Statistical Learning Can Help the Judiciary Fulfill Its Gatekeeping Role Over Expert Witnesses

Andrew Baker[1]

ABSTRACT

This article examines the historical challenges of expert testimony in the American legal system and proposes a forward-looking reform grounded in modern statistical learning techniques. Tracing the evolution from court-appointed experts to partisan witnesses, the paper highlights how adversarial practices and scientific complexity have strained judicial gatekeeping, particularly under the *Daubert* standard of judicial review of expert testimony. The paper argues that shifting from traditional model-driven estimation methods to data-driven, algorithmic approaches can improve the reliability and transparency of expert evidence. Through empirical examples in securities litigation and corporate valuation, it demonstrates how statistical learning methods can reduce expert discretion and aid judicial decision-making. The proposed reforms offer a practical pathway for courts to enhance the quality and fairness of expert testimony in modern litigation.

> Indeed, it is difficult to conceive of language within the bounds of decent and temperate criticism, which ought to be regarded as excessively severe in commenting upon the expert testimony nuisance as it has, of late years, been infesting our courts. In the way of wasting the public time, in the way of burdening litigants with expense, and in the way of beclouding the real issues to be tried and effecting miscarriages of justice, it has grown to the proportions of an offensive scandal. Instead of being an aid in the administration of the law, it has become a positive hindrance to it. Instead of assisting in the approximation of the truth, it has become the means of obscuring it.
>
> Judge Gustav Endlich, 1896

DOI: https://doi.org/10.15779/Z384J0B041

1. University of California, Berkeley School of Law. I thank attendees of the 2024 Institute for Law and Economic Policy Symposium for helpful feedback, in particular my discussants Katherine Lubin Benson, Caley Petrucci, and Christin Hill. I also thank the members of the Berkeley Business Law Journal for their helpful assistance in revising this draft. A special thanks to Joe Ross, Tiana Wang, and Elise Wurtman for their generosity.

TABLE OF CONTENTS

1. INTRODUCTION

Expert testimony has long been a cornerstone of adjudication in complex legal disputes, serving as a bridge between complex topics and judicial decision-making. However, courts have struggled with challenges in overseeing such testimony since its inception, ranging from questions about the reliability of methodologies employed to the potential for partisanship in expert opinions. This article offers a historical perspective on the evolution of expert testimony and a forward-looking proposal for reform rooted in statistical learning and algorithmic modeling. Building upon case law and interdisciplinary insights, I provide one potential improvement in ensuring the judiciary's gatekeeping function and improving the quality of expert evidence put before the court.

This article opens by tracing the historical evolution of expert testimony, emphasizing its deep roots in common law courts. From the 14th-century testimonies of surgeons determining "mayhem" to the landmark 1782 decision in *Folkes v. Chadd*, courts have long relied on specialized knowledge to inform their decisions. However, the rise of adversarialism in the 18th and 19th centuries transformed the role of experts, shifting from court-appointed neutral advisors to partisan witnesses employed by litigants. This shift introduced credibility of scientific testimony in the eyes of the judiciary and the public.

In contemporary commercial litigation, expert testimony has grown not only in prevalence but also in complexity. This article highlights how experts are now pivotal in high-stakes disputes involving sophisticated financial instruments, global commerce, and advanced technologies. Fields such as antitrust, securities litigation, and employment discrimination increasingly rely on expert analyses to bridge the gap between legal principles and technical realities. However, this reliance has amplified concerns about the reliability of expert evidence, particularly because methodologies have become more intricate and less transparent to lay judges and juries.

A critical turning point in the judiciary's engagement with expert testimony came with the adoption of the *Daubert* standard in 1993, which superseded the earlier *Frye* standard. By emphasizing scientific validity, testability, peer review, and error rates, *Daubert* established a more rigorous framework for admitting expert evidence. However, this framework has also placed a significant burden on judges, who must navigate increasingly complex scientific and technical matters to fulfill their gatekeeping responsibilities. This article concludes by proposing one concrete reform to enhance the provision and evaluation of expert testimony. Central to this proposal is the integration of statistical learning and algorithmic modeling—approaches that prioritize predictive accuracy and minimize subjective discretion in model selection. By shifting the focus from traditional parametric models, which rely heavily on *a priori* assumptions, to more data-driven methodologies, this article argues for a more objective and reliable framework for expert analyses.

In practical terms, this article illustrates how these data-driven methodologies can be applied to specific legal contexts, such as securities litigation. For instance, it demonstrates how penalized regression models can improve the accuracy of event studies by systematically selecting industry peers based on predictive performance rather than subjective judgment. This approach not only enhances the credibility of expert testimony but also provides judges with a more transparent and administrable tool for evaluating complex evidence.

2. Expert Work and Commercial Litigation

Expert testimony is a critical tool for courts in clarifying complex issues that arise in disputes between businesses. Expert witnesses provide specialized knowledge that helps the court understand intricate technical issues, industry standards, or specific data that are beyond the common knowledge of judges and juries.[2] Given their importance to the disposition of civil suits, studies demonstrate that experts are consistently present in the majority of litigated cases.[3] Even twenty years ago, famed district court judge Jack Weinstein noted

2. *See* Jack B. Weinstein, *Improving Expert Testimony*, 20 U. RICH. L. REV. 473, 473-75 (1986).

3. *See, e.g.*, Andrew W. Jurs, *Expert Prevalence, Persuasion and Price: What Trial Participants Really Think About Experts*, 91 IND. L. J. 353, 367–69 (2016) ("Forty-two of thirty-six (86%) civil jury trials in Polk County, Iowa in 2012 contained at least one expert witness endorsement."); Anthony Champagne, Daniel Shurman & Elizabeth Whitaker, *An Empirical Examination of the Use of Expert Witnesses in American Courts*, 31 JURIMETRICS J. 375, 380 (1991) (63% of civil trials examined in Dallas, Texas, in 1988 included expert testimony.); Samuel Gross, *Expert Evidence*, 1991 WISC. L. REV. 1113, 1119 (1991) (86% of the 529 cases reported in *Jury Verdicts Weekly* between 1985 and 1986 involved

how "[t]he law's use of expert witnesses ha[d] expanded at a pace reflective of society's reliance on specialized knowledge", and that "[h]ardly a case of importance is tried today in the federal courts without the involvement of a number of expert witnesses."[4] Given the role played by expert witnesses in resolving cases, lawyers now recognize that "experts can make or break a case."[5]

Though often viewed as a modern problem, "scientific expert testimony in common law courts has a long and rich history . . . the putative problems of scientific expert testimony in common law courts have existed since science was first introduced into the adversarial courtroom."[6] In the 14th century, surgeons testified in common law courts about whether a wound amounted to "mayhem."[7] By the 16th century, courts understood the necessity of bringing in scientific advice where they lacked the required knowledge or expertise to settle disputed facts.[8] Originally, experts were not distinguished from other lay witnesses, who were often allowed to testify as to their opinions based on direct knowledge of the facts at issue in the dispute.[9] This gradually changed as part of a larger transformation in the English common law system known as the "Adversarial Revolution."[10]

This "revolution" in legal practice has historically been associated with an increase in the presence of lawyers in criminal proceedings.[11] Before the 18th century, judges controlled criminal proceedings, directly examining the parties and witnesses without the presence of legal representatives. Defense counsel began to appear in regular criminal proceedings by the 1730s, perhaps in response to an expansion in criminal prosecutions by the Crown.[12] Previously, courts had summoned and controlled experts, but as courts adopted a more neutral position, and as the litigants assumed the responsibility for their arguments, parties started hiring their own experts. With the rise of this "partisan" provision of expert testimony, courts gradually began to grapple with the issue of ensuring reliable expert guidance when the jury needed it.[13]

---

expert testimony.); Shari Seidman Diamond, *How Jurors Deal With Expert Testimony and How Judges Can Help*, 16 J. L. & POL'Y 47, 56 (2007) (also finding that 86% of the cases in her sample — civil trials in Arizona that were videotaped as part of a study on jury behavior — included expert testimony.).

4. Weinstein, *supra* note 2, at 473.

5. Michelle Garcia & Nichole C. Patton, *Experts and Opinions: The Pitfalls and Possibilities of Expert Witness Testimony*, 24 PASS IT ON 1 (Fall 2014) (https://www.americanbar.org/content/dam/aba/publications/pass_it_on/experts_opinions_witness_testimony_PIO_F14.pdf).

6. Tai Golan, *Revisiting the History of Scientific Expert Testimony*, 73 BROOK. L. REV 879, 936 (2008).

7. *See generally* 9 W. S. HOLDSWORTH, A HISTORY OF ENGLISH LAW 212 (1926).

8. *See* Buckley v. Rice Thomas, 1 Plowden 118, 124, 75 Eng. Rep. 182, 192 (1554) (Saunders, J.)

9. *See* 4 JOHN HENRY WIGMORE, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW, at 101–03 (2d ed. 1923).

10. Golan, *supra* note 6, at 882.

11. *Id.*

12. *Id.* at 882–83.

13. *Id.* at 885.

Concomitant with the rise of legal adversarialism was a growth in the "culture of science" and the social significance of its practitioners, who called themselves "Newtonian philosophers," reasoning from first principles rather than through specific training or experience.[14] The struggle to deal with scientific evidence and partisan witnesses culminated in *Folkes v. Chadd*, [15] a 1782 civil dispute over the cause of harbor decay on the Norfolk coast of England. Lord Mansfield's opinion in the case has been called "the foundation of the rules governing expert evidence,"[16] clarifying the status of evidence adduced from "those skilled in matters of science, who, though they personally knew nothing about the circumstances of a particular case, might yet, perhaps by way of exception, give their opinion on the matter."[17] In his ruling, Lord Mansfield accepted the testimony of John Smeaton, a civil engineer who was considered the utmost authority on harbors in the kingdom at the time, over the objection that he was testifying as to his scientific opinion rather than personal knowledge of the harbor. Lord Mansfield thus recognized the importance of "a new class of witnesses, skilled in matters of science, who could give opinions that were not based directly on the traditional trustworthiness of the senses."[18]

The role of the partisan scientific expert, established formally in *Folks v. Chadd*, became increasingly central to English common law during the expansion of science and technology into industry and other institutions. During the early years of the 19th century, an increasing cast of scientists, including chemists, geologists, and engineers, began appearing in courtrooms. These experts were hired to explain the underlying science behind nascent industries, from mining to insurance, energy, and toxicology.[19] However, the combination of the rise in adversarialism with the advent of the scientific expert witness generated novel difficulties, both for the court and for the scientists. It led to the now-common experience of leading experts aggressively contradicting each other on the witness stand—a habit that gradually called into question the integrity of science and its practitioners in the eyes of the legal profession and the public.[20] As a harbinger of future frustration, courts became increasingly

14. *Id.* at 886.

15. 3 Doug. 157, 99 Eng. Rep. 589 (1782).

16. Anthony Kenny, *The Expert in Court*, 99 LAW Q. REV. 197, 199 (1983). *See also* James Bradley Thayer, *A Selection of Cases on Evidence at the Common Law* 666 (1892) (arguing that the case created the practice of calling experts as partisan witnesses before juries); Stephan Landsman, *Of Witches, Madmen, and Products Liability: A Historical Survey of the Use of Expert Testimony*, 13 BEHAV. SCI. & LAW. 131, 141 (1995) (contending that *Folkes* represented courts' "seal of approval on the whole adversarial apparatus including contending experts"); Tal Golan, however, persuasively argues that these claims are oversold—courts had already begun calling experts as partisan witnesses before juries well-before *Folkes*. Golan, *supra* note 6, at 898.

17. 4 JOHN HENRY WIGMORE, A TREATISE ON THE ANGLO-AMERICAN SYSTEM OF EVIDENCE IN TRIALS AT COMMON LAW, at 666-7 (1892).

18. Golan, *supra* note 6, at 902.

19. *Id.* at 905.

