



The Negative Effect Fallacy: A Case Study of Incorrect Statistical Reasoning by Federal Courts

Ryan D. Enos, Anthony Fowler,* and Christopher S. Havasy

This article examines the negative effect fallacy, a flawed statistical argument first utilized by the Warren Court in *Elkins v. United States*. The Court argued that empirical evidence could not determine whether the exclusionary rule prevents future illegal searches and seizures because “it is never easy to prove a negative,” inappropriately conflating the philosophical and arithmetic definitions of the word *negative*. Subsequently, the Court has repeated this mistake in other domains, including free speech, voting rights, and campaign finance. The fallacy has also proliferated into the federal circuit and district court levels. Narrowly, our investigation aims to eradicate the use of the negative effect fallacy in federal courts. More broadly, we highlight several challenges and concerns with the increasing use of statistical reasoning in court decisions. As courts continue to evaluate statistical and empirical questions, we recommend that they evaluate the evidence on its own merit rather than relying on convenient arguments embedded in precedent.

I. INTRODUCTION

The law has an uneasy relationship with statistical and scientific evidence.¹ Legal history is ripe with examples of lawyers and judges relying on controversial or incorrect interpretations of empirical evidence. Perhaps most famously, the Supreme Court in *Brown v. Board of Education*² relied on dubious psychological evidence in order to justify its holding that racial segregation violates the Equal Protection Clause of the Fourteenth

*Address correspondence to Anthony Fowler, 1155 E. 60th St., 165, Chicago, IL 60637; anthony.fowler@uchicago.edu. Enos is Associate Professor, Department of Government, Harvard University; Fowler is Associate Professor, Harris School of Public Policy, University of Chicago; Havasy is Ph.D. Candidate, the Department of Government, Harvard University.

We thank Scott Ashworth, Richard Fallon, Jon Gould, Cody Gray, Pablo Montagnes, Nick Stephanopoulos, Matthew Stephenson, and Susannah Barton Tobin for insightful comments and helpful guidance.

¹See David L. Faigman, To Have and to Have Not: Assessing the Value of Social Science to the Law as Science and Policy, 38 Emory L.J. 1005, 1008 (1989). Cf. David Reisman, Some Observations on Law and Psychology, 19 U. Chi. L. Rev. 30, 32 (1951).

²347 U.S. 483 (1954).

Amendment.³ Although some have argued that this psychological evidence was largely inconsequential to the Court's decision in *Brown*,⁴ the Court's apparent reliance on it has been widely criticized.⁵

Perhaps this state of affairs between law and statistics is not surprising; a standard legal education does not include rigorous training in statistics or the evaluation of scientific evidence. Federal judges and their clerks bring their limited statistical experience with them to the bench, and they have little incentive to develop these skills. As a result, statistical errors in one case can propagate to others through precedent. Allison Orr Larsen finds that legal reasoning based on empirical information can form "factual precedents" through "the tendency of lower courts to over-rely on Supreme Court opinions and to apply generalized statements of fact from old cases to new ones."⁶

In the case of the flawed psychological evidence used in *Brown*, the wider legal profession eventually corrected itself—similar research is no longer used as factual precedent by federal courts. But other mistakes in scientific reasoning have received less scrutiny by the legal profession, despite their widespread consequences. In this article, we examine an incorrect statistical argument from an exclusionary rule case in 1960,⁷ which then propagated over time and across various different legal domains. In subsequent decades the same incorrect reasoning was applied repeatedly by the Supreme Court to cases regarding the exclusionary rule, voting rights, and free speech.⁸ Most recently, the Court used the argument to justify its ruling in *Arizona Free Enterprise*, in which the Court struck down the matching funds provision of Arizona campaign finance laws.⁹ Proliferating further, lower federal circuit courts have widely adopted this fallacy across different legal domains.¹⁰

The specific error in question involves the empirical task of "proving a negative" or, more appropriately, determining whether a law in question decreases a particular

³Id. at 494 n.11.

⁴See James E. Ryan, The Limited Influence of Social Science Evidence in Modern Desegregation Cases, 81 N.C. L. Rev. 1659, 1665 (2003); Jack M. Balkin, Rewriting *Brown*, in Jack M. Balkin and Bruce A. Ackerman, eds., *What Brown v. Board of Education Should Have Said* 51 (2002).

⁵For early criticisms of the footnote, see Charles L. Black, Jr., The Lawfulness of the Segregation Decisions, 69 Yale L.J. 421, 427 (1960); Edmund Cahn, Jurisprudence, 30 NYU L. Rev. 150, 167 (1955). For recent discussions of the controversy surrounding the footnote, see Michael Heise, *Brown v. Board of Education*, Footnote 11, and Multidisciplinarity, 90 Cornell L. Rev. 279, 292–95 (2005); Sanjay Mody, Note, *Brown* Footnote Eleven in Historical Context: Social Science and the Supreme Court's Quest for Legitimacy, 54 Stan. L. Rev. 793, 803–09 (2002).

⁶Allison Orr Larsen, Factual Precedents, 162 U. Pa. L. Rev. 59, 62 (2013).

⁷*Elkins v. United States*, 364 U.S. 206, 218 (1960).

⁸See Section III for a list of these cases.

⁹*Arizona Free Enter. v. Bennett*, 131 S. Ct. 2806, 2823 (2011).

¹⁰See Section III for a list of these cases.

outcome of interest. For example, as was the case in *Arizona Free Enterprise*, a court may want to know whether a campaign finance law inhibits various forms of private political speech. Social scientists have developed empirical tools for answering these kinds of questions, but the federal courts have ignored this evidence by writing “it is never easy to prove a negative.” Of course, estimating the effects of laws is difficult, but there is no reason that negative effects are harder to detect than positive ones. Courts have conflated the philosophical and arithmetic definitions of the word *negative*. Although it is often difficult to prove that something does not exist (e.g., can we *prove* that Santa Claus does *not* exist?), there is no reason that we cannot show evidence that the effect of a law on an outcome of interest is arithmetically negative. We refer to this error as the “negative effect fallacy.” In short, relevant statistical evidence has been repeatedly ignored by federal courts because of an elementary but contagious error of language and logic.

As we discuss, the negative effect fallacy appears to have several adverse consequences. In some cases, the fallacy may have been consequential for the decision, as it would otherwise have been difficult for judges to square their assertions with the empirical evidence that the fallacy allowed them to dismiss. And even when the fallacy is not pivotal in a decision, perhaps because ideologically motivated judges use the fallacy when convenient for their preferred conclusions, it still allows judges to ignore relevant evidence and obscure the true rationale for their decisions. In addition, every time a judge utilizes the fallacy, he or she further expands its reach into precedent, making it easier for future judges to propagate these adverse effects into other cases. For these reasons, we believe the problem identified in this article is a serious one that should be eradicated.

Beyond the specific topic of the negative effect fallacy, our investigation highlights several challenges associated with the increasing use of statistical evidence in federal courts. After explaining and documenting the history of the negative effect fallacy, we discuss several potential reasons for why judges may make statistical mistakes, and we provide several recommendations for avoiding similar mistakes in future cases.

II. EXPLAINING THE NEGATIVE EFFECT FALLACY

Many have heard the adage that *you can't prove a negative*. One might prefer a weaker version of the statement such as *it is difficult to prove a negative*. What do we mean when we say this? Typically, we are referring to the idea that some statements are harder to prove than others, and negative statements are often of this sort. Consider the following two statements:

1. A Jewish person was at the party.
2. No Jewish person was at the party.

The first statement is clearly easier to prove than the second. The proof of (1) would require that we find one Jewish person who was at the party. The proof of (2), on the

other hand, would require that we assemble a complete list of everyone at the party and confirm that none of them are Jewish. Alternatively, we could account for the whereabouts of every Jewish person in the world at the time and confirm that none were at the party. The first statement is positive and the second statement is negative, where *negative* means that the statement is characterized by the absence or nonexistence of something rather than its presence or existence. For convenience, we will refer to this as the *philosophical* definition of the word *negative*. These are the kinds of examples we have in mind when we say that it is difficult to prove a negative.

Philosophers, logicians, and linguists will point out that there is nothing special, in general, about positive and negative statements.¹¹ Consider two more statements:

3. Not everyone at the party was a gentile.
4. Everyone at the party was a gentile.

Statements (3) and (4) are logically equivalent to (1) and (2), respectively, but now (3) is negative and (4) is positive. Clearly, we cannot say that positive statements are universally easier to prove than negative statements. Furthermore, the statement that it is difficult to prove a negative is not a good rule of thumb because we could always rewrite a positive statement as a negative one and vice versa with enough grammatical acrobatics. Nonetheless, the notion that proving a negative is difficult is common in our rhetoric. The most favorable interpretation of the adage is that it reminds us that inductive reasoning does not produce certain conclusions. We can never be absolutely certain that Santa Claus and flying unicorns do not exist, although we might be highly confident based on theory and evidence. Similarly, positive conclusions reached through induction are also uncertain. We cannot be sure that the sun will rise tomorrow, although we might conclude that this phenomenon is highly likely given previous observations and our understanding of physics.

To the extent that there is any usefulness to the adage that proving a negative is difficult, it disappears entirely if we use the *arithmetic* definition of *negative* rather than the *philosophical* definition. Consider the following two statements, which are arithmetically positive and negative, respectively.

5. You will make money, in expectation, if you play the lottery.
6. You will lose money, in expectation, if you play the lottery.

