and Dec. 1, 2015. Lexis Advance Res., https://advance.lexis.com (last visited July 30, 2016). Washington and Lee's website allows the compilation of lists of student-edited journals and peer-edited or refereed journals. *Law Journals: Submission and Ranking, 2008-2015*, Wash. & Lee. Sch. L. L. Library, http://lawlib.wlu.edu/LJ/ (last visited July 27, 2016). A quick review of the lists suggests that student-edited journals are much more likely to contain "law review" or "law journal" in their title relative to peer-edited or referred journals. Lexis counts 678 articles published during the period 2006-2015, an average of sixty-eight articles per year (the distribution is roughly uniform across the years). Lexis Advance Res., *supra*. The same search over publication years 2011-2015 produces 379 hits, an average of seventy-six articles per year. *Id.* To the extent reviewing only a portion of published articles is feasible, society committees could limit reviews to studies published in main journals, studies published in journals with the highest impact according to the Washington and Lee ranking, or a random selection of articles published in all student-edited journals.

72. A grading system likely will be most effective the more widely publicized it is. Efforts could be made to get annual grade reports into the hands of the law journal editorial boards, law school administrators, law school rank and tenure committees, judges, and other policymakers.

73. Authors would have an opportunity to issue a written response. The academic society committees would edit and publish the responses along with the reviews. While implementing a double-blind system would be impossible because the reviewer would know the author of the published study, reviewers could remain anonymous.

most common errors that study authors might adopt as a checklist to reduce the likelihood of such errors in their own work. Eventually such a checklist might be used to generate a handbook that the committees could distribute to incoming editorial boards through the school's expert faculty liaison.[74]

Once the post-publication review process gets underway, assigning grades to student-edited law journals would be fairly straightforward. Grades would be based on a combination of the quality of published empirical studies,[75] the quality of any pre-publication peer-review process the journal employs, and the journal's data and procedure disclosure policy and its implementation.[76] Grade assignment would be transparent, and the society committees would update them annually.[77]

If the grading system is perceived as legitimate, it could work to motivate student editorial boards to take better care in selecting and editing empirical studies.[78] It might also justify intervention by faculty or law school administrators despite the strong norm of student editor independence. Empirical scholars might be compelled to submit articles only to the top-ranked journals to bolster their reputations and to signal the quality of their work to those unable to independently assess quality.[79] In addition, if authors believe that readers evaluate articles at least in part based on journal ranking, authors might lean toward submitting only to highly ranked journals to increase the potential impact of the study.

74. Such a handbook might also include guidance on construction of tables and figures and standards for database citation, publication, and access that track the guidelines promoted by Data Access & Research Transparency. *See TOP (Transparency and Openness Promotion) Guidelines,* Data Access & Res. Transparency, http://www.dartstatement.org/#!2015-cos-top-guidelines/kye04 (last visited July 30, 2016).

75. Reviewers could be asked to evaluate each study along some number of dimensions using Likert scales, and the scores could be used (carefully) to assess the overall quality of studies published by each journal. The reviewers could be directed to focus on objective measures of quality related to the sorts of basic issues noted *supra,* note 30 and accompanying text. If the idea of posting public critiques is disfavored, evaluations of published studies could be used solely in the determination of journal grades. The downsides are the lost opportunity to educate students and authors on the proper use of methods and the lost strong incentive for authors to ensure quality to avoid negative public evaluations.

76. The societies might offer journals advice on data and procedure disclosure policies and how to implement them, following the lead of organizations such as DA-RT.

77. The society committees might explore the possibility of incorporating the grades into Washington and Lee's ranking system.

78. While transferring the functions of selection and editing from student editors to methods experts would be optimal, I'm pessimistic about law school administrators agreeing to go along with this sort of radical change.

79. Michele Landis Dauber offered a similar prediction related to a movement toward peer review. Dauber, *supra* note 30, at 1914 n.69 ("Status differentiation among law journals in favor of those which are peer reviewed would provide in law, as in other disciplines, a way to differentiate the quality of scholarship published in the field by the locus of publication, thus conferring reputational benefits on scholars who chose to submit their work to the rigors of peer review.").

To the extent that public post-publication reviews compel student editors to seek pre-publication peer reviews, the academic society committees could help streamline the process. Given that authors simultaneously submit to potentially hundreds of law reviews, duplication of effort could be vastly reduced if student editors filtered requests for reviews through the society committees. Once a journal requests peer review of a submitted study, the committees would log it, and all student editors would have access to the peer reviews produced for the first requester.[80] If upon receipt of the peer reviews the editorial board decides to issue an offer for publication conditional on implementation of changes recommended by the reviews, the editors would have the option of submitting a revised version to the reviewers for approval. Editors might indicate in the publication that the article was peer-reviewed and approved.