20. *Id.* at 912.

disturbed and disillusioned by the lack of consensus generated by the partisan experts. Famed judge and legal historian James Fitzjames Stephen noted at the time that "[n]o one expects an expert, except in the rarest possible cases, to be quite candid. Most of them—for there are a few exceptions—are all but avowedly advocates, and speak for the side which calls them."[21]

The use of partisan scientific experts crossed the pond by the middle decades of the 19th century.[22] Similar to their English counterparts, scientific experts found lucrative opportunities to testify across areas of litigation in American courts.[23] And, again common to the English experience, their introduction to the legal process inevitably resulted in scientists disagreeing with each other on the witness stand, casting doubt on the integrity of the burgeoning scientific community.[24] This lack of consensus on scientific evidence from partisan scientific experts called into question the credibility of this new form of testimony, with some judges discounting it entirely.[25] The U.S. Supreme Court Chief Justice Morrison Remick Waite wrote in 1874 that "whoever has read the reports of trials or been present at them, in which experts are seen arrayed against each other, prostituting at times the science which they professed to represent, . . . need not be told, that the subject of expert testimony as now understood, is one of no ordinary importance."[26]

As business transactions became more sophisticated, the use of expert testimony in United States courts expanded into commercial litigation. Courts today increasingly rely on expert witnesses to bridge the gap between legal principles and the detailed factual underpinnings of commercial disputes.[27] And, as industries became more specialized and the legal environment more intertwined, experts from a wider range of fields, including economics, finance, and accounting, have been called to provide testimony.[28] The rise of global commerce, digital technologies, and complex financial instruments has further driven the need for expert testimony to explain the complexities involved in modern commercial litigation.[29]

21. JAMES FITZJAMES STEPHEN, A GENERAL VIEW OF THE CRIMINAL LAW OF ENGLAND, 199 (London, MacMillan & Co., 2nd ed. 1890).

22. Golan, *supra* note 6, at 915.

23. FRANCIS WHARTON, A COMMENTARY ON THE LAW OF EVIDENCE IN CIVIL ISSUES §§ 434–451, at 394–421 (Phila., Kay & Bro., 3rd ed. 1888).

24. J. SNOWDEN BELL, THE USE AND ABUSE OF EXPERT TESTIMONY 28–34 (Phila. Rees Welsh & Co. 1879).

25. *Expert Testimony*, 5 AM. L. REV. 227, 228 (1871).

26. Morrison R. Waite, *Testimony of Experts*, 8 W. JURIST 129, 134–35 (1874).

27. Raymond Kolls & Jeffrey Stec, *Why Expert Witnesses Are Key to Navigating Complex Litigation*, BLOOMBERG L., Jan. 5, 2023, https://news.bloomberglaw.com/us-law-week/why-expert-witnesses-are-key-to-navigating-complex-litigation.

28. Roman L. Weil, et al., LITIGATION SERVICES HANDBOOK: THE ROLE OF THE FINANCIAL EXPERT, 4 (5th ed. 2014).

29. *See* Michael J. Mandel, *Going for Gold: Economists as Expert Witnesses*, 13 J. ECON. PERSP. 113, 114 (1999).

One key area where expert testimony has seen significant growth is in damages calculations.[30] Historically, damages were often calculated using basic methods, but as litigation in sectors, such as antitrust, intellectual property, and securities fraud has increased, courts require more precise models to understand potential losses or financial harm.[31] Economists and financial experts are now frequently hired to create complex models that assess the impact of alleged misconduct, lost profits, or market manipulations. These experts can provide clarity by offering nuanced insights into causation and quantifying harm in ways that were not previously possible.[32]

Another factor contributing to the increase in the use of expert testimony is the expanding scope of regulatory environments. With regulatory bodies like the Securities and Exchange Commission (SEC) and the Federal Trade Commission (FTC) playing an increasingly significant role in enforcing business practices, litigation involving regulatory compliance has grown.[33] Experts in securities law, environmental regulations, and telecommunications standards are frequently hired to explain whether a company's conduct meets or violates established legal standards. Their testimony often becomes pivotal in determining the outcome of cases, particularly when there is a need to interpret new and evolving regulations that require deep subject matter expertise.[34]

Over time, the role of expert testimony has also been shaped by the increasing complexity of commercial relationships, particularly those involving cross-border disputes or multinational corporations.[35] Experts in international trade,

30. Robert Thornton & John Ward, *The Economist in Tort Litigation*, 13 J. ECON. PERSP. 101, 101 (1999). ("Over the past two decades, the participation of economists as consultants and expert witnesses in civil tort actions has grown rapidly. This involvement has taken the form of applying the theory and methodology of economics to the measurement of damages in litigation involving mainly personal injury, wrongful death, employment discrimination, and commercial disputes.").

31. *See, e.g.*, Comcast Corp. v. Behrend, 569 U.S. 27 (2013) (holding that a plaintiff's damages model must measure only those damages attributable to the specific theory of harm that survives class certification. In doing so, the Court reinforced that courts will subject damages methodologies to "rigorous analysis," effectively raising the bar on the precision and reliability required of expert models in large-scale antitrust, intellectual property, and securities fraud cases.).

32. *See generally* Mark A. Allen et al., *Reference Guide on Estimation of Economic Damages*, in REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 425–9 (Fed. Jud. Ctr. 3d ed. 2011).

33. *See, e.g.*, Skadden, Arps, Slate, Meagher & Flom LLP, *FTC Enforcement Trends*, in *2024 Insights: Enforcement and Litigation* (Dec. 2023), https://www.skadden.com/insights/publications/2023/12/2024-insights/enforcement-and-litigation/ftc-enforcement-trends; Michael Ewens et al., *Regulatory Costs of Being Public: Evidence from Bunching Estimation*, 153 J. FIN. ECON. 2 (2024) (finding that regulatory compliance costs amount to 4.3% of the market capitalization for a median US public firm).

34. *See, e.g.*, S.E.C. v. Johnson, 525 F. Supp. 2d 70, 75 (D.D.C. 2007) (discussing the use of a certified public account to testify to accounting standards and regulations in a case concerning an alleged fraudulent scheme to materially and improperly inflate revenue figures).

35. For example, a 2014 survey of multinational corporations by Hogan Lovells found that complex and costly cross-border legal disputes are projected to grow significantly. Hogen Lovells, *Survey: Cross-Border Litigation on the Rise; Many Corporations Identify Legal Systems in the U.S. and China as the Most Challenging*, PR NEWSWIRE, Feb. 11, 2014, https://www.prnewswire.com/news-releases/survey-cross-border-litigation-on-the-rise-many-corporations-identify-legal-systems-in-the-us-and-china-as-the-most-challenging-244892911.html.

global finance, and cross-jurisdictional regulatory compliance have become invaluable in cases where national legal systems intersect. For example, in disputes involving international mergers or allegations of anticompetitive behavior, courts often rely on expert testimony to assess how business activities in one jurisdiction affect markets in another.[36] This expansion of expert testimony in international commercial litigation reflects the global nature of modern commerce, where legal, financial, and economic issues are deeply intertwined across borders.[37]

The credibility and reliability of expert testimony have been notionally reinforced by the heightened standards established by legal precedents, namely the *Daubert* standard.[38] This standard, which governs the admissibility of expert testimony in federal courts, requires that experts use reliable methods and base their opinions on sufficient data, which has further solidified the role of experts in commercial litigation.[39] Courts have come to expect rigorous and well-reasoned testimony, leading to an increased demand for highly credentialed experts who can withstand judicial scrutiny. As a result, the selection of expert witnesses has become a strategic decision for attorneys, with significant resources being invested in finding and vetting individuals who possess the knowledge and credibility to persuade a judge or jury.[40]

The growth in demand for expert testimony created a nascent and profitable new industry providing partisan expert witnesses in litigation. Charles River Associates (CRA), one leading litigation consulting firm, was founded in 1965 and gained prominence as a member of IBM's antitrust defense team. CRA developed a business model in which prominent academics affiliate exclusively with a consulting practice, a practice now copied by other firms in industry like Analysis Group and Compass Lexecon.[41] It is difficult to generate an accurate estimate of the total size and profitability of the litigation consulting industry, given that many practices are subsidiaries of larger consulting firms. However, CRA alone had an estimated revenue range of $670 to $685 million for fiscal

36. *See* Hartford Fire Ins. Co. v. Cal., 509 U.S. 764, 796–99 (1993) (recognizing the extraterritorial reach of U.S. antitrust laws and discussing complex, multi-jurisdictional issues); OECD, *Cross-Border Merger Control: Challenges for Developing and Emerging Economies*, 9–10 (2011), https://www.oecd.org/content/dam/oecd/en/publications/reports/2011/09/cross-border-merger-control_cf19d571/b6efd932-en.pdf at (discussing the need for specialized expertise to evaluate the impacts of multijurisdictional mergers).

37. McKinsey&Co.,*GlobalFlows:TheTiesThatBindinanInterconnectedWorld* (Nov. 15, 2022), https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/global-flows-the-ties-that-bind-in-an-interconnected-world.

38. *See infra* Section 3.

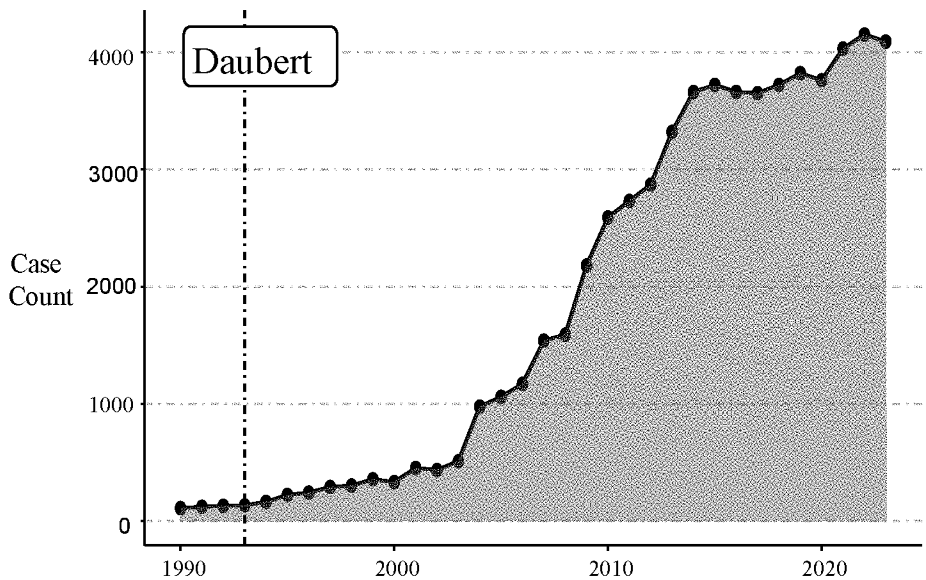
39. *See, e.g.*, Margaret A. Berger, *What Has a Decade of Daubert Wrought?*, 95 AM. J. PUB. HEALTH 559, 564–65 (2005), available at https://ajph.aphapublications.org/doi/pdf/10.2105/AJPH.2004.044701 (last visited Feb. 11, 2025).

40. Maria Salgado, *A Primer on When to Use Expert Witnesses and How to Find Them*, BLOOMBERG L. (Jan. 14, 2013), https://news.bloomberglaw.com/us-law-week/a-primer-on-when-to-use-expert-witnesses-and-how-to-find-them.

41. Mandel, *supra* note 29, at 114.

year 2024.[42] One crude attempt to explore the importance of expert work over time is to analyze measures of its empirical frequency. Figure 1 reports the number of state and federal judicial opinions referencing an "expert report" in the Google Scholar Cases database from 1990 to 2023. Consistent with popular commentary and anecdotal evidence, the growth in such work has ballooned over this period, with roughly 4,000 cases a year referencing expert work by the end of the sample.[43]

Figure 1: Google Case Citations to "Expert Report" Over Time



This figure shows the number of hits from the Google Scholar "Cases" database for "expert report" over time.

In conclusion, expert testimony in commercial litigation has grown in both importance and complexity, reflecting the evolving nature of business and legal disputes. From providing clarity on technical issues to offering detailed economic analyses, expert witnesses now play a central role in shaping the outcomes of high-stakes commercial cases. As industries continue to advance and legal frameworks grow more intricate, the reliance on expert testimony will

42. Charles River Assocs., *An Overview of Charles Rivers Associates, Q3 FY2024*, https://crainternationalinc.gcs-web.com/static-files/0b265a92-00d5-43ab-86b6-40e1dc564f86 (last visited Jan. 15, 2025).

43. I note that this is not meant to be read literally, as part of the increase could be due to an increase in the use of the term "expert report" rather than other terms. Nevertheless, it is consistent with practitioners and judges who have discussed the rise in expert work.

likely continue to increase, making it a cornerstone of modern commercial litigation.