How would we go about proving statement (5)? We would calculate the expected value of the lottery by enumerating every possible outcome, multiplying the net earnings for each outcome by its probability, and summing these products. If this expected value is greater than zero, that is, positive, then statement (5) is proven true. How would we go about proving statement (6)? The methodology is identical. If at the end of our

¹¹See generally, e.g., Steven D. Hales, Thinking Tools: You Can Prove a Negative, 10 *Think* 109 (2005); Kevin W. Saunders, The Mythic Difficulty in Proving a Negative, 15 *Seton Hall L.R.* 276 (1984).

calculation, the expected value is less than zero, that is, negative, then statement (6) is proven true. In this case, the negative statement is no more difficult to prove, and this is typical of most arithmetic calculations. The way you would go about proving that the sign of a numerical result is positive or negative is the same. In either case, you would simply calculate the result and compare it to zero.

Another way to see that there is nothing special about arithmetically positive or negative statements is that, just as in the case of philosophical statements, we can always flip the sign:

7. The state lottery fund will lose money, in expectation, if you play the lottery.
8. The state lottery fund will make money, in expectation, if you play the lottery.

If all the money won and lost in state lotteries comes from or goes to the state lottery fund, then statements (7) and (8) are equivalent to (5) and (6), but now (7) is arithmetically negative and (8) is arithmetically positive. Clearly, there is nothing special about arithmetically negative statements because they can often be rewritten as arithmetically positive statements. However, this observation misses the bigger point about arithmetically negative and positive statements. When a mathematical problem has a unique, numerical solution, one can determine its sign by computing the solution and comparing it to zero—a task that is equally difficult regardless of sign of the result.

Federal courts rarely consider arithmetic problems like the one above where there is a unique and uncontroversial numerical solution. Instead, as in the case of *Arizona Free Enterprise v. Bennett*,¹² they often consider statements of the following form:

9. The matching funds provision in Arizona's campaign finance law decreases private political contributions.

This is a counterfactual statement about the effect of a law. It posits that if a particular law were not in place, then private political contributions would be greater. In other words, the statement posits that the effect of a law on private contributions is arithmetically negative. Of course, no counterfactual statement can be proven with certainty because we can never observe what would have happened in the counterfactual world where Arizona did not pass its particular campaign finance law. Nonetheless, there may be good ways to estimate the effect of interest.

In the case of *Arizona Free Enterprise*, a team of social scientists attempted to estimate the effect of the law in several ways, concluding that the effect is close to zero.¹³

¹²131 S. Ct. 2806 (2011).

¹³Brief for Costas Panagopoulos, Ph.D., Ryan D. Enos, Ph.D., Conor M. Dowling, Ph.D., and Anthony Fowler as Amici Curiae Supporting Respondents, *Arizona Free Enterprise v. Bennett*, 131 S. Ct. 2806 (2011) (No. 10-238, 10-239), 2011 WL 686404. See also generally Conor M. Dowling, Ryan D. Enos, Anthony Fowler & Costas Panagopoulos, Does Public Financing Chill Political Speech? Exploiting a Court Injunction as a Natural Experiment, 11 Election L.J. 302 (2012).

Most notably, they showed that in a particular year when matching funds in Arizona were deactivated as a result of a court injunction, there was no increase in private political spending in Arizona relative to other states as one would expect if statement (9) is true. Regardless of the quality or direction of the empirical results, Chief Justice Roberts, writing the majority opinion, dismissed the statistical evidence, declaring “it is never easy to prove a negative.”¹⁴ He further explains that *negative* refers to the proposition “that candidates and groups did not speak or limited their speech because of the Arizona law.”¹⁵ In other words, the Court conflated the philosophical and arithmetic definitions of the word *negative*, arguing that because proving a negative is difficult, it should be difficult to show evidence that a law decreases speech.

Just as there is nothing special about proving an arithmetically negative statement, there is nothing special about providing empirical evidence that a particular effect is negative. One way to see this is to write equivalent statements that are arithmetically positive:

10. Campaign finance laws with no matching funds provision increase private political contributions relative to laws with matching funds.
11. The repeal of a matching funds provision increases private political contributions.
12. The matching funds provision in Arizona’s campaign finance law increases the amount of money saved or consumed that would have otherwise been contributed for political purposes.

Statements (10), (11), and (12) are essentially equivalent to statement (9), but they are arithmetically positive. Another way to see this is to think about whether it would have been easier to find evidence of a positive or negative effect. If the social scientists wanted to test whether matching funds increase private political contributions, they would have implemented the same empirical tests, which could have produced arithmetically negative or positive results.

Arithmetically negative effects are no more difficult to detect than positive effects. To the weak extent that philosophically negative statements are more difficult to prove or show evidence for, this rule does not apply to the arithmetic definition of the word *negative*. In the case of *Arizona Free Enterprise*, researchers investigating the effect of matching funds could have found that Arizona’s law decreased political speech, providing strong evidence in support of statement (9), but they did not. Evidence that was apparently relevant to the case at hand was ignored because of a fundamental error in logic and statistical reasoning. In the next section, we demonstrate that this particular error, which we refer to as the negative effect fallacy, has a long history in federal courts across multiple legal domains.

¹⁴*Arizona Free Enter.*, 131 S. Ct. at 2823.

¹⁵*Id.*

III. THE NEGATIVE EFFECT FALLACY IN FEDERAL COURT OPINIONS: 1960–PRESENT

When executing a search warrant unrelated to the pending federal prosecution of James Elkins for illegal wiretapping, city police officers seized evidence that turned out to be related to his federal prosecution.¹⁶ That evidence, which was later ruled to have been unlawfully collected by Oregon state courts, was used in the federal prosecution of Elkins under the “silver platter doctrine,”¹⁷ which allowed federal prosecutors to use evidence that was illegally gathered by state police. Elkins argued that the evidence should have been thrown out under the exclusionary rule,¹⁸ as laid out in *Weeks*¹⁹ and partially incorporated to the states in *Wolf*.²⁰ When the case reached the Supreme Court, a divided Court found the silver platter doctrine in violation of the Fourth Amendment.²¹

The *Elkins* opinion is a mix of reasoning from precedent and from practical necessities of police enforcement and evidence gathering. The Court first argues that the incorporation of the Fourth Amendment to the states in *Wolf* undermined any constitutional difference between state and federal police officers when evidence is presented in federal court.²² Then, the Court shifts away from abstract reasoning toward a pragmatic consideration of the effects of the silver platter doctrine, weighing the potentially adverse effects of the doctrine for future investigations against “the general need for untrammelled disclosure.”²³ The Court asserts that criticisms of the exclusionary rule have not addressed the “basic postulate” that “the rule is calculated to prevent, not to repair.” According to the Court, the purpose of the exclusionary rule “is to deter—to compel respect for the constitutional guaranty in the only effectively available way—by removing the incentive to disregard it.”²⁴ The Court then pivots to the difficulties of empirically demonstrating the deterrent effects of the exclusionary rule and argues that there are no statistics to show that there are fewer unlawful seizures in states that follow the exclusionary rule, as compared to states that do not follow the exclusionary rule.²⁵

¹⁶*Elkins v. United States*, 364 U.S. 206, 207 (1960).

¹⁷*Id.* at 208.

¹⁸*Id.* at 209–10.

¹⁹*Weeks v. United States*, 232 U.S. 383, 398 (1914).

²⁰*Wolf v. Colorado*, 338 U.S. 25, 26–27 (1947).

²¹*Elkins*, 364 U.S. at 208.

²²*Id.* at 213–15.

²³*Id.* at 216.

²⁴*Id.* at 217.

²⁵*Id.* at 218.

In this discussion, the Court falls prey to the negative effect fallacy: “Since as a practical matter it is never easy to prove a negative, it is hardly likely that conclusive factual data could ever be assembled.”²⁶ This sentence is crucial in the Court’s argument, as the majority uses the lack of empirical evidence regarding the deterrence effect of the exclusionary rule as justification to then turn to other evidence, including the way federal courts have operated under the exclusionary rule since *Weeks* was decided in 1914.²⁷ According to the Court, the fact that the FBI has not been rendered “ineffective” since the exclusionary rule was introduced and that some states have adhered to the exclusionary rule is enough pragmatic evidence to justify the continued validity of the exclusionary rule doctrine.²⁸ Ironically, on the question of whether the exclusionary rule decreases future illegal searches, the Court invokes the negative effect fallacy; but on the question of whether the exclusionary rule decreases the effectiveness of the FBI, the Court claims to have compelling evidence that the effect is not negative.

In *Elkins*, the Court dismissed the consideration of empirical evidence because of the negative effect fallacy. At the time of the case, there may have been no compelling statistical evidence on the effect of the exclusionary rule on illegal searches and seizures, but the Court had little justification for failing to examine and consider any available evidence. There is no reason, in principle, that empirical evidence could not be assembled to test whether the exclusionary rule decreases illegal searches and seizures. For example, researchers could have compared states with and without exclusionary rules, controlling for other important differences. Even better, they could have examined states that adopted or repealed the exclusionary rule, comparing changes in police activity in these states to that in other states that did not change their law. Of course, counterfactual inferences are difficult, but the Court dismissed the possibility that such statistical evidence could be collected in order to justify deciding the case on other grounds. If the exclusionary rule decreases or increases undesirable police activity, then social scientists have tools for estimating these effects, whether they are arithmetically positive or negative, and courts should consider this kind of evidence when an empirical claim is so important to a case.

The *Elkins* Court may have intended to limit the application of the negative effect fallacy to the exclusionary rule. The language of the Court in *Elkins*—“it is hardly likely that conclusive factual data could ever be assembled”²⁹—could be interpreted as pertaining to only one specific empirical question within the exclusionary rule jurisprudence. However, the negative effect fallacy was employed again in a different context eight years later in *Harrison v. United States*.³⁰

²⁶Id.

