Empirical Health Law Scholarship: The State of the Field

Michelle M. Mello & Kathryn Zeiler†

Abstract

The last three decades have seen the blossoming of the fields of health law and empirical legal studies and their intersection, empirical scholarship in health law and policy. Researchers in legal academia and other settings have conducted hundreds of studies using data to estimate the effects of health law on accident rates, health outcomes, health care utilization and costs, and other outcome variables. Yet the emerging field of empirical health law faces significant challenges—practical, methodological, and political. The purpose of this Article is to survey the current state of the field by describing commonly-used methods, analyzing enabling and inhibiting factors in the production and uptake of this type of research by policymakers, and suggesting ways to increase the production and impact of empirical health law studies. In some areas of inquiry, high-quality research has been conducted and the findings successfully imported into policy debates and used to inform evidence-based lawmaking. In other areas, the level of rigor has been uneven and the best evidence has not translated into policy making. Despite challenges and historical shortcomings, empirical health law studies can and should have a substantial impact on regulations designed to improve public safety, increase access to and quality of health care and foster technological innovation.

† Mello is the C. Boyden Gray Associate Professor of Health Policy and Law in the Department of Health Policy and Management of the Harvard School of Public Health. Zeiler is Professor of Law at the Georgetown University Law Center. The authors thank Kevin Davis, Daniel Fox, Glenn Cohen, Lawrence Gostin, and Charles Silver for helpful comments, and Isaac J. Wheeler and Alexander Parachini for excellent research assistance.
INTRODUCTION

In recent years, the field of empirical legal studies has grown exponentially. Although empirical methods have been used to study some areas of law—corporate law, securities regulation, antitrust, family law, and criminal law, for example—more than others, nearly every area of law has been explored using empirical methods. Undoubtedly, health law and policy has been among the most widely studied areas. Examples include, but certainly are not limited to, studies estimating the influence of statutory damages caps on medical-malpractice insurance premiums and physician location decisions, the influence on job mobility of tax laws that encourage employers to provide health insurance as an employee benefit, the impact of regulation of prices and investment in hospitals on hospital investment and quality of care, the impact of government regulation on tobacco consumption, and the effects of basing Medicare physician payment on performance on quality of care.

Although the field of empirical health law (EHL) is still in its adolescence, the increased visibility of this research and its ever-increasing policy salience make it timely to survey the field. In this Article, we take up this task, describing commonly-used methods, analyzing enabling and inhibiting factors in the production and uptake of EHL research by policy makers, and suggesting ways to increase the volume, rigor, and impact of EHL studies.

We define the field of EHL to include quantitative and qualitative investigations of the effects of enacted and proposed health laws, regulations, and policies on economic, social, and health outcomes. Two key features distinguish methods used in EHL research from other approaches to legal scholarship: utilization of data of some kind and employment of the scientific method as opposed to legal argumentation or other forms of advocacy. Research designs used to conduct EHL research range from pure qualitative studies to controlled observational studies to


5 For a review of the EHL literature on this topic, see David S. Salkever, Regulation of Prices and Investment in Hospitals in the United States, in HANDBOOK OF HEALTH ECONOMICS, VOLUME 2 (A.J. Culyer & J.P. Newhouse eds. 2000).


7 For a review of EHL studies on this topic, see David A. Hyman & Charles Silver, You Get What You Pay For: Result-Based Compensation For Health Care, 58 WASH. & LEE L. REV. 1427 (2001).
We assess how well these methods are used in EHL research and discuss major methodological challenges, including the dearth of available data, difficulties in controlling for confounding variables, and obstacles to communicating results to policy makers.

We also assess whether and how EHL scholarship has impacted policy. We present two case studies—motor vehicle safety related laws and medical malpractice reform—that demonstrate the successful and unsuccessful application, respectively, of social-science research to public policy. Studies of motor vehicle safety regulation provided rigorous and timely evaluations of initial experiments with new laws relating to young drivers and identified desirable laws. Political interest groups helped to disseminate study findings to policy makers. As a result, effective laws were adopted widely. Studies of medical malpractice have not had this effect. This is partly because the research is not uniformly of high quality, partly because the studies are too methodologically complex for easy digestion by non-experts, and partly because some interest groups have undermined rather than promote the uptake of rigorous, scholarly work by policy makers.

Although some barriers that hinder EHL’s impact on policy are beyond researchers’ control, EHL scholars can take steps to increase the impact of research findings on policy. Among these steps are attracting more funding to obtain useful data, encouraging interdisciplinary collaborations, and training emerging legal scholars in scientific and empirical methods. Additionally, scholars can strive to increase the rigor of EHL studies, improve dissemination efforts, and help policymakers to become educated consumers of empirical legal research. Our aim is to propose a roadmap for those in the academy interested in increasing the impact of EHL scholarship and to encourage legal scholars interested in using empirical data to influence health policy to transform their scholarly approach, moving away from the role of the expert-advocate and toward the social-science model of scholarship.

In Part I of this Article, we offer an account of why empiricism has gained popularity in health law and policy. In Part II we describe methodological approaches and challenges. In Part III we examine enabling and inhibiting factors in EHL research through the lens of case studies in motor vehicle safety laws and medical malpractice reform. Finally, in Part IV, drawing on lessons learned from the case studies, we offer suggestions for increasing EHL’s rigor, prevalence, and impact on policy.

I. **Why Empiricism Has Come to Health Law**

Empirical scholars have investigated the connections between legal rules and outcomes over a wide range of legal topics, but health law has been a particularly active field. This development is a positive one given that empirical investigations of health law are a necessary component of sound policy making. Policy proposals often emanate from theories of how

---

changes in policy will impact various outcome measures or claims about the status quo (e.g., the impact of large jury awards on medical-malpractice insurance premiums). Empirical verification of the theoretical predictions, assumptions, and rhetorical claims on which policy proposals are based increases lawmakers’ confidence in the likely efficacy of proposals. It is critical, therefore, for empiricists to provide information about the relationship between health law and outcomes.

Several reasons account for the popularity of health law as a topic of investigation among empiricists. First, data produced by and gathered from health care industries have become increasingly available, along with a great deal of administrative and survey data on health status and conditions. Second, the dominance of state regulation in health care has generated variation and natural experiments that can be exploited to estimate the effects of various health laws on outcomes. Third, in recent years, law schools have hired a greater number of faculty trained in both law and empirical methods, which has increased the capacity for and interest in empirical legal studies generally. Fourth, the audience interested in the findings of EHL research has expanded, increasing the demand for it. Finally, demand has also been expanded by the positive impact the EHL field has had on policy and by growing concern among policy makers about health care spending and perceived crises in several corners of the health care system, including the medical malpractice liability system and health care insurance markets. We cover each of these in turn in more detail below.

A. Data on Outcome Variables

Despite the recent explosion in empirical legal research, many corners of the field are plagued by a paucity of accessible data that could be used to estimate the effects of legal rules on outcomes of interest. Although EHL scholars also grapple with difficulties in obtaining data useful for advancing our knowledge of the connections between health law and outcomes, there has been relatively good access, in some cases for decades, to various datasets on health care utilization and health outcomes. Regulators continue to open useful datasets to the research community, and researchers have often convinced market actors to do the same. The data usually come with strict confidentiality agreements, limiting the potential for other research teams to test the validity and robustness of the reported results, but the data have allowed

---

9 See, e.g., Gillian K. Hadfield, Judging Science: An Essay on the Unscientific Basic of Beliefs about the Impact of Legal Rules on Science and the Need for Better Data about Law, 14 J.L. & POL’Y 137 (2006) (discussing the general dearth of data necessary to study the impact of legal rules and policy); Michael Heise, The Importance of Being Empirical, 26 PEPP. L. REV. 807, 822 (1999) (pointing to underinvestment in data production as the reason for the dearth of data).

10 See infra Part IV.

11 Examples include the Healthcare Cost and Utilization Project (available since 1988), the Longitudinal Study of Aging (available since 1984), and the National Health Interview Survey (available since 1957). On the other hand, developments in privacy law have made it more difficult for researchers to access individual patients’ medical records for purposes of doing chart review studies (the Harvard Medical Practice Study, see supra note 56, is an example of such a study).

12 Examples of datasets that have more recently come on line include the National Practitioner Data Bank (reporting data on medical malpractice payments and various sanctions against practicing physicians since 1990), the Dartmouth Atlas of Health Care (reporting on hospital expenditures and costs, hospital capacity, physician and specialist work forces, frequency of diagnostic and surgical procedures, and Medicare payments for care provided by hospitals since 1996), and the Employer Health Benefits Survey (providing data on employer sponsored health benefits since 1998).
researchers to explore interactions between markets and legal rules that formerly resided in black boxes.

**B. State-Based Law**

A second competitive advantage for EHL researchers is that states have been the primary loci of lawmaking around important aspects of health care markets, including insurance coverage, regulation of health care delivery and tort law. The state-based nature of legal rules confers two specific advantages for research. First, adoption of legal reforms generally happens in waves rather than in all states at once. This pattern provides the cross-sectional and longitudinal variation researchers need for robust analysis of the effects of changes in legal rules on health outcomes. Second, that health care markets are made up of a large number of moving pieces increases the likelihood that legal rules are adopted for reasons unrelated to the outcome variable of interest—an important precondition for empirical models to produce unbiased estimates of the effects of the rules on the outcome variable.

**C. Human Resources**

Historically, the bulk of EHL research has been conducted outside law schools, in public health and public policy schools, business schools, and social science departments, including economics, political science and psychology departments. This is unsurprising in light of the fact that this type of research arises out of scientific and social-scientific traditions.

In recent years, however, we have witnessed an increase in EHL research performed within the legal academy and published in law journals. Several factors may account for this. Pioneering work by “bridge” scholars such as Dan Kessler and Troy Brennan, who are dually trained in law and empirical methods, has been important in establishing prominence and respectability for this research within the legal academy. Health economists’ interest in regulatory questions—propelled by Kenneth Arrow’s influential article, which helped shape health economics as a field—spurred recognition of potential synergies of interdisciplinary collaborations between legal scholars and economists. Interest in these same health care regulatory questions within the legal academy has led to the creation of interdisciplinary health law research programs at leading universities, including Georgetown/Johns Hopkins and Harvard.

---

13 Examples of legal rules related to health care adopted at the state level include tort reform and legal standards of care enforced by tort law, safety regulations such as seatbelt and workplace smoking laws, information disclosure regulations such as mandated disclosure of the terms of contracts between managed care organizations and physicians, regulation of health insurance coverage, taxation and other forms of regulation related to vices such as tobacco and alcohol, and regulation of access to unhealthful foods and beverages in schools. Of course, empirical researchers in other areas of law, including criminal law, contract law, and property law, enjoy similar benefits bestowed by gradual and partial adoption of reforms by states.

14 When there is a two-way relationship between the legal reform adopted and the outcome it aims to improve—for example, states with high malpractice insurance premiums are more likely to adopt malpractice damages caps, and damages caps have the effect of reducing premiums—a problem of endogeneity results, necessitating the use of sophisticated (and sometimes infeasible) analytical methods. For a more technical explanation, see JEFFREY M. WOOLDRIDGE, INTRODUCTORY ECONOMETRICS: A MODERN APPROACH 521-29 (3d ed. 2006).


which provide dedicated institutional support for collaborative, empirical research. Finally, carried by this momentum, a number of law schools recently have appointed faculty members trained in both law and empirical methodology and interested in EHL research.  

D. Expanding Audience

The move to increase production capacity for EHL research has been fueled by an increasing demand for this type of scholarship in several quarters. Physicians and institutional health care providers, who highly value quantitative research and evidence-based decision making, are regular consumers of EHL research as a source of information about how the law potentially impacts their practice and industry. This has recently become true of other health care market actors as well, including medical malpractice insurers, health insurers, employers who provide health insurance as a benefit to their employees, and government agencies that finance health care purchases for subsets of the population.

Public health communities also are regular consumers of EHL studies. These stakeholder groups value studies focused on potential legal solutions to, and sometimes law-related causes of, public health problems (e.g., obesity, tobacco- and alcohol-related illnesses, limited access to care). In addition, EHL studies focused on the impact of various legal institutions on the health and safety of at-risk populations are of interest to public health organizations (e.g., studies shedding light on how tort reform impacts subsets of potential victims of medical malpractice). Empirical studies of legal rules intended to control the spread of infectious diseases, including compulsory vaccination and compulsory treatment, are also of special interest.

Legal system insiders, including advocates, judges, and legislators, have also increased their demand for EHL research. More than ever before, lawyers rely on empirical studies to support legal arguments made to courts, arbiters and administrative agencies in health-related matters. For example, the Supreme Court brief filed by the respondents in Ashcroft v. Raich (later known as Gonzales v. Raich) cited a General Accounting Office study of the experience with medical marijuana laws in four states. Likewise, judges, themselves, refer to empirical research to justify their holdings. For example, in another recent Supreme Court decision, Justice Souter cited an empirical study to support claims regarding the influence of managed care contracts with physicians on physician satisfaction. Legislators also find these studies instructive for shaping policy and useful as ammunition to support their positions.

E. Growing Policy Significance

In 2004, the U.S. spent $1.9 trillion on health care, an average of $6,280 per person and 16% of the country’s Gross Domestic Product (compared to 5% in 1960, 12% in 1990 and 13%.

17 In addition to one of the authors, recent hires include Jill Horwitz (University of Michigan Law School), Anup Malani (University of Chicago Law School), Jonathan Klick (Florida State University Law School), Ronen Avraham (Northwestern University Law School), and Daniel Ho (Stanford Law School).
20 See, e.g., 105th Congress; 2nd Session, 98 CIS J. Print 93224 (citing a Congressional Budget Office report reviewing the empirical literature on the connections between cigarette price increases and tobacco consumption).
Most ($1.6 billion) of this spending is devoted to personal health care costs. The federal government financed 34% of personal health expenditures and state and local governments covered 11%. The sheer size of the industry and the growing burden on taxpayers have ignited public demand for policy makers to take affirmative steps to ease this burden and fix other health care market problems. The availability of data, advances in techniques for estimating the effects of legal interventions, and public intolerance for policy failures in this critical area have pushed policy making toward a more evidence-based approach.

A number of political interest groups have responded by churning out their own empirical research in an effort to sway the political process in their favor. In many cases, the research is simplistic and, by academic standards, of poor methodological quality. This has raised the stakes for EHL scholars to contribute a better alternative.

These developments illustrate the potential for growth and impact in the field of EHL. We further explore the status of the field by considering the state of the methods used to conduct EHL research and the uptake of research by the policy community, as illustrated by two case studies.

II. METHODS OF EMPIRICAL HEALTH LAW RESEARCH

The methods used in EHL research are as diverse as the topics investigated. In this Part, we discuss major methodological approaches and their relative representation in the field, and then examine some of the significant challenges that empirical legal scholars confront in researching health law issues.

A. The Spectrum of Analytic Methods

A wide spectrum of methods has been used to empirically study developments in health law. We review the most common methodologies, ordered from least to most scientifically rigorous, below.

First, case studies occasionally have been used to examine health law problems, particularly problems that are unique or for which examination of extreme examples is of interest. Because case studies are not controlled analyses and the example selected for study may be unrepresentative of the broader population, case studies have limited generalizability. They can, however, provide a rich, vivid, and detailed lens on a particular example of a problem.
or policy innovation. A well-known example is the fascinating chronicle of the federal government’s response to a perceived threat of epidemic swine flu in 1977, including how it addressed vaccine manufacturers’ liability concerns.\footnote{Richard E. Neustadt & Harvey V. Fineberg, The Swine Flu Affair: Decision-Making on a Slippery Disease (1978), cited in David C. Hoaglin et al., Data for Decisions: Information Strategies for Policymakers 119-23 (1982).}

A second, quite prevalent form of analysis in EHL research is qualitative analysis of judicial opinions.\footnote{See Michael Heise, The Importance of Being Empirical, 26 Pepp. L. Rev. 807, 824 (1999) (“Much of the existing empirical legal scholarship falls loosely into one or more of three broad categories: judicial opinion coding or case content analysis, descriptive, and inferential.”).} This approach bears some resemblance to traditional legal scholarship, which focuses on deep doctrinal analysis of a small number of judicial decisions. It differs in two key respects: expansion of the sample to include a larger number of cases, and use of systematic methods of content analysis. A typical research design involves random selection of a sample of cases meeting certain subject-matter criteria, application of a detailed coding and content-analysis methodology, and drawing of inferences based on both thematic analysis and frequency counts of various case characteristics.\footnote{See Kevin M. Clermont & Theodore Eisenberg, Litigation Realities, 88 Cornell L. Rev. 119, 125-26 (2002); Mark A. Hall & Ronald F. Wright, Systematic Content Analysis of Judicial Opinions. Wake Forest Univ. Legal Studies Paper No. 913336 (2006), http://ssrn.com/abstract=913336.}

This design offers a broader and more systematic examination of judicial decision making than more traditional forms of doctrinal analysis, and is very useful for getting the “lay of the land” in an area of litigation that has not yet been broadly characterized. Like all studies that rely on published judicial decisions, it is limited by the potential unrepresentativeness of published decisions.\footnote{Clermont & Eisenberg, supra note 28, at 125-26 (discussing several ways in which published decisions may be unrepresentative of the broader population of judicial decisions); William B. Sage, Judicial Opinions Involving Health Insurance Coverage: Trompe L’Oiel or Window on the World?, 31 Ind. L. Rev. 49, 61-68 (1998) (describing a number of limitations of content analyses of judicial decisions).} Content analysis has also been criticized for being too subjective.\footnote{See, e.g., Heise, supra note 27, at 825.}

Another application of qualitative methods in health law research is the collection of survey and interview data used to examine the impacts of health laws or litigation. In addition to performing content analysis, such studies may analyze survey and interview data quantitatively. Major challenges that compromise the quality of these studies are the difficulties in avoiding selection bias and in obtaining an adequate response rate, and concerns about respondents’ self-interest in the outcome of the study.

Occasionally, researchers have conducted descriptive studies in health law involving quantitative analysis of data from outside the legal system. The aim typically is to contribute suggestive evidence about the impact of a law or to test assumptions or claims policymakers and other researchers rely on to argue a position or develop theories. Simple descriptive statistics such as the mean and standard deviation, or frequency and proportion, are presented, and may be supplemented by bivariate\footnote{Bivariate analyses involve two variables and typically consist of cross-tabulations comparing values of an outcome variable in groups that have different values of an explanatory variable—for example, rates of mammography in states that do and do not mandate that insurers cover routine mammography. Univariate analyses give summary statistics about a single variable—for example, the mean rate of mammography across all states and the standard deviation.} analyses estimating associations between variables. Univariate analyses have no explanatory power, but can provide useful information on the nature of a health law problem or be used to support or call into question conventional wisdom—for example, the

\footnote{\textit{Richard E. Neustadt} & \textit{Harvey V. Fineberg}, \textit{The Swine Flu Affair: Decision-Making on a Slippery Disease} (1978), cited in \textit{David C. Hoaglin et al.}, \textit{Data for Decisions: Information Strategies for Policymakers} 119-23 (1982).\footnote{See \textit{Michael Heise}, \textit{The Importance of Being Empirical}, 26 Pepp. L. Rev. 807, 824 (1999) (“Much of the existing empirical legal scholarship falls loosely into one or more of three broad categories: judicial opinion coding or case content analysis, descriptive, and inferential.”).\footnote{See \textit{Kevin M. Clermont} & \textit{Theodore Eisenberg}, \textit{Litigation Realities}, 88 Cornell L. Rev. 119, 125-26 (2002); \textit{Mark A. Hall} & \textit{Ronald F. Wright}, Systematic Content Analysis of Judicial Opinions. Wake Forest Univ. Legal Studies Paper No. 913336 (2006), http://ssrn.com/abstract=913336.\footnote{Clermont & Eisenberg, supra note 28, at 125-26 (discussing several ways in which published decisions may be unrepresentative of the broader population of judicial decisions); \textit{William B. Sage}, \textit{Judicial Opinions Involving Health Insurance Coverage: Trompe L’Oiel or Window on the World?}, 31 Ind. L. Rev. 49, 61-68 (1998) (describing a number of limitations of content analyses of judicial decisions).\footnote{See, e.g., \textit{Heise}, supra note 27, at 825.\footnote{Bivariate analyses involve two variables and typically consist of cross-tabulations comparing values of an outcome variable in groups that have different values of an explanatory variable—for example, rates of mammography in states that do and do not mandate that insurers cover routine mammography. Univariate analyses give summary statistics about a single variable—for example, the mean rate of mammography across all states and the standard deviation.}}}}
extent to which physicians are personally exposed to malpractice settlements in excess of their policy limits. The major weakness of bivariate studies is that, although some analyses may be stratified, they otherwise do not control for confounding variables. This limits their usefulness in supporting inferences about causal connections between the variables. Therefore, although these methods can produce very useful information for policymakers and theorists, if limitations are not fully and clearly disclosed, readers may incorrectly attribute observed differences between comparison groups to the variables studied, when in reality other factors drive the differences.

A more sophisticated design, controlled observational or quasi-experimental studies using multivariate regression techniques, constitutes the bulk of empirical legal studies conducted by law-and-economics scholars. Such studies attempt to overcome the problems with drawing causal inferences from univariate and bivariate analyses by employing techniques to control for confounding variables. The simplest variant is the pre/post analysis, sometimes called event studies, that employs no explicit controls, but simply measures the outcome variable before and after the intervention. This design’s major weakness is that it does not control for exogenous changes in the environment over time that may introduce confounding variables.

Alternatively, multivariate regression techniques may be applied in cross-sectional or time-series analyses. Longitudinal designs may be employed to maximize the amount of information that can be gleaned from a limited number of cross-sectional units (such as states) and analyze changes over time. These studies take advantage of variation in legal rule adoption and outcomes over units of observation and time, and implement controls to rule out alternative explanations, allowing the researchers confidently to use the findings as evidence either of support or rejection of claims and theories positing causal connections. While these studies are useful as evidence for or against particular theories of causal connections, however, they should not be viewed as conclusive or touted as such.

---

33 A stratified analysis is one that divides observations into groups, or strata, based on their value of some potentially confounding variable and then performs bivariate analyses within each stratum. For example, a researcher might be interested in the relationship between state mandates for physical education in schools and the prevalence of overweight and obesity among schoolchildren, but suspect that the effect will be different for girls and boys. A simple way to investigate that hypothesis would be to stratify by sex and analyze the relationship within each stratum.
34 If an analysis does not control for confounding variables, then variation in an outcome variable may be incorrectly attributed to the explanatory variable of interest. For example, tort reforms are often implemented as packages that include several different types of reforms (e.g., damages caps plus joint-and-several liability reform). If joint-and-several liability reforms are not included in the analysis, observed reductions in the costs of malpractice claims might appear to be due to damages caps when, in fact, they may be caused in part by the other reforms.
36 See Martin Eccles et al., Research Designs for Studies Evaluating the Effectiveness of Change and Improvement Strategies, 12 QUALITY & SAFETY IN HEALTH CARE 47, 50 (2003).
37 Cross-sectional designs measure many units (e.g., health care consumption for all 50 states) at one point in time, while time-series data captures just one unit (e.g., health care consumption in Texas) at different points in time. For an in-depth discussion of these designs, see THOMAS D. COOK & DONALD T. CAMPBELL, QUASI-EXPERIMENTATION: DESIGN AND ANALYSIS ISSUES FOR FIELD SETTINGS (1979).
Virtually absent from the EHL literature are experimental designs. Randomized experiments are the gold standard for testing the effects of interventions; the use of random treatment assignment is ideal for reducing the influence of confounding variables. Although experimental designs have occasionally been used to test the effect of changes in health care organizational policy that could be implemented as public policy—for example, imposing copayments on enrollees in a private or public insurance program—for obvious reasons they are almost never an option for evaluating the effects of health laws.

In evaluating the rigor of EHL scholarship as a whole, it is useful to consider evaluative tools that reviewing organizations use to judge the scientific rigor of various study designs. Among the most prominent typologies of research designs is that of the U.S. Preventive Services Task Force. In considering whether an adequate evidence base exists to support recommending a particular preventive health measure, the Task Force applies the following rating system, which flows from best to least good:

I. Evidence obtained from at least one properly designed randomized controlled trial.
II-1. Evidence obtained from well designed controlled trials without randomization.
II-2. Evidence obtained from well-designed cohort or case-controlled analytic studies, preferably from more than one center or research group.
II-3. Evidence obtained from multiple time series with or without the intervention.
III. Opinions of respected authorities, based on clinical experience, descriptive studies, or reports of expert committees.

This typology emphasizes the importance of randomized studies, places little weight on descriptive studies, and views quasi-experimental designs as a heterogeneous group, with some designs superior to others. To our knowledge, the overwhelming bulk of EHL research conducted to date can best be classified as category “II-3” or “III” studies. In addition to the paucity of controlled trials, another notable gap in the EHL literature is the scant use of techniques of meta-analysis or systematic review. Meta-analysis uses sophisticated statistical

---

38 See Archie L. Cochrane, Effectiveness and Efficacy: Random Reflections on Health Services (1979); see also Rachel Croson, Why and How to Experiment: Methodologies from Experimental Economics, 2002 U. ILL. L. REV. 921, 945 (2002) (“Experiments can accomplish objectives that other forms of analyses cannot. They can offer clean tests of economic theories by constructing experiments that meet the assumptions of the theories, and observing the outcomes. They can investigate alternative causes of observed anomalies...by independently manipulating factors that in reality are confounded.”).

39 See Willging, supra note 35, at 1127; Eccles et al., supra note 36, at 48.

40 See William Manning et al., A Controlled Trial of the Effect of a Prepaid Group Practice on the Use of Services, 310 NEW ENG. J. MED. 1505 (1984).

41 Notable exceptions exist. See, e.g., Joseph P. Newhouse, Free For All? Lessons from the RAND Health Insurance Experiment (1996) (studying the impact of randomizing patients to health insurance plans with different levels of patient cost-sharing on health care consumption and health outcomes); Lauren A. McCormack et al., Health Plan Decision Making With New Medicare Information Materials, 36 HEALTH SERVICES RES. 531 (2001) (reporting results from controlled experiments to determine the impact of Medicare informational materials on new and experienced Medicare beneficiaries).

methods to aggregate and collectively analyze data from multiple studies on the same topic using similar designs. Systematic reviews apply a structured process to compare and evaluate the relative strength of findings from a number of studies examining the same phenomenon. These approaches are not without methodological controversy, but they have the advantage of harnessing the power of multiple studies. They are, however, rarely used to study the effects of legal rules on outcomes. As with all empirical studies, meta-analyses and research syntheses have limitations, but they can help to provide broader views of treatment effects identified in different studies especially when the studies report divergent results. Overall, the field of EHL lags well behind the standard of evidence that researchers and policymakers aim for in other areas of evaluation of health interventions.

B. Major Methodological Challenges

Empirical health law studies confront several significant methodological challenges that may explain why the field has not advanced farther. Three issues that cut across many areas of inquiry within health law are particularly troubling.

First and foremost is the problem of data availability. Although there are excellent datasets on health care utilization and many health outcomes, the investment that government and private research sponsors have made in building health datasets for research purposes is not mirrored in the realm of health law. The government’s most significant commitment to empirical legal research, the Bureau of Justice Statistics, has made only occasional forays into health law, all limited to descriptive studies of medical malpractice claims, and relying on data aggregated from other sources. There are few examples of government-funded or government-run projects aimed at building databases of health laws. The best known recent instance may be the National Science Foundation’s grant to Ronen Avraham at Northwestern University Law School to create a comprehensive database of state malpractice tort reform laws. Some private organizations, most notably the National Conference of State Legislatures, maintain useful lists of pending and enacted state laws on health topics, but these are not organized into a database, nor do they provide detailed information about the laws.

Without databases custom-built for research, EHL scholars must rely on administrative data collected for other purposes, or build their own datasets from the ground up. For research

---

43 See, e.g., Peter C. Gotzsche et al., Data Extraction Errors in Meta-analyses That Use Standardized Mean Differences, 298 JAMA 430 (2007).
44 Recently published studies suggest that they may be gaining traction in EHL research, though. See, e.g., Caroline M Fichtenberg & Stanton A Glantz, Effect of Smoke-Free Workplaces on Smoking Behaviour: Systematic Review, 325 BMJ 188 (2002) (systematically reviewing primary empirical studies to compare the efficacy of workplace smoking bans and tax increases); Stephen Kisely et al., Randomized and Non-Randomized Evidence for the Effect of Compulsory Community and Involuntary Out-Patient Treatment on Health Service Use: Systematic Review and Meta-Analysis, 37 P SYCH. MED. 3 (2007).
45 Of course, the explanation rests in part on the difficulties with conducting experimental studies in health law, as we discuss further in Part II.B.
46 See medical malpractice studies collected at http://www.ojp.usdoj.gov/bjs/pubalp2.htm#M.
47 This is a particular challenge for EHL researchers attempting to identify the impact of tort reforms on medical-malpractice insurance premiums and other relevant outcome variables. See infra Part III.B.
topics related to health or liability insurance, administrative data held by private insurers, state departments of insurance, state medical boards, and state health departments are an attractive option. Unfortunately, such databases can be difficult to access; states and insurers hold them tightly, due to a combination of legal concerns and anxieties about the potential political fallout from adverse study findings. Additionally, data-collection processes and data fields often are not standardized across states. Single-state datasets are rarely nationally generalizable, and datasets from multiple states cannot easily be aggregated to support national studies.

In the medical liability realm, the National Practitioner Data Bank and jury verdicts databases offer two alternatives, though each captures only a portion of all filed claims and has limited data fields. The Data Bank most likely suffers from underreporting and restricts the data fields the public may access. Jury verdict reports have also been criticized for their unrepresentativeness and incompleteness.

Administrative datasets also have the obvious shortcoming of not having been constructed with researchers’ needs and objectives in mind. They may omit data fields that researchers view as crucial to a rigorous analysis, and the data may not have been collected with the degree of completeness and accuracy that researchers desire. Finally, researchers external to the organization that collected the data may find it challenging to become familiar with the dataset and its quirks and quality problems. Unless an organizational insider is willing to spend the time to brief the researchers on the dataset and respond to their queries, important data-quality issues and definitional issues may go unnoticed.

Engaging in primary data collection to build new datasets is a possibility in many areas of empirical health law research, but it can be prohibitively expensive, and few extramural research sponsors are interested in funding such enterprises. Additionally, projects that require on-the-ground data collection (for example, reviewing medical charts to identify instances of malpractice or reviewing insurance claims to identify wrongful denials) require substantial infrastructure to train, house, and support a large data-collection staff. Few law schools can afford to meet these demands, so scholars tend to rely on readily available but very limited administrative datasets, and what they can cull from judicial opinions and electronic databases of statutes and regulations. Data-collection efforts are more common in schools of public health and

---

50 See Neil Vidmar, Medical Malpractice and the Tort System in Illinois: A Report to the Illinois State Bar Association 5-6 (2005) (reporting that the Illinois Department of Insurance refused to grant permission to analyze its claims data); Sage, supra note 29, at 59 (“health insurers often assert proprietary interests in their coverage standards”).

51 See Michelle M. Mello et al., Changes in Physician Supply and Scope of Practice During a Malpractice Crisis: Evidence From Pennsylvania, 26 HEALTH AFFAIRS w425 (2007) (describing an instance where an insurance department expressed concerns about violating its confidentiality obligations under state law).


55 For a discussion of the ways in which databases maintained by health insurers fall short of providing sufficient information to evaluate medical appropriateness decisions, see, for example, Maxwell J. Mehlem, Getting a Handle on Coverage Decisions: If Not Case Law, Then What? Comments on a Paper by Professor William Sage, 31 IND. L. REV. 75, 75-78 (1998). On the other hand, administrative data may be more accurate than data collected for research purposes because the information providers may have stronger incentives to treat the task seriously and provide truthful responses. HOAGLIN ET AL., supra note 26, at 159.
medicine, but their heavy reliance on extramural funding can make it difficult for faculty to pursue such projects if grants are not available to support them.

The second major methodological challenge is the difficulty of controlling for confounding variables—a problem common to all observational studies. As previously discussed, the ideal method to implement such controls is through the use of experimental designs including randomized studies. Although there has been some use of experimental designs in other areas of empirical legal research focusing on judicial processes, there are few opportunities to conduct pure experimental studies in health law. Judicial processes are of less interest, and the agendas and actions of state legislatures cannot be dictated by a randomization process. Thus, as in other areas of social science, researchers must rely on carefully controlled quasi-experimental designs.

Controlling for confounding variables is particularly hard in studies of innovations in state law, due to the profound and multifaceted cross-sectional variation. Finding appropriate data to control for variables that might be correlated with state enactment of health laws can be a difficult, and sometimes impossible, task. Fixed-effects models offer a possible solution, but only where the laws in question exhibit a reasonable amount of variation over the study period, the number of cross-sections (e.g., states) is small relative to the number of time periods covered by the data, and the unobservable heterogeneity is constant over time. Difference-in-differences models are more attractive in some cases, but have rarely been used in health law studies, and those researchers who have used them do not appear to have incorporated appropriate adjustments for the inconsistency of standard errors that arises from serially-correlated data.

---

56 The most famous example may be the Harvard Medical Practice Study, a review of approximately 30,000 medical records that measured the incidence of medical negligence and malpractice claiming. The main study findings are collected at Paul Weiler et al., Medical Malpractice on Trial (1993).
58 See, e.g., Willging, supra note 35, at 1128-29 (describing an experiment with mandatory attorney conference prior to briefing and arguing appeals in the U.S. Court of Appeals for the Second Circuit); Croson, supra note 38, at 935 (describing experiments that tested the effect of a damages cap law on pretrial bargaining among parties in a hypothetical personal injury case).
59 Among the barriers to randomized experiments in social-science research generally are ethical concerns about randomization in some circumstances, differences between experimental and actual delivery of the intervention, time and money costs, and controlling the integrity of the experiment in the field. See Peter H. Rossi & Howard E. Freeman, Evaluation: A Systematic Approach (5th ed. 1993) at 290-93.
60 Fixed-effects models are used to control for unobservable heterogeneity that is constant over time. See Cheng Hsiao, Analysis of Panel Data 30-34 (2d ed. 2002).
61 Given the limited time horizon of many studies, this often will not be the case. For example, medical malpractice tort reforms in the states were largely time-invariant between the late 1980s and the early 2000s. See Y. Tony Yang et al., A Longitudinal Analysis of the Impact of Liability Pressure on the Supply of Obstetrician-Gynecologists, at 6 (2007) (on file with authors).
62 Difference-in-differences models are used to distinguish effects caused by an “exogenous shock” such as a legal reform from differences attributable to time trends driven by other causes.
63 Among the rare instances in which such models have been used are Daniel P. Kessler et al., Impact of Malpractice Reforms on the Supply of Physician Services, 293 JAMA 2618 (2005); Kessler & McClellan, supra note 15; Klick & Stratmann, supra note 3; and William E. Encinosa & Fred J. Hellinger, Have State Caps on Malpractice Awards Increased the Supply of Physicians?, w5 Health Affairs 250 (2005).
64 Year-to-year correlations in the dependent variable may bias standard errors downward, resulting in overestimation of significance levels. See Marianne Bertrand et al., How Much Should We Trust Differences-in-
Due to these complexities, there is a substantial risk of omitted variable bias in studies of many kinds of health laws. By properly controlling for this bias, however, parameter estimates can be used to draw causal inferences.

The third major challenge is how to model state laws that may have similar aims and general approaches but exhibit considerable heterogeneity in their specific provisions. For example, some laws that aim to ensure that women will be protected from “drive-through” deliveries set a specific minimum hospital stay, while others provide that the length of stay shall be determined by the woman and her physician rather than her insurance company.\(^{65}\) There is no consensus within the field about whether such diverse laws should be modeled using a single dummy variable (indicating whether a law of that general type is present or absent) or a series of variables that compares different variations of the law to one another. The latter approach avoids the dilution effects that may occur if some variants of the law are more effective than others but all are lumped together in the analysis. But in some cases, as in the mastectomy law example, not enough states have adopted laws of that type to permit such detailed classification. Grouping heterogeneous legal mandates together can lead to greater parsimony, but at the price of less helpful results.

These challenges require considerable expertise and resources to surmount. For reasons we discuss later in the Article, scholars in law schools may struggle to find them. In summary, the field of EHL presents a wealth of opportunities for scholars to make important contributions to expanding the relatively small knowledge base; however, these opportunities are costly to pursue. The EHL literature to date has been dominated by descriptive analyses and quasi-experimental designs. The challenge for empirical legal researchers is to find the proper balance between expediency and rigor\(^ {66}\) in selecting a study design, keeping in mind the quality of evidence they would like to contribute and the constraints of the research environment in which their work is conducted. It is also to ensure that their findings are interpreted—in their own reports and in policy debates—in a way that reflects the limitations of their data and analytical methods.

With this survey of the field in mind, we now turn to the question of how EHL scholarship has impacted policy.

### III. The Impact of Empirical Health Law Scholarship on Policy: Two Case Studies

In this Part, we juxtapose two policy histories that illustrate, respectively, the promise and the pitfalls of building a reliable evidence base for legal reforms and ensuring that policymakers respond to it. As an example of successful, evidence-based health law, we offer the history of laws aimed at reducing motor vehicle injuries to young drivers. As an illustration of where

---


policymaking can depart significantly from the best evidence about what is likely to succeed, we offer the case of medical malpractice reform during the most recent insurance “crisis.”

A. Motor Vehicle Safety Laws

Injuries from motor vehicle accidents are the leading cause of death among youth aged sixteen to twenty in the United States. In 2000, there were 5,600 adolescent fatalities in traffic crashes, and the economic cost of accidents involving drivers under age 21 was $32.8 billion. The high burden of injury, coupled with the long history of state and federal regulation of highway safety, has made vehicular injuries involving young people an obvious target for legal interventions. These interventions have been plentiful (and not restricted to adolescents). The best known are seatbelt laws, motorcycle helmet laws, speed limits, laws relating to alcohol consumption and driving, and driver’s education, testing, and licensing provisions.

Motor vehicle safety policy is an EHL research success story. The Centers for Disease Control and Prevention has identified the massive reduction in motor vehicle fatalities in the latter half of the twentieth century as one of the century’s ten greatest public health achievements, and credited public health laws as a major contributor. Not only have public health laws in this area proved effective, but evidence of their effectiveness has influenced their spread across the states. There are many areas in which EHL research contributed to the widespread adoption of efficacious motor vehicle laws, but we will focus on two of the best-documented successes: graduated driver licensing (GDL) and minimum legal drinking age (MLDA) laws, both of which aim to reduce traffic accidents involving young drivers.

1. Graduated Driver Licensing

To address the risks arising from driver inexperience, GDL laws provide that young drivers initially will be issued restricted licenses and will gain full privileges only after they have acquired more experience. Typical GDL provisions require holding of a learner’s permit for a minimum amount of time (during which time the learner may only drive with an older person present), restrict the number of passengers that young drivers may carry, prohibit freeway, highway, and nighttime driving, and require zero blood alcohol content. In some countries, GDL schemes apply to all new drivers regardless of age, but laws in the U.S. focus on young drivers.

The story of GDL laws illustrates how empirical research can play an integral role in both policy generation and policy evaluation, leading to successful, evidence-based lawmaking. In

---

68 Christine Branche et al., Graduated Licensing for Teens: Why Everybody’s Doing It, 30 J.L. MED. & ETHICS 146, 146 (2002).
70 We thank David Hemenway for suggesting these examples and referring us to pertinent research, which he summarizes in his book, DAVID HEMENWAY, WHILE YOU WERE SLEEPING: SUCCESS STORIES IN INJURY AND VIOLENCE PREVENTION (forthcoming 2008).
brief, GDL laws emerged from public health research highlighting a risk factor for serious injury. In response to these findings, experimentation with GDL laws ensued in a small number of jurisdictions. The pioneer laws were evaluated rigorously and in a timely fashion. The findings were impressive and led other jurisdictions to adopt similar laws. This, in turn, led to additional evaluations and more widespread adoption of the laws. This iterative process of research and policy making constitutes an ideal “lifecycle” for a health law (Figure 1).73

Figure 1. Lifecycle of an empirical law “success story”

The first stage in this lifecycle was the identification of a health risk factor based on sound epidemiological research. In 1971, the North Carolina Highway Commission reported findings indicating that young drivers were disproportionately represented in motor vehicle accidents.74 The risk was particularly elevated at night and when young drivers were accompanied by young passengers in the front seat. These findings served as the original inspiration for GDL proposals.75

The earliest proposals were made in the late 1970s by a transportation safety researcher, and the National Highway Transportation Safety Administration developed a model GDL program in 1977.76 However, the idea did not immediately take root among policymakers; only a

73 This lifecycle bears some resemblance to the framework for evaluating complex interventions for improving health suggested by Michelle Campbell and colleagues, which also has four stages: development of the theoretical basis for an intervention, definition of components of the intervention, exploratory studies to develop further the intervention and plan a definitive evaluative study, and definitive evaluative study. Michelle Campbell et al., Framework for Design and Evaluation of Complex Interventions to Improve Health, 321 BRIT. MED. J. 694 (2000).
75 Id. at 18.
few states implemented GDL legislation over the next decade. Meanwhile, other countries were also considering GDL schemes. In 1985, an international symposium on young drivers held in Canada endorsed the idea of GDL systems. The notion received further support after publication in 1990 of a highly influential Canadian study that definitively linked increased driver experience with decreased collision rates.

The second stage of the lifecycle—pioneering policy experiments—emerged around this time. In 1987, New Zealand implemented a GDL program in response to data demonstrating that young drivers were at a highly elevated risk for accidents and suggestions from a leading traffic safety researcher that a GDL system would be helpful. This pioneering law was studied from its earliest days in order to ascertain its effectiveness. Anecdotal evidence of reductions in traffic accidents began to surface in New Zealand almost immediately. A series of more rigorous evaluations, the first of which was published in 1992, showed dramatic results.

This third stage in the lifecycle, evaluation of policy experiments, led swiftly to the fourth, emulation of successful experiments. Ontario, Canada adopted a GDL program in 1994 based on the New Zealand results. Nova Scotia did the same later that year, and four more provinces followed over the next five years.

These laws, too, were evaluated and shown to be highly effective. In 1998, a pre-post analysis of the Ontario law showed a 31 percent decrease in the collision rate for novice drivers, a 27 percent drop in alcohol-related collisions, a 62 percent drop in collisions between midnight and 5:00 a.m., and a 61 percent decline in freeway crashes. Controlled studies of the Canadian programs were available by the year 2000, showing reductions in Nova Scotia’s per-capita crash rate in the 24 to 26 percent range and a 5 percent reduction in fatalities and 14 percent reduction in injuries in Quebec.

---

77 Simpson, supra note 71, at 27. There was some limited state action in response to the NHTSA report—Maryland adopted a nighttime driving restriction in 1979 and California and Oregon introduced provisional licenses in 1983 and 1989, respectively—but comprehensive GDL schemes did not emerge until the 1990s. See Allan F. Williams, *Earning a Driver’s License*, 112 PUB. HEALTH REP. 453 (1997).


80 Dorothy Begg & Shaun Stephenson, *Graduated Driver Licensing: the New Zealand Experience*, 34 J. SAFETY RES. 99, 99 (2003), citing NEW ZEALAND MINISTRY OF TRANSPORT, A GRADUATED DRIVER LICENSING SYSTEM (1985) (reporting that drivers aged 15-19 accounted for 27% of all drivers in traffic accidents notwithstanding the fact that they drove only 8% of the total mileage driven in New Zealand).

81 Simpson, supra note 71, at 29, reports that the initial evaluations of the New Zealand program found reductions in the per-capita crash rate on the order of 25%, while later studies found longer-term effects around 7%. See also Begg & Stephenson, supra note 80, at 100-103 (summarizing study findings).

82 Simpson, supra note 71, at 28.


84 Daniel R. Mayhew et al., *Impact of the Graduated Driver Licensing Program in Nova Scotia*, 2 J. CRASH PREVENTION & INJURY CONTROL 179 (2001), cited in Simpson, supra note 71, at 29. Subsequent studies confirmed this impact, but found that most of it occurred during the learner’s permit stage rather than among drivers who held a restricted license. Simpson, supra note 71, at 29.

The Canadian laws appear to have sparked interest in GDL in the U.S. In 1996, Florida became the first U.S. state to adopt GDL.\textsuperscript{86} By 1997, 6 states had adopted GDL laws. Around this time, the Insurance Institute for Highway Safety and the Traffic Injury Research Foundation of Canada issued a joint report, which subsequently was updated several times, outlining detailed recommendations for GDL programs.\textsuperscript{87} By 1999, 17 states had adopted them.

The early state laws were evaluated in relatively short order, and a series of studies published around 2001 reported uniformly significant results.\textsuperscript{88} The magnitude of the decreases in rates of fatal crashes, injury crashes, and all crashes varied across states, but in some states was staggering: North Carolina, for example, experienced a 57 percent drop in the per-capita fatal crash rate among 16-year-old drivers.\textsuperscript{89} Based on data reported in a 2004 review of 13 studies evaluating seven GDL schemes in the US and 6 in foreign countries,\textsuperscript{90} the median reduction in per-capita crashes was 31 percent and the median reduction in injury crashes was 28 percent. Statistically significant effects of the laws have been observed in all states that have been studied.\textsuperscript{91} Many of these studies used quite sophisticated designs to control not only for confounding variables, but also uncertainty about the time frame in which the effects of the laws would be seen.\textsuperscript{92} Overall, the evidence for the laws’ effectiveness is very strong.\textsuperscript{93}

Over the past decade, the results of the evaluations have led 47 states and the District of Columbia to adopt some kind of GDL law.\textsuperscript{94} GDL researchers have noted that “the primary driving force” in the proliferation of GDL laws across states has been not merely “the cascading effect of follow the leader,” but the scheme’s “proven effectiveness.”\textsuperscript{95} Thus, in GDL laws we see initial policy experiments maturing into a dominant policy approach through an iterative cycle of evaluation and emulation.\textsuperscript{96}

2. Minimum Legal Drinking Age Laws

\textsuperscript{86} Simpson, \textit{supra} note 71, at 28.
\textsuperscript{87} The most recent version is \textsc{Insurance Institute for Highway Safety & Traffic Injury Research Foundation, supra} note 72.
\textsuperscript{89} Robert Foss et al., \textit{Initial Effects of Graduated Driver Licensing on 16-year-old Driver Crashes in North Carolina}, 286 \textsc{JAMA} 1588 (2001).
\textsuperscript{90} Lisa Hartling et al., \textit{Graduated Driver Licensing for Reducing Motor Vehicle Crashes Among Young Drivers}, 2 \textsc{Cochrane Database of Systematic Reviews} CD003300 (2004).
\textsuperscript{91} Shope & Molnar, \textit{supra} note 88, at 67.
\textsuperscript{92} \textit{See, e.g.}, Michael R. Elliott & J T. Shope, \textit{Use of a Bayesian Changepoint Model to Estimate Effects of a Graduated Driver’s Licensing Program}, 1 \textsc{J. Data Sci.} 43 (2003); John D. Langley et al., \textit{An Evaluation of the New Zealand Graduated Driver Licensing System}, 28 \textsc{Accid. Anal. & Prev.} 139 (1996). \textit{See also} Hartling et al., \textit{supra} note 90, at 6 (summarizing study designs in a comprehensive literature review).
\textsuperscript{93} \textit{See} Hartling et al., \textit{supra} note 90, at 9.
\textsuperscript{94} \textsc{Insurance Institute for Highway Safety, U.S. Licensing Systems for Young Drivers} (2007), http://www.iihs.org/laws/state_laws/pdf/us_licensing_systems.pdf. This resource provides a detailed summary of the features of each state’s GDL law.
\textsuperscript{95} Simpson, \textit{supra} note 71, at 29.
\textsuperscript{96} In reviewing the literature on GDL laws, we also examined whether empirical legal research was influential in judicial decision making; for example, in adjudicating constitutional challenges to the laws. We found that GDL laws were not challenged in the courts.
The MLDA is the best-studied alcohol policy in the United States, and another example of a success story for health law and EHL research. Aimed at reducing alcohol consumption by young adults, the MLDA has been set by state law and historically has varied from state to state between 18 and 21. In 1970, 33 states set the age at 21. Between 1970 and 1975, 29 states lowered their MLDA to 18, 19 or 20. This occurred as part of the broader social recognition of 18-to-20-year-olds as legal adults that occurred with the passage of the Twenty-Sixth Amendment in response to the Vietnam War draft.

When the MLDA was first lowered, it provoked considerable social controversy concerning the effect on alcohol-related motor vehicle accidents involving young drivers. The available evidence at this point in time consisted largely of small, uncontrolled studies, fuelling debate about whether observed increases in crash rates in states that lowered the MLDA were due to the legal change or broader secular forces, such as improved police reporting practices.

By the mid-1970s, controlled studies had become available. These generally confirmed the initial study findings: lower MLDA was significantly associated with elevated risk of alcohol-related crashes among young drivers and increased alcohol consumption. Effect sizes varied somewhat across studies, but clustered in the range of a ten to thirty percent increase in crashes.

In light of this evidence, many states moved back in the other direction. Maine was the first, in 1977, and several others quickly followed. The experience of these early movers was promptly studied, and the evaluations demonstrated a reduction in alcohol-related crashes. By 1983, 16 states had moved to a MLDA of 21. Researchers continued to evaluate these changes, and by 1983, 5 studies had been published on the effect of raising the MLDA on traffic crashes; 4 demonstrated statistically significant effects on at least one type of crash.

Political interest groups, particularly Mothers Against Drunk Driving, used these findings to press the case for reform at the national level. Their advocacy was extremely effective, and in April 1982, President Reagan established a Presidential Commission on Drunk Driving at the urging of more than 300 members of Congress. While the Commission was carrying out its work, the National Transportation Safety Board recommended a national minimum drinking age of 21. This recommendation appeared in the Commission’s interim report in December 1982 and headlined its final report in November 1983.

---


99 Id.

100 See Wagenaar & Toomey, supra note 97, at 207, 210 (listing studies on lowering the MLDA and their findings).

101 Id. (summarizing study findings and citing several other reviews of the literature); see also Wagenaar & Toomey, supra note 97, at 209, 212-13, 215 (tallying outcomes among 132 studies).

102 See Wagenaar & Toomey, supra note 97, at 210.


105 Id. at 3.

106 PRESIDENTIAL COMMISSION ON DRUNK DRIVING, FINAL REPORT (1983).
The Commission’s work quickly led to congressional action. In October 1982, the Congress enacted the Alcohol Traffic Safety and National Driver Registration Act, which mandated that the Secretary of Transportation consider a state MLDA of 21 as a criterion for awarding supplemental grants to states. In December, it inserted a statement in its highway funding act that “The Congress strongly encourages each State to prohibit the sale of alcoholic beverages to persons who are less than 21 years of age.”

When the states did not immediately adopt laws enshrining this recommendation, Congress sought to increase its leverage. Initial legislation proposed a national minimum drinking age to be imposed on states by Congress under its interstate commerce power. The Reagan Administration had originally opposed a federal mandate, preferring to inspire states to act on their own, but quickly reversed its position. The initial legislation nonetheless failed, but Congress responded with an alternative proposal. In 1984, it passed the National Minimum Drinking Age Act (“NMDA”), which, rather than imposing a national drinking age outright, tied it to federal highway funds. The Act withheld 5 percent of highway funding from states that did not raise their MLDA to 21, increasing to 10 percent beginning in fiscal year 1988.

Empirical legal research played prominently in hearings on both versions of the NMDA legislation. Witnesses offering testimony cited research by University of Michigan researchers, the Insurance Institute for Highway Safety, the National Highway Traffic Safety Administration, and law professors at the University of Chicago. Although there was some dispute among witnesses about how to interpret and credit the studies, the Congress ultimately appeared to find the evidence persuasive. The NMDA was passed, and in response, all states moved their MLDA to 21 by 1988.

---


110 SEGAL, supra note 104, at 3.


114 Wagenaar & Toomey, supra note 97, at 206.
Research continued to be conducted after implementation of the NMDA, and these studies confirmed the earlier findings.\(^{115}\) A comprehensive evaluation of the evidence base requested by Congress and reported in 1987 by the General Accounting Office ("GAO")\(^{116}\) raised the profile of the evidence and dispelled any lingering doubts about the effectiveness of raising the drinking age. The GAO’s thorough, 111-page review concluded that “the evidence is persuasive” that raising the MLDA has significant effects on alcohol-related crashes among 18-to-20-year-olds, and that the observed effects were consistent across studies in different states and with different designs and methods.\(^{117}\)

The GAO scrutinized the methodologies of the existing studies and found that although the studies varied in strength, a number of reliable, high-quality analyses were available. The designs of MLDA studies have ranged from simple pre/post studies to cross-sectional studies to the more tightly controlled, longitudinal and time-series analyses. The best analyses\(^{118}\) exploit not only variation over time as the law changed, but also variation across age groups (18-to-20-year-olds, who were affected by the law, to 21-and-over groups, who were not).

MLDA laws have been challenged infrequently in court, but the three cases that have been litigated have produced memorable results. The earliest case was an equal-protection challenge to a Michigan constitutional amendment raising the MLDA from 19 to 21. The federal district court’s opinion in *Felix v. Milliken*\(^{119}\) contained an exhaustive discussion of two studies of Michigan’s law by Dr. Richard Lee Douglass on the effect of lowering the drinking age on traffic accidents in Michigan.\(^{120}\) The court’s exposition of the studies is remarkably detailed, and after this review of the evidence, the court had little trouble in concluding that Michigan’s voters had a reasonable basis for changing the law.

The second case was South Dakota’s famous challenge to the NMDA, which culminated in the U.S. Supreme Court’s decision in *South Dakota v. Dole*,\(^{121}\) upholding the statute against challenges under the Spending Clause and the Twenty-First Amendment. The Court’s Spending Clause analysis cited findings of the Presidential Commission, which had concluded that varying drinking ages across the states fostered drunk driving by encouraging young people to commute to states where the drinking age was lower,\(^{122}\) and concluded that the Act pursued “the general welfare”\(^{123}\) and directly related to the federal purpose for which highway funds were spent.

---

\(^{115}\) See *id.*, at 211-12 (collecting studies). Some of these studies have focused on the interactions of the MLDA law change with other alcohol-related laws. See, e.g., William R. Ponicki et al., *Joint Impacts of Minimum Legal Drinking Age and Beer Taxes on US Youth Traffic Fatalities, 1975 to 2001*, 31 **ALCOHOL CLIN. EXP. RES.** 804 (2007) (examining interaction with beer taxes); Robert B. Voas et al., *Assessing the Effectiveness of Minimum Legal Drinking Age and Zero Tolerance Laws in the United States*, 35 **ACCIDENT ANAL. PREV.** 579 (2003) (examining interaction with 0.2% blood alcohol concentration limit for drivers under age 21). Twenty-four evaluations of MLDA laws were published in the two years following the adoption of the NMDA alone. U.S. GENERAL ACCOUNTING OFFICE, **DRINKING-AGE LAWS: AN EVALUATION SYNTHESIS OF THEIR IMPACT ON HIGHWAY SAFETY**, GAO/PEMD-87-10, at 11 (1987).

\(^{116}\) *id.*

\(^{117}\) *id.* at 39.


\(^{120}\) Michigan lowered the drinking age to 18 in 1972, and then raised it to 19 in 1978. *Id.* at 1364.

\(^{121}\) 483 U.S. 203 (1987).

\(^{122}\) *id.* at 209, citing PRESIDENTIAL COMMISSION ON DRUNK DRIVING, **FINAL REPORT** 11 (1983).

\(^{123}\) *id.* at 206.
The third case, *Manuel v. Louisiana*, is a relatively recent challenge to a post-NMDA state law. Ruling on an equal-protection challenge, the Louisiana Supreme Court held that the MLDA survived even intermediate scrutiny due to the extensive evidentiary record that the higher MLDA substantially furthered the state’s important interest in reducing alcohol-related crashes involving youth. The court’s decision on rehearing overturned its initial holding on the basis that it had been based on the wrong piece of statistical data. The rehearing opinion contained an extensive discussion of the evidence base on MLDA laws in an appendix, specifically referencing the GAO’s report as proof that the efficacy of MLDA laws was “widely accepted by experts.” In summary, the courts, as well as state and federal legislatures, have made good use of high-quality empirical research on the efficacy of MLDA laws.

Overall, MLDA laws have had a lifecycle similar to GDL laws, characterized by iterative interactions between lawmaking and empirical research. The MLDA story is even more dynamic than the history of GDL laws, however, because of the presence of not one but two natural experiments. Twenty-nine states first lowered their GDL laws, then sixteen states reversed that decision within a short period of time. After the drinking age was lowered, empirical evaluations provided evidence that the change in the law created a health risk. A commonsense solution was clear, and many states adopted it. These changes in the law were then extensively evaluated and the positive findings disseminated to state and federal policymakers through the efforts of political interest groups. They were deemed sufficiently compelling to support federal action, which led to universal adoption by the states. Finally, research and evaluation continued well after universal adoption, providing a continually growing evidence base that can be used to head off efforts to reverse the laws or challenge them on constitutional grounds.

This is an unqualified success story for EHL research. As a result of higher MLDA, the National Highway Traffic Safety Administration has estimated that 17,359 traffic fatalities were averted between 1975 and 1997.

3. Enabling and Inhibiting Factors

What accounts for the success of EHL research in the area of motor vehicle safety? Four enabling factors appear to have considerable explanatory force: the presence of natural experiments, the ready availability of data, the uncontested framing of the issue as a public health problem, and the facilitating role of political interest groups.

First, natural experiments with changes to GDL and MLDA laws provided opportunities for controlled testing of hypotheses about the efficacy of these changes. This was particularly true for MLDA laws, because the pioneering experiments took place in the United States and because states moved their MLDA from high to low and back again, providing two points at which to observe the effects of this change.

---

125 *Id.* at 125.
127 The Louisiana court originally held that the state had not demonstrated the efficacy of its scheme because the number of alcohol-related crashes was higher among drivers aged 21 to 24 than among those aged 18 to 20. On rehearing, the court decided that the correct statistic was the proportion of alcohol-related traffic crashes attributable to each group. Finding that the younger group was proportionally the most dangerous group, it upheld the law. *Manuel*, 677 So. 2d at 122-23.
128 *Id.* at 128-32.
Second, researchers were able to take advantage of these natural experiments because of the ready availability of good data on the outcome variables of interest. The primary outcome variable for GDL laws is the overall rate of motor vehicle crashes involving young drivers. Secondary outcomes include rates of injury crashes, fatal crashes, hospitalizations related to crashes, crashes during nighttime, alcohol-related crashes, and crashes involving property damage. In MLDA studies, the outcome variables have included traffic crashes involving alcohol, rates of alcohol consumption among under-21-year-olds, alcohol sales levels, and rates of driving while under the influence of alcohol.

For some of these measures, researchers have had to rely on self-reports on youth surveys, impressions of police officers responding to accidents, or statistical surrogates for alcohol involvement in a crash (for example, crashes that occur at nighttime and involve an unaccompanied male driver). In general, however, the quality of the data has been good, and its availability has been very good, compared to other areas of health law. State departments of public safety, police departments, or divisions of motor vehicles maintain electronic databases of motor vehicle accidents. Characteristics of drivers are often available from police reports or by searching driver’s license databases. Reliable measures of alcohol involvement in crashes—for example, breath tests and blood-alcohol content tests conducted by police officers and coroners—have often been available.

Since the 1980s, the federal government and many states have invested over time in maintaining and improving their traffic crash databases. This has enabled researchers to overcome some of the criticisms of early studies—for example, the objection that many MLDA studies relied on proxy measures of alcohol involvement in crashes rather than actual evidence of driver impairment. Advances in statistical modeling and computing in the 1980s and 1990s also made it possible for researchers to execute more sophisticated models that addressed concerns about confounding in studies from the 1970s and took full advantage of the available data.

A third enabling factor had to do with the framing of the issue in policy debates. From the earliest days of the debates, motor vehicle accidents involving young people were framed as a public health problem. Research had clearly established the high rate of crashes involving young drivers and the devastating injuries that resulted. Injury prevention was an established subfield of public health, and prevention of harm to children a core concern of the field. As a result, it was natural to look to public health research for policy solutions. Indeed, framing the issue as a public health problem attracted the interest of a number of skilled public health researchers. The overwhelming bulk of the empirical scholarship on GDL and MLDA laws was conducted not by legal scholars, but by researchers in the public health community. Finally, the clear public health ramifications of the problem drew the interest of influential and relatively apolitical federal agencies, such as the NHTSA, which provided funding and a receptive ear for empirical research.

The fourth enabling factor was the facilitative role played by political interest groups. In the case of GDL laws, no influential interest groups emerged to oppose the policy change or dispute the evidence base that supported it. This was partially related to the framing issue: characterizing the problem as a public health threat and a pediatric problem made it a highly

---

130 Hartling et al., supra note 90, at 3.
131 Wagenaar & Toomey, supra note 97, at 207-212 (listing outcome measures in all MLDA studies published between 1960 and 1999); GENERAL ACCOUNTING OFFICE, supra note 116, at 21-23.
132 Id. at 24-25.
133 GENERAL ACCOUNTING OFFICE, supra note 116, at 13.
sympathetic cause. No interest group contested that reducing the rates of crashes due to driver inexperience or alcohol use among young people was a worthy target for policy interventions, though there were constitutional challenges. This paved the way for rapid policy responses. The only industry group that played prominently in the policy debate was the Insurance Institute for Highway Safety, which favored GDL laws and funded and participated in studies to build the evidence base for them. In the case of MLDA laws, there was organized opposition to raising the drinking age, primarily alcohol industry groups. But interestingly, the influence of these organizations may actually have helped promote high-quality research. They did not produce a raft of poor-quality studies of their own, but rather dedicated themselves to criticism of academic research demonstrating a public health benefit to raising the MLDA. Researchers were aware that these groups would try hard to discredit their studies and, as a result, took a great deal of care in designing them, so that it would be difficult to find scientific flaws. Political interest groups that favored raising the MLDA were also crucial to the success story of MLDA laws. Groups such as Mothers Against Drunk Driving partnered with researchers who were committed to ensuring that their findings reached a policy audience, resulting in a highly successful dissemination effort.

To be sure, these enabling factors operated in the face of several factors that inhibited the successful production of reliable empirical legal research. These were primarily methodological problems. A key shortcoming of the empirical base regarding GDL laws is the studies’ consistent focus on evaluating GDL laws one state (or foreign country) at a time. Single-state studies suffer from clear generalizability issues. Perhaps more important, because each state’s GDL scheme consists of a number of different features, finding that one state’s scheme had a bigger effect than another state’s provides little information about exactly which features were most important. GDL laws are highly variable across states—for example, in the age groups covered and the particular restrictions imposed. This heterogeneity may be one factor explaining the variability in the magnitude of effects found among studies focusing GDL laws in different states. Along with variations in study methodologies, it makes it difficult to directly compare individual states’ evaluation results.

For MLDA, one challenge for empirical evaluation of the laws has been that studies have found very low rates of enforcement actions (such as compliance checks on alcohol retailers) in

---

134 In contrast, motorcycle helmet laws, which have also been proven to be highly effective in preventing injuries, have faced strong opposition on antipaternalism grounds, probably because they affect primarily adults. See e.g., City of Adrian v. Poucher, 247 N.W.2d 798 (Mich. 1976). Indeed, the recent trend in state motorcycle-helmet laws has been to scale them back so that they cover only young people.

135 Cf. Simpson, supra note 71, at 28 (noting that the Institute “had played a dominant leadership role for years, not only in the research that provided the groundwork for graduated licensing but also in actively promoting its adoption”).

136 In other debates, industry groups have produced their own studies to counter those produced by organizations advocating changes disfavoring them. See e.g., notes 168-170 and accompanying text.

137 Wagenaar, supra note 98, at 80S.

138 Id.

139 Shope & Molnar, supra note 88, at 67 (“Understanding which components of GDL have contributed to the crash reductions is difficult because programs are implemented as a total package.”).

140 See INSURANCE INSTITUTE FOR HIGHWAY SAFETY & TRAFFIC INJURY RESEARCH FOUNDATION, supra note 94, at 1-8 (outlining several dimensions of variation).

141 Simpson, supra note 71, at 33.

142 See Shope & Molnar, supra note 88, at 64.
some states.\textsuperscript{143} If alcohol sellers are aware that the probability of an enforcement action is low, they may not comply with MLDA laws. Assuming that this is likely the case, effect sizes observed in MLDA studies probably underestimate the potential effects of raising the MLDA on traffic crashes were states to enforce the laws vigorously.

Despite these methodological issues, the empirical research base around MLDA and GDL laws is impressive, and policymakers’ response has been equally impressive. Motor vehicle safety laws are one area that demonstrates well the power of empirical legal research to contribute to enormous reductions in the burden of injuries.

B. Medical Malpractice Reform

1. The Production of Scholarship and Policy

In striking contrast to the story of motor vehicle safety laws, the story of medical malpractice reform laws exemplifies a missed opportunity to use rigorous EHL research to develop evidence-based policy. Legislative consideration of liability-limiting tort reforms for medical malpractice has a long and cyclical history\textsuperscript{144}—interest in reform waxes and wanes with trends in the cost of medical malpractice liability insurance. Despite the accretion of a substantial base of scholarly research studying the effects of these reforms over time,\textsuperscript{145} the policy debates have had the same, disappointing tone, focus, and outcomes over three successive malpractice “crises,” which are characterized by spikes in malpractice insurance premiums and/or reductions in the supply of malpractice insurance.

During the first malpractice crisis, in the mid-1970s, the empirical research base was extremely thin. Little was known about the extent to which malpractice reforms achieved their four primary aims: stabilizing liability insurance premiums, controlling the frequency of malpractice claiming, reducing the severity (average cost) of malpractice claims, and preventing physician behaviors that are believed to be correlated with high-risk malpractice environments, such as leaving practice and practicing defensive medicine.

The first serious attempt to synthesize the literature to aid policy making came in 1973 with the report of the Secretary of Health, Education and Welfare’s Commission on Medical Malpractice,\textsuperscript{146} and most of the studies it identified were not of sufficient rigor or scope to generate useful broad conclusions about the malpractice environment.\textsuperscript{147} In the late 1970s, the National Association of Insurance Commissioners advanced the empirical base significantly by collecting and tabulating three years’ worth of insurance-carrier data on closed malpractice

\textsuperscript{143} Wagenaar & Toomey, supra note 97, at 218.

\textsuperscript{144} Rogan Kersh, Medical Malpractice and the New Politics of Health Care, in MEDICAL MALPRACTICE AND THE U.S. HEALTH CARE SYSTEM 40-67, 45-49 (2006), provides an excellent summary of this history.

\textsuperscript{145} To bind our discussion, we will focus on the role that empirical studies of malpractice reforms have played, especially during the most recent crisis, from 2000 to 2006.


The availability of this database made malpractice a more attractive and feasible topic for empirical scholars to study. Although Patricia Danzon and other analysts contributed a handful of carefully conducted, controlled analyses of the effects of tort reforms over the next decade, when the next malpractice crisis struck in the mid-1980s, policymakers still had relatively little in the way of empirical research to illuminate the path to effective reforms. The federal government again appointed a task force, which issued its findings as the crisis began to wane. Notwithstanding this contribution, the policy debates tended, as in the 1970s, to be dominated by political interest groups representing the medical profession on the one hand, and the legal profession on the other. Scholars and their research findings rarely played a role. The results were similar to those of the mid-1970s round: through the end of the 1980s and into the early 1990s, legislatures reached for a limited set of reforms, such as damages caps and collateral-source offsets, making few innovations. Notably, the few studies of damages caps available at this time had determined that although they reduced the average value of awards, they did not significantly affect the frequency of claiming or the level of malpractice insurance premiums.

Over the next fifteen years, the body of scholarly literature evaluating the effects of tort reforms grew significantly. A number of top-flight economists, law professors, and public-policy scholars contributed studies that controlled (more or less successfully) for variations in state environments that might explain gaps in premium levels, physician supply, or claiming levels and costs. When the third malpractice crisis hit, around the turn of the twenty-first century, scholarly interest in the topic redoubled. Various arms of the federal government published studies of varying quality in the early years of this crisis. The lifecycle of academic studies was considerably longer, but by 2004, scholarly studies capturing data through the early 2000s began to surface.

The academic studies of this period have been, as a group, quite rigorous and carefully conceived. They have made it possible to draw a few fairly definitive conclusions about damages...
caps and other conventional tort reforms. Noneconomic damages caps appear to have statistically significant effects on premiums and physician supply, but the effect sizes are quite modest. Caps have a larger effect on award amounts, as might be expected, but there is no evidence that they affect claims frequency. Joint-and-several liability reform has been found to be significantly negatively associated with premium levels, but not other outcome variables. Study findings concerning statutes of limitations and repose are mixed, but some have found an effect on premiums and claims frequency. The weight of the evidence suggests that attorney-fee limits, collateral-source rule reform, pretrial screening panels, and periodic payment have little to no effect on relevant outcome variables.

High-quality empirical research on malpractice reforms continues to be produced at a high rate. The question is, is anyone listening? Unfortunately, the answer appears to be no for legislators. Although the evidence taken as a whole does not support a clear regulatory solution, broad conclusions about the connection between malpractice reforms and outcome variables could be helpful for policymakers as they attempt to fashion sensible reforms. Despite this, little of this research has found its way into policy debates over reforms during the latest malpractice crisis. To the extent that it has, policymakers have tended to use the studies more to support their *ex ante* policy preferences than to determine optimal regulation.

The Congress has been focused narrowly on legislation proposing a federal cap on noneconomic damages. The House and Senate held several hearings on medical liability in 2003-2006 to consider this legislation (and, on a smaller number of occasions, bills to authorize grants to support demonstration projects of experimental tort reforms such as administrative compensation programs and early-offer programs). The hearings on caps were remarkably consistent in the composition of the panel of invited witnesses, the kinds of arguments and evidence offered by the witnesses, and the tenor of the dialogue between witnesses and committee members. The panels usually consisted of one or more malpractice “victims” who told their personal stories, one or more representatives of physician or hospital organizations, an insurance industry representative, a representative of the trial bar, and a representative of a consumer organization (often one with ties to the bar). Occasionally, a lone academic would appear.

Many witnesses did cite data in their testimony, but they almost uniformly relied on descriptive studies, eschewing the more complex, controlled studies produced by academics. Testimony was liberally peppered with anecdotes about persons affected by medical errors or

---


high insurance costs. Questioning of the witnesses by the committee members rarely addressed data, and when it did, it typically focused on analyses conducted by interest groups rather than academics.

The policy process at the state level was somewhat better, in that a number of state governors and legislators appointed study commissions to investigate the evidentiary basis for a range of proposed reforms. These commissions, however, tended to fall back on familiar strategies of information-gathering—namely, assembling a panel of witnesses, most of who represented the various stakeholder groups. Little independent investigation and review of scholarly literature appears to have been conducted.

The experience of the Florida Governor’s Select Task Force on Healthcare Professional Liability Insurance is illustrative. Appointed by Jeb Bush in 2002, the Task Force was directed to examine Florida’s medical liability environment and make recommendations for improving it. The Task Force had a blue-ribbon roster, including former Secretary of Health and Human Services, Donna Shalala, and ended up recommending to the legislature that it adopt a package of tort reforms. The proposed package included a $250,000 noneconomic damages cap, expert-witness qualifications, periodic payment, joint-and-several liability reform, and small-scale demonstration projects of administrative compensation. It took its work seriously, meeting ten times over five months, holding several hearings, producing a 349-page report, and submitting thirteen volumes of supporting materials along with it. Despite this seemingly comprehensive investigation and the fact that scholarly studies have generated evidence in support of many of the reforms the Task Force recommended, nothing indicates that the Task Force considered the academic research. Its discussion of expert-witness qualifications, joint-and-several liability reform, and periodic payment contains no citations to empirical studies of any kind and no references to information provided by academics. Its findings concerning administrative compensation also cited no scholarly studies of such systems, but did at least reference a report by the Institute of Medicine and testimony from a handful of scholars familiar with the literature. Its discussion of damages caps is, from an academic’s perspective, highly unsatisfactory. The Task Force referenced not a single published study of the effects of damages caps, relying instead on proprietary analyses performed by a consultant, William Hamm, under a contract from a patient-rights organization.

Similar processes and conclusions played out in other states. The dominance of uncontrolled, descriptive studies conducted by political interest groups in these debates has been striking. A typical report compares premiums in California, which has long had a $250,000 cap

---

162 One colorful example is an exchange between Rep. John Conyers and an insurance association executive, in which Conyers administered a “true or false” exam in an attempt to force the executive to acknowledge the findings of a consumer group, the Center for Justice Democracy, whose work he believed to be of poor quality. See U.S. Representative James Sensenbrenner (R-WI) Holds Hearing on Medical Liability Issues, FDCH Political Transcripts, Mar. 4, 2003.
163 Id. at viii-xiii.
165 See id. at 235-39, 260-61, 276-77.
166 Id. at 159-70.
167 See id. at 193-200. The Hamm reports are not publicly accessible, so it is difficult to ascertain the rigor of the analyses.
on noneconomic damages, to premiums in other states that do not have caps, without controlling for any of the ways in which the states may differ.\(^{168}\) These studies suffered from numerous other methodological problems including inappropriate comparison groups, failure to adjust for inflation, and selective reporting of data.\(^{169}\) They were, however, presented simply, in straightforward language, with vivid charts and graphs, in short briefing papers that were far more accessible (physically and cognitively) to policymakers than scholarly journal articles.\(^{170}\) The academic articles described complicated multivariate regression models in lengthy and pointy-headed presentations—hardly fodder for headline-grabbing hearings.

As a result, to the extent that enacted policies reflected empirical findings, it was largely a matter of luck rather than deliberate process. Examining the relative prevalence of various kinds of tort-reform legislation at the state level over the past two years (Table 1), the most encouraging observation is that the reform with the greatest amount of empirical support from well-designed studies, damages caps, has been considered by more states than any other reform. On the other hand, many other reforms with little or no evidence to support them have been repeatedly considered and, on several occasions, enacted. Similar observations can be made from an examination of the prevalence of enacted state reforms over time (Figure 2).

<table>
<thead>
<tr>
<th>Evidence of efficacy</th>
<th>2005 Considered</th>
<th>2005 Enacted</th>
<th>2006 Considered</th>
<th>2006 Enacted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Damages caps</td>
<td>Some</td>
<td>28</td>
<td>5</td>
<td>14</td>
</tr>
<tr>
<td>Shortening statute of limitations</td>
<td>Some</td>
<td>13</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Joint-and-severalliability reform</td>
<td>Some</td>
<td>12</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Collateral-source rule reform</td>
<td>Virtually none</td>
<td>9</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Screening panels</td>
<td>Virtually none</td>
<td>13</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Periodic payment</td>
<td>None</td>
<td>11</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Attorney-fee limits</td>
<td>None</td>
<td>12</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Expert witness standards</td>
<td>Not studied</td>
<td>17</td>
<td>5</td>
<td>7</td>
</tr>
</tbody>
</table>


\(^{169}\) For a full discussion, see Mello, supra note 158, at 22-23; and Mello & Studdert, supra note 156, at 20-30.


At the federal level, the tort-reform debate during the last malpractice crisis has been “much sound and fury, yielding no substantive policy change.” The House repeatedly passed the federal damages cap legislation, and the Senate repeatedly rejected it, until the Democrats took control of the Congress following the 2006 election and definitively killed the initiative. Overall, in assessing the results of the federal reform effort, Rogan Kersh has written that “the contrast between expectations and legislative outcome has never been starker in U.S. malpractice politics.” A sunny reading of this history is that a litany of single-minded lobbyists and second-rate analyses failed to sway the Congress. A realist view is that the advocates, and the legislators sympathetic to them, quickly came to irremediable gridlock.

What about judicial decision making? This scholarship certainly is relevant to questions of whether tort reforms reasonably advance a state interest, as required by equal-protection and due-process analysis. Yet, empirical legal studies are rarely discussed in these judicial

172 Data come from Avraham, supra note 48, at 8. Year-to-year variations represent both new adoptions of laws and existing laws being struck down by courts.
173 Kersh, supra note 144, at 49.
174 Id.
175 See Ronen Avraham, An Empirical Study of the Impact of Tort Reforms on Medical Malpractice Payments, (October 2006). Northwestern Law & Econ Research Paper No. 06-07 Available at: http://ssrn.com/abstract=912922 (finding little evidence of an association between reforms and medical malpractice insurance premiums and arguing that courts should consider the empirical evidence when evaluating the constitutionality of reforms). Most states require only a showing of a reasonable relationship, but a few apply intermediate scrutiny and require that tort reforms substantially advance the state’s interest. See Michelle M. Mello & Carly Kelly, Are Medical Malpractice Damages Caps Constitutional? An Overview of State Litigation, 33 J.L. MED. & ETHICS 515, 522-23 (2005).
opinions. For example, in rejecting an equal-protection challenge to Louisiana’s $500,000 damages cap in 1992, the Louisiana Supreme Court considered no empirical evidence of the cap’s effectiveness whatsoever. Applying rational-basis review, it simply concluded, without marshaling facts or argument, that the legislation “furthers the state’s purpose of compensating victims” by increasing the likelihood physicians would be able to maintain insurance. A dissenting justice, in contrast, criticized the majority for overlooking several empirical studies showing that caps had not had their intended effects.

An exception to the general trend is the Wisconsin Supreme Court’s 2005 decision in Ferdon v. Wisconsin Patients Compensation Fund holding that the state’s noneconomic damages cap violated equal protection. The court cited several descriptive studies of damages caps that suggested caps were unlikely to have much of an effect because only a small proportion of all awards were high enough to implicate the cap and also discussed limitations of other studies investigating the connection between caps and malpractice insurance premiums. It also cited descriptive studies concerning caps’ effects on physician migration. The court’s discussion of these studies is relatively sophisticated, compared to other judicial opinions, but that sophistication is undermined by the fact that in selecting studies for citation, it excluded so much wheat and included so much chaff. Scrutiny of the citations reveals the key selection criterion for studies: the court cited what it could read full-text on the internet or in Westlaw or LexisNexis. Other works in medical and health policy journals do not appear.

2. Enabling and Inhibiting Factors

When we look broadly at this history, several factors appear to have encouraged the production of empirical malpractice scholarship and held promise for its potential incorporation into policy making, but these factors have been overshadowed by a number of substantial barriers.

On the research side, a key catalyst for growth in the field was the development of closed-claim datasets, such as the National Practitioner Data Bank and state insurance department databases. This spurred a cottage industry in malpractice research among a cadre of well-known economists. The high visibility and policy salience of the issue may also have helped attract highly-qualified scholars to the field.

177 Id. at 527 (Dennis, J., dissenting).
178 Ferdon v. Wisconsin Patients Compensation Fund, 701 N.W.2d 440 (2005)
179 Id. at 469-71.
180 Id. at 485-86.
181 For example, the court relied heavily on two student papers and cited studies commissioned by interest groups, see id. at 448, 470, 473, 474, 483, 484, citing Kevin J. Gfell, Note, The Constitutional and Economic Implications of a National Cap on Non-Economic Damages in Medical Malpractice Actions, 27 IND. L. REV. 773 (2004); Elizabeth Stewart Poisson, Comment, Addressing the Impropriety of Statutory Caps on Pain and Suffering Awards in the Medical Liability System, 82 N.C. L. REV. 767 (2004); Health Insurance Association of America, Issue Brief: Why Do Health Insurance Premiums Rise (Sept. 2002), yet utterly dismissed a controlled, multivariate analysis by a well-regarded health economist on the basis that it included states that capped total damages as well as states that capped noneconomic damages only, see id. at 505, citing Kenneth E. Thorpe, The Medical Malpractice ‘Crisis’: Recent Trends and the Impact of State Tort Reforms, w4 HEALTH AFFAIRS 20 (2004). The inclusion of those four states may have affected the study findings, but the Thorpe study certainly still had more explanatory power than the other cited sources. Moreover, the court omitted any mention of a large number of other controlled studies, see Mello, supra note 158, at 24-25 (collecting studies).
Second, as with motor vehicle safety laws, malpractice reforms offer a large number of natural experiments that create ready opportunities for study. The long history of states’ reform efforts has generated substantial variation in the explanatory variable both across time and across states, facilitating powerful (if complicated) models.

Third, there has been funding available to support malpractice research. The Robert Wood Johnson Foundation, for example, has made a major financial commitment in this area. The fact that the topic has clear implications for core areas of health services research, such as the financing and quality of health care, has made it possible for researchers to compete for funding under general solicitations for health policy and health services research projects as well. The availability of extramural funding has made it possible for scholars in soft-money environments to enter the field and for researchers to carry out large data-collection efforts.

Less in evidence are enabling factors for the use of this research in malpractice reform efforts. Although academics have sometimes been invited to give testimony in state and federal hearings and brief legislative staff, they have not had the kind of access that lobbyists for stakeholder organizations have had.

One way to counter this problem is to make EHL research more accessible to the policy community. Perhaps the most important means to this end has been the dedication of research sponsors and organizations to generating and disseminating high-quality malpractice research for consumption by policymakers. The Robert Wood Johnson Foundation and AcademyHealth have been at the leading edge of these efforts, employing skilled communications staff to translate complex research findings into easily digestible policy briefs that are then aggressively disseminated to the policy community. RWJF’s Synthesis Project, for instance, produces “concise and thought-provoking briefs and reports that translate research findings on perennial health policy questions” by “synthesizing what is known, while weighing the strength of the research evidence and exposing gaps in current knowledge.” The Synthesis Project engaged one of us to synthesize the literature on tort reforms, and then created briefs and forums for the results to be disseminated, as well as making the products freely downloadable on the internet. Few other areas of empirical health law research have received so great a commitment to dissemination.

A number of challenges have made it difficult for EHL research to achieve its scientific and policy aims, however. Many of the richest datasets on claims exist at the insurer or state level and are not clearly generalizable or integrable with one another. All existing datasets that have supported secondary analyses have significant limitations—for example, jury verdicts exclude settlements, the Data Bank excludes unpaid claims, and state closed-claim databases omit unpaid claims and some kinds of health care providers, may have suboptimal data accuracy,
The best data remain proprietary: the records of liability insurers and the providers they insure. These data are difficult to gain access to, and cumbersome and expensive to extract. This, too, leads to datasets that are limited in geographic scope (it is not practicable to review thousands of litigation records in every state). When only a few states are represented, there are very limited possibilities for evaluating state tort reforms. For this reason, although studies based on insurers’ litigation records and/or medical chart review have made seminal contributions to describing the epidemiology of medical injury and malpractice claiming, they have not evaluated tort reforms.

There are also real methodological challenges to modeling the complex tort-reform environment. Controlling for other state characteristics which may affect the outcome variables can be difficult, and fixed-effects models are not a possibility for some periods when tort reforms were time invariant in nearly all states. There are also concerns about endogeneity, since tort reforms are most likely to be passed by states that are experiencing volatility in insurance premiums or claims costs, and questions about how to construct variables to represent a large and heterogeneous group of reforms.

The main barrier to the productive use of malpractice research in policy making has been that the involved political interest groups have not played a facilitative role. Throughout the past three malpractice crises, the policy debate as been dominated by two interest groups, trial lawyers and health care providers, and their allies. Their focus has been single-minded: are noneconomic damages caps a good idea or a bad idea? Although empirical legal scholars have repeatedly tried to shift policymakers’ focus to reforms that are more likely to address fundamental and well-documented problems with the medical liability system, few interest groups have followed their lead. Groups have instead honed in on the narrow, if important, problem of spiraling liability costs, and the reform that most directly addresses the size of malpractice awards.

These groups have been fiercely adversarial in their postures and have appeared to believe that maintaining those postures requires the utter rejection of empirical evidence that does not support their position. This has resulted in a highly polarized debate with none of the nuance that characterizes evidence-based policy making, and in which “each side portrays the other in very harsh language and portrays the other side’s position as disastrous for society.”

Rather than nurturing policymakers’ interest in data, these interest groups have tended to strongly emphasize anecdotes featuring sympathetic victims of malpractice and the malpractice

---

187 The best known such study is the Harvard Medical Practice Study, discussed supra note 56.
188 See Kersh, supra note 144, at 54-55 for a historical account of this adversarial policy debate.
189 William M. Sage, Malpractice Reform as a Health Policy Problem, in Medical Malpractice and the U.S. Health Care System 30-42, 31 (2006) (“Both long-standing antipathy between doctors and lawyers and the macro-politics of tort law tend to fixate the malpractice debate on lowest-common-denominator reforms of the legal process that are familiar to the medical profession and that translate readily to nonmedical litigation.”).
crisis. These narratives have strong appeal to legislators and journalists because they are easy to understand and have a human face.\textsuperscript{194} Statistical studies, in contrast, require a great deal of cognitive work and often lead to a grey, rather than black or white, view of the world.

Scholars have not had the incentives that interest groups have had to make participation in and influence over the policy debate a priority.\textsuperscript{195} While professors have struggled to find time and money to attend briefings and hearings, well-financed advocates have been working full time to press their agendas in those settings and through extensive behind-the-scenes lobbying.\textsuperscript{196} It is not surprising that it is their position papers that get read, and that these rarely include data from scholarly studies.

A second problem has been that over time, the evidentiary field has become quite crowded. The number of analyses by academics, government agencies, and political interest groups and their consultants has skyrocketed during the latest malpractice crises. There are now so many studies of tort reforms that it requires considerable effort to find and review them all, much less evaluate what is of good and poor quality. In other areas of health policy, communities of scientists have come together to rigorously catalog and analyze the literature on topics of great importance for health care providers and policymakers, but there are few such efforts in health law.\textsuperscript{197}

The problem of evaluation is exacerbated by the great complexity of many academic studies of tort reforms. Some legislators struggle with even basic statistical concepts,\textsuperscript{198} and as methodological innovations in econometrics have led to ever more sophisticated analyses, it has become increasingly difficult for non-experts to decipher and critically evaluate their findings. Adversaries in the policy debate use this to their advantage by attempting to discredit the studies,\textsuperscript{199} leading to further confusion.

\textsuperscript{194} Id. at [not yet available]. [EDITOR: look right before section headed “2. The Scholars”].
\textsuperscript{195} Haltom & Michael, supra note 170, at 101. Haltom and McCann also note that “their very claim to authority as professional experts has limited the scholars’ public visibility,” id. at 17, presumably because it requires heavy investment of time to maintain the necessary base of expertise and scholars may consider it unseemly to descend into political fights rather than focus on perfecting their craft.
\textsuperscript{196} Sebok, supra note 193, at [not yet available] (“The sociolegal scholars, as full-time academics, had neither the time nor the resources to promote themselves like the tort reformers at the Manhattan Institute”), citing Haltom & McCann, supra note 170, at 101.
\textsuperscript{197} The Cochrane and Campbell Collaborations conduct systematic reviews of the evidence on matters of clinical care and select areas of social policy (including a few public health interventions), respectively. See Cochrane Collaboration, Cochrane Reviews and the Cochrane Library – an Introduction, http://www.cochrane.org/reviews/clibintro.htm; Campbell Collaboration, About the Campbell Collaboration, http://www.campbellcollaboration.org/About.asp. Although systematic reviews potentially are useful for policymakers in deciding which health care services should be funded by government programs, this method is not commonly used to study the impacts of specific legal rules on outcomes.
\textsuperscript{198} An example is a member of Congress’s attempt to cross-examine a witness at a committee hearing using data from the National Practitioner Data Bank. The witness, an insurance association representative, had offered insurer data indicating that the mean malpractice payout was $496,000. Rep. Bart Stupak asserted that that figure did not “jive” with Data Bank data indicating that the median payout was $125,000, and appeared not to understand the witness’s explanation that the mean and median are two different statistics. See U.S. Representative Michael Bilirakis (R-FL) Holds Hearing on Medical Liability Issues, FDCH Political Transcripts, Feb. 27, 2003.
\textsuperscript{199} See, e.g., Don Finley, Study Says Medical Malpractice Crisis Is a Myth, SAN ANTONIO EXPRESS-NEWS, Mar. 11, 2005, at 3B (quoting Dr. Bohn D. Allen, president of the Texas Medical Association, as saying “The authors of this study used data that was inflation-adjusted, population-adjusted, and health-care-cost adjusted. It was so adjusted that it became truth-adjusted.”). The study to which Allen referred reported time trends in medical malpractice filing rates and payouts size. The authors wished to examine whether out-of-control juries were behind increases in medical liability insurance premiums; therefore, they adjusted their estimates to eliminate obvious alternative
Finally, the way in which malpractice has been framed as a policy issue has inhibited deliberative lawmaking. Framing the problem of rising liability insurance costs as a “crisis,” however defensible from the perspective of those who have to pay those costs, has driven policymakers toward a narrow goal—stabilizing premiums—and a narrow solution—damages caps. The atmosphere of emergency has made it difficult to engage policymakers in deliberations about persistent, fundamental problems with the malpractice system and far-reaching reforms that could address those problems along with the problem of insurance costs. Rather, policymakers have been attracted to seemingly simple fixes that can satisfy powerful interest groups’ demands to do something. It is one of the great paradoxes of malpractice politics that the conditions that maximize the chances that legislators can mobilize enough support to pass legislation also minimize the chances that the legislation will reflect careful, thorough consideration of what the best evidence suggests will be a successful reform. Together, these factors explain why empirical research on malpractice reforms has not lived up to its promise for shaping policy.

C. Summary

Viewed together, the two case studies help to identify a common set of factors that affect the likelihood EHL research will influence policy decisions. In general, EHL research is more plentiful and impactful when (1) there is a considerable amount of variation in the health laws in question across jurisdictions and over time, which strengthens researchers’ ability to statistically estimate the effects of the laws; (2) data on laws, their outcomes, and other factors that might confound analysis of the relationship between the laws and the outcomes are readily available; (3) the issue is convincingly framed as a public health issue and visible enough to attract skilled empirical researchers and funders of both primary studies and systematic reviews; (4) infrastructures exist within the academy to support the production and dissemination of empirical research; (5) interest groups act as facilitators rather than barriers to basing policy decisions on the best available evidence; and (6) policymakers understand and have confidence in the methods used to produce relevant empirical findings.

When these stars do not align, empirical research, if any exists, has less impact on policy decisions. Although some of the factors surely are out of the control of the legal academy, steps can be taken inside the academy to increase the likelihood that the enabling conditions are satisfied. In the following Part we offer specific suggestions for ways in which the legal academy can both increase the production of high-quality EHL studies and increase the impact of these studies on the implementation of health law and policy.

IV. Future Directions

The case studies examined in Part III serve as a useful jumping-off point to consider more generally the barriers that impede the production of EHL research and limit its impact on explanations, including increases in damages due to rising health care costs and general inflation and increases in the quantity of medical services provided due to population growth. See Bernard Black et al., Stability, Not Crisis: Medical Malpractice Claim Outcomes in Texas, 1988-2002, 2 J. EMPIRICAL LEG. STUD. 207 (2005).

See Kersh, supra note 144, at 50 (discussing the influence of framing on the politics of malpractice reform).


health law and policy. While many impediments apply to EHL research conducted both within and outside the legal academy, several are unique to legal academia. Pushing the field forward will require that these barriers be addressed. We offer a few suggestions along these lines, some of which emanate from the case studies we have analyzed.

A. Ways to Increase Production of High-Quality EHL Studies

Although production of EHL studies has increased substantially over the past three decades and Part III presents two cases in which EHL research was plentiful, a broader view of the field suggests that much more research is needed to develop a fuller picture of how health law and policy influence outcomes. Several reasons account for the gaps in knowledge and the fact that a relatively small portion of this literature is produced by scholars within the legal academy. In particular, production of EHL studies is hindered by a lack of easy and inexpensive access to reliable data, institutional barriers, and a general lack of acceptance of the scientific approach within law schools. These inhibiting factors present challenges, but interested parties within the legal academy can take specific steps to break down these barriers and increase the number of high-quality EHL studies produced by legal scholars. In particular, they can work to attract more funding to obtain data, encourage collaborations between legal scholars and methodologists, and educate law students and faculty on empirical methods and the merits of employing a scientific approach.

1. Increase Access to and Collection of Reliable Data

As we discussed in Part II.B., a major factor inhibiting the production of EHL research is the difficulty of accessing reliable data.  

Although EHL studies often employ data that directly measure the variables of interest, empirical findings are often based on proxies—variables the researchers hope are positively correlated with the true variable of interest—because data on the variables of interest are not available or are too expensive to collect. In the most troubling cases, connections between legal rules and outcomes go unexplored because data allowing either direct measurements or proxy-based measurements are unavailable.

For EHL research to evolve, availability of data and data-collection methods need to evolve. We offer two suggestions. First, EHL scholars should more aggressively seek out funding to collect or obtain data not currently available.  

Research sponsors will be more likely to view EHL study proposals favorably if these proposals are not anomalies among the submissions they receive, and if the proposals make a strong case that such research is important and can make a difference to public health. As we increase our production of impactful research, it will become easier to demonstrate this.

Second, scholars should also encourage government agencies that collect data for regulatory purposes, as well as private organizations in the health care market, to make them

---


204 See id. at 1360 (“Many academics are strikingly clueless about how to market their scholarship to funders and the media. Writing effective grant proposals and press releases is an acquired skill—and one not taught in law schools. It should be taught, at least to faculty.”).
more available and useful to researchers. These entities should move toward collecting data with the same contemplation of their possible research uses as occurs in other fields, such as labor and health care. We need to move from a world where researchers peck around at datasets collected for other purposes to a world where data are collected systematically with research in mind. We also need to better address agencies’ and companies’ concerns about data confidentiality and security. Methods widely in use in other areas of health research, such as development of limited, deidentified (or partially deidentified) datasets for public use and execution of data-use agreements could be applied here.

2. Break Down Institutional Barriers

Legal scholarship traditionally has been produced under a model emphasizing theoretical analysis and solo authorship. Although it has begun to expand outside this model, general adherence to and preference for the traditional approach has created institutional barriers within the legal academy to the production of empirical legal research. Because health law enjoys little prestige among legal scholars, EHL studies is especially marginalized, lying at the intersection of two nontraditional areas of scholarship.

One challenge is overcoming the traditional model’s low esteem for collaborative research. A large portion of the most compelling EHL research has been produced though collaborations between legal scholars and scholars from other fields who contribute disciplinary grounding and methodological expertise. Collaborations are more likely to form when academic environments create incentives for faculty to seek them out (for example, in “soft money” settings, collaborations are essential because strong teams have a competitive advantage in grant seeking). At present, these incentives do not exist in law schools, and the emphasis on sole-authored work in the promotion process is a strong disincentive to engage in collaborative work. There are signs of change in some law schools—for example, establishing centers for health law research with faculty affiliates from around the university, sponsoring interdisciplinary workshops and conferences and providing grants for collaborative research. One recent example is the involvement of academics in the debate over the National Association of Insurance Commissioners’ model law requiring disclosure by medical malpractice insurers of closed-claim data. See http://www.naic.org/committees_c_sitf.htm.

See Heise, supra note 9, at 812 (“[T]he amount of theoretical and doctrinal scholarship that presently fills many law review pages overwhelms the amount of empirical scholarship.”); Tracey E. George & Chris Guthrie, Joining Forces: The Role of Collaboration in the Development of Legal Thought, 52 J. LEG. EDUC. 559 (2002) (“[M]ost law review articles seem to be written by one author.”)

Thomas S. Ulen, The Unexpected Guest: Law and Economics, Law and Other Cognate Disciplines, and the Future of Legal Scholarship, 79 CHI.-KENT L. REV. 403, 413 (2004) (“[T]here seems to be a widely held belief in the legal academy that the scientific method is inappropriate for the study of law.”).


George & Guthrie, supra note 206 at 571 (“Collectively, the comprehensive citation-count studies demonstrate that articles by coauthors have had less impact than we would have expected given the rate at which coauthored articles have been published.”)

See id. at 560 (“[T]here are practical impediments to collaboration in law schools that may not exist in law practice. For example, law schools often give faculty only token credit for coauthored works.”).

Harvard Law School’s Petrie-Flom Center for Health Law Policy, Biotechnology and Bioethics and Georgetown Law Center’s O’Neill Institute for National and Global Health Law are two recently-established examples.

The Petrie-Flom Center has sponsored a number of such conferences and runs an annual semester-long health law and policy workshop.
We would submit, however, that the balance of incentives will not tip until coauthored work comes to enjoy greater legitimacy. Various factors account for this disparity, but chief among them are a lack of familiarity with empirical methods, a lack of appreciation for the ways in which observation and measurement can enrich (rather than replace) theory, and an understanding of the concept of intellectual creativity that focuses on the development of innovative ideas rather than discoveries about the world. Closing this gap requires educating legal scholars in empirical methods, clarifying the role of empirical research in theory development, demonstrating the impact of empirical research on policy and human welfare, and nurturing the idea that such real-world impact is an important measure of the excellence of scholarship. As it becomes clearer that the usefulness of theory hinges, in part, on whether a theory’s predictions are consistent with observations from the field or the laboratory, legal scholars will come to appreciate more fully the value of empirical work. As more law schools offer courses in empirical research methods, there will be a larger cadre of scholars able to conduct it.

As familiarity with and appreciation for empirical work grows, the preferred model of legal scholarship will continue to evolve. Furthermore, as empirical and collaborative research becomes more common, legal scholars will gain experience evaluating the work for purposes of hiring and promotion. This will further break down the barriers that discourage legal scholars from engaging in this type of research.

Of course, the proposals set out here are not meant to imply that empirical studies should replace non-empirical research or that resources should be shifted away from non-empirical research. Clearly, empirical methods are not well suited for all important investigations. Empirical research depends in many cases on theoretical and doctrinal research for grounding. Although a large portion of EHL research is purely description, much of it is performed to test theoretical predictions and assumptions about causal connections on which theorists and policymakers commonly rely. Our intent is to argue for an increase in the production of high-

---

213 For example, Georgetown Law Center provides research grants for collaborative research to cover costs of travel and other expenses.
214 Others have suggested collaborations as one way to help mitigate the difficulties in meshing legal scholarship and social science. See, e.g., Jeremy A. Blumenthal, Law and Social Science in the Twenty-First Century, 12 S. CAL. INTERDISC. L.J. 1, 37-38.
215 Robert W. Gordon, Lawyers, Scholars and the “Middle Ground,” 91 Mich. L. Rev. 2075, 2085 (“[empirical legal research] remains to this day the most neglected and ridiculously undervalued as well as the most potentially fruitful branch of legal studies.”); Craig Allen Nard, Empirical Legal Scholarship: Reestablishing a Dialogue between the Academy and the Profession, 30 Wake Forest L. Rev. 347 (offering reasons why empirical legal research has taken a back seat to other forms of legal scholarship).
217 See Nard, supra note 215, at 363 (finding that “twenty-five percent of non-tenured professors were concerned that empirical research would not be viewed favorably by their respective tenure committees.”); Ulen, supra note 207, at 418 (“Junior law faculty may perceive making a commitment to empirical techniques as a high-risk, and, therefore, unattractive strategy for promotion….“).
quality empirical research while at the same time preserving the important role of and respect for theoretical and doctrinal analysis.

B. Ways to Increase the Impact of EHL Studies on Health Law and Policy

Increasing production of EHL studies is but a first step. These studies create minimal value if they have little or no impact on the structure of law and the shape of policy. As the case studies presented earlier demonstrate, high-quality EHL studies do not always impact law and policy to the extent they should. Several reasons account for this, including lack of rigor, failure to aggressively disseminate new findings, and difficulties policymakers have in reading, interpreting and critiquing empirical studies. As with the production barriers, interested parties within the academy can take steps to mitigate or eliminate these barriers.

1. Strengthen Rigor

Although much EHL research can be classified as rigorous according to scientific standards, this is not uniformly the case. As EHL scholarship produced by legal scholars continues to converge toward social-science research with respect to the methodology employed and the presentation of results, publishers and consumers of EHL research will demand that it meet the standards of scientific rigor to which other kinds of social-science research are held.219 The legal academy can help strengthen the rigor of EHL studies by fine-turning hiring and tenure standards and making greater use of peer-review processes.

Law schools traditionally have not been in the business of hiring empiricists; therefore, law faculties have relatively little experience evaluating the quality of empirical studies produced by job candidates. Promotion committees and processes also do not make much effort to understand the dynamics of collaborative research, and thus may misappre hend or underappreciate the unique intellectual contributions that faculty have made in collaborative projects. Even such basic matters as how the order of authors listed on articles in different disciplines reflects the authors’ contributions may be mysterious. Constituting hiring and promotion committees such that candidates know that good-quality research and important contributions will be recognized and rewarded, and sloppy studies and free riding in collaborations duly noted, in their decisions would do a great deal to spur the kind of EHL research we want and need. Involving reviewers from other departments within the university would be a low-cost way to bring in additional expertise. Hiring committees can also encourage the growth of EHL research by favoring candidates who have formal methods training or a demonstrated record of engaging in collaborations with methodologists that produce high-quality, impactful EHL studies.

Consideration should also be given to how to more fully reap the advantages of peer-review processes to improve empirical legal scholarship. The formal standards (for example, criteria for authorship) and informal norms (for example, minimum response rates for surveys) of peer-reviewed journals in other fields may be appropriate to apply to some kinds of EHL studies. The most rigorous journals require authors to submit data and logs of procedures used to derive

219 Although this section contains suggestions for interventions to increase the rigor of empirical health law research, some have argued that the marketplace of ideas will naturally separate the methodologically sound studies from those reporting seriously flawed findings. See, e.g., McGinnis, supra note 8, at __.
reported results. After the paper is published, some journals post the data and result logs on their websites. This makes replication and further investigation easier for other researchers.

EHL scholars should also be encouraged to submit their work to peer-reviewed journals in both the legal and health fields and serve as referees for these journals. Undergoing the peer-review process, though time consuming and often burdensome for authors, can markedly improve the quality of a research paper, as well as facilitating the match between the quality of the research and the prestige of the journal in which it appears. This match is important because journal reputation has a signaling effect for consumers of research, indicating the quality and reliability of the research published there. We readily concede that peer-review processes and their results are far from perfect, but submit that they are more reliable than the methods by which most student-run law journals screen articles for publication.

Law reviews and journals can also take steps to ensure the rigor and reliability of results they publish. They can request reviews by faculty with empirical-methods training for portions of submitted articles that student editors are not qualified to evaluate. The journals also can require submission of data and results logs or encourage authors to post this information on their own websites.

2. Increasing Wide-Spread and Effective Dissemination of New Findings

Increasing the quality of EHL studies is the first step toward increasing their impact on policy. Once produced, however, the studies will have little influence on policy unless the results end up in the hands of policymakers. The first obvious step to disseminate findings is publication. As the New York Times recently reported, however, legislators and judges might not read scholarly journals as often as we might hope. Therefore, we need to take additional steps to increase our impact on policy debates.

Medical and public health schools have established dedicated personnel in offices of communications to disseminate faculty research findings. For example, faculty members at the Harvard School of Public Health have a press officer assigned to them who becomes familiar with the research...
with their portfolio of work. When a new paper is about to be published, faculty can send the galleys to their officer and discuss the message the authors would like to convey. If appropriate, the press officer then drafts a press release, contacts reporters who might be interested in the work, and arranges interviews. He also disseminates the research on the School’s website and through podcasts and listservs. These arrangements are expensive, but highly effective, because they utilize media professionals and minimize demands on faculty members’ time. Building such structures in law schools would do a great deal to raise the public profile of EHL research.

Widespread dissemination is necessary for EHL researchers to impact policy, but may not be sufficient. A second measure that our case studies clearly suggest would be helpful is for scholars to engage in efforts to make their empirical research more comprehensible to policy makers and judges. Legal scholars now have more to offer to policymakers than mere learned opinion, but scientific results can be difficult to explain and understand, and can be challenged on many different bases. By describing results in ways law audiences can understand and providing clear warnings regarding limitations, researchers can help policymakers import the studies into debates in appropriate, useful and defensible ways. They might also consider creating “policy maker’s digests”—short issue briefs accompanying research findings relevant to active policy debates. In health research, study sponsors routinely produce such briefs and disseminate them to policy makers, making liberal use of electronic newsletters as well. Finally, participation in legislative briefings and hearings, though time consuming and subject to receiving an invitation, can be an important means of clarifying the message of one’s research, and the motor vehicle safety law case study suggests that it can be effective in shaping policy. Similarly, scholars can help ensure that study findings are not “spun” too much for political purposes by responding to such misuse through op-eds and other outlets.

The story of medical malpractice reform also indicates that for some policy issues, it is critical for scholars to help policy makers sort through a large body of conflicting empirical analyses of varied quality, fit individual studies into the larger policy picture, and understand what policy prescriptions emerge from the empirical findings. Otherwise there is a danger that they will not rely on the best available evidence or will draw faulty implications from the results of narrow studies. One way to assist policymakers is to publish literature reviews with an emphasis on determining what conclusions can be drawn from available body of evidence and identifying areas where there is insufficient data to draw reliable conclusions.

In addition, systematic reviews modeled on structured methods would be invaluable in helping policymakers draw conclusions from bodies of EHL research. Such reviews have been

229 Id. at 538. Espstein et al., On the Effective Communication of the Results of Empirical Studies, Part I, 59 VAND. L. REV. 1811 (2006). To reach judges, further measures may be required, such as greater representation of empirical scholars in amicus briefs.
230 See, for example, the collection of Synthesis Project briefs at http://www.rwjf.org/pr/synthesis/index.html.
232 See, e.g., Mello, supra note 158; Randall R. Bovbjerg & Robert A. Berenson, Surmounting Myths and Mindsets in Medical Malpractice, URBAN INSTITUTE (Oct. 2005), http://www.urban.org/UploadedPDF/411227_medical_malpractice.pdf. Because many courts look only to law reviews, EHL scholars should also aim to publish literature summaries in law reviews. As courts gain the capacity to search additional databases, publication venue will become less of a concern.
powerful and useful tools for advancing clinical and public health practice. The reviews produced by the Cochrane Collaboration and Campbell Collaboration, for example, rigorously evaluate the available evidence regarding a clinical issue or social program, respectively. Their rigor is such that they are widely recognized as the definitive statements of the evidence on a particular issue. Similarly, the Centers for Disease Control and Prevention has convened an independent Task Force on Community Preventative Services to produce systematic reviews of primary research studying the effectiveness, economic efficiency, and feasibility of interventions to promote community health and disease prevention. Among the interventions evaluated to date have been public health laws and policies aimed at reducing tobacco use, encouraging physical activity, and reducing firearm injuries. Such evaluations could be much more widely applied to health laws.

A second way researchers can help is to explicitly bridge individual study results into the policy debates. Primary EHL studies usually are insufficiently broad to generate specific policy recommendations because they tend to focus on narrow components of multi-faceted issues. By including a discussion of the implications individual results have for the larger policy debate, researchers can help policymakers understand the power and the limitations of narrow results.

How can law schools facilitate these initiatives? Again, tweaking the incentive structure for faculty is key. Hiring and tenure standards can be adjusted to encourage empirical scholars to take active steps to inject their research into policy debates. Current standards emphasize a candidate’s publication record and number of citations to the candidate’s work by other legal scholars over evidence that a candidate’s scholarship has been read, cited, and relied upon by policy makers and judges. Readiness for promotion should be evaluated not only by reference to how the individual’s work has advanced her field of scholarship, but also how it has impacted the world outside the legal academy. Adding this standard to the criteria for promotion and tenure decisions will create incentives for empirical scholars both to increase the rigor of their work and to disseminate their findings widely.

3. Educating Consumers of EHL Research

Finally, in addition to strengthening the rigor of EHL research and improving dissemination and communication, we can increase the impact of our research by helping to educate potential audiences for EHL about how to read, interpret and evaluate empirical studies. Some in the legal academy have been engaged in this endeavor for many years, but as the
production of empirical studies increases, we likely will see an increase in the demand for education on the part of legislators, legislative staff, judges and clerks, and law faculty.

Because lawyers dominate all of these groups of constituents for EHL research, one approach is to include this type of training as part of the standard legal education curriculum. While many law schools offer quantitative-methods courses, they are not universally available, much less required. Those that do not offer such courses should consider hiring faculty or adjunct staff to teach them. Short courses, similar to those offered to train judges in scientific research methods so that they can serve as quality gatekeepers at trial, can also be an effective way to give jurists a basic familiarity with empirical research and how to find it.

The politics of health policy decisions will continue to create pressures for lawmakers to focus on the desires of special interest groups and their lobbyists even when sound empirical evidence counsels that policy should move in a different direction. For some issues and some policy makers, increases in the funding, volume, visibility, quality, comprehensibility, and accessibility of empirical research will be insufficient to change this focus. But addressing these conditions will reduce the time costs and cognitive barriers, if not the political barriers, to greater use of evidence in policy making. We believe that many in the policy community will respond positively.

C. Summary

Our proposals, in some ways, amount to no less than a call for a culture shift within the legal academy. We suggest that legal scholars transform themselves from expert-advocates into scientists, who approach inquires through hypotheses, grounded in theory or commonly-held assumptions, rather than a particular position; who give equal weight to all available evidence, whether favorable or unfavorable to the hypotheses; who report all empirical results, regardless of whether they confirm or call into question their initial position; and who affirmatively disclose the shortcomings of their work. We further suggest that legal scholars should acquire additional expertise (either personally or through collaborations) to accomplish this objective, even though their investment will not be rewarded (at least initially) in the usual processes through which the legal academy recognizes scholarly achievements. Despite the challenges rooted in the different and sometimes conflicting methodologies of legal scholars and social scientists, we can do much to bridge the gap and increase the impact of empirical legal studies on policy.


239 Eisenberg, supra note 216, at 1741-46.
240 See Blumenthal, supra note 214, at 47, summarizing Craig Haney, Psychology and Legal Change: On the Limits of a Factual Jurisprudence, 4 LAW & HUM. BEHAV. 147, 159-68 (1980) (“(1) social science is innovative, while law resists innovation, (2) social science is based on data and observation, while law is based on precedent and hierarchy, (3) social science seeks an objective answer to problems, while law seeks an adversarial victory, (4) social science is descriptive, while law is prescriptive, (5) social science is nomothetic, while law is idiographic, (6) social science conclusions are probabilistic and tentative, while legal conclusions are irrevocable and must appear certain, (7) social science is proactive, while law is reactive, and (8) social science is abstract, while law deals with concrete issues.”)
Clearly, this is not for everyone, or even every law school. And indeed, it is critical to preserve what is distinctive and valuable about traditional modes of legal scholarship—for example, the depth of study and exposition, the consideration of political and moral dimensions of policy problems, and the normative recommendations that point judges and policy makers in a clear direction. Our claim is simply that empirical legal research should be afforded the same respect and nurtured to the same extent as more traditional forms of inquiry. This involves recognizing the unique challenges that empirical legal studies face, as well as the special contributions it has made and can make toward betterment of public policy.

V. CONCLUSION

The field of empirical health law is in its infancy, but it is growing quickly. Several forces have aligned to support its continued expansion. Unless law schools undergo substantial changes, however, this expansion will continue mostly outside the legal academy with limited involvement of legal scholars, who have the potential to bring a deep understanding of law and legal institutions to the enterprise. Empirical health law research stands to benefit greatly from changes within the legal academy to foster an increase not only in the production of this research but also in the impact it has on enactment of effective legislation and dispute resolution. Despite the challenges we have outlined, we remain optimistic that the field of empirical health law will continue to grow and mature, and that the great public health achievements of the twenty-first century, like many of those of the last century, will owe much to its contributions.