# DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD

# Building An Infrastructure for Empirical Research in the Law

*(Article begins on next page)*

# Building an Infrastructure for Empirical Research in the Law

## Lee Epstein and Gary King

In every discipline in which "empirical research" has become common-place, scholars have formed a subfield devoted to solving the methodological problems unique to that discipline's data and theoretical questions. Although students of economics, political science, psychology, sociology, business, education, medicine, public health, and so on primarily focus on specific substantive questions, they cannot wait for those in other fields to solve their methodological problems or to teach them "new" methods, wherever they were initially developed. In "The Rules of Inference,"[1] we argued for the creation of an analogous methodological subfield devoted to legal scholarship. We also had two other objectives: (1) to adapt the rules of inference used in the natural and social sciences, which apply equally to quantitative and qualitative research,[2] to the special needs, theories, and data in legal scholarship, and (2) to offer recommendations on how the infrastructure of teaching and research at law schools might be reorganized so that it could better support the creation of first-rate quantitative and qualitative empirical research without compromising other important objectives. Published commentaries on our paper,[3] along with citations to it, have focused largely on the first—our application of the rules of inference to legal scholarship.[4] Until now, discussions of our second goal—suggestions for the improvement of legal scholarship, as well as our argument for the creation of a group that would focus on

**Lee Epstein** <epstein@artsci.wustl.edu> <http://www.artsci.wustl.edu/~polisci/epstein> is the Edward Mallinckrodt distinguished university professor of political science and professor of law at Washington University. **Gary King** <King@harvard.edu> <http://GKing.Harvard.edu> is the David Florence professor of government at Harvard University and director of the Harvard-MIT Data Center.

1. The Rules of Inference, 69 U. Chi. L. Rev. 1 (2002). Portions of the present article are reprinted from "The Rules of Inference," at 114–33, a complete copy of which can be found at <http://gking.harvard.edu/files/abs/rules-abs.shtml>.

2. See Gary King et al., Designing Social Inquiry: Scientific Inference in Qualitative Research 46 (Princeton, 1994).

3. Frank B. Cross et al., Empirical Research and the Goals of Legal Scholarship: Above the Rules: A Response to Epstein and King, 69 U. Chi. L. Rev. 135 (2002); Jack Goldsmith & Adrian Vermeule, Empirical Research and the Goals of Legal Scholarship: Empirical Methodology and Legal Scholarship, 69 U. Chi. L. Rev. 153 (2002); Richard L. Revesz, Empirical Research and the Goals of Legal Scholarship: A Defense of Empirical Legal Scholarship, 69 U. Chi. L. Rev. 169 (2002).

4. Peter J. Hammer & William M. Sage, Antitrust, Health Care Quality, and the Courts, 102 Colum. L. Rev. 545, 561 (2002).

methodological problems unique to law—have been relegated to less public forums, even though, judging from the volume of correspondence we have received, they seem to be no less extensive.

We are thus especially grateful to the distinguished participants in this symposium for helping to make public what has until now been more of a private discussion. Law schools already attempt to achieve a diverse set of goals, and so figuring out how to add empirical research to this long list will not be obvious or easy. Since many legal academics, faculty and administrators alike, are now looking for ways to accomplish this task, we hope that this symposium will contribute to their ongoing discussion. If the quality of empirical scholarship produced in and disseminated by the nation's law schools is to improve, adherence to the rules of inference alone is surely insufficient. Science requires more than being scientific; it requires community and community discussion. Only by adapting the infrastructure so that it more fully supports, encourages, and enhances the ability of scholars to carry out empirical research and the ability of lawyers, judges, and students to consume it will deep and lasting change come.

The recommendations to which participants in this symposium respond are aimed at facilitating the development of such an infrastructure. In what follows we briefly highlight and excerpt (with a bare minimum of footnotes) the four sets of proposals most pertinent to this symposium: those geared toward (I) law school students and (II) law school faculty, and those that may be of interest to the entire legal community, (III) law reviews and (IV) data archiving and documentation.

One final note before turning to the recommendations. In order to convey them as clearly as possible, we lay them out with a certain degree of specificity. But as we highlighted in our article, they are certainly not the only way to proceed. Indeed, these recommendations are based only on our *hypothesis* that implementing them would improve empirical analyses in the law. We obviously are not certain that any of our ideas will work as intended at any particular law school, and we have conducted no analyses to evaluate them. Such studies surely should be done. At the same time, our experience in other disciplines suggests support for the general direction of these proposals. Whether we have appropriately extracted the right general principles and adapted them appropriately to the culture of legal scholarship and law schools in general may be less important than how such proposals might be adapted further to the unique local conditions at individual schools. The same qualifications of course also apply to our participants; we did not ask for, and do not expect, controlled experiments at this stage, only that they bring their extraordinarily deep experience to bear on this question. Whatever the ultimate fate of empirical research in the community of legal scholars, we hope the vibrant discussion they spark will continue—both in the pages that follow here and in discussions to follow.