²⁷For a discussion of the interplay between empirical and practical reasoning in the Court’s exclusionary rule jurisprudence, see generally Yale Kamisar, Does (Did) (Should) the Exclusionary Rule Rest on a “Principled Basis” Rather Than an “Empirical Proposition”? 16 *Creighton L. Rev.* 565 (1982).

²⁸*Elkins v. United States*, 364 U.S. 206, 218–19 (1960).

²⁹Id. at 218.

³⁰392 U.S. 219 (1968).

Harrison concerned the admissibility of testimony of the defendant from a previous trial that was given only in response to the introduction of a confession by the prosecution that was later deemed to be inadmissible by the state courts.³¹ The Court ruled that testimony resulting from the introduction of an inadmissible confession is itself also inadmissible in a subsequent trial.³² Justice White in dissent argued that such testimony should be admissible because the inadmissibility of such testimony does not deter future violations.³³ The majority countered Justice White by citing the negative effect fallacy from *Elkins*.³⁴ The Court argued that a deterrent effect cannot be proven, dismissing the possible use of empirical evidence to answer the question of the efficacy of the exclusionary rule and therefore justifying its decision on the basis that it promotes judicial integrity.³⁵

The Court again used the negative effect fallacy in an exclusionary rule case in *United States v. Janis* in 1976.³⁶ At issue was whether evidence determined to be illegally collected during a criminal investigation could be used in a civil proceeding brought by the United States.³⁷ A search had been conducted in Janis's home for bookmaking materials, during which police found betting records and nearly \$5,000 in cash.³⁸ Based solely on this evidence, the IRS assessed penalties against Janis for failure to pay taxes on the earnings.³⁹ However, Janis subsequently had the search warrant quashed by the district court, which excluded the evidence from his criminal trial.⁴⁰

The Court ruled that the evidence was not excluded from the civil proceedings of the IRS.⁴¹ In refusing to extend the exclusionary rule, the Court engaged in an extensive analysis of the academic empirical literature on the effectiveness of the exclusionary rule on deterring illegal law enforcement conduct.⁴² Finding these studies flawed and unconvincing, the Court boldly declares that “[w]e find ourselves, therefore, in no

³¹*Id.* at 220–22.

³²*Id.* at 225–26.

³³*Id.* at 232.

³⁴*Id.* at 224, n.10.

³⁵*Id.*

³⁶428 U.S. 433 (1976).

³⁷*Id.* at 434.

³⁸*Id.* at 436.

³⁹*Id.* at 436–37.

⁴⁰*Id.* at 438–39.

⁴¹*Id.* at 459–60.

⁴²*Id.* at 449–53.

better position than the Court was in 1960.⁴³ At this point, the *Janis* Court cites the negative effect fallacy from *Elkins*.⁴⁴ However, unlike in *Elkins*, where the Court uses the negative effect fallacy to justify its support of the exclusionary rule, the *Janis* Court concludes that the exclusionary rule cannot be empirically justified and therefore should not be expanded.⁴⁵

The negative effect fallacy does not reappear in another Supreme Court decision until a 1997 voting rights case. *Reno v. Bossier School Parish Board*⁴⁶ concerned the relationship between Sections 2 and 5 of the Voting Rights Act.⁴⁷ After the Bossier School Parish Board submitted a proposed redistricting for preclearance under its Section 5 obligations, the redistricting proposal was rejected by the Attorney General on Section 2 grounds because the plan “unnecessarily limit[ed] the opportunity for minority voters to elect their candidates of choice”⁴⁸ compared to an alternative proposal that was put forth by the NAACP.⁴⁹ The Court ruled that Section 5 is not automatically violated when Section 2 is violated because Section 5 only addresses effects from redistricting that are retrogressive in nature.⁵⁰ In rejecting the argument that Section 5 is violated when Section 2 is violated, the Court argued that the burden of proof for a jurisdiction subject to preclearance is already high when it attempts to redistrict under Section 5 because a covered jurisdiction must show that its proposal does not decrease minority representation.⁵¹ Then, the Court cites the negative effect fallacy from *Elkins* as evidence to show that this burden of proof is already sufficiently high, and that requiring a covered jurisdiction to litigate whether its plan was dilutive prior to its implementation would be an unacceptable increase in the burden of proof due to the resulting federalism costs.⁵²

The final two uses of the negative effect fallacy by the Supreme Court are in the domain of the Court’s free speech jurisprudence. In 2001, the negative effect fallacy was employed in Chief Justice Rehnquist’s dissent in *Bartnicki v. Vopper*.⁵³ In *Bartnicki*, the

⁴³Id. at 450.

⁴⁴Id. at 453.

⁴⁵Id. at 454.

⁴⁶520 U.S. 471 (1997).

⁴⁷Id. at 474.

⁴⁸Id. at 475.

⁴⁹Id. at 474–76.

⁵⁰Id. at 480.

⁵¹Id.

⁵²Id.

⁵³532 U.S. 514 (2001).

Court ruled that an illegally intercepted private telephone communication regarding a matter of public concern could be broadcast by a third party not involved in the illegal interception, in this case the radio media.⁵⁴ One relevant empirical question was whether preventing the dissemination of illegally intercepted information would decrease the prevalence of illegal interceptions in the future, which the Court refers to as the “dry-up-the-market” theory.⁵⁵ The Court cited evidence suggesting little deterrent effect, justifying its decision to allow the broadcast of such information. In his dissent, Chief Justice Rehnquist dismisses this evidence embraced by the majority by calling it “voodoo statistics.”⁵⁶ He goes on to question whether it would be possible for Congress to empirically assess the effects of preventing the dissemination of illegally intercepted information by citing the negative effect fallacy from *Elkins*.⁵⁷

Lastly, as noted in Section II, the negative effect fallacy arose in the 2011 campaign finance case *Arizona Free Enterprise*. At issue was a matching funds provision in Arizona state campaign finance law. After a certain dollar amount, each dollar spent by a privately financed candidate triggered direct payment of public money of an equal amount to the publicly financed candidate.⁵⁸ Arizona attempted to justify the matching funds provision based on anti-corruption grounds, but the Court, under strict scrutiny analysis, held that the state’s anti-corruption interest did not justify the “substantial burden” on political speech.⁵⁹

One important question in this analysis was whether the matching funds provision substantially burdens political speech. If so, strict scrutiny would be applied to the provision.⁶⁰ The Court cites the negative effect fallacy, arguing that it is difficult to prove that groups did not limit their donations based on the Arizona law.⁶¹ The Court then uses qualitative evidence, including witness testimony,⁶² and legal reasoning to argue that the matching funds provision did substantially burden speech and therefore strict scrutiny must be applied to the provision.⁶³ Nowhere in the Court’s analysis did it directly engage with the quantitative evidence submitted to the Court that showed

⁵⁴*Id.* at 535.

⁵⁵*Id.* at 550.

⁵⁶*Id.* at 552, n.9.

⁵⁷*Id.* at 552.

⁵⁸*Arizona Free Enter. v. Bennett*, 131 S. Ct. 2806, 2815 (2011).

⁵⁹*Id.* at 2824.

⁶⁰*Id.* at 2820.

⁶¹*Id.*

⁶²*Id.*

⁶³*Id.* at 2824.

private contributors do not alter their behavior in response to matching funds.⁶⁴ Instead, the Court appears to have dismissed this evidence as a result of the negative effect fallacy.

The Supreme Court's use of the negative effect fallacy in *Arizona Free Enterprise* is emblematic of a regular pattern. In five of the six cases that utilize the negative effect fallacy, the fallacy is used to dismiss the applicability of empirical evidence and justify deciding the case on alternative grounds.⁶⁵ In *Elkins*,⁶⁶ *Harrison*,⁶⁷ and *Janis*,⁶⁸ the Court uses the negative effect fallacy to assert that factual evidence could not be assembled to address whether the exclusionary rule deters unlawful searches. Justice Rehnquist used the negative effect fallacy in dissent in *Bartnicki* to disprove the "voodoo statistics" discussed by the majority opinion.⁶⁹ Most recently, Justice Roberts for the majority in *Arizona Free Enterprise* used the fallacy in conjunction with witness testimony to discredit quantitative evidence that private funders did not substantially alter their funding behavior under Arizona's matching funds regime of public financing for state elections.⁷⁰

Strikingly, the negative effect fallacy was used by the Court to justify opposing positions in the exclusionary rule cases. In *Elkins* and *Harrison*, the liberal group of the Warren Court and swing Justice Stewart, who wrote both majority opinions, used the negative effect fallacy in order to expand the scope of the exclusionary rule.⁷¹ However, in *Janis*,⁷² the ascendant Nixon-appointed conservative plurality and the swing vote of Justice Powell used the fallacy at a crucial stage of their argument in order to stop the further expansion of the exclusionary rule. These cases demonstrate the danger and malleability of the negative effect fallacy. If it can be used to justify any position, it cannot be a compelling argument.

Paradoxically, the 1960s saw a burgeoning movement to seek quantified evidence regarding criminal justices policies in the United States, including the

⁶⁴See Section II for a discussion of the quantitative evidence presented to the Court.

⁶⁵These cases are *Elkins*, *Harrison*, *Janis*, *Bartnicki*, and *Arizona Free Enterprise*.

⁶⁶*Elkins v. United States*, 364 U.S. 206, 218 (1960).

⁶⁷*Harrison v. United States*, 392 U.S. 219, 224 n.10 (1968).