It's possible that students will resist the peer-review process given that it will substantially slow down the acceptance process. On a positive note, a few of the top main law journals now ask authors for exclusive submission for a period of time to allow a review.[81] Even if journals decide not to slow down the process, all hope might not be lost if authors who wish to avoid critical post-publication reviews are motivated to submit only high-quality studies. Some might decide to limit themselves to producing only nonempirical studies. Others might decide to team up with a methods expert and jointly conduct empirical studies.[82] The society committees might also implement a pre-submission peer-review process. Authors could submit completed drafts, and the society committees could facilitate a review and approval process similar to the process for journal editors.[83] If reviewers decide the study is

80. These reviews would not be posted publicly, and the committees might request that student editors keep them private and encourage the author to include mention of peer review in the study's initial footnote. Societies might ask their relevant board members to agree to peer-review some number of submissions each year and to submit them on a timely basis.

81. As of June 2016, websites for three of the top main law reviews (as described in footnote 57, *supra*) request that authors submit exclusively to that journal for at least a period of ten days. Harvard and Stanford "strongly recommend" authors submit exclusively to the respective journal for at least ten days, while Yale recommends submitting exclusively for at least two weeks. *See Harvard Submissions, supra* note 57; *Stanford Submissions*, *supra* note 57. None of the other top law reviews (Columbia, University of Chicago, NYU, Penn, UC Berkeley, Michigan, UVA, Duke, Northwestern, Cornell, Georgetown, Texas (Austin), Vanderbilt, Washington (St. Louis), USC, Boston University, Iowa, Minnesota, Fordham, Notre Dame, UCLA) discusses exclusivity in its online submission guides.

82. The society committees might consider implementing some sort of system that would help nonexperts find experts with compatible subject interests and an interest in considering co-authorships.

83. The number of reviews the system can produce is, of course, limited by the number of volunteers willing to conduct reviews. If an insufficient number of volunteers come forward, only some studies will get through the review process, and top student-edited law journals will compete for those studies. This might make it difficult for other journals to improve their grades. These journals, however, will have options beyond publishing empirical studies of questionable quality. They can choose not to publish empirical studies. Alternatively, they might seek out alternative ways to obtain expert reviews such as requesting them

methodologically sound, the committees would post the approved draft on a webpage administered by the societies. The author could then point student editors to the webpage to verify that the societies approved the submitted draft, and the author could note this in the study's initial footnote.[84]

The society committees also might offer ongoing assistance to student editors by recruiting at least one faculty member at each law school with expertise in empirical methods (if any) to serve as a liaison between the school's editorial boards and the society committees. Those expert faculty members might meet with (or, at a minimum, communicate with) incoming board members to caution them on the difficulties inherent in assessing the quality of empirical submissions and provide editors with a list of potential peer reviewers broken down by substantive area and inform them of the services provided by the society committees.

Conclusion

In singing the praises of empirical legal scholarship Professors McAdams and Ulen write, "[E]mpirical and experimental methods have already contributed a great deal to legal scholarship, [and] . . . there is a good deal more that those methods could contribute to the scholarly understanding of the law. . . ."[85] I wholeheartedly agree with this sentiment. They go on, however, to note that "the techniques necessary to become adept at these methods are not so daunting that legal scholars should be hesitant to make them a routine part of their toolkits."[86] This characterization of what's required to conduct sound empirical work might have led us down a primrose path. While it is certainly not impossible for inexpert scholars to get up to speed, a short summer course might do more harm than good. Unfortunately, these courses provide training that allows for the production of results using empirical models that are difficult to master in just a handful of days.

The current crises plaguing psychology, economics and other scientific fields certainly illustrate that peer review is insufficient to guarantee high quality. The field of empirical legal studies, however, is facing much more basic quality issues. We should consider joining other scientific fields in efforts to increase the quality of published work, for the sake of the scientific endeavor and the

from capable members of their faculties or from capable members of their schools' other departments.

84. Another option is to use a badge system similar to one the Association for Psychological Science adopted in 2014. See *Open Practices*, Ass'n for Psychol. Sci., http://www.psychologicalscience.org/index.php/publications/journals/psychological_science/badges (last visited July 27, 2016). Journals could award badges to studies approved after peer review. The society committees might consider working with the editors of The Bluebook to recommend or require a more general badge system to recognize compliance with best practices, such as peer review, data and procedure disclosure, and publication of data to an existing repository such as The Dataverse Project.

85. Richard H. McAdams & Thomas S. Ulen, *Introduction*, 2002 U. Ill. L. Rev. 791, 792.

86. *Id*.

promise of better policymaking. Although many past suggestions have either fallen flat or potentially exacerbated the problem, reforms that appeal to the interests of editors and authors (and, one hopes, the experts) might motivate reform.

The suggestions included here are merely that—ideas about how we might capitalize on the interests of involved actors to generate incentives for change. Myriad methods for tapping into relevant actors' interests exist. The main point of this essay is that passive suggestions haven't worked and likely won't work to effect change. We need something more. Implementation of some or all of the suggestions on a grand scale might be impossible, at least at the start. The hope is that the basic framework acts as a starting point for getting some institutional changes underway. What is certain, however, is that change will require sustained effort by experts. This essay is a call to empirical legal scholars and to academic societies to work together to heighten the standards of published empirical scholarship and to join the ranks in other fields who are working toward the same goal.