## I. Offer courses in empirical research for law students.

Our first recommendation is that law schools incorporate into their curriculum at least one course on empirical research—a course that covers quantitative and qualitative approaches to research design and evaluation. It certainly should be required for students serving on the schools' law reviews (a subject to which we return in part III below) and probably for most or all others as well. Whether law schools should make this a part of the first-year curriculum or reserve it for second- and third-year students depends on a school's particular needs and goals. One relevant comparison is to graduate programs in the social sciences, which typically encourage students to take tool-oriented courses (for example, research design, methods, foreign languages, game theory) as early as possible in their academic careers.

We offer this recommendation not because all students will necessarily be conducting empirical research of their own. Students at most law schools will never prepare law review articles, either in law school or afterwards. Nevertheless, they will need the skills to evaluate such research, whether for clients, senior members of their law firms, or judges; whether in criminal or civil suits. This is true today, and it may become even more so as judges increasingly make demands on lawyers to meet particular legal standards, to question experts, or to back up specific claims with credible empirical support.

In addition to meeting the needs of students and the legal community, training students in the standards and norms of empirical research has at least two happy byproducts. First, again given the increasing demand for data, students with these skills will be more marketable than those without them.[5] Second, faculty will benefit enormously. Offering empirical courses will require law schools to hire a scholar trained in empirical methodology who, in turn, could serve as a resource for faculty—one that they may have been unable to hire but for curriculum needs.

We have more to say about the infrastructure requirements of faculty below, but let us first address two obvious concerns: who might these methodologists be, and from what academic fields should they come? Starting with the first, surely the methodologist should be a dedicated scholar and teacher, well versed in the rules of inference and the norms and standards for conducting empirical research. But more than that is necessary. The selected methodologist also should be able to teach students and faculty how to analyze *their* data and thus should possess technical skill sets.

This methodologist could, on the one hand, hail from any number of academic disciplines. Because empirical research in law has methodological problems that overlap with those in biology, chemistry, economics, medicine and public health, political science, psychology, and sociology, methods from those other disciplines can be adapted to the study of the law. On the other hand, in virtually every discipline that has begun to develop a serious empirical research program, scholars discover methodological problems that are unique to the special concerns in that area. Each new data source, as it turns out, often requires at least some adaptation of existing methods, and some-

5.   For more on this point, see Epstein & King, *supra* note 1, at 124–25.

times the development of new methods altogether. There is bioinformatics within biology, biostatistics and epidemiology within medicine and public health, econometrics within economics, chemometrics within chemistry, political methodology within political science, psychometrics within psychology, sociological methodology within sociology, and so on.

Thus, to encourage serious, enduring, and continually improving empirical research, the legal community should foster the development of a subfield of methodology within law. To accomplish this, law schools should hire scholars who have deep training in empirical methods in their chosen discipline. But schools also should select a methodologist who has, or at least is interested in developing, an understanding of the kinds of problems that interest, and the sorts of data available to, legal scholars. Certain academic disciplines regularly turn out Ph.D.s who fit this description (for example, economics and political science). And law faculty can help themselves out by inculcating in these methodologists an even greater appreciation of their concerns. This could come about through coteaching courses, which would work to the benefit of students and faculty. It also might evolve via collaborative research—a subject we discuss in more detail in the next section. Either way, new scholarly links would be created. Some new empirical legal methods would then be developed by, say, the political scientist or economist—which is fine, given the increasingly interdisciplinary nature of law—but others would be developed by law teachers. And eventually law schools would not need to contract out methodological concerns. The field—the empirical methodology of legal scholarship—would flourish on its own.

## II. Enhance opportunities for faculty to conduct high-quality empirical research—and then disseminate it quickly.

When it comes to their research, legal scholars seem to have developed a norm of timeliness—a norm that is useful, because scientifically valid input into current debates about public policy can make highly important and dramatically influential contributions. Legal scholars have proven time and time again that they are uniquely situated to take on this task at least with regard to speed; where they have not completely succeeded is with the high-quality aspect of the task. This failure is unfortunate *because they can do both.* That is, they can conduct first-rate research that they can create and disseminate rapidly. Even if time, information, and resources are limited, there are ways to produce credible results by merely doing the best you can and appropriately reporting the uncertainty in your estimates.

To help law schools do this, we offer two sets of recommendations. The first centers on fostering the development of the skill sets necessary for faculty to do high-quality research so that they can respond better in the time available; the second is aimed at building an infrastructure to allow law teachers to produce credible research results as quickly as possible.

### A. Help build methodological skills.

Legal academics will require additional training to implement the rules of inference and to master skills associated with the analysis of data—whether of the qualitative or the quantitative variety. How can they develop them?

Individual faculty can proceed in any number of ways, with three rather obvious. First, they can take an empirical research course. This is a regular occurrence in many cognate disciplines, wherein scholars—whether tenured or untenured, whether beginners or senior scholars, whether to brush up on their skills or learn entirely new ones—regularly take technical courses. Second, they can get training at various institutes, including the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan and Washington University's Workshop on Empirical Legal Scholarship, which are geared specifically to law teachers. Third, a law teacher can learn on the job by entering into collaborations with a methodologist in the law school or, for example, a social scientist colleague with an interest in law. This is, perhaps, the easiest and most efficient way—and one used quite often in other academic disciplines—for legal academics to develop an appreciation of empirical methods and to learn the skills necessary to carry out such inquiries on their own.

Law schools can also facilitate each of these activities. For faculty who would like to take an empirical research course and ultimately demonstrate a mastery of the skills (perhaps in the form of a research presentation to the faculty or a published article), their schools could provide some release time from teaching. For those who would like to attend the program at Michigan or the one at Washington University, the law schools could pay their tuition, as many graduate programs in other fields currently do. And for legal academics interested in entering into collaborations with empirically skilled colleagues, their schools should provide incentives to turn interest into action—perhaps in the form of seed grants for the proposed project or other forms of support.

However they proceed, *law schools must acknowledge that the point of an academic research career is to make the maximum contribution to a scholarly literature and the world.* Whether that contribution is single- or coauthored should not matter so long as the contribution is there. Even more to the point, if including coauthorship in one's repertoire can help a researcher generate a larger total contribution—as is the case for scholars in many other fields— then it should be strongly encouraged. At the very least, law schools should not punish faculty (for example, by denying tenure) for coauthoring articles.

*B. Save time by improving resources.*

Scholars can conduct serious empirical research no matter how limited the time or resources. But if both time and resources are tightly constrained, they will pay a price in the form of less certain findings. (In other words, the less time and resources, the smaller the number of observations that can be collected, or the less reliable the measurement procedures that can be used, and hence the greater the inefficiency of the resulting inferences.) Since at least some law teachers want to produce research results that they can disseminate as quickly as possible while also ensuring that they are as informative as possible, taking more time is not in the cards. Increasing resources, however, could make a nontrivial difference. We see at least four ways law schools can help.

First, they can ensure that faculty conducting empirical research have computers and software up to the task, along with the technical support they need to use those resources. As a rough calculation, computers should be replaced every three years, and software upgraded approximately every year. Staff support could take many different forms but normally includes network administrators, systems operators, user-support personnel, and clerical assistance. Our suggestion, in addition to maintaining excellence in this area, is to supplement the existing information technology group with experts who can perform specific research tasks, such as a specialist in statistical software programs and another in graphic design.

A second way law schools can help their faculty make the best use of their time is to provide additional person power in the form of research assistants, who will enable scholars to collect data as quickly and efficiently as possible. Academic departments accomplish this in various ways—including fellowships, stipends, and course credit for students providing research assistance—all of which would be feasible and to some extent already exist in many law schools. Perhaps some RAs can be specifically unassigned at the start of the semester so they can be used for especially timely projects as they arise.

Third, as we already have mentioned, law schools should encourage their faculty to enter into collaborations with scholars who know how to conduct serious empirical research. In addition to the reasons we offered earlier, collaborative empirical work is faster to conduct. Legal academics need not waste precious time learning the details of every possible new skill and instead can rely on coauthors, who presumably would benefit from the substantive expertise that law faculty bring to the table.

Finally, to conduct empirical research, scholars often require funding: they may need to acquire a particular data set, field a survey, hire interviewers, and so on. We recommend that law schools and their associated centers follow the lead of many other academic units and supply seed money to credible projects. Such funding would enable scholars to conduct pilot studies that they could, in turn, use to inform public policy debates or to demonstrate the worthiness of their research to various outside funding agencies, foundations, and donors. External funding certainly has benefits for individual research projects, but it also has positive implications for law schools. Surely deans would not turn down reimbursements for indirect costs that would flow into their coffers if more of their faculty obtained support from the National Science Foundation's Law and Social Science Program. As a further incentive, law schools might follow the path of research units that pass back some fraction of indirect cost reimbursement to the faculty generating the funds in the first place.

## III. Move to an alternative model of scholarly journal management.

In the law world, students run and edit their school's flagship journal, though they often consult informally with faculty before making decisions about particular articles. We have read various accounts of how this norm came about, and we appreciate the tradition. At the same time, certain aspects of it are problematic, such as its failure to conform to a critical aspect of

empirical research, that it not be ad hominem, that the focus be on the work and not the person. Without some form of external reviewing, separating the person from the product is difficult. Also problematic is that students (and indeed any one person) may lack the expertise necessary to evaluate the submissions that cover complex and technical areas of the law or employ sophisticated statistical or qualitative methods. Finally, the lack of blind peer review in most law journals may put legal academics at a distinct disadvantage vis-à-vis the rest of the university.

Legal academics are, of course, too well aware of these problems. But, as far as we can tell, switching wholesale to the full blind-peer-review model used in academic journals throughout the natural and social sciences is infeasible. The large number of journals would create an enormous increase in the workload for law faculty serving as anonymous reviewers. Moreover, the work that goes into reviewing is ordinarily accompanied by a prohibition against submission to multiple journals (so that the editors' and their reviewers' efforts are not wasted); accordingly, a switch to peer review could also slow publication—an especially undesirable outcome given the norm of speed. Other difficulties have also been identified, but suffice it to say that the full-blown version of the traditional blind-peer-review model does not seem to fit with the norms, needs, and goals of the legal community.

We thus propose an alternative model—one that enables law schools to continue the existing norm, while enhancing it by taking advantage of some features of the peer-review system. Other possibilities exist, of course, but this one, which is similar to that which many university (book) presses follow, may best fit with the traditions in law. Our model would work as follows.

- Students would continue to serve, as they do now, as law review editors and members. But law schools would expand editorial boards to include faculty.

- As they receive manuscripts, students—like university press editors—can reject manuscripts for whatever reasons they think valid, just as they do now. But for any manuscript that they deem potentially publishable, they must get at least one outside peer review. The reviewer should be an expert in at least some aspect of the subject or methods in question. In most situations, this means a law professor (ideally, but not always, from another law school), although occasionally it may mean a student who has written a thesis on a related subject or, possibly, a Ph.D. in another area. The key is that the editors' choice of a reviewer should be based on the reviewer's expertise, not status.

- After receiving the external evaluation, students would be free to reject the manuscript. But if they want to publish the essay, they must bring the anonymous peer review and any internal student evaluations to the editorial board for final approval. If desired, the student law review editor may assign someone to write a response to the anonymous review, solicit a response (or even a revised version of the article) from the author, and include this material in the information that the editorial board reviews.

Whatever the exact procedure, the point is that the law review would publish only articles that (1) have been reviewed by at least one external expert and (2) have been approved by the editorial board. The editorial board would serve as a check on the student law review editor, but in the vast majority of cases expectations would be clear enough that the editorial board would support the law review editor's decision. Indeed, what happens at most university presses, and what would in all likelihood happen if this model were appropriately adapted to law, is that the editorial board operates to empower the editor, a position that would be as autonomous as it is now. The new system would make it easier for the editor to say no to senior faculty who may hold some influence over their future careers ("I'm sorry, the editorial board did not approve your article . . . "), and it would add substantial credibility to the decision-making process, to the prestige of the law review, and to the scholarly value of its content. Student editors are already aware of some of these advantages, as their practice of consulting faculty informally attests.

In offering this model, we recognize that following it—or some variation of it—may add to the burden of students and faculty. Students, with faculty guidance, must begin to develop a pool of external referees. Deans will need to persuade faculty to sit on editorial boards; and faculty will, on occasion, be asked to serve as manuscript evaluators. Moreover, scholars who are accustomed to relatively rapid turnaround time will (perhaps) have to wait slightly longer for decisions on manuscripts sent out for review.

To us, none of these costs seems terribly onerous or problematic.[6] Both students and faculty accrue an advantage of which scholars in other disciplines are only too well aware: reading and judging manuscripts, while a chore, is a great way to learn about the state of the literature, and to do so even before publication. This is one of the reasons why scholars in other disciplines are willing to take on the burdens associated with reviewing and serving on editorial boards. Moreover, our alternative model provides a mechanism— the editorial board—to facilitate faculty-student interaction, to break down the hierarchy that seems more severe and entrenched in law schools than in many other academic programs. We think the only relevant hierarchy in an academic discipline is based on knowledge, and sometimes students have this knowledge and faculty do not. Indeed, while the opinions of outside experts can help determine whether the article in question is "right," sometimes the person who knows more than most anyone else about a subject is a student who has researched it; sometimes the person with the best idea about a research topic is someone who has not been "biased" by years of operating within the standard paradigm, and this too sometimes may be a student. Having faculty and students make joint decisions has enormous benefits for all involved. It invites students to become a part of the academic community, to be socialized into a world where learning never stops, where expertise is shared, where the norms of the free exchange of academic information are inculcated, and where new ideas are developed.

6.   We address these in *id.* at 128–29.

Building an Infrastructure for Empirical Research in the Law

## IV. Develop standards for data archiving.

One of the strongest norms in legal publishing is the norm of textual documentation: law review editors and authors are, to a greater extent than most others in academia, obsessed with footnotes. We realize that this norm has come under attack from many quarters and, from some perspectives, is a waste of effort. But, from the perspective of empirical research, it has two important advantages. First, it connects the extant scholarship to existing literatures and, second, elaborate footnotes enable readers to locate any text cited in an article and learn about the content of that text. And if the text is unpublished, scholars are able to get it from the author or from the law reviews themselves, who ask authors to provide unpublished materials for storage in their archives. We certainly cannot say the same of the norms in almost any other academic discipline.

Given the importance (and value) of this norm of documentation to the legal community, it is surprising that violations are rampant when it comes to nontextual sources of information—most relevant here, quantitative or qualitative data analyzed in empirical research. So, for example, while the law reviews regularly get unpublished material from authors, they do not typically store qualitative or quantitative data or documentation necessary to replicate the studies that they publish. Along the same lines (and with few notable exceptions), we found it very difficult, if not impossible, to obtain data used in law review articles from public sources, from the law reviews, or from the authors of the articles directly.

The upshot of these practices is that the very basis of the most important documentary evidence in empirical law review articles is forever lost. This monumental waste of resources should not continue. How can the scholarly community evaluate such work? How can future scholars build on it? For that matter, how can even the original author conduct followup research? How can the scholarly community correct mistakes, improve its methods, or benefit from the most important advantages of having a scholarly community in the first place?

We recommend that law reviews, at a minimum, require documentation of empirical data with as much specificity as they do for textual documentation. And, just as for textual documentation, this should be a prerequisite for publication. This means simply making it possible for any reader to traverse the chain of empirical evidence amassed to support the conclusions published. Citing public-use data sets is one way to comply with this rule, but in virtually all situations the only way to ensure full compliance is to require researchers to deposit their original data, and all information necessary to replicate their results, in some *public archive*. This may sound like an unusual idea, but scholars in every field who have tried to replicate another's empirical work know how hard that task is to accomplish without the original data. Even those using public data sets would normally need to deposit at least the full details of their calculations—how they moved from the publicly available data to their numerical results—and exact information about which version of the public data set they analyzed. By the same token, researchers conducting

surveys should deposit the individual-level responses to their questions (removing only information necessary to protect the identity of the respondents) and any calculations performed (for example, how missing data were handled). Investigators coding cases would deposit their data sets, complete coding rules, and the precise connections between their numerical data and the original cases from which they were coded. Scholars studying speeches of legislators would deposit the texts of the speeches (if they were not easily retrieved from other sources), or detailed citations to all speeches consulted.

Many public archives exist, but a healthy procedure would be for law reviews, individually or collectively, to establish their own data archives, so that they could keep empirical evidence and satisfy the norms of the legal profession. The Virtual Data Center <http://TheData.org/> project provides easy public domain tools that journals and others can use to set up their own archives, as well as the exact standards for citing empirical data and verifying that they will always exist.

This recommendation centers on the law reviews. Another pertains to legal scholars themselves: those who comply with this rule ought to receive credit for it. Legal academics should list the data sets they have made publicly available on their vitae, just as they now list published articles. Hiring, tenure, and promotion committees, and other sanctioning bodies, need to recognize the contribution that publicly available data make to the scholarly community.

We realize that following this recommendation and the others we have offered will confront law schools and their faculties with a host of challenges. Meeting them should not be too difficult. After all, the interest in empirical research and the norms supporting documentation are in place. It is now a matter of making productive use of them.