3. JUDICIAL GATEKEEPING OF EXPERT EVIDENCE

From its inception, the American legal community recognized the need to grapple with the problems endemic to scientific testimony. By the end of the 19th century, it was abundantly clear that our evidence laws were incapable of adequately controlling the problem without reform.[44] The first full-fledged judicial attempt to get a handle on the problem came in a 1923 D.C. Court of Appeals opinion about lie-detector technology.[45] In *Frye v. United States*, the defendant attempted to introduce expert witness testimony from one of the inventors of the lie detector to prove his innocence.[46] The trial court refused to admit the new technology into evidence, and Frye appealed on the grounds that his choice in scientific expert was improperly excluded.[47] At the time, the traditional evidentiary criteria for inclusion of evidence was the "logical relevancy" of the evidence and its usefulness to the trier of fact, as well as the qualifications of the expert witness.[48] Given the difficulty in excluding the testimony on traditional grounds, the appellate court proffered a novel exclusionary rule—that lie detection based on systolic blood pressure had not yet "gained such standing and scientific recognition among physiological and psychological authorities as would justify the courts in admitting expert testimony deduced from the discovery, development, and experiments thus far made."[49] The *Frye* "general acceptance" standard augured a trend towards increased judicial scrutiny of evidence that would persist through the second half of the 20th century.[50]

The rise of judicial scrutiny under *Frye* was not met without criticism, namely that it deprived jurors of their right to decide on the usefulness of evidence[51] and that it was excessively vague.[52] The decision in *Frye* was ultimately superseded at the federal level with the enactment of the *Federal Rules of Evidence* ("FRE") in 1975. These rules allowed for the opinion testimony of experts qualified by "knowledge, skill, experience, training, or education" if the knowledge provided will "assist the trier of fact to understand the evidence or to

44. Golan, *supra* note 6, at 923.

45. Frye v. United States, 293 F. 1013, 1013 (D.C. Cir. 1923).

46. Kenneth J. Weiss, Clarence Watson, & Yan Xuan, Frye*'s Backstory: A Tale of Murder, a Retracted Confession, and Scientific Hubris*, 42 J. AM. ACAD. PSYCHIATRY & L. 226, 227 (2014).

47. *Id.* at 1013–14.

48. Paul R. Rice, *Peer Dialogue: The Quagmire of Scientific Expert Testimony: Crumping the Supreme Court's Style*, 68 MO. L. REV. 53, 56 (2003).

49. Frye, 293 F. at 1014.

50. Golan, *supra* note 6, at 930.

51. *See* CHARLES T. MCCORMICK, HANDBOOK OF THE LAW OF EVIDENCE § 14, at 363 (1954).

52. David E. Bernstein, Frye, Frye *Again: The Past, Present, and Future of the General Acceptance Test*, 41 JURIMETRICS 385, 390 (2001).

determine a fact in issue."[53] The new approach under FRE Rule 702 is considered a relaxation of the traditional standard of review of expert evidence, and was ultimately held by the Supreme Court to be inconsistent with the "austere" *Frye* standard.[54] Initially, however, courts were unsure how to unify *Frye* and the *Federal Rules of Evidence*, and some considered the "general acceptance" standard to survive as a pre-condition for the admissibility of scientific experts.[55] Ultimately, the Supreme Court formally overturned *Frye* in a case brought against the pharmaceutical corporation Merrell Dow over birth defects blamed on the anti-nausea drug Bendectin.[56] This case, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, created a new standard consistent with the new FRE, affirming the central role played by judges in gatekeeping evidence to the jury.

The standard in *Daubert* remains the legal rule governing the admissibility of expert testimony in U.S. federal courts, particularly in relation to scientific and technical evidence. *Daubert* set forth the criteria that federal judges must use to determine whether proffered expert testimony is sufficiently reliable and relevant to be presented to a jury. The overarching goal of *Daubert* is to ensure that expert evidence is grounded in scientific validity rather than speculation or unreliable methodologies.[57] The decision represented a break from the earlier *Frye* standard, with its focus on the validity of the proffered evidence for the specific purpose of the case, rather than the general acceptance of the methodology within a relevant scientific community.[58]

Under *Daubert*, judges are required to evaluate several factors to determine the admissibility of expert testimony. These factors include whether the theory or technique employed by the expert can be (and has been) tested; whether it has been subjected to peer review and publication; the known or potential rate of error; the existence and maintenance of standards governing the methodology's operation; and whether the theory or technique is generally accepted within the relevant scientific community.[59] These considerations guide the court in ensuring that evidence introduced to the court is scientifically grounded.

Judges serve a crucial role as gatekeepers when applying the *Daubert* Standard. It is their responsibility to assess whether the methodology underlying the expert's testimony is not only scientifically valid, but also relevant to the case at hand. This requires the judge to move beyond simply evaluating an expert's credentials or field of expertise; they are required to scrutinize the reasoning and

53. FED. R. EVID. 702 (1975).
54. Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579, 588–89 (1993).
55. Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *A Half-Century Later*, 80 COLUM. L. REV. 1197, 1228–31 (1980).
56. Daubert, 509 U.S. at 582.
57. *Id.* at 594–95.
58. *Id.* at 591.
59. *Id.* at 593–94.

processes that lead to the expert's conclusions.[60] Even if a method is reliable as a general principle, it must be shown to have direct relevance to the facts in dispute for it to be admitted.[61] Judges also exercise significant discretion in determining which of the *Daubert* factors are most applicable in each case and how heavily to weigh them. The standard does not require that all factors be met, but it does provide a framework for ensuring that expert testimony is grounded in reliable scientific principles.[62]

In upholding their gatekeeping role, judges often hold *Daubert* hearings as part of the pretrial process to assess the admissibility of expert testimony.[63] *Daubert* hearings provide both sides an opportunity to argue for or against the use of a particular expert, and they offer judges a venue to explore the scientific foundations of the proposed evidence. The rulings made during these hearings often significantly shape the course of a trial, as the exclusion of expert testimony can weaken a party's case or change the dynamics of the evidence presented to the jury.[64]

The *Daubert* standard applies in federal courts, but state courts are free to follow their own rules regarding expert evidence. As of 2024, only six states continue to use a *Frye* standard: California, Illinois, Minnesota, New York, Pennsylvania and Washington.[65] While a majority of states have adopted *Daubert*, some states have adopted modified versions of *Daubert*.[66] Even in states that have not formally adopted *Daubert*, it has been argued that "Daubert's shadow" impacts the decision whether to admit expert testimony.[67] Regardless of the precise legal standard governing expert evidence, judges in every state are required to play some role in gatekeeping evidence provided by experts to juries.

Daubert has arguably had a profound impact on the use of expert witnesses in courtrooms, placing a greater burden on those experts to demonstrate not only

60. *Id.* at 592–93.

61. *Id.* at 591.

62. Carl F. Cranor et al., *Judicial Boundary Drawing and the Need for Context-Sensitive Science in Toxic Torts After* Daubert v. Merrell Dow Pharmaceuticals Inc., 16 VA. ENV'T. L. J. 1, 8 (1996).

63. G. Michael Fenner, *The* Daubert *Handbook: The Case, Its Essential Dilemma, and Its Progeny*, 29 CREIGHTON L. REV. 939, 948 (1996).

64. D. Alan Rudlin, *The Judge as Gatekeeper: What Hath* Daubert-Joiner-Kumho *Wrought?*, 29 PROD. SAFETY & LIAB. REP. (BL) 329, 336 (2001) ("[T]he *Daubert* hearing and ruling have effectively become virtually as case outcome determinative as a class certification hearing and ruling: once decided, a case either shrivels up and goes away, or becomes more dangerous to try. *Daubert* hearings are often every bit as case dispositive, practically speaking, as a summary judgment hearing. Thus, practitioners whose cases rely in any material way on expert testimony must . . . be prepared for a full-blown 'trial within a trial' that the *Daubert* hearing often becomes.").

65. DAMIAN D. CAPOZZOLA, EXPERT WITNESSES IN CIVIL TRIALS § 2:54 (2024–2025 ed. 2024).

66. For example, in Iowa courts are encouraged to apply *Daubert*, but they are not required to do so. *See* Leaf v. Goodyear Tire & Rubber Co., 590 N.W.2d 525, 532–33 (Iowa 1999) (holding that while the use of the *Daubert* factors may be helpful to the trial court when assessing the reliability of expert testimony, it is not required under Iowa law.)

67. DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY § 23:21 (2024–2025 Ed.).

their expertise but also the scientific rigor of their methodologies. By emphasizing factors such as testability, peer review, and error rates, the standard filtered out so-called "junk science" from influencing court decisions. At the same time, it increased the responsibility placed on judges, who must now have a degree of understanding in scientific and technical matters to effectively evaluate expert evidence.

This increased responsibility has not come without costs. As Chief Justice Rehnquist argued in partial dissent in *Daubert*, the ruling forced trial judges "to become amateur scientists" to fulfill their gatekeeping role.[68] In the intervening period, cases have increased in scope and complexity, and these burdens have only increased. According to the historian Tai Golan:

> Consequently, at the beginning of the twenty-first century, lay judges find themselves deeper than ever in the strange land of biostatistics, confidence levels, meta-analysis, and falsifiability, charged with the difficult task of weighing the merit of highly specialized scientific claims. How well the lay judges can meet these challenges and whether their gate-keeping role will lead to better adjudication are questions that will bear careful watching.[69]

By this point in the development of *Daubert* and its progeny, most commentators would conclude that lay judges have struggled to meet the challenge.

4. PROPOSALS TO ENHANCE THE PROVISION OF EXPERT TESTIMONY

At the advent of partisan scientific testimony, eighteenth-century judges relied on a gentlemanly code of honor for believing that men of science could be trusted on to give unbiased testimony when called upon.[70] "The status of the gentleman—his economic independence, the freedom of his actions, the moral discipline he imposed on himself—guaranteed the credibility of his word."[71] As explained *infra* Section 3, the expansion of science and technology into commerce and society ensconced scientists within the adversarial legal system, denting the credibility of science and its adherents within the legal community. This growing mistrust of the scientific community represented a threat to carrying out justice and the public image of the scientific community, generating reform proposals within and outside of the scientific community.[72]

68. Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

69. Golan, *supra* note 6, at 942.

70. *Id.* at 903.

71. Golan, *supra* note 6, at 903 (citing to PHILIP MASON, THE ENGLISH GENTLEMAN: THE RISE AND FALL OF AN IDEAL (1982), SIMON RAVEN, THE ENGLISH GENTLEMAN (1961), and Peter Dear, Totius in Verba: *Rhetoric and Authority in the Early Royal Society*, 76 ISIS 145 (1985)).

72. *Id.* at 913.

One recurring proposal for reforming expert testimony was to use court-appointed experts.[73] If the problem with partisan scientific expert testimony is partisanship rather than science, then removing the ability of litigants to select their own experts would solve it. During the first Victorian debates on expert testimony, nearly all of the scientific proposals for reforming legal practice agreed that courts should be allowed to use their own independent witnesses.[74] These calls were repeated during the early-American experience with partisan expert witnesses, with reform proposals calling for the selection of experts by the court, unassisted or made from an official roster of selected experts.[75] In fact, Michigan passed a statute in 1905 mandating that courts nominate their own experts in murder trials, which was ultimately struck down as unconstitutional by the Michigan Supreme Court.[76]

The authority to appoint independent experts by federal courts is set forth in FRE 706,[77] as well as being inherent in the power of courts to take actions required for their decision-making function.[78] Experts retained under FRE 706 are chosen by the judge after consultation with both parties, and the fees and other costs are typically borne equally by both parties.[79] However, the use of court-appointed experts in federal courts is rare.[80] Many judges have a severe reluctance to appoint experts because it feels contrary to our adversarial system.[81] Even if theoretically justified, it arguably leads to an unconstitutional delegation of the judiciary's Article III authority,[82] and a lack of objectivity on behalf of the judiciary.[83] Regardless of the reason for this judicial reticence, centuries of discussion and proposals for court-appointed experts have had minimal impact on the practice of court-appointed expert testimony in American courts.

Another frequent proposal is to rely on the professional communities that govern experts to clean up witness practice. During the twentieth century, most

73. *See, e.g.*, Joe S. Cecil & Thomas E. Willging, *Accepting* Daubert*'s Invitation: Defining a Role for Court-Appointed Experts in Assessing Scientific Validity*, 43 EMORY L. J. 995, 998 (1994).

74. TAL GOLAN, LAWS OF MEN AND LAWS OF NATURE 120 (2004).

75. A.M. Kidd, *The Proposed Expert Evidence Bill*, 3 CALIF. L. REV. 216, 223 (1915).

76. People v. Dickerson, 129 N.W. 199, 200–01 (Mich. 1910).

77. FED. R. EVID. 706.

78. FED. R. EVID. 706 advisory committee's note; *see also* United States v. Green, 544 F.2d 138, 145 (3d Cir. 1976) ("the inherent power of a trial judge to appoint an expert of his own choosing is clear"); Scott v. Spanjer Bros., Inc., 298 F.2d 928, 930 (2d Cir. 1962) ("Appellate courts no longer question the inherent power of a trial court to appoint an expert under proper circumstances.").

79. Daniel L. Rubinfeld & Joe S. Cecil, *Scientists as Experts Serving the Court*, 147 DAEDALUS 152, 154 (2018).

80. *Id.* at 155.

81. Douglas H. Ginsburg, *Appellate Courts and Independent Experts*, 60 CASE W. RES. L. REV. 303, 304
(2010).

82. *Id.*

83. Tahirih V. Lee, *Court-Appointed Experts and Judicial Reluctance: A Proposal to Amend Rule 706 of the Federal Rules of Evidence*, 6 YALE L. & POL. REV. 480, 497–98 (1988).

professions created associations that developed codes of ethics and minimum professional standards through examinations carried out by the relevant community or through state boards of examiners.[84] As legal scholars at the time argued, a successful campaign to increase the honesty of expert witnesses would need to come from within the respective professions, rather than top-down reforms from courts or legislatures which had failed in the past.[85] We see similar calls for community oversight today; Luigi Zingales, a professor of finance at the University of Chicago, has written that "[a]lthough academic writings are scrutinized during expert testimonies, expert testimonies are not scrutinized by the academic community. It is time for this to start."[86] One difficulty that this renewed interest in professional oversight might face today is the frequent sealing of expert reports in federal court.[87] A nascent movement to roll back the trend in federal court over-sealing would make successful professional oversight more likely.[88]

Other commonly proposed reforms to the provision of expert testimony include the use of baseball-style arbitration incentive mechanisms and concurrent expert evidence hearings, also known colloquially as "hot-tubbing." Professional baseball implemented a dispute resolution procedure that has been considered successful in decreasing the costs of arbitration and expediting the time-to-resolution of pay disputes.[89] Under baseball-style arbitration, each side submits a proposed resolution to the dispute, and an independent arbiter may choose only one party's proposals.[90] Also known as "last best offer," this method is intended to moderate bargaining positions, as extreme proposals are likely to be rejected by the arbiter.[91] It has been used to resolve tax[92] and construction industry disputes,[93] and has been proposed as a possible alternative to the battle

84. Golan, *supra* note 6, at 926.

85. Lee M. Friedman, *Expert Testimony, Its Abuse and Reformation*, 19 YALE L. J. 247, 252 (1910).

86. Luigi Zingales, *Preventing Economists' Capture*, SSRN Scholarly Paper No. 2353489 at 3 (Nov. 15, 2013), https: //papers.ssrn.com/sol3/papers.cfm?abstract_id=2353489.

87. Leslie Brueckner & Beth Terrell, *When it Comes to Sealing Court Records, the Presumption of Public Access Requires that You "Just Say No"*, PUB. JUST. (Jul. 6, 2017), https://www.publicjustice.net/comes-sealing-court-records-presumption-public-access-requires-just-say-no/ (last visited Feb. 11, 2025) ("[C]ourt records in this jurisdiction—as elsewhere—are sealed all too often without any showing of any need for secrecy at all, much less the type of compelling need for secrecy required by the First Amendment.").

88. Heather Abraham, Jonathan Manes & Alex Abdo, *Judicial Secrecy: How to Fix the Over-Sealing of Federal Court Records*, KNIGHT FIRST AMEND. INST. AT COLUM. UNIV. (Oct. 21, 2021), https://knightcolumbia.org/blog/judicial-secrecy-how-to-fix-the-over-sealing-of-federal-court-records.

89. Lochlin B. Samples, *Resolving Construction Disputes Through Baseball Arbitration*, Am. Bar Ass'n,
Under Construction (Mar. 12, 2019), https://www.americanbar.org/groups/construction_industry/publications/under_construction/2019/spring/resolving-dispute-baseball.

90. Luis Flavio Neto, *Baseball Arbitration: The Trendiest Alternative Dispute Resolution Mechanism in International Taxation*, 2019 INT'L TAX STUD. 2, 2 (2019).

91. *Id.* at 3.

92. *Id.*

93. Samples, *supra* note 89.

of experts in costly valuation proceedings.[94] While baseball arbitration is theoretically and empirically appealing as a moderating mechanism for experts, it is challenging to see how it would fit into our adversarial legal system. Under FRE and *Daubert*,[95]judges are instructed to allow scientifically valid expert witness testimony, and it is the jury's decision how much weight to give evidence.[96]

Finally, concurrent expert evidence refers to an Australian practice where competing experts are sworn in and presented as witnesses at the same time. The experts remain on the stand together and the testimonial dialogue ensures that experts address the same issues under the same assumptions simultaneously, allowing differences of opinion to be clarified or explained. Experts can promptly address any misunderstandings or questions from the judge or counsel. This approach enables the judge to compare opposing experts' evidence in real time, rather than weeks or days later through pleadings and depositions. Concurrent expert testimony enhances the quality, precision, and clarity of technical communication, while highlighting and sharpening any existing differences between the experts.[97] While not historically common in the United States, it has been used recently in high-profile antitrust litigation.[98]

All of these proposals are worthwhile, either in isolation or conjunction. The use of court-appointed experts makes obvious sense from an incentive perspective. However, judges have not been receptive to the idea of replacing partisan experts with court-appointed ones as a general practice,[99] and the odds of this changing in the near-future seem low. The professional communities from which testifying experts are drawn should have an interest in safeguarding their reputations with courts and the legal community. This is challenging given the overly permissive approach taken by many federal courts in sealing expert work product, and the legal community should pressure judges to pull back from the practice. Baseball-style arbitration is promising, but likely a bad fit with our civil litigation regime, and while concurrent testimony seems on the rise, it does little to change the underlying incentive system that has long plagued the use of partisan expert witnesses. In the next section, I propose a modification to the

94. *See* Keith Sharfman, *Valuation Averaging: A New Procedure for Resolving Valuation Disputes*, 88 MINN. L. REV. 357, 365–66 (2003).

95. Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

96. David L. Faigman, *Evidence: Admissibility vs. Weight in Scientific Testimony*, 1 THE JUDGES' BOOK 45, 45
(2017).

97. *Is There Room in American Courts for an Australian Hot Tub?*, Jones Day Insights (Apr. 26, 2013), https://www.jonesday.com/en/insights/2013/04/room-in-american-courts-for-an-australian-hot-tub.

98. *See* Dan Papscun, *Courtroom 'Hot Tub' Puts Google Trial Experts to Stress Test*, BLOOMBERG LAW (Oct. 6, 2023), https://news.bloomberglaw.com/antitrust/courtroom-hot-tub-puts-google-trial-experts-to-stress-test.

99. *See* Lee, *supra* note 83.

form of a subset of, largely economic, evidence submitted for litigation purposes, which could work in conjunction with any of the reforms mentioned here.

5. MODEL-DRIVEN VS. DATA-DRIVEN ESTIMATION

In two prior papers with my coauthors Jonah Gelbach and Eric Talley,[100] we propose an alternative approach to improving the reliability and administrability of expert testimony in commercial litigation that builds upon an established literature in statistics, computer science, and economics on "statistical learning." Across two substantive practice areas—securities litigation and corporate valuation—we show how data-driven estimation strategies are both more accurate and less susceptible to expert discretion than conventional practices. We argue that expert practice would improve if judges requested that experts provide such evidence, even simply as a benchmark comparison to their other testimony.

Our argument is not particularly complicated or even novel: it was in fact made decades ago in a similar setting by statistician Leo Breiman.[101] When using statistical modeling to generate conclusions or impressions from data, there are two distinct approaches. The "data modeling culture" assumes that the data is generated from a given data generating process and estimates the values of the parameters that best fit the model from a sample of the data. A model of this type is of the form:[102]

$$response\ variable = f(predictors, random\ noise)$$

Here, the analysis models the outcome variable as a (usually linear) function of a specified set of inputs (also known as predictors, or independent variables), allowing for some random (or "stochastic") noise in the relationship.[103] Historically, this is how most statisticians used data models, and it is still the conventional approach used by social scientists (statisticians, economists, sociologists, accountants, etc.) in most expert testimony today. Wages are assumed to be a linear function of experience and years of education, firm stock returns or valuations are linear functions of market and industry factors, etc.

100. *See* Andrew Baker & Jonah B. Gelbach, *Machine Learning and Predicted Returns for Event Studies in Securities*
Litigation, 5 J. L. FIN. & ACC. 231 (2020); Andrew C. Baker, Jonah B. Gelbach, & Eric Talley, *Validating Valuation: How Statistical Learning Can Cabin Expert Discretion in Valuation Disputes*, (unpublished manuscript,
SSRN 2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4849281.

101. *See* Leo Breiman, *Statistical Modeling: The Two Cultures*, 16 STATISTICAL SCIENCE 199 (2001).

102. *Id.* at 199.

103. *Id.*

As Breiman noted in 2001, "[t]his enterprise has at its heart the belief that a statistician, by imagination and by looking at the data, can invent a reasonably good parametric class of models for a complex mechanism devised by nature. Then parameters are estimated and conclusions are drawn."[104] Experts can ostensibly discern between competing models by analyzing goodness-of-fit measures like R^2, as is still done in some litigation today, even though there are well-documented limitations to this approach.[105] Breiman was frustrated by the dominance of this paradigm; after a stint outside of academia as a paid consultant, at times to government agencies, he felt that the standard approach was a straight-jacket that led to "questionable scientific conclusions" rather than allowing the research to "[f]ocus on finding a good solution—that's what consultants get paid for."[106]

Another approach to statistical modeling exists. Rather than explicitly defining the stochastic data model, the "algorithmic modeling culture" based on the practice of statistical learning considers the mapping from inputs (e.g. education, age, training) to outputs (e.g. wages) complex and unknown, and instead looks for a function that best predicts the response. Rather than using goodness-of-fit measures that are potentially biased and subject to manipulation, model selection is done using prediction error and cross-validation.[107] Comparing competing models on a straightforward measure like out-of-sample prediction error, rather than wading into a murky battle over the asymptotic properties of differentially-specified parametric models, is also much easier to explain to a lay judge or jury.

Looking back with 25 years of hindsight, Breiman decisively won the battle in academia, industry, and policy. Algorithmic models—from regression trees to random forests and neural nets—now dominate data analysis in practice and in are pervasive in leading academic journals.[108] Trillions of dollars are being invested into generative artificial intelligence companies, which mine seemingly infinite computing resources to glean insights from massive datasets.[109]

---

104. *Id.* at 202.

105. *Id.* at 202–04 ("[D]ifferent models, all of them equally good, may give different pictures of the relation between the predictor and response variables. The question of which one most accurately reflects the data is difficult to resolve. One reason for this multiplicity is that goodness-of-fit tests and other methods for checking fit give a yes-no answer. . .There is no way, among the yes-no methods for gauging fit, or determining which is the better model.").

106. *Id.* at 199–201.

107. *Id.* at 204. Cross-validation refers to the practice of estimating the model on a portion of the data and testing the prediction error on the held-out sample.

108. *See* Foster Provost & Tom Fawcett, *Data Science and its Relationship to Big Data and Data-Driven Decision Making*, 1 BIG DATA 51, 51 (2013); Susan Athey, *The Impact of Machine Learning on Economics*, *in* THE ECONOMICS OF ARTIFICIAL INTELLIGENCE: AN AGENDA 507, 507-08, 516–17 (Ajay Agrawal, Joshua Gans & Avi Goldfarb eds., 2019), https://www.nber.org/system/files/chapters/c14009/c14009.pdf.

109. Goldman Sachs, *Will the $1 Trillion of Generative AI Investment Pay Off?*, GOLDMAN SACHS (Aug. 5, 2024), https://www.goldmansachs.com/insights/articles/will-the-1-trillion-of-generative-ai-investment-pay-off.

Moreover, the algorithmic approach of statistical learning clearly satisfies the evidentiary standards of litigation—it is well-accepted by the relevant academic communities[110] and is verifiable, with clearly-defined error rates. The frontier of research in the social sciences from which experts are largely drawn use these methods extensively, from econometrics,[111] to predicting stock returns in finance,[112] modeling wage gaps in labor economics,[113] and detecting cartels in the field of industrial organization.[114] However, the practice has made little inroads in scientific expert witness testimony for commercial litigation.

A shift towards the use of more statistical learning in expert testimony would limit the scale of the differences between experts in some contentious disputes.[115] While statistical learning still requires discretion over the potential inputs into the model, the importance of these choices in litigation will be less important because of the data-driven, rather than researcher-driven, mapping from the inputs to the outputs. For similar reasons, statistical learning has also been proposed as a partial remedy to the frequent and increasingly problematic use of specification searches, colloquially referred to as "p-hacking," in empirical academic research.[116] The use of prediction error (where applicable), rather than the ad-hoc and subjective comparison metrics used today, will also aid courts in comparing the analyses of competing experts.

Statistical learning is not a panacea for all that ails the production and adjudication of partisan expert testimony. The algorithmic approach works best when the goal is prediction, rather than learning directly about the parameters of a given model. In a previous paper[117],I argued that many tasks currently undertaken by experts can be framed as prediction exercises; however, this will not always be the case. Many modern statistical learning approaches are conceptually complex and opaque to varying degrees. For this reason, we have largely proposed the use of a straightforward and interpretable algorithm for testimonial purposes—penalized regression models.[118] At this juncture, courts are intimately familiar with the concept of regression analysis across litigation

110. For example, the leading introductory casebook on statistical learning has nearly 25,000 Google Scholar citations as of the time of writing. Gareth James, Daniela Witten, Trevor Hastie, & Robert Tibshirani, AN INTRODUCTION TO STATISTICAL LEARNING (2013).

111. Sendhil Mullainathan & Jann Spiess, *Machine Learning: An Applied Econometric Approach*, 31 J. ECON. PERSP. 87, 87 (2017).

112. Bryan Kelly & Dacheng Xiu, *Financial Machine Learning*, 13 FOUNDS. & TRENDS IN FIN. 205, 206-10 (2023).

113. Marina Bonaccolto-Töpfer & Stephanie Briel, *The Gender Pay Gap Revisited: Does Machine Learning Offer New Insights?*, 78 LAB. ECON. 1 (2022).

114. Martin Huber & David Imhof, *Machine Learning with Screens for Detecting Bid-Rigging Cartels*, 65 INT'L J. INDUS. ORG. 278 (2019).

115. *See, e.g.*, Baker et al., *supra* note 100, at 35.

116. *See, e.g.*, Victor Chernozhukov, Christian Hansen, & Martin Spindler, *Valid Post-Selection and Post-Regularization Inference: An Elementary, General Approach*, 7 ANN. REV. ECON. 649, 650 (2015).

117. Baker & Gelbach, *supra* note 100, at 270.

118. *See id.* at 246-47; Baker et al., *supra* note 100, at 30.

areas. Penalized (or "regularized") models are natural outgrowths of standard regression models, with the only difference in the objective function being the inclusion of optimally-chosen penalty values for the inclusion of independent variables in the model.[119] To the extent that courts are comfortable accepting and inspecting conventional regression models, our proposal does not add much to, and may in fact subtract from, the cognitive burden on the judiciary.

It should also be noted that some of the limitations to the other reform proposals discussed in Section 4 may also apply here. It is hard to see why the expert witness community, which profits from the discretion afforded by the conventional approach, would willingly adopt an approach that limits discretion. Courts of equity generally have wide latitude in fashioning remedies and could almost surely incorporate the approach proposed here. For example, in *In re Mirant Corp.*, a bankruptcy proceeding discussed *infra* Section 6.2.2, the judge refused to accept the analysis of either expert following a valuation hearing and instructed the parties to "recalculate the value of Mirant Group based on necessary changes in data and assumptions."[120] There does not appear to be any reason why a similarly-situated court could not also instruct the parties to use a data-driven estimation procedure. However, courts sitting in law rather than equity may be reticent to impinge on the ability of each litigant to present the evidence of their choosing to the jury.

6. EMPIRICAL EXAMPLES

This section documents through a series of empirical examples how statistical learning can cabin expert discretion in a judicially administrable manner.

6.1 Event Studies and Securities Litigation: The Case of Halliburton

6.1.1 Event Studies and Securities Litigation

Securities class action lawsuits are governed by the SEC's Rule 10b-5, promulgated under Section 10(b) of the Securities and Exchange Act of 1934. Rule 10b-5 makes it unlawful for any person to make an untrue statement of a material fact, or to omit to state a material fact necessary to make other statements not misleading, in connection with the purchase or sale of security.[121] In a securities fraud suit brought under Rule 10b-5, plaintiffs are required to prove the existence of a material misrepresentation or omission that is made with scienter (or a mindset embracing an intent to deceive). In addition, plaintiffs bear

119. *See* Baker & Gelbach, *supra* note 100, at 246 for a more detailed explanation of how penalized regression models operate.

120. *In re* Mirant Corp., 334 B.R. 800, 824 (Bankr. N.D. Tex. 2005).

121. 17 C.F.R. § 240.10b-5(b) (2024).

the burden of proving reliance, which, building upon the common law of deceit, requires the plaintiff to have actually and justifiably relied on the misrepresentation in causing them to transact in the security in question. Under the Supreme Court's test in *Basic v. Levinson*,[122] there is a presumption of reliance where the defendant makes a material representation in an informationally efficient market. Finally, plaintiffs must prove loss causation—that the defendant's wrongful act was the proximate cause of the plaintiff's loss—and be able to prove class-wide damages in justiciable manner.

Antifraud cases brought under our securities laws, particularly those brought pursuant to Rule 10b-5, represent an area of commercial litigation where expert-provided evidence is often outcome-determinative. Occasionally, experts are asked to opine on business and industry facts that can help a court determine the materiality of a specific piece of information.[123] In addition, experts sometimes instruct the court on what a reasonable investor would intuit from a given disclosure. In nearly every case,[124] experts are hired to conduct an "event study" analysis linking specific misstatements and disclosures to the firm's stock price.

An event study is an empirical technique used to identify the effect of an event on the value of a firm's security (typically, though not always, the value of its common equity). Event study evidence is used, and often de-facto required, to support multiple of the elements of a plaintiffs cause of action; including reliance,[125] materiality,[126] loss causation,[127] and damages.[128] Each of these elements is *critically dependent* on the provision of a reliable event study by a

122. 485 U.S. 224, 247 (1988).

123. *See, e.g.*, Ark. Tchr. Ret. Sys. v. Goldman Sachs Grp., Inc., 77 F.4th 74, 96–97 (2d Cir. 2023) (where Dr. Laura Starks "opined that the alleged misrepresentations were 'unlikely, in a vacuum, to consciously influence investor behavior'").

124. At least those cases that make it past a motion for summary judgment.

125. Event studies are often used to determine whether the market for a firm's stock is informationally efficient, by analyzing whether the stock responds in a consistent and statistically significant manner to news regarding the firm's prospects.

126. Following the Supreme Court's holding in *Halliburton II*, defendants are entitled to an opportunity to rebut "price impact" at the class certification stage. Although materiality does not need to be established for a class to be certified, defendants are now allowed to present evidence rebutting the materiality of alleged misrepresentations based on a price impact analysis. Halliburton Co. v. Erica P. John Fund, Inc. (*Halliburton II*), 573 U.S. 258, 285 (2014).

127. Loss causation is the plaintiff's burden to establish a direct connection between the alleged fraud and the economic harm to the shareholders. This harm is measured at two points in time–when the security was purchased and when the fraud was disclosed. Dura Pharm. Inc. v. Broudo, 544 U.S. 336, 345–46 (2005). Evidence of price distortion nearly always requires a formal event study analysis to disentangle the return on the security from other contemporaneous market changes.

128. Damages in securities class actions follow the "out-of-pocket" damages established in *Affiliated Ute Citizens v. United States*, 406 U.S. 128, 155 (1972). Under *Ute*, defrauded purchasers of a security are entitled to the difference between the price paid for the security and the price it would have traded at had the material misrepresentation or omission not occurred. This determination also nearly always requires an event study to disentangle the effect of normal variation in returns from the fraud-induced change.

qualified expert.[129] As some have argued, "the law governing event studies has become inseparable from the substantive law governing securities fraud litigation" because "[c]ourts have effectively collapsed securities fraud actions into a single question: Whether the defendant's misrepresentation or omission created a disparity between the transaction price of a security and its true value measured by the precise reaction of the market price to the disclosure of the concealed information."[130] It is thus unsurprising how much litigant and judicial effort (and expense) has been dedicated to conducting and attacking event studies in securities disputes.

To conduct an event study, an expert first specifies an econometric model that relates the return of a security to the corresponding return on market factors. The expert defines an "estimation period", over which the model will be estimated, and an "event window", which is the period over which the effect of the event on the security will be analyzed. After estimating the model, an expert can determine the potential magnitude and significance of an event by comparing the actual return of the security over the event window to the return predicted by the model's estimated parameters.[131] If this difference is sufficiently large in magnitude in comparison to the model's typical estimation error,[132] then the expert will testify that the returns cannot be explained by normal patterns in the data. Note that, in addition to the quantitative results provided by the model, the expert still has a qualitative role to play in convincing the finder of fact that there were no other firm-specific events that occurred at the same time as the alleged misstatement or correct disclosure that could explain the residual portion of the firm's return.

A key ingredient in constructing an event study analysis is specifying the model that links the expected returns to other contemporaneous returns in the market. In support of their analysis, experts will frequently cite to academic work. However, the event study was created by financial economists as an empirical technique to assess the impact of a general *type of event*— such as mergers or dividend announcements—on the value of a *set* of securities. In such a setting, modeling errors, if uncorrelated with treatment timing, can be expected to average out in the aggregation process. However, in a litigation setting we are almost always dealing with an event that only impacted one firm, without any ability to margin out prediction errors.[133]

129. *See generally* Andrew C. Baker, *Single-Firm Event Studies, Securities Fraud, and Financial Crisis*, 68 STAN. L. REV. 1207 (2016).

130. Michael J. Kaufman & John M. Wunderlich, *Regressing: The Troubling Dispositive Role of Event Studies in Securities Fraud Litigation*, 15 STAN. J. L. BUS. & FIN. 183, 186 (2009).

131. Baker, *supra* note 129, at 1226–31.

132. Models will never perfectly predict the security's return over the estimation period, so we compare the unexplained portion of the return in the event window to the *typical* unexplained component of the return over the estimation period.

133. This core limitation to the single-firm event study has now been extensively addressed in the literature, and presents difficulties for both generating adequate predictions, and conducting valid

Initial academic work used simple expected return models—including the constant mean return model (where the predicted return is equal to the average firm return over the estimation window) and the market adjusted model (where the predicted return is simply the contemporaneous return on a market index). Under one particularly influential theory in academic finance, the Capital Asset Pricing Model (CAPM), the return on a stock is solely a function of its systematic risk, measured as the covariance between its return and the return on a market portfolio.[134] This theory led naturally to the "market model" event study approach, which estimates predicted returns using a linear regression of the firm's returns on the market index over the estimation period. The market model identifies two parameters, α, which is the expected return on the stock when the market return is zero, and β, which measures the firm's systematic risk.[135] Unfortunately, the CAPM assumptions don't hold: β is not the only risk that explains returns. Later models, including that of Fama and French and its extensions, supplement the market risk factor with portfolio return risk factors meant to capture the effect of size, valuation, and momentum.[136] An important takeaway from this shift in the literature is that the field of empirical asset pricing largely moved away from structural, *a priori* model-based estimation to a predictive exercise in finding risk factors, or anomalies, that can predict firm returns. The state-of-the-art methods in the literature now include using non-linear, machine learning methods to forecast predicted returns.[137]

6.1.2 Halliburton

Erica P. John Fund, Inc. v. Halliburton Co. was a long-running securities class action brought under Rule 10b-5 that twice made its way to the Supreme Court. Plaintiffs alleged that Halliburton and its executives issued material misrepresentations regarding the company's potential liability in asbestos litigation, its expected revenue from a series of construction projects, and the benefits of a merger.[138] The defendants initially argued that plaintiffs had not met their burden to invoke *Basic*'s reliance presumption because they could not adequately plead loss causation. After winning on that theory at the district and

statistical inference, especially in the presence of changes in time-varying volatility. *See, e.g.*, Baker, *supra* note 129, at 1226; Jonah B. Gelbach et al., *Valid Inference in Single-Firm, Single-Event Studies*, 15 AM. L. & ECON. REV. 495, 499 (2013); Edward G. Fox et al., *Economic Crisis and the Integration of Law and Finance: The Impact of Volatility Spikes*, 116 COLUM. L. REV. 325 (2016).

134. Eugene F. Fama & Kenneth R. French, *The CAPM is Wanted, Dead or Alive*, 51(5) J. FIN. 1947, 1948 (1996).

135. Baker, *supra* note 129, at 1230.

136. *See generally* Eugene F. Fama & Kenneth R. French, *Common Risk Factors in the Returns on Stocks and Bonds*, 33 J. FIN. ECON. 3 (1993); Mark M. Carhart, *On Persistence in Mutual Fund Performance*, 52 J. FIN. 57, 61 (1997).

137. *See generally* Shihao Gu et al., *Empirical Asset Pricing via Machine Learning*, 33 REV. FIN. STUD. 2223 (2020).

138. *Halliburton II*, 573 U.S. at 264.

circuit courts, the Supreme Court overturned, holding that proving loss causation is a separate inquiry from reliance, and not a requirement at the class certification stage.[139] On remand, Halliburton argued that class certification was inappropriate because the event study evidence provided by their expert to disprove loss causation also demonstrated a lack of "price impact"—i.e. proof that "the alleged misrepresentations affected the market price in the first place."[140] The lack of price impact, arguably, "sever[ed] the link between the alleged misrepresentation and . . . the price received (or paid) by the plaintiff,"[141] rendering the presumption of reliance from *Basic* inapplicable. Halliburton lost on this secondary argument at the lower courts, with the Supreme Court again stepping in, this time in (partial) support of the company to find that "defendants must be afforded an opportunity before class certification to defeat the presumption through evidence that an alleged misrepresentation did not actually affect the market price of the stock."[142]

After the Supreme Court vacated the lower court judgements and remanded the case for further class certification proceedings, the district court ordered additional briefing on price impact and its relation to class certification. Both parties submitted additional expert reports centered around their event study evidence[143] and, as expected, a dispute among the two experts (Chad Coffman for the funds and Lucy Allen for the company) arose. As the court noted, "[t]he determination of whether lack of price impact ha[d] been shown largely turns on the competing methodologies of the parties' experts."[144] The two experts disagreed on a number of methodological issues—including the relevant dates to analyze, the correct estimation period based on the testing dates, the use of one-day or two-day event windows, and whether, and how, to adjust for multiple testing. However, here I will focus on a more fundamental difference between the two experts that has arisen in several securities suits—how to risk-adjust within the market model.

Under the CAPM, β completely captures the explainable portion of a stock's return. A common event study specification, frequently used in litigation, that builds upon this model is:

$$r_t = \alpha + \beta M_t + \epsilon_t$$

where r_t is the return on the company's stock on date t, α is the model intercept that captures the expected return when the market return is zero, M_t is the return on a broad market index (like the S&P 500), and β is the measure of the firm's

139. Erica P. John Fund, Inc. v. Halliburton Co. (*Halliburton I*), 563 U.S. 804, 813 (2011).
140. *Id.* at 814.
141. *Basic*, 485 U.S. at 248.
142. *Halliburton II*, 573 U.S. at 284.
143. Erica P. John Fund, Inc. v. Halliburton Co., 309 F.R.D. 251, 256–57 (N.D. Tex. 2015).
144. *Id.* at 262.

systematic risk. ϵ_t is the model error and reflects the fact that we can never perfectly capture the expected return on a security.

It is common to supplement this specification in litigation, where the event study is estimated for only a single security, with the inclusion of a second index designed to capture industry-specific trends in returns.[145] Consistent with the CAPM not fully explaining the cross-section of expected returns, Baker and Gelbach (2020)[146] shows through a simulation analysis that the inclusion of a simple industry index based on two-digit SIC codes increases the out-of-sample predictive power of the market model. The experts in the *Halliburton* litigation disagreed about the proper way to adjust for the industry component of Halliburton's return prediction.

Lucy Allen, the defendant's expert, estimated an event study that controlled for the company's two primary lines of business: energy services and engineering and construction (E&C). She used the S&P 500 Energy Index to control for the former, and a bespoke equally-weighted index of composed of firms in the Fortune 1000 that are classified as being in the E&C industry for the latter.[147] Chad Coffman, the funds' expert, argued that the Allen model incorrectly controlled for Halliburton's primary business, because the S&P 500 energy index was driven in large measure by energy *producers* rather than energy *servicers*. Coffman created a separate index based off the listed peers in Halliburton's analyst reports, another common way to generate industry indices.[148]

The results of both event study models are reported in Table 1. The first two columns present my best attempt at replicating the models as described in the reports,[149] and the second two columns provide the reported values. For each estimated coefficient I report the estimate, standard error (in parenthesis), t-Statistic, and corresponding p-value (in that order). As can be seen from the model results, I closely, though not exactly, match the results submitted to the court. The difference in industry controls generate a dispute between the experts over one disclosure date in particular—December 4, 2001—when Halliburton announced an adverse judgment in a Texas case regarding its asbestos liability.[150] The as-reported Allen model generates an excess return estimate of -2.9%, with a p-value of 0.20, while the Coffman model leads to a

145. *See* David I. Tabak & Frederick C. Dunbar, *Materiality and Magnitude: Event Studies in the Courtroom* 8 (Nat'l Econ. Rsch. Assocs., Working Paper No. 34, 1999).

146. Baker & Gelbach, *supra* note 100, at 269.

147. Expert Report of Lucy Allen at ¶ 20, Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co., (No. 3:02-CV-1152-M), 2014 WL 4479528 (N.D.Tex.).

148. Expert Report of Chad Coffman at ¶ 30, Archdiocese of Milwaukee Supporting Fund, Inc. v. Halliburton Co., (No. 3:02-CV-1152-M), 2014 WL 4479528 (N.D.Tex.).

149. One reason why I do not perfectly match the estimated model coefficients in the reports is because they remove Halliburton from the S&P 500 Energy Index. Unfortunately, this requires the industry weights, which is a proprietary dataset that I currently don't have access to.

150. Erica P. John Fund, Inc. v. Halliburton Co., 309 F.R.D. 251, 274–75 (N.D. Tex. 2015).

-3.7% excess return with a p-value of 0.02, below the conventional cutoff for statistical significance.

This is the type of methodological dispute that a generalist judge is ill-suited to adjudicate. Both experts provide a plausible story for the inclusion (or lack of inclusion) of different industry controls, without a clear way to resolve the dispute. The difference matters for the resolution of the case, as the inclusion or exclusion of different disclosure dates changes the effective class period and estimates of class-wide damages. Moreover, while Coffman argues for his model based on the superior adjusted-R^2 (a measure of a model's explanatory power), it is not clear that is the right way to do model selection in this setting.

Table 1: Expert Model Results and Predictions

	Replication		Report Values	
	Allen	Coffman	Allen	Coffman
Model Results				
Intercept	0.00	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.00)
	-1.18	-1.70	-0.09	-0.97
	0.24	0.09	0.93	0.33
S&P Energy Index	1.40	0.28	1.38	0.28
	(0.06)	(0.07)	(0.06)	(0.06)
	23.99	4.31	23.30	4.57
	0.00	0.00	0.00	0.00
Fortune E&C Index	0.13	0.08	0.16	0.12
	(0.06)	(0.04)	(0.06)	(0.04)
	2.40	1.85	2.78	2.87
	0.02	0.06	0.01	0.00
Industry Peer Index		0.90		0.85
		(0.04)		(0.04)
		22.70		24.04
		0.00		0.00
Predictions for December 4, 2001				
Excess Return	-0.027	-0.035	-0.029	-0.037
t-Statistic	-1.156	-2.007	-1.290	-2.270
p-value	0.248	0.045	0.198	0.024

This table reports the event study model estimates from the Lucy Allen and Chad Coffman Reports in the Halliburton Securities Lawsuit. The first two columns represent my best attempt at replication, while the second two columns present the results are shown in their reports. I also report the event study results for December 4, 2001, including the excess return and associated t-statistic and p-values from the alternative approaches.

Baker and Gelbach (2020) argues for a re-framing of the question in the litigation context. Rather than view an event study as a way to determine the best model of expected return—surely a fool's errand—we can instead view an event study as an example of a prediction problem. The opposing experts are attempting to generate a prediction of the expected return, considering contemporaneous returns in the market, on a set of pre-determined dates. The experts and the court are not concerned with the model's parameters—namely, the weights placed on different market and industry indices—but instead are solely concerned with generating accurate predictions of the counterfactual return over the event window. Viewed from that angle, a natural way to adjust for the return on a firm's industry is to use a data-driven procedure to select peer firms based on a constructive notion for their use—the extent to which a given peer firm's returns assist in generating a valid prediction of the target firm's returns. This avoids dealing with the non-probative question of which firms qualify as a valid industry peer.

In Baker and Gelbach (2020), we present one intuitive and interpretable manner for doing such a prediction exercise. Rather than create bespoke indices of peer firms, we use a penalized regression model to predict the return on a given target stock based on the returns on the market index and the returns of each individual peer firm. There are multiple ways to penalize the inclusion of additional factors in the model—from lasso to ridge and the elastic net, which is a combination of the two—and the penalization parameters can be optimized using cross-validation or leave-one-out prediction error.[151] The advantage of this approach in the context of securities litigation is that it transforms the debate from a relatively subjective one (what is the correct industry and set of peers based on the business attributes of the company) to a comparatively objective one (which combination of firms and weights seems to best predict the return of the stock during the estimation period using out-of-sample prediction methods).

Table 2 shows the model coefficients from different forms of penalization for the event study in the Halliburton case. Similar to the Allen and Coffman models, I use the class period as the estimation period, omitting the days where the plaintiffs allege an affirmative misstatement or a corrective disclosure was made, and include the returns on the S&P Energy Index and all of the firms with a full trading data over the period that enter either of the two indices used by Coffman in his report. Leave-one-out cross validation determines the penalty value λ that minimizes the root mean squared prediction error over this period. The first column reports the value from using the L_1 norm as a penalty parameter

151. Baker & Gelbach, *supra* note 100, at 245–46.

(or penalizing the absolute value of size of the coefficient on each index or firm return); this type of model is typically used for model selection and will "shrink" the estimated coefficients towards zero. As shown in the table, the lasso model drops the returns of twenty of the twenty-nine potential peer firms. The second column reports the estimated coefficients from the ridge model, which uses the L_2 norm, penalizing the square of each coefficient; ridge models tend to shrink each of the estimated coefficients towards each other rather than towards zero, and we see far fewer firms dropping entirely out of the model. The elastic net model finds the optimal combination of each form of penalization; in this case the optimal combination value, called α, is equal to 0.1, so the elastic net and lasso models generate very similar estimated coefficients.

Table 2: Penalized Regression Weights on Peer Firms

Index/Company	Lasso	Ridge	Elastic Net
(Intercept)	0.00	0.00	0.00
S&P Energy Index	0.24	0.24	0.24
Baker Hughes	0.15	0.13	0.14
Beazer Homes	0.00	-0.02	0.00
BJ Services	0.12	0.11	0.11
Centex	0.00	0.01	0.00
Champion Enterprises	0.00	0.00	0.00
Clayton Homes	0.00	0.01	0.00
Comfort Systems	0.03	0.04	0.04
Cooper Cameron	0.07	0.10	0.09
DR Horton	0.00	-0.01	0.00
Emcor Group	0.00	0.00	0.00
Fluor	0.00	0.01	0.00
Foster Wheeler	0.00	0.01	0.00
Granite Construction	0.00	-0.01	0.00
IT Group	0.00	0.01	0.00
Jacobs	0.00	-0.03	0.00
Lennar	0.00	-0.03	0.00
McDermott Intl	0.00	0.02	0.01
MDC Holdings	0.01	0.03	0.02
NVR Inc	0.00	0.01	0.00
Oakwood Homes	0.00	0.01	0.00
Oceaneering	0.04	0.06	0.05
Pulte	0.00	0.03	0.01

Ryland Group	0.00	0.01	0.00
Schlumberger Ltd	0.23	0.18	0.20
Smith Intl	0.07	0.10	0.09
Standard Pacific	0.00	0.00	0.00
Toll Brothers	0.00	-0.01	0.00
URS Corp	0.00	0.02	0.01
Weatherford Intl	0.13	0.11	0.12

This table reports the coefficient values on the peer firms and energy index using different forms of penalized regression. The outcome variable is the log return for Halliburton and the features that enter the regression are the log returns on the index and the peer firms. We use daily data over the class period, omitting the misstatement and disclosure dates, and optimize the tuning parameter using leave-one-out cross-validation.

The coefficient values from Table 2 can be used to predict the returns for the target firm during the event window. Table 3 reports the predictions and confidence levels from the expert reports, along with the corresponding predictions and confidence from the statistical learning models. One noteworthy feature of regularization-based estimates is that they are very close to each other, generating predicted excess returns of -3.2% to -3.3%, regardless of how you shrink the estimates. This is an advantage of a data-driven approach to conducting an event study: the models will typically pick up some low-dimensional set of factors that predict returns, rather than over-fitting based on the subjective design choices made by experts. From the perspective of a fact finder, this is particularly appealing: they can focus their attention on how the question is framed and the answer will be driven by the data, not by expert discretion. In this case, the statistical learning models produce estimates that are close to, but smaller (in magnitude) than the Coffman model. The statistical significance for the estimate is slightly above 5%, which is higher than the threshold set by many, but not all, courts.[152]

152. It's not clear that 5% is the correct benchmark for litigation purposes at any event. *See* Jonah Gelbach, *Estimation Evidence*, 168 U. PA. L. REV. 549, 564 (2020).

Table 3: Expert and Statistical Learning Predictions

	Expert Predictions		Statistical Learning Approach		
	Allen (D)	Coffman (P)	Lasso	Ridge	Elastic Net
Excess Return	-0.029	-0.037	-0.033	-0.032	-0.033
t-Statistic	-1.290	-2.270	-1.900	-1.835	-1.922
p-value	0.198	0.024	0.058	0.067	0.055

This table presents the excess return calculations using the values from the expert reports, as well as the data-driven predictions from the statistical learning models. I report the excess return, as well as the associated t-Statistic and p-value. The optimal value of α for the elastic net model is 0.1.

6.2 Valuation Disputes

Litigation of firm valuation, or adjudicating disputes over the fundamental value of a firm, is another area of common disagreement among experts that frequently calls for judicial oversight. Initially founded in corporate and securities litigation, financial valuation now plays an increasingly pivotal role in nearly all areas of high-stakes commercial litigation.[153] As a result, in some litigation areas, much of the judicial burden in commercial litigation has become, dominated by valuation disputes that hinge on complex financial economics—including bankruptcy, tax, family law, fiduciary duties, and garden-variety questions in tort, property and contract law.

6.2.1 The Use of Valuation in Commercial Litigation

In both courtrooms and boardrooms, financial valuation is primarily driven by three competing methodologies: Comparable Companies (CC), Comparable Transactions (CT), and Discounted Cash Flow (DCF) analyses. These approaches are commonly used, often in conjunction, to assess the value of a company or financial asset, especially in complex merger litigation and bankruptcies. Experts occasionally also employ other techniques—such as historical premium analysis, analyst forecasts, or leveraged buyout evaluations—but these are typically supplemental to the three core methods.

153. Baker et al., *supra* note 100, at 3.

Comparable Companies (CC)[154]

The Comparable Companies method does largely what it says: it uses publicly available financial data from actively traded companies to generate a counterfactual valuation for a target firm. This method compares a company's financials to those of similar firms that are publicly listed and traded, offering an advantage in terms of data availability when compared to other approaches (like Comparable Transactions). Stock prices are often considered a proxy for a company's economic value, providing a robust dataset for generating comparable firm valuations.

The CC process begins by identifying comparable firms in the same industry, of similar size, and with similar capital structures. Analysts then convert the firm's valuation to enterprise value and apply valuation multiples—most commonly the ratio of the "enterprise value" (EV) to earnings before interest, tax, depreciation, and amortization (EBITDA). EV represents the total value of a firm's equity and debt, accounting for differences in financing structures. The key analytical difficulty, from the perspective of a neutral fact finder, is that CC engenders a substantial amount of discretion, even in applying the selection of the appropriate multiple. The options for multiples include the last fiscal year's earnings, the last twelve months, or projections of future earnings.[155]

An advantage of the CC approach is the volume of available data. Since stock prices for publicly traded companies are observable daily, large sets of comparable companies can be built, unlike the more limited transaction data available in the CT approach. However, the CC method anchors a company's valuation to stock market prices, which may not reflect intrinsic value, particularly in illiquid or volatile markets. Additionally, when CC is used to value private companies, the liquidity premium that comes with publicly traded firms must be considered. Most importantly for our purposes, there is little guidance on how to identify potential peer firms, how many peer firms to consider, and how to calculate the target ratio from the identified peers.

Comparable Transactions (CT)[156]

The Comparable Transactions approach mirrors how real estate appraisers use recent home sales to estimate property value. The idea is to identify analogous assets that were recently sold under similar conditions and use those sale prices to estimate the value of the company in question. For companies, this means looking at sales of firms in similar industries, regions, or with similar capital structures.

154. For a longer discussion of the Comparable Companies methodology *See* Baker et al., *supra* note 100, at 11-12.

155. *See id.*

156. For a longer discussion of the Comparable Transaction approach, *see generally id.* at 8–11.

The first step in CT is to identify the appropriate comparable firms. Ideally, these are companies of similar size, industry, and capital structure. Analysts then adjust the purchase prices to reflect both the equity and debt structure of the firms, converting the sales price into enterprise value to standardize comparisons. Analysts will then typically apply valuation multiples to normalize the data after calculating the enterprise value. The most common metric is again the EV/EBITDA multiple, which provides a proxy for cash flow. For less mature companies, other metrics, such as revenue multiples, may be used. However, EBITDA-based multiples are favored in most cases, as they are considered more reliable for mature firms. Normalizing the value of the firm by a measure of profits allows an expert to generate comparisons for a target firm even among comparable firms of different scale.

The CT approach faces two notable constraints. First, finding sufficient data can be challenging, as genuine arm's-length sales within a particular industry may be rare, forcing analysts to work with a small pool of comparable companies. Second, transaction prices often include a control premium—the added value paid for acquiring a controlling interest in the company. This control premium can distort the pure cash flow value of the company, and analysts must adjust for it if the valuation's purpose is to exclude such a premium (as in an appraisal action).

Discounted Cash Flow (DCF)

The DCF approach diverges from the comparative nature of CC and CT by focusing on the company's expected future cash flows. Instead of looking for similar firms, the DCF model estimates the intrinsic value of a company by calculating the present value of its future free cash flows, discounted at an appropriate rate to account for risk. The DCF formula can be expressed as follows:

$$FMV = PV(Cash\ Flows) = \sum_{t=1}^{T} \frac{FCF_t}{(1+WACC)^t} + \frac{S_T}{(1+WACC)^t}$$

Here, FCF_t represents projected free cash flows, S_T is the terminal value at the end of the forecast horizon, and *WACC* is the Weighted Average Cost of Capital, a risk-adjusted discount rate. [157]

DCF models require careful forecasting of cash flows, often based on internal company projections, management estimates, or external financial forecasts. These projections typically cover a period of 5-10 years, after which a terminal value is calculated to represent the company's remaining value. The terminal value can be determined by assuming the firm will grow indefinitely at a constant rate (using the growing perpetuity formula) or by reverting to a valuation

157. *See, e.g.*, In re Vanderveer Ests. Holding, LLC, 293 B.R. 560, 578 (Bankr. E.D.N.Y. 2003).

multiple based on comparable companies, effectively blending CC and DCF methods.

The DCF approach offers a more fundamental analysis of a company's value but is also more technically demanding and sensitive to the assumptions used for cash flow projections, discount rates, and terminal values.[158] Each component of the DCF model introduces its own complexities. For example, determining the appropriate discount rate requires careful estimation of the company's cost of equity and debt, often derived from asset pricing models such as the Capital Asset Pricing Model (CAPM). Similarly, cash flow projections can be influenced by broader market trends or company-specific factors.

Summary

Each of the three valuation methodologies—CT, CC, and DCF—provides different insights and comes with its own set of challenges. CT and CC offer market-based valuations but can be constrained by data availability and the need for careful adjustments, such as removing control premiums. DCF, while offering a more granular and intrinsic valuation, requires complex forecasting and careful discretion in applying assumptions. While DCF is often viewed as the "gold standard" in valuation practice for litigation, it is not at all clear that this presumption is warranted. One academic has argued that DCF "is a speculative exercise disguised in the trappings of mathematical rigor but squarely within the domain of pseudoscience."[159] Moreover, there is substantial evidence that the actual valuation of firms in the market is done through comparing multiples—essentially the Comparable Companies analysis—rather than discounting cash flows.[160] In practice, analysts often use a combination of these approaches to create a more comprehensive valuation, as each method compensates for the limitations of the others.

6.2.2 In re Mirant Corp.

Mirant Group was a company that produced and marketed electric power, and their revenue was largely derived from long-term contract sales of power to utilities and from sales of power and capacity in the wholesale energy market. Most of the company's facilities were put in operation while Mirant Group was controlled by its parent-firm TSC. Unfortunately, the company overbuilt its generation facilities and found itself in financial straits following a downturn in

158. *See* Verition Partners Master Fund Ltd. v. Aruba Networks, Inc., 210 A.3d 128, 136–137 (Del. 2019) ("Dell's references to market efficiency focused on informational efficiency—the idea that markets quickly reflect publicly available information and can be a proxy for fair value . . . .").

159. J.B. Heaton, *Why Does Pseudoscience Still Thrive Under Daubert? The Case of Discounted Cash Flow Valuation*, ONE HAT RESEARCH LLC (Oct. 14, 2024), https://ssrn.com/abstract=4976642.

160. Itzhak Ben-David & Alex Chinco, *Expected EPS × Trailing P/E.* (Nat'l Bureau of Econ. Rsch., Working Paper No. w32942, 2024).

the energy market in 2001 and 2002. Its debtors sought relief under Chapter 11 of the Bankruptcy Code after the company failed to accomplish an out-of-court workout with their creditors.[161]

The debtors proposed a restructuring plan based on the assumption that unsecured creditors would not receive full satisfaction from the enterprise value of Mirant Group. The equity holders of Mirant were to receive only the *potential* right to receive distributions after paying off the Mirant creditors and other beneficiaries of subordinated debt. The equity committee for the shareholders filed a complaint contending that debtors had undervalued the firm in their plan, directly to the harm of existing shareholders. Given the latent dispute over valuation, the court called a valuation hearing with interested parties.[162]

The valuation hearing lasted for 27 days over 11 weeks, and included numerous expert reports, with the parties placing into evidence a total of 454 exhibits.[163] Following the hearing, the court adjudicated the merits of the competing reports and ordered that the value of Mirant Group be recalculated in accordance with its stipulated changes. In ordering the re-valuation of the firm, the court registered an exasperation with the practice of valuation in litigation and felt the need to comment "on the questionable reliability of [the] valuation methods."[164] In noting its disapproval, the court cited prior judicial claims about the limitations of the valuation exercise, which rested less on scientific certitude than subjective judgments.[165] According to the court:

> "[a]t best, the valuation of an enterprise like Mirant Group is an exercise in educated guesswork. At worst, it is not much more than crystal ball gazing. There are too many variables, too many moving pieces in the calculation of value of Mirant Group for the court to have great confidence that the result of the process will prove accurate in the future. Moreover, the court is constrained by the need to defer to experts and, in proper circumstances, to Debtors' management."[166]

While the court admittedly had misgivings about the accuracy of valuation analysis, "let alone a valuation subject to inherent methodological weaknesses and assumptions unsupported by history," they felt constrained by the law and their comparative disadvantage at the task. At the end of the day, expert

161. *In re* Mirant Corp., 334 B.R. 800, 806 (Bankr. N.D. Tex. 2005).

162. *See id.* at 807.

163. *Id.* at 809–10.

164. *Id.* at 818.

165. *Id.* (quoting *In re* Beker Indus. Corp., 58 B.R. at 739 ("[Deciding] going concern value is hardly elementary. It involves consideration of what Shelley in 'A Defense of Poetry' called 'the gigantic shadows which futurity casts upon the present.' Those who would prepare future cash flow analyses and discount them to present values are not oracles. The opinion evidence they present . . . should be taken as a set of assumptions that are factored into a model and critical analysis then employed to test those assumptions. The evidence in the exercise is hardly clear, is highly judgmental and consists largely of inferences.").

166. *Id.* at 848.

testimony and conventional valuation approaches were "the tools available to the court in its task." [167] Although the result of the valuation exercise would inevitably be uncertain and "soft," the court needed to exercise its discretion in establishing some range of values for Mirant Group that would inevitably include or exclude equity participation under the proposed bankruptcy plan. [168]

The court refused to take an average of the two valuation estimates, because the range of values was simply too large, from $7.2 billion (Houlihan for the debtors) to $13.6 billion (PJSC for the equity committee). "[F]or the court to simply average these numbers—derived based on varying assumptions and data—would make a mockery of the valuation process and would be terribly unfair to parties whose rights are thereby disposed of."[169] Consequently, the court concluded that the parties had to recalculate the value of Mirant Group based on stipulated changes in data and assumptions. For the equity participants to get any recovery in the bankruptcy, the valuation estimates for Mirant Group had to reach or exceed $11 billion.[170]

6.2.3 A Better Way to Value Firms in Litigation

Baker, Gelbach, and Talley (2024)[171] proposes an alternative valuation process, building off the Comparable Companies approach, that uses statistical learning to automate the subjective portion of the valuation process. Rather than have experts disagree about which of a group of peers is truly a "comparable firm" for the target, we use a data-driven procedure to select peers based on the objective ability of the comparable set to predict the target firm's valuation in a clean period. Like in the event study context, we use the weights from this exercise to create a counterfactual value as of the valuation date. It is noticeable that a valuation approach based on penalized regression is precisely the type of weighted estimate that the *Mirant* court suggested would be appealing:[172]

> In this regard[,] the court is compelled to note that [the] weighting of comparable companies based on their similarity to the subject being valued [has] some appeal. The experts the court questioned about this rejected the idea, and the court therefore will not adopt such an approach; it may be that raising the question here will prove useful in future valuations.

Both sides in *Mirant* issued expert reports that used Comparable Companies analysis to value the firm. Blackstone issued a report for the debtors, and selected

167. *Id.* at 820.
168. *Id.*
169. *Id.* at 824.
170. *Id.* at 820.
171. Baker et al., *supra* note 100, at 6.
172. *In re* Mirant Corp., 334 B.R. at 838.

four comparable: AES, Reliant, NRG, and Dynegy.[173] The equity committee urged the court to also consider Calpine as a peer, which the court ultimately declined to do because of its "precarious financial condition" that made its stock price more of an option than "a true reflection of equity value."[174] In addition, multiple witnesses testified that operational differences between Calpine and Mirant Group were "sufficient, when considered together with Calpine's relatively weak financial condition, to disqualify its use as a comparable."[175] The court also noted that AES was substantially larger than Mirant and had more of an international focus, and that the countries where AES operated were generally more stable.[176] The court entertained using NRG, clearly the closest competitor, as the sole comparable, but ultimately held that it did not believe it appropriate to rely solely on one company in formulating a value by the Comparable Method.[177]

This discussion reflects the limitations to gatekeeping a dispute over Comparable Companies analysis. There is very little information to guide a judge or jury in how to consider which expert has selected a more appropriate peer set of firms. In part, this is because the objective is not clear. The similarity between firms is relevant only insofar as it assists in predicting the valuation, or valuation multiple, of the target firm. Lengthy investigations into the similarity of business lines, geographic regions, and financial position are at best a questionable use of scarce court time, and at worst a hopeless diversion from the true underlying question.

As in the event study exercise above, we estimate the valuation of Mirant from regularized regressions with its peers, using a data-driven procedure to select the peers and their weights. Instead of using firm returns as the outcome variable, we use the firm's market capitalization (the product of equity price and shares outstanding).[178] A predicate decision under this approach is to determine an estimation window for the model. Determining the window would be an appropriate exercise for the court to decide after relevant testimony from the experts, as it involves selecting a period where the valuation is untainted by the allegations in the complaint, but which is close enough in time to the valuation date for the weights to remain accurate. Given data limitations in this case with peer firms also entering bankruptcy themselves, the period used was from May 1, 2001, to December 31, 2001 to get the estimated weights.[179]

173. *Id.* at 836.

174. *Id.* at 837.

175. *Id.*

176. *Id.*

177. *Id.* at 837–38.

178. In Baker et al., *supra* note 100, at 42, we show how one can use returns and an event study framework to calculate equity market value. However, in this example the length of time between the estimation window and valuation date is long enough that we stick with market capitalization as the outcome variable.

179. Mirant went into bankruptcy protection in mid-2003. *In re* Mirant Corp., 334 B.R. at n.10.

The results of this exercise are reported in Table 4. The model intercept captures the expected valuation of Mirant if the peer firms went to zero. The other values reflect the marginal increase (in thousands of dollars) for Mirant's equity that arises from a thousand-dollar increase in the peer. As mentioned earlier, the lasso regression model is typically used for model selection, as it will tend to "drop" predictors that don't sufficiently explain the outcome. Given the dispute regarding the inclusion (or exclusion) of Calpine as a peer, it is noteworthy that the lasso model does not drop the firm, suggesting that it *does* help in explaining Mirant's valuation. However, AES, perhaps for the reasons explained by the court, is given zero weight and thus is arguably not a useful peer for valuation purposes.

Table 4: Penalized Regression Weights on Peer Firms (Thousands)

Company	Lasso	Ridge	Elastic Net
(Intercept)	$757,561.82	$554,096.59	$345,549.94
AES	$0.00	$0.04	$0.03
Calpine	$0.38	$0.19	$0.25
NRG	$1.01	$1.76	$1.61
Reliant	$0.12	$0.19	$0.14
Dynegy	$0.38	$0.40	$0.43

This table reports the coefficient values on the peer firms using different forms of penalized regression for the valuation of Mirant. The outcome variable is the market capitalization for Mirant and the features that enter the regression are the market capitalization values for the peer firms. I use daily data for Mirant and the peer firms from May 1, 2001 to December 31, 2001, and optimize the tuning parameter using leave-one-out cross validation. The units are in thousands of USD.

Table 5 reports the predicted equity valuation for Mirant using the regularized models. In this example, given that the outcome variable (valuation) is in levels rather than returns and the long period of time between model estimation and valuation, the valuation range is substantially larger than in the event study example, with a lower bound of $5.7 billion and an upper bound of $8.3 billion. After adding back in Mirant's last reported debt levels before bankruptcy of $3.7 billion, these valuation estimates suggest a total enterprise value range of $9.4 to $12.0 billion. The range is between the two values provided by each respective side and could potentially support a (small) recovery for the plaintiffs.

Table 5: Statistical Learning Predictions for Equity Value on June 27, 2005 (Millions)

	Lasso	Ridge	Elastic Net
2005-06-27	$5,703.05	$8,281.79	$7,420.19

This table presents the predicted market capitalization using the data-driven predictions from the statistical learning models. The units are in millions of USD.

6.3 Other Areas of Litigation

In our prior work, we focused on securities litigation and valuation as two areas where statistical learning in expert testimony would work well. However, for similar reasons, its use would benefit the court in other practice areas that frequently rely on expert testimony, including:

Employment Discrimination: Experts are frequently engaged in employment discrimination disputes to support or dispute the presence of illegal pay disparities within firms, universities, or government agencies.[180] In such cases, the expert for the plaintiffs will often use regression analysis to demonstrate that there are unexplainable differences between the wages of, for example, black and white employees at a firm. "In effect, the regression controls for the explanatory variables—those factors that one would expect to influence pay—and then compares the wages of white and black employees."[181] The defendants will typically hire their own expert, who will often argue that the plaintiff's regression failed to control for a critical variable that determines wages.[182] Courts understandably struggle to determine whether the experts have controlled for the "major factors" that determine the wage structure,[183] and the use of a principled approach to variable selection would make the court's jobs easier.

A point of caution is warranted here—frequently the experts will also disagree about whether a given control variable is "tainted" by the same

180. *See* Joni Hersch & Blair Druhan Bullock, *The Use and Misuse of Econometric Evidence in Employment Discrimination Cases*, 71 WASH. & LEE L. REV. 2365, 2368 (2014).

181. Morgan v. United Parcel Serv. of Am., Inc., 380 F.3d 459, 466 (8th Cir. 2004).

182. *See, e.g.*, *id.* at 466 ("All three experts performed regression analyses, and all agreed that this form of statistical analysis was proper. But the experts came to different conclusions because each of them included different explanatory variables."); Dukes v. Wal-Mart Stores, Inc., 222 F.R.D. 137, 159 (N.D. Cal. 2004) ("Defendant also contends that Dr. Drogin's statistical analysis should be rejected because it fails to account for a variety of factors, or control elements, that could be responsible for the disparities in question—referred to as 'omitted variable bias."'); Melani v. Bd. of Higher Educ. of City of New York, 561 F. Supp. 769, 779 (S.D.N.Y. 1983) ("Finally, defendant claims that plaintiffs' regression analyses are flawed by their failure to include a variable reflecting academic department and thereby to account for differing market conditions characterizing each department.").

183. Bazemore v. Friday, 478 U.S. 385, 400 (1986).

discriminatory practices that drove the plaintiff's complaint.[184] If so, including the tainted or "inappropriate variable" in the regression will bias the analysis against finding a discriminatory effect, even if one were to exist.[185] This type of "bad controls" problem is a challenge for the design and interpretation of any empirical analysis of discriminatory effect, and nothing inherent to statistical learning solves the problem. Some have even argued that it can make it worse.[186] These statistical learning techniques are not designed to displace experts in the litigation process—but merely direct the court's attention to more fruitful avenues of investigation like this based on actual institutional knowledge of the causal question at issue.

Antitrust: Economists are almost always retained in cases brought under the Sherman Antitrust Act.[187] While expert testimony in this area does not necessarily rely on simple regression analysis, it sometimes does. For example, in *In re Processed Egg Products Antitrust Litigation*, the expert for the plaintiffs used a regression model to determine the effect of an anticompetitive conspiracy on the price of eggs.[188] Defendants challenged the expert's testimony on the grounds that it failed to control for many of the important factors that drive the price of eggs.[189] While the court refused to exclude the testimony in that case, it cited others where the failure to control for relevant factors was so significant as to render the entire analysis unreliable.[190] Again, determining which variables are "critical" to control for in a regression analysis used for an adversarial proceeding is a deeply subjective and challenging task, which would at minimum be aided by the results of a statistical technique designed for the task.

Death Penalty Litigation: In a non-commercial setting, the use of statistical learning would also assist the fact finder in certain constitutional challenges to the practice of the death penalty. Studies have shown that race often influences sentencing outcomes, with defendants of color disproportionately receiving

184. *See* Valentino v. U.S. Postal Serv., 674 F.2d 56, 73 at n.30 (D.C. Cir. 1982) ("Absent clear, affirmative evidence that promotions were made in accordance with neutral, objective standards consistently applied, there is no assurance that level or rank is an appropriate explanatory variable, untainted by discrimination.").

185. *See, e.g.*, Michael O. Finkelstein, *The Judicial Reception of Multiple Regression Studies in Race and Sex Discrimination Cases*, 80 COLUM. L. REV. 737, 738–42 (1980).

186. Paul, Hünermund, Beyers Louw, & Itamar Caspi, *Double Machine Learning and Automated Confounder Selection: A Cautionary Tale*, 11 J. OF CAUSAL INFERENCE 1, 2 (2023).

187. 15 U.S.C. §§ 1-7 (2024).

188. 81 F. Supp. 3d 412, 429 (E.D. Pa. 2015).

189. *Id.* at 430.

190. *Id.* at 431–32 (citing to Multimatic, Inc. v. Faurecia Interior Sys. USA, Inc., 358 F. App'x 643, 654 (6th Cir. 2009) ("Perceived flaws in an expert's opinion go to weight only if they fall within the accepted norms of the discipline and have a non-speculative basis in fact."); Blue Cross & Blue Shield United of Wisconsin v. Marshfield Clinic, 152 F.3d 588, 592 (7th Cir. 1998) ("Any nonconspiratorial factors *likely* to have made the prices change . . . had to be taken into account"); *In re* Wireless Tel. Servs. Antitrust Litig., 385 F. Supp. 2d 403, 428 (S.D.N.Y. 2005) ("[Expert's] failure to test for these *obvious* and *significant* alternative explanations renders [expert's] analysis essentially worthless.").

death sentences, especially when the victim is white.[191] Geographic inconsistencies further highlight that capital punishment is applied unevenly across jurisdictions, raising concerns about arbitrary enforcement.[192] These empirical findings have been instrumental in shaping legal arguments and judicial scrutiny, as seen in cases like *Furman v. Georgia*[193] and *McCleskey v. Kemp*,[194] where data-driven insights on arbitrariness and bias formed the basis of constitutional challenges.[195] When providing evidence of the discriminatory impact of the death penalty, experts typically need to control for "an array of legitimate factors relevant to the crime," which again can be aided by using a disciplined manner to select the required control variables.[196]

7. CONCLUSION

Expert testimony plays a crucial role in modern litigation, bridging the gap between technical expertise and legal decision-making. However, its use is not without significant challenges. The historical evolution from the presumption of professional agreement and alignment to the rigorous scrutiny required under the *Daubert* standard reflects the ongoing effort of the judiciary to ensure reliability and integrity in expert evidence. While advancements such as the adoption of statistical learning and other objective methodologies hold promise for reducing expert discretion and partisan bias, courts still face obstacles in implementing these innovations. The complexity of legal disputes, particularly in commercial litigation, demands that judges engage deeply with technical methodologies—an expectation that strains judicial capacity and resources.

To address these challenges, reforms must focus on enhancing judicial tools for evaluating expert testimony, encouraging collaboration between professional organizations and the judiciary, and leveraging modern analytical techniques. This article proposes one actionable framework for improving the reliability of expert evidence in high-stakes litigation—shifting from a model-driven to a data-driven approach to uncovering relationships in data. Although the adversarial system inherently complicates efforts to standardize expert practices, targeted reforms that align with evidentiary standards and judicial goals can pave the way for more transparent, consistent, and equitable outcomes. By embracing these

191. *E.g.*, Catherine M. Grosso, Jeffrey Fagan, & Michael Laurence. *The Influence of the Race of Defendant and the Race of Victim on Capital Charging and Sentencing in California*, 21 J. EMP. L. STUD. 482, 503–05 (2024).

192. John J. Donohue III, *An Empirical Evaluation of the Connecticut Death Penalty System Since 1973: Are There Unlawful Racial, Gender, and Geographic Disparities?*, 11 J. EMP. L. STUD. 637, 637 (2014).

193. *See* 408 U.S. 238, 256–57 (1972) (Douglas, J., concurring).

194. *See* 481 U.S. 279, 279, 289 (1987).

195. *E.g.*, *id.* at 286–87.

196. Donohue III, *supra* note 192, at 637.

innovations, the judiciary can better fulfill its gatekeeping role and foster greater trust in the legal process.