⁶⁸*United States v. Janis*, 428 U.S. 433, 453 (1976).

⁶⁹*Bartnicki v. Vopper*, 532 U.S. 514, 552 n.9 (2001).

⁷⁰*Arizona Free Enter. v. Bennett*, 131 S. Ct. 2806, 2833 (2011).

⁷¹*Elkins* was a 5–4 decision, with Chief Justice Warren and Justices Brennan, Black, Douglas, and Stewart in the majority and Justices Frankfurter, Clark, Harlan, and Whitaker in dissent. *Harrison* was 6–3, with Chief Justice Warren and Justices Brennan, Douglas, and Stewart joined by Justices Fortas and Marshall in the majority, with Justices Black, Harlan, and White in dissent.

⁷²*Janis* was a 5–4 decision with Justice Blackmun writing the opinion joined by Chief Justice Burger and Justices Powell, White, and Rehnquist. Justices Brennan, Marshall, and Stewart dissented, and Justice Stevens did not take part in the case.

exclusionary rule.⁷³ *Wolf v. Colorado* appeared to be a catalyst for this research because Justice Murphy signaled the Court's openness to empirical evidence. Justice Murphy himself attempted to determine how police training practices of 38 large metropolitan cities differed between those who did and did not adopt the exclusionary rule.⁷⁴ Between 1949 and the mid-1960s, social scientific methodologies improved and more rigorous studies were conducted in order to gauge the effect of the exclusionary rule, including Stuart S. Nagel's 1963 survey⁷⁵ and Michael Katz's 1966 study⁷⁶ after *Mapp v. Ohio*.⁷⁷ Additional surveys prior to⁷⁸ and after⁷⁹ *Mapp* provide additional evidence on the effects of the exclusionary rule. This initial research created a groundswell in the academic literature during the 1960s and 1970s regarding the proper methods to determine whether the exclusionary rule had an effect on policing.⁸⁰

This discussion shows that empirical evidence regarding the deterrent effects of the exclusionary rule was available when *Elkins*, *Harrison*, and *Janis* were decided by the Court. Further, advances in applied statistics that were available at this time would have allowed researchers to design studies that could have identified a negative effect of the

⁷³For a thorough review of the existing empirical literature on the exclusionary rule up to 1970, see Dallin H. Oaks, *Studying the Exclusionary Rule in Search and Seizure*, 37 U. Chi. L. Rev. 665, 678–701 (1970).

⁷⁴*Wolf v. Colorado*, 338 U.S. 25, 44–46 (1949).

⁷⁵Stuart S. Nagel, *Testing the Effects of Excluding Illegally Seized Evidence*, 1965 Wis. L. Rev. 283, 283–86 (1963).

⁷⁶Michael Katz, *Supreme Court and the State: An Inquiry into Mapp v. Ohio in North Carolina*, 45 N.C. L. Rev. 119 (1966).

⁷⁷367 U.S. 643 (1960).

⁷⁸Comment, *Search and Seizure in Illinois: Enforcement of the Constitutional Right of Privacy*, 47 Nw. U. L. Rev. 493, 498 (1952). The author summed up her findings on the exclusionary rule as follows: "the rule has failed to deter any substantial number of illegal searches. . . . These figures . . . may indicate that the exclusionary rule is most effective in discouraging illegal searches in cases involving serious offenses, where conviction is important. Conversely, where the police believe that a policy of harassment is an effective means of law enforcement, the exclusionary rule will not deter their use of unlawful methods." *Id.* at 498.

⁷⁹See Dallin Oaks & Warren Lehman, *A Criminal Justice System and the Indigent: A Study of Chicago and Cook County* 88–89 (1968).

⁸⁰For examples of papers from the 1960s and 1970s that attempted to determine the effectiveness of the exclusionary rule through different methodologies, see generally, e.g., Bradley C. Canon, *Testing the Effectiveness of Civil Liberties Policies at the State and Federal Levels: The Case of the Exclusionary Rule*, 5 Am. Pol. Q. 57 (1977); James E. Spiotto, *Search and Seizure: An Empirical Study of the Effectiveness of the Exclusionary Rule and its Alternatives*, 2 J. Leg. Stud. 243 (1973); Bradley C. Canon, *Is the Exclusionary Rule in Failing Health? Some New Data and a Plea Against a Precipitous Conclusion*, 62 Ky. L. J. 681 (1973); Neil A. Milner, *Supreme Court Effectiveness and the Police Organization*, 36 L. & Contemp. Probs. 467 (1971); Comment, *Effect of Mapp v. Ohio on Police Search-and-Seizure Practices in Narcotics Cases*, 4 Colum J. L. & Soc. Probs. 87 (1968); Nagel, *supra* n.75; Katz, *supra* n.76.

exclusionary rule on future illegal searches.⁸¹ In pointing this out, we do not contend that *Elkins*, *Harrison*, and *Janis* should have been decided differently or should have been decided on empirical evidence alone. However, the Court was incorrect to assert in *Elkins*,⁸² and then repeat in both *Harrison* and *Janis*, that “it is hardly likely that conclusive factual data could ever be assembled,”⁸³ nor should the Court have hid behind fallacious reasoning in order to decide the case entirely on alternative grounds. This pessimism, which we believe is unwarranted, allowed the Court to ignore highly relevant empirical evidence and turn to other forms of reasoning.

Outside the U.S. Supreme Court, the federal circuit and district courts have also cited *Elkins* for the proposition that it is difficult to prove a negative in a number of different jurisprudential domains. Starting in the late 1970s, the fallacy from *Elkins* has been cited in at least 12 cases at the circuit level, by seven different circuits, and involving seven different legal domains.⁸⁴ In addition, the *Elkins* reasoning has been cited at least eight times by district courts across the country.⁸⁵

Citations to the negative effect fallacy in the circuit courts began with three tax cases during the late 1970s and early 1980s.⁸⁶ In two cases, the fallacy was used to

⁸¹Statistical methods for testing empirical hypotheses in observational data like regression and instrumental variables have been available since at least the 1920s and became refined and more commonly used throughout the 20th century. For reviews, see James H. Stock & Francesco Trebbi, *Retrospectives: Who Invented Instrumental Variable Regression?* 17 *J. Econ. Persp.* 177, 178–82 (2003). See also generally John Aldrich, *Correlations Genuine and Spurious in Pearson and Yule*, 10 *Stat. Sci.* 364 (1995).

⁸²*Elkins v United States*, 364 U.S. 206, 218 (1960).

⁸³*United States v Janis*, 428 U.S. 333, 453 (1976) (quoting *Elkins*, 364 U.S. at 218); *Harrison v United States*, 392 U.S. 219, 224 (1968) (same).

⁸⁴*United States v Wilgus*, 638 F.3d 1274, 1289 (10th Cir. 2011) (religious freedom—RFRA); *Overby v. National Ass’n of Letter Carriers*, 595 F.3d 1290, 1294 (D.C.C. 2010) (sufficiency of the evidence in employment benefits—ERISA); *United States v. Cortez-Rivera*, 454 F.3d 1038, 1042 (9th Cir. 2006) (Fifth Amendment—right to indictment by grand jury); *Sissoko v. Rocha*, 440 F.3d 1145, 1162 (9th Cir. 2006) (*Bivens* action alleging detention in violation of the Fourth Amendment); *American Boat Co. v. Unknown Sunken Barge*, 418 F.3d 910, 914 (8th Cir. 2005) (service of process/notice); *Oxford Capitol Corp. v. United States*, 211 F.3d 280, 287 (5th Cir. 2000) (tax); *Andrew Crespo Gallery, Inc. v. Commissioner*, 16 F.3d 1336, 1342 (2d Cir. 1994) (tax); *Portillo v. Commissioner*, 932 F.2d 1128, 1133 (5th Cir. 1991) (tax); *Tirado v. Commissioner*, 689 F.2d 307, 310 (2d Cir. 1982) (tax); *Weimerskirch v. Commissioner*, 596 F.2d 358, 361 (9th Cir. 1979) (tax); *United States v. Fearn*, 589 F.2d 1316, 1323 (7th Cir. 1978) (sufficiency of evidence in a criminal fraud conviction); *Flores v. United States*, 551 F.2d 1169, 1176 (9th Cir. 1977) (tax).

⁸⁵*United States v. Harry*, 20 F. Supp.3d 1196, 1243 (D.N.M. 2014) (testimony admissibility in criminal sexual assault); *Thomsonian v. Wells Fargo Bank, N.A.*, 2014 WL 1244892 *1, *19 (D. Or. Mar. 25, 2014) (Fair Credit Reporting Act); *Texas v. Holder*, 888 F. Supp.2d 113, 123 (D.D.C. 2012) (Section 5 of the Voting Rights Act); *Doe v. Nestle, S.A.*, 748 F. Supp.2d 1057, 1140 (C.D. Cal. 2010) (Alien Tort Statute); *United States v. Yannotti*, 415 F. Supp.2d 280, 290 n.62 (S.D.N.Y. 2005) (federal criminal racketeering); *Varry White Music v. Banana Joe’s of Akron, Inc.*, 2002 WL 32026609, *5 (N.D. Ohio Oct. 28, 2002) (copyright infringement); *Simon v. Ward*, 80 F. Supp.2d 464, 467 n.5 (E.D. Pa. 2000) (venue in § 1983 action); *United States v. All Right, Title & Interest in Real Prop. & Appurtenances Thereto Known as 163 Renwick St., Newburgh, N.Y., Listed as Section 45, Block 11, Lot 2, Defendant-in-rem*, 859 F. Supp. 93, 96 (S.D.N.Y. 1994) (civil forfeiture).

⁸⁶See *Tirado*, 689 F.2d at 310; *Weimerskirch*, 596 F.2d at 361; *Flores*, 551 F.2d at 1175.

determine the burden of proof regarding the sufficiency of evidence,⁸⁷ while the third case is an application of the exclusionary rule issue from *Janis*.⁸⁸ Interestingly, these early circuit cases utilize the philosophical rather than the arithmetic definition of the word *negative*. For example, regarding the burden of proof in *Flores v. United States*, the Ninth Circuit used the precedent to justify placing the burden of persuasion on the government to show that a levy is wrongful due to a taxpayer having no interest in the property in question.⁸⁹ The Ninth Circuit reasoned that if the burden was not on the government, then the defendant would be forced to prove a negative fact—she has no relation to the property in question—which, when combined with the precedent of *Elkins*, is found by the Ninth Circuit to be contrary to “[p]rinciples of fair play and common sense.”⁹⁰ In effect, the Ninth Circuit isolated the negative effect fallacy from its original context and placed it in *Flores* to make a burden of proof argument using the philosophical definition of proving a negative.

From these cases, the negative effect fallacy has repeatedly been used by circuit courts on issues related to probable cause and burdens of proof in tax proceedings.⁹¹ After 2000, the fallacy spread to other area domains in both circuit and district courts. At the circuit level, the fallacy appeared in a number of disparate legal doctrines, including service of process,⁹² illegal detention under the Fourth Amendment,⁹³ the Fifth Amendment right to a grand jury,⁹⁴ the Religious Freedom and Restoration Act,⁹⁵ and even an ERISA case.⁹⁶ At the district level, the fallacy has been used in cases involving copyright infringement,⁹⁷ the Alien Tort Statute,⁹⁸ and the constitutionality of Section 5

⁸⁷See *Weimerskirch*, 596 F.3d at 360–61; *Flores*, 551 F.2d at 1175.

⁸⁸*Tirado*, 689 F.2d at 310–14.

⁸⁹*Flores*, 551 F.2d at 1175.

⁹⁰*Id.* at 1175–76.

⁹¹See *Oxford Capitol Corp. v. United States*, 211 F.3d 280, 287 (5th Cir. 2000); *Andrew Crespo Gallery, Inc. v. Commissioner*, 16 F.3d 1336, 1342 (2d Cir. 1994); *Portillo v. Commissioner*, 932 F.2d 1128, 1133–34 (5th Cir. 1991).

⁹²*American Boat Co. v. Unknown Sunken Barge*, 418 F.3d 910, 914 (8th Cir. 2005).

⁹³*Sissoko v. Rocha*, 440 F.3d 1145, 1162 (9th Cir. 2006).

⁹⁴*United States v. Cortez-Rivera*, 454 F.3d 1038, 1042 (9th Cir. 2006).

⁹⁵*United States v. Wilgus*, 638 F.3d 1274, 1289 (10th Cir. 2011).

⁹⁶*Overby v. National Ass’n of Letter Carriers*, 595 F.3d 1290, 1294 (D.C.C. 2010).

⁹⁷*Varry White Music v. Banana Joe’s of Akron, Inc.*, 2002 WL 32026609, *5 (N.D. Ohio Oct. 28. 2002).

⁹⁸*Doe v. Nestle, S.A.*, 748 F. Supp.2d 1057, 1140 (C.D. Cal. 2010).

of the Voting Rights Act,⁹⁹ among other domains. Some of these decisions have transmuted the precedent from *Elkins* into the philosophical definition, while others have made the same negative effect fallacy.

Although these cases span different legal domains, many utilize the negative effect fallacy for the same purpose: to argue that one side did not or could not meet its evidentiary burden. For example, in *Texas v. Holder*,¹⁰⁰ which would later be vacated and remanded¹⁰¹ based on *Shelby County v. Holder*,¹⁰² the District Court of DC had to determine whether Texas's newly enacted voter ID law had retrogressive effect on racial minorities under Section 5 of the Voting Rights Act.¹⁰³ In setting the legal framework of the case, the district court highlighted that Texas bore the burden of proof under Section 5 of the Voting Rights Act to prove that the law lacked both a discriminatory purpose and retrogressive effect.¹⁰⁴

Crucially, the district court then cites the negative effect fallacy in *Reno* and *Elkins* to emphasize the "difficult burden" for Texas to prove a negative.¹⁰⁵ It is through this high evidentiary burden that the district court then analyzes Texas law. Interestingly, Texas submitted its own quantitative evidence that the law did not depress voter turnout, but the United States countered with evidence to show that this question was still an open empirical issue within the social sciences, including quantitative research that directly countered Texas's empirical claim.¹⁰⁶ The court additionally takes it upon itself to criticize at length the methodologies of the quantitative evidence presented by both Texas¹⁰⁷ and the United States.¹⁰⁸ This internal inconsistency within the social sciences then allows the district court to find that Texas had not met its burden in showing the law had no retrogressive effect because of its "failure to prove a negative,"¹⁰⁹ and that the Texas law would, in fact, depress voter turnout.¹¹⁰

⁹⁹*Texas v. Holder*, 888 F. Supp.2d 113, 123 (D.D.C. 2012).

¹⁰⁰888 F. Supp. 113 (D.D.C. 2012).

¹⁰¹*Texas v. Holder*, 133 S. Ct. 2886 (2013).

¹⁰²*Shelby County v. Holder*, 133 S. Ct. 2612 (2013).

¹⁰³*Texas v. Holder*, 888 F. Supp.2d at 114.

¹⁰⁴*Id.* at 123.

¹⁰⁵*Id.*

¹⁰⁶*Id.* at 127.

¹⁰⁷*Id.* at 129; *id.* at 134–37.

¹⁰⁸*Id.* at 133–34.

¹⁰⁹*Id.* at 138.

¹¹⁰*Id.*

IV. POTENTIAL EXPLANATIONS FOR THE USE OF THE NEGATIVE EFFECT FALLACY

The previous section documented the use of the negative effect fallacy in the federal courts over the past 50+ years. But why was the fallacy initially used by the Supreme Court, and why has it managed to proliferate in the Court and lower federal courts since its first use? This section offers several potential explanations for the pervasiveness of the negative effect fallacy. Of course, we cannot know for sure why a judge utilized a particular argument in any individual case,¹¹¹ but we consider four potential reasons why the negative effect fallacy, while misguided, might be particularly appealing. These explanations are admittedly speculative, and they are not necessarily exhaustive or mutually exclusive. The prevalence of the fallacy is likely attributable to a combination of factors. In any case, we hope this discussion will be useful for developing remedies to reduce the prevalence of the negative effect fallacy and similar errors.

Justices rely on the briefs of the parties before them in order to understand the case at hand. For any potential empirical analysis important to a legal issue, one litigant would like to discredit the analysis, and may be willing to make any argument that can accomplish this goal. If the negative effect fallacy is a persuasive way to discredit evidence, perhaps because judges are overly willing to rely on an old adage about proving a negative, then they will be eager to introduce the fallacy.¹¹² Considering this, we might expect the negative effect fallacy to appear in a brief by the litigants or *amici curiae*¹¹³ during the course of their advocacy before the Court.

¹¹¹We acknowledge upfront the inherently speculative nature of attempting to study and interpret judicial decision making. For discussions of the intricate and difficult study of judicial decision making, see generally Thomas J. Miles & Cass R. Sunstein, *The New Legal Realism*, 75 *U. Chi. L. Rev.* 831 (2008); Theodore W. Ruger, Pauline T. Kim, Andrew D. Martin & Kevin M. Quinn, *The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking*, 104 *Colum. L. Rev.* 1150 (2004); Frank B. Cross, *Decision-Making in the U.S. Circuit Courts of Appeals*, 91 *Cal. L. Rev.* 1457 (2003); Tracey George & Lee Epstein, *On the Nature of Supreme Court Decision Making*, 86 *Am. Pol. Sci. Rev.* 323 (1992). See also Barry Friedman, *Taking Law Seriously*, 4 *Perspectives on Polit.* 261, 261 (2006).

¹¹²For discussions of the importance of litigant briefs for judges, see *Practitioner's Handbook for Appeals to the United States Court of Appeals for the Seventh Circuit* 117–18 (2014 ed.), available at: <https://www.ca7.uscourts.gov/Rules/handbook.pdf>; Paul M. Michael, *Effective Appellate Advocacy*, 24 *Litig.* 19, 21 (1998); Fred I. Parker, *Appellate Advocacy and Practice in the Second Circuit*, 64 *Brook. L. Rev.* 457, 461 (1998).

¹¹³*Amicus* briefs are increasingly prevalent in the federal court system, and it is not unusual for the Supreme Court to utilize arguments made in an *amicus* brief. See, e.g., *Teague v. Lane*, 489 U.S. 288, 300 (1989) (O'Connor, J., plurality); *Mapp v. Ohio*, 367 U.S. 643, 646 n.3 (1961). However, some have argued that *amicus* briefs do not alter the opinions of justices. See Donald R. Songer & Reginald S. Sheehan, *Interest Group Success in the Courts: Amicus Participation in the Supreme Court*, 46 *Pol. Sci. Q.* 339, 350 (1993). But see generally Paul M. Collins Jr. & Wendy L. Martinek, *Friends of the Circuits: Interest Group Influence on Decision Making in the U.S. Courts of Appeals*, 91 *Soc. Sci. Q.* 397 (2010); Paul M. Collins, Jr., *Friends of the Court: Examining the Influence of Amicus Curiae Participation in U.S. Supreme Court Litigation*, 38 *Law & Soc. Rev.* 807 (2004).

However, the negative effect fallacy was not used by litigants or amici curiae in *Elkins*.¹¹⁴ Therefore, the Court appears to have generated and incorporated the fallacy sua sponte into the *Elkins* opinion. Neither do the litigant and amici curiae briefs from *Harrison*,¹¹⁵ *Janis*,¹¹⁶ or *Arizona Free Enterprise*¹¹⁷ discuss the fallacy. *Bartnicki* appears to be the only Supreme Court case where litigants or amici curiae discuss the negative effect fallacy. The Brief for the Petitioners argues in favor of the fallacy in a footnote that discusses the dry-up-the-market theory, but the brief does not quote any of the exclusionary rule cases that previously mentioned the fallacy.¹¹⁸ The Brief for the United States directly quotes from the section of *Elkins* that first uses the negative effect fallacy in order to reject any evidentiary showing being required in order to establish a deterrent effect of criminalizing certain forms of speech,¹¹⁹ and the fallacy is used in at least one of the amici briefs regarding the dry-up-the-market theory.¹²⁰

An analysis of the available litigant briefs in the circuit cases and the pleadings and motions in the district cases that use the fallacy displays a similar trend—in only one of the circuit and district cases in which the briefs or pleadings are available did

¹¹⁴Petitioners' Reply Brief, *Elkins v. United States*, 364 U.S. 206 (1960) (No. 126); Brief for the Petitioners, *Elkins v. United States*, 364 U.S. 206 (1960) (No. 126); Brief for the United States, *Elkins v. United States*, 364 U.S. 206 (1960) (No. 126).

¹¹⁵Petitioners' Reply Brief, *Harrison v. United States*, 392 U.S. 219 (1968) (No. 876); Brief for the United States, *Harrison v. United States*, 392 U.S. 219 (1968) (No. 876); Brief for the Petitioner, *Harrison v. United States*, 392 U.S. 219 (1968) (No. 876).

¹¹⁶Brief for the Respondent, *United States v. Janis*, 428 U.S. 433 (1976) (No. 74-958); Brief for the Petitioners, *United States v. Janis*, 428 U.S. 433 (1976) (No. 74-958). The Brief for the Petitioners in *Janis* does make brief mention of the empirical uncertainty around the deterrence effect of the exclusionary rule and argues that this uncertainty should caution against the extension of the exclusionary rule to civil tax cases. Brief for Petitioners at 33, *United States v. Janis*, 428 U.S. 433 (1976) (No. 74-958).

¹¹⁷Reply Brief for Petitioners *Arizona Free Enter. v. Bennett*, 131 S. Ct. 2806 (2011) (Nos. 10-238, 10-239); Reply Brief for Petitioners, *McComish v. Bennett* (2011) (Nos. 10-238, 10-239); Brief of State Respondents, *Arizona Free Enter. Club's Freedom Club PAC v. Bennett*, 131 S. Ct. 2806 (2011) (Nos. 10-238, 10-239); Brief of Respondents Clean Elections Inst., Inc., *Arizona Free Enter. Club's Freedom Club PAC v. Bennett*, 131 S. Ct. 2806 (2011) (Nos. 10-238, 10-239); Brief of Petitioners *Arizona Free Enter. Club's Freedom Club PAC, et al., Ariz. Free Enter. Club's Freedom Club PAC v. Bennett*, 131 S. Ct. 2806 (2011); Brief for Petitioners, *Arizona Free Enter. Club's Freedom Club PAC v. Bennett*, 131 S. Ct. 2806 (2011) (Nos. 10-238, 10-239). In addition, none of the 25 amici curiae briefs in *Arizona Free Enterprise* mention the negative effect fallacy or cite to cases that previously discussed the fallacy.

¹¹⁸Brief for Petitioners Gloria Bartnicki & Anthony F. Kane, Jr. at 37 n.18, *Bartnicki v. Vopper*, 532 U.S. 514 (2001) (Nos. 99-1687, 99-1728).

¹¹⁹Brief for the United States, *Bartnicki v. Vopper*, 532 U.S. 514 (2001) (Nos. 99-1687, 99-1728).

¹²⁰See Brief of Amicus Curiae Representative John A. Boehner in Support of Petitioners at 25-26, *Bartnicki v. Vopper*, 532 U.S. 514 (2001) (Nos. 99-1687, 99-1728).

a litigant brief discuss the fallacy.¹²¹ Therefore, it appears that the same pattern regarding sua sponte generation of the fallacy appears at all three levels of federal courts.

Considering that litigant or amici briefs in most cases do not discuss the fallacy, many courts probably found the fallacy through their own research into past precedent. Judges and clerks may have purposefully looked for this kind of argument in order to ignore empirical evidence and justify their preferred reasoning. This may explain why the fallacy spread from the expansionary rule line of cases into *Reno* and *Arizona Free Enterprise*.¹²² Over time, a similar pattern emerges at all levels of the federal courts, whereby the fallacy spreads into an increasing number of legal domains. These examples illustrate that an overreliance on precedent can have adverse consequences. Once a fallacy is available in a prominent opinion, it provides a ripe opportunity for future courts to reuse the fallacy to justify their choice to ignore relevant evidence.

The previous discussion focuses on *how* judges incorporated the negative effect fallacy into their opinions but not *why*. One plausible explanation is motivated reasoning, a well-established psychological phenomenon across a broad range of decision making whereby individuals accept or reject reasoned evidence in order to support decisions they have already made or to be consistent with a psychological worldview,¹²³ potentially but not necessarily due to positive or negative emotional affect.¹²⁴

¹²¹Litigant briefs were found in Westlaw for the following circuit court cases: *United States v. Wilgus*, 638 F.3d 1274, 1289 (10th Cir. 2011); *Overby v. National Ass'n of Letter Carriers*, 595 F.3d 1290, 1294 (D.C.C. 2010); *United States v. Cortez-Rivera*, 454 F.3d 1038, 1042 (9th Cir. 2006); *Sissoko v. Rocha*, 440 F.3d 1145, 1162 (9th Cir. 2006); *Oxford Capitol Corp. v. United States*, 211 F.3d 280, 287 (5th Cir. 2000); *Portillo v. Commissioner*, 932 F.2d 1128, 1133 (5th Cir. 1991). Litigant pleadings and motions were found in Westlaw for the following district court cases: *Thomasian v. Wells Fargo Bank, N.A.*, 2014 WL 1244892 *1, *19 (D. Or. Mar. 25, 2014); *Texas v. Holder*, 888 F. Supp.2d 113, 123, D.D.C. (2012); *Doe v. Nestle, S.A.*, 748 F. Supp.2d 1057, 1140 (C.D. Cal. 2010); *United States v. Yannotti*, 415 F. Supp.2d 280, 291 (S.D.N.Y. 2005). Of these cases, the negative effect fallacy was discussed in a litigant brief only in *Portillo*. Brief for Appellants, *Portillo v. Commissioner*, 932 F.2d 1128, 1133 (5th Cir. 1991) (No. 92-4526). However, in *Portillo*, the fallacy was transformed into a philosophical definition of proving a negative, and not the empirical definition we argue against the use of in this article.

¹²²Remember that Justice Stewart wrote both *Elkins* and *Harrison*, while *Janis* ceased the expansion of the exclusionary rule that had begun during the 1960s by the Warren Court. See Jerold H. Israel, *Criminal Procedure, the Burger Court, and the Legacy of the Warren Court*, 75 Mich. L. Rev. 1319, 1404 (1977). In addition, *Bartnicki* had extensive briefing regarding the fallacy, and so the statistical issues surrounding the dry-up-the-market theory were likely to be already live for the justices when deciding the case.

¹²³See generally Ziva Kunda, *The Case for Motivated Reasoning*, 108 Psych. Bull. 480 (1990). See also Dan Kahan, *Forward: Neutral Principles, Motivated Cognition, and Some Problems for Constitutional Law*, 125 Harv. L. Rev. 1, 19 (2011).

¹²⁴We do not take a position on the question of whether motivated reasoning largely stems from emotion. For the argument that it does, see generally George E. Marcus, *The Sentimental Citizen: Emotion in Democratic Politics* (2002); Jonathan Haidt, *The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 Psych. Rev. 814 (2001). For the argument that emotions and reasoning may occur through distinct cognitive processes, see generally Keith Frankish, *Dual-Process and Dual-System Theories of Reasoning*, 5 Phil. Compass 914 (2010); Johnathan St. B.T. Evans, *Dual-Process Accounts of Reasoning, Judgment, and Social Cognition*, 59 Ann. Rev. Psych. 255 (2008); Daniel Kahneman & Shane Frederick, *Frames and Brains: Elicitation and Control of Response Tendencies*, 11 Trends in Cog. Sci. 45 (2007).

In the cases in which judges used the fallacy, the judge or her clerks may have been motivated to find support for their intuition regarding the legal holding. Justice Stewart wrote both *Elkins* and *Harrison* and used the fallacy in both cases in order to argue for his holding by saying that empirical information could not be assembled to refute his extension of the exclusionary rule. Similarly, Justice Blackmun used the fallacy to make the opposite claim in *Janis*—because the empirical information regarding the effectiveness cannot be assembled, caution is required before extending the exclusionary rule into other domains.¹²⁵ Most recently, in *Arizona Free Enterprise*, Justice Roberts and his clerks had different forms of evidence to support both potential holdings of the case—to uphold or strike down Arizona’s public financing scheme. One set of evidence (witness testimony) supported his preferred holding, while the other set of evidence (statistical studies) went in the opposite direction. The negative effect fallacy allowed Justice Roberts to discredit the statistical studies in favor of the witness testimony that supported his legal holding in the case.

A more charitable interpretation of the negative effect fallacy, and the preferred interpretation of some of our academic colleagues, is that this argument is used as a stand-in for the following argument: empirical research on the effects of laws on behavior—the type of research most directly relevant to many of these cases—is hard. It requires a compelling research design that can provide a reasonable estimate of a counterfactual scenario. Researchers themselves often disagree with one another about what constitutes a compelling research design, and it may be difficult for judges to read and evaluate this kind of evidence.¹²⁶ Judges may be reluctant to wade into an area outside their technical areas of expertise, so the negative effect fallacy is an attractive argument for a judge who prefers not to have to make sense of the complicated and sometimes conflicting empirical evidence.

If critics of our article would like to defend the courts that have employed the negative effect fallacy, we suspect this is their best argument. In other words, one plausible counterargument is that the negative effect fallacy is not such a problem because the judges employing the fallacy had good reasons to dismiss the empirical evidence but simply articulated their reasons poorly. Our response is that we see little sign that the courts employing the fallacy engaged with the relevant empirical evidence in a rigorous way,¹²⁷ but if they did, they should more clearly articulate their reasons for dismissing

¹²⁵United States v. Janis, 428 U.S. 433, 453 (1976).

¹²⁶There are a multitude of different research design debates in the social sciences. For a few noteworthy recent debates, see generally Andrew C. Eggers, Anthony Fowler, Jens Hainmueller, Andrew B. Hall & James M. Snyder Jr., On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races, 59 Am. J. Pol. Sci. 259 (2015); Jasjeet S. Sekhon & Rocio Titiunik, When Natural Experiments Are Neither Natural Nor Experiments, 106 Am. Pol. Sci. Rev. 35 (2012).

¹²⁷For our discussion of the analytical work being done by the fallacy in the Supreme Court cases that cite to it, see Section III. Note that the fallacy is being used in these cases not simply to point to the difficulty of empirical research, but to also allow the Court to negate the strength of potential counterevidence and to switch to methods of legal analysis that favor the legal holding of the case.

the evidence. Citing a logical fallacy should not be sufficient grounds for dismissing statistical evidence.

V. PROPOSALS TO IMPROVE THE USE OF STATISTICAL REASONING BY FEDERAL COURTS

The previous sections have been retrospective—examining previous federal cases and analyzing the motives, reasoning, and rhetoric used by the courts in regard to the fallacy. Here, we offer some suggestions as to how federal courts should proceed when statistical and empirical evidence is relevant for a particular case. Our proposals begin by focusing on the use of the negative effect fallacy *per se* and then expand to the broader topic of the evaluation of empirical evidence in legal settings.

We focus first on our most straightforward proposal—the removal of the negative effect fallacy from federal jurisprudence. Remember that the concept of “proving a negative” has two meanings that are distinct from each other—one philosophical and the other empirical.¹²⁸ The philosophical definition of “proving a negative” has been subject to discussion by other commentators¹²⁹ and bears its own logical shortcomings, but the negative effect fallacy involves the arithmetic use of the word *negative*. In this subsection we focus on a straightforward proposal: federal courts should stop using the negative effect fallacy.

We showed in Section II that there is nothing special about arithmetically negative statements. They can be statistically rejected or supported just as easily as arithmetically positive statements. When an empirical claim is relevant for a case, judges should consider evidence for that claim on its own merits. Judges might find the evidence unconvincing, and they might choose to ignore the evidence on well-justified grounds, such as the evidence being inconclusive or the analysis not being credible. However, the negative effect fallacy is not a valid reason to ignore statistical evidence.

Despite the logical incoherence of the negative effect fallacy, it has been repeatedly used by the Supreme Court and lower federal courts in numerous legal domains. We propose that no federal courts of any level should cite to *Elkins* for the negative effect fallacy or otherwise utilize the fallacy in order to claim that it is difficult or impossible to empirically prove a negative statistical proposition. Judges should instead engage with the empirical evidence (or lack thereof) instead of citing to the fallacy if they wish to argue that an arithmetically negative proposition is supported or not.

The complete cessation of the use of the negative effect fallacy will have several benefits for the federal courts system. First, stopping the fallacy will stop the use of a logically flawed argument, which is important in its own right. Federal courts should fairly consider all relevant empirical evidence in order to improve the quality and

¹²⁸See Section III.

¹²⁹For extended discussions on the use of the philosophical definition of proving a negative, see note 11.

validity of court decisions,¹³⁰ signal to citizens that the courts are proper adjudicators of disputes involving empirical information, and signal to other cognate disciplines that the legal system can fairly adopt empirical modes of reasoning.¹³¹ The eradication of the negative effect fallacy will also make it more difficult for the courts to shroud value-based reasoning behind the cloak of apparent necessity. The Supreme Court has repeatedly used the fallacy over the past 50+ years in order to justify its rejection of reasoning grounded on empirical facts so that it could embrace value-based reasoning. Although this rhetorical move may provide a salve to the justices by allowing them to argue for their holding from more familiar grounds, it does a disservice to both the legal community and the citizenry.¹³²

Although we cannot know with certainty whether previous decisions would have been different if courts had not utilized the negative effect fallacy, the error appears to have had significant consequences in several cases. For example, in *Arizona Free Enterprise*, the empirical assertion that matching funds chill private contributions was central to the majority opinion, yet the most credible evidence on this question suggested no systematic effect. Therefore, had the Court been unable to utilize the fallacy, it would have had little choice but to consider the evidence, and upon reviewing the evidence, it would have had three choices: (1) alter its decision, (2) provide a more justifiable reason for disregarding the evidence, or (3) provide a different rationale for its decision.

Even if the negative effect fallacy has not changed the results of court decisions, it unnecessarily hides the true rationale of the court for the judgment in a particular case and alters the contours of legal doctrine. If one believes part of the function of judicial opinions is to serve an explanatory function to the citizenry or to provide guidance to other actors in the political system,¹³³ then these functions are impaired when a justice uses a fallacy as justification. When a justice hides her rationale of a holding, or performs a rhetorical sleight of hand to justify how she reached her holding, a justice denies the legal community and the wider citizenry the true justification for why she has reached that particular outcome. Not only does such action by the justice undermine the important democratic value of judicial transparency for its own sake, but it also makes it more difficult for litigators and the wider legal community

¹³⁰For discussions regarding the importance of the quality of judicial opinions, see Richard A. Posner, *The Federal Courts: Challenge and Reform* 36 (1985); Walter V. Schaefer, *Precedent and Policy*, 34 *U. Chi. L. Rev.* 3, 10 (1966).

¹³¹For the importance of judicial opinions as signals to other political institutions and individuals, see note 162.

¹³²See Robert A. Leflar, *Quality in Judicial Opinions*, 3 *Pace. L. Rev.* 579, 581 (1983).

¹³³See Thomas E. Baker, *Rationing Justice on Appeal: The Problems of the U.S. Courts of Appeals* 119–20 (1994); Michael Wells, *French and American Judicial Opinions*, 19 *Yale. J. Int'l L.* 81, 85–91 (1994); Edwin W. Patterson, *The Case Method in American Legal Education: Its Origins and Objectives*, 4 *J. Leg. Educ.* 1, 17 (1951).

to understand and utilize such arguments in subsequent cases.¹³⁴ Therefore, avoiding the fallacy will have several benefits for current and future cases, including improving the transparency of the judiciary, educating the public, and improving the quality of future litigation.

Beyond the specific case of the negative effect fallacy, our investigation produces recommendations regarding the general use of statistical evidence in courts. There is already a voluminous literature on the use of statistics in courts,¹³⁵ and many commentators have proposed large-scale institutional reforms to improve the use of statistics in courts.¹³⁶ We acknowledge upfront that a comprehensive study of empirical evidence and judicial reasoning is beyond the scope of this article. We briefly mention three recommendations that may improve the way in which courts evaluate empirical evidence and incorporate this evidence into their decisions. Specifically, we suggest (1) increased statistical and empirical education and training for judges, clerks, and other judicial officers, (2) hiring empirical experts as staff in the courts, and (3) a general shift in standard practices such that judges consider all relevant empirical evidence and evaluate it according to its internal credibility.

The general structure of the federal courts may limit the ability of judges to effectively assess statistical information. First, judges must rely on their own background knowledge and that of their clerks while researching and analyzing the cases before them, meaning that different judges may have varied ability to parse and

¹³⁴For discussions on judicial transparency, see Hon. T.S. Ellis, III, Sealing, Judicial Transparency, and Judicial Independence, 53 *Vill. L. Rev.* 939, 940 (2008); Lynn M. LoPucki, Court-System Transparency, 94 *Iowa L. Rev.* 481, 494–513 (2008); Stephen A. Smith, Contract Theory 25 (2004). For a more general discussion on the importance of transparency for a democratic government, see generally Dennis F. Thompson, Democratic Secrecy, 114 *Pol. Sci. Q.* 181 (1997).

¹³⁵For notable examples of areas of the law in which legal and statistical commentators have criticized the statistical reasoning of courts, see, e.g., D. James Greiner, The Quantitative Empirics of Redistricting Litigation: Knowledge, Threats to Knowledge, and the Need for Less Districting, 29 *Yale L. & Pol'y Rev.* 527, 536–38 (2011); Wendy K. Tam Cho & Albert H. Yoon, Strange Bedfellows: Politics, Courts, and Statistics: Statistical Expert Testimony in Voting Rights Cases, 10 *Cornell J. L. & Pub. Pol'y* 237, 239 (2001); Bernard Grofman, Multivariate Methods and the Analysis of Racially Polarized Voting: Pitfalls in the Use of Social Science by the Courts, 72 *Soc. Sci. Q.* 826, 827–28 (1991); Thomas J. Campbell, Regression Analysis in Title VII Cases: Minimum Standards, Comparable Worth, and Other Issues Where Law and Statistics Meet, 36 *Stan. L. Rev.* 1299, 1302–05 (1984); Thomas J. Sugrue & William B. Fairley, A Case of Unexamined Assumptions: The Use and Misuse of the Statistical Analysis of Castaneda/Hazelwood in Discrimination Litigation, 24 *B.C. L. Rev.* 925, 935–37 (1983).

¹³⁶For commentators who discuss comprehensive reform to the structure of the federal judiciary for the purpose of improving the ability of the judiciary to understand and use empirical information, see, e.g., Adrian Vermeule, Should We Have Lay Justices? 59 *Stan. L. Rev.* 1569, 1581 (2007); Arti K. Rai, Specialized Trial Courts: Concentrating Expertise in Fact, 17 *Berkley Tech. L.J.* 877, 889–90 (2002); Edward V. Di Lello, Note, Fighting Fire with Firefighters: A Proposal for Expert Judges at the Trial Level, 93 *Colum. L. Rev.* 473 (1993); John W. Osborne, Judicial/Technical Assessment of Novel Scientific Evidence, 1990 *U. Ill. L. Rev.* 497, 540–43 (1990); Ellen R. Jordan, Specialized Courts: A Choice, 76 *Nw. L. Rev.* 745, 745 (1981); James A. Martin, The Proposed “Science Court,” 75 *Mich. L. Rev.* 1058, 1058 (1977).

analyze statistical information.¹³⁷ Second, some judges assign varying levels of opinion-writing duties to their clerks, which increases the reliance of the judge on her clerks' background knowledge.¹³⁸ Third, adversarial amicus briefs that present complex statistical information may serve to confuse, rather than enlighten, judges regarding the empirical issues embedded within a case.¹³⁹ Finally, there is no formal fact-checking mechanism, so judges must rely on their own knowledge, the knowledge of their clerks, and perhaps on that of their fellow judges in order to avoid statistical errors in their opinions.¹⁴⁰

One potential solution to mitigate some of these problems is improved statistical education and training for judges, clerks, and other judicial officers. Before newly confirmed district court judges take the bench, they often attend an orientation seminar hosted by the Federal Judicial Center (FJC), which is affectionately called "Baby Judges

¹³⁷This being said, judges often self-specialize during their time on the bench by choosing to write opinions in specific domains and therefore acquire increasingly detailed knowledge regarding those jurisprudential areas. See Edward K. Cheng, *The Myth of the Generalist Judge*, 61 *Stan. L. Rev.* 519, 533–40 (2008).

¹³⁸For a discussion of the practice of clerks writing judicial opinions at the circuit level, see Richard Posner, *Will the Federal Courts of Appeals Survive Until 1984? An Essay on Delegation and Specialization of the Judicial Function*, 56 *S. Cal. L. Rev.* 761, 768–69 (1983). This reliance on clerks by federal judges to help analyze cases and write opinions has led Justice Breyer to suggest that judges hire law clerks with specialized background knowledge and training to help narrow the scientific questions that are in dispute in a case. *General Elec. Co. v. Joiner*, 522 U.S. 136, 149 (1997) (Breyer, J., concurring).

¹³⁹Amicus briefs are becoming increasingly prevalent in the federal courts, and they present the opportunity to explain specialized information to the court that the judge(s) may lack. See Joseph D. Kearney & Thomas W. Merrill, *The Influence of Amicus Curiae Briefs on the Supreme Court*, 148 *U. Penn. L. Rev.* 743, 752 (2000). This has led Justice Breyer to publicly invite amici briefs on empirical information that may be important for the justices to know when they decide cases involving such information. Justice Breyer *Calls for Experts to Aid Courts in Complex Cases*, *N.Y. Times* (Feb. 17, 1998). However, the majority of amicus briefs are written to be persuasive for the brief writer's preferred legal or policy outcome. See Linda Sandstrom Simard, *An Empirical Study of Amici Curiae in Federal Court: A Fine Balance of Access, Efficiency, and Adversarialism*, 27 *Rev. Litig.* 669, 676 (2008); Paul M. Collins, Jr., *Lobbyists Before U.S. Supreme Court: Investigating the Influence of Amicus Curiae Briefs*, 60 *Pol. Res. Q.* 55, 64–65 (2007). As a result, judges have begun to look at amicus briefs skeptically. See *Jaffee v. Redmond*, 518 U.S. 1, 35–36 (1996) (Scalia, J., dissenting); *Ryan v. Commodity Futures Trading Comm'n*, 125 F.3d 1062, 1063 (7th Cir. 1997) (Posner, J.).

¹⁴⁰Instead, error correction in the federal courts often happens *ex post* after the opinion is released by the court, as other judges and legal commentators may criticize an opinion for being incorrect on various grounds, including its misuse of statistical or other forms of empirical reasoning. This is what eventually happened to famous footnote 11 of *Brown v. Board of Education*, which has subsequently become perhaps "the most dispute-laden footnote in American constitutional law." Paul L. Rosen, *History and State of the Art of Applied Social Research in the Courts*, Remarks from Panel Discussion at Washington, D.C., Conference on the Use/Nonuse/Misuse of Applied Social Science Research in the Courts (Oct. 1978) in *The Use/Nonuse/Misuse of Applied Social Research in the Courts* 9 (Michael J. Saks & Charles H. Baron eds. 1980). Richard Lazarus has recently discussed the practice of Supreme Court justices revising their opinions after the opinion is published, which allows for a justice to alter an opinion if such error correction takes place in the relatively narrow window of a few months to a few years after the opinion is first published. Richard J. Lazarus, *The (Non)Finality of Supreme Court Opinions*, 128 *Harv. L. Rev.* 540, 543–44 (2014).

School.¹⁴¹ This voluntary seminar is meant to prepare the new district judges to handle many of the aspects of being a federal judge, including the use of scientific evidence.¹⁴² In addition, the FJC occasionally puts out reference manuals on technical evidence, including scientific evidence.¹⁴³

This training and preparation could include more background and technical information on the evaluation of statistical results. The FJC appears to have little relevant information and no reference manual for social scientific evidence available for judges, such as how statistics is specifically applied to social science research.¹⁴⁴ While new judges are given a few hours of education on scientific evidence in Phase II Orientation, we suspect this is insufficient. For judges who will deal with complex statistical information across many different legal domains, including employment discrimination, criminal procedure, voting rights, and equal protection cases,¹⁴⁵ additional training in statistical reasoning could prove useful.

In addition, we are aware of no similar introductory seminar for law clerks or staff attorneys, and we suspect such a seminar would prove useful given judicial reliance on these institutional actors during the research and opinion-writing process. Furthermore, law schools—where all future judges are trained—could take a leading role and encourage their students to take more statistics and econometrics.¹⁴⁶ This particular reform is already underway,¹⁴⁷ and while it may take several decades for these recent graduates to

¹⁴¹“Baby Judges School” Jump Starts Learning Process, 37 Third Branch 1, 10 (2005), available at: <http://www2.fjc.gov/sites/default/files/2014/IJR00020.pdf>; Rex Bossert, A Week at Boot Camp for Judges: Rookie Jurists Get a Crash Course and Swap Court Tips at Baby Judge School, Nat’l L.J. A1 (Jul. 7, 1997); Thomas B. Russell, Bridging the Gap: Between a Trial Lawyer’s Experience and Becoming a Good Judge Is a Distance that Goes Beyond Ordinary Measurements, 27 Judges J. 16, 17 (1988); Beverly Blair Cook, The Socialization of New Federal Judges: Impact on District Court Business, 1971 Wash. U. L.Q. 253, 263–66 (1971).

¹⁴²“Baby Judges School” Jump Starts Learning Process, 37 Third Branch 1, 10 (2005), available at: <http://www2.fjc.gov/sites/default/files/2014/IJR00020.pdf>.

¹⁴³See generally Federal Judicial Center, Reference Manual on Scientific Evidence (3d ed. 2011).

¹⁴⁴The *Reference Manual on Scientific Evidence* does contain over 200 pages on statistics, multiple regression, and survey research, but these contain only general information, and do not discuss the unique difficulties of applying statistical methodologies to social science research, such as political science or psychological research. Id. at 211–423.

¹⁴⁵See note 135 for some examples of criticism regarding the use of statistical evidence in multiple distinct legal domains.

¹⁴⁶For more detailed proposals regarding the introduction of quantitative methods into legal curricula, see Lee Epstein & Gary King, Building an Infrastructure for Empirical Research in the Law, 53 J. Leg. Educ. 311, 313–14 (2003); Lee Epstein & Gary King, The Rules of Inference, 69 U. Chi. L. Rev. 1, 116–18 (2002). For a sympathetic response, see generally Howell E. Jackson, Analytical Methods for Lawyers, 53 J. Leg. Educ. 321 (2003).

¹⁴⁷Indeed, 16 of the *U.S. News & World Report’s* top 20 law schools appear to offer at least one class on statistics and quantitative methods for law students. There are also multiple casebooks dedicated to analytical methods for law students. See generally Michael O. Finkelstein & Bruce Levin, *Statistics for Lawyers* (3d ed. 2015); Howell E. Jackson, Louis Kaplow, Steven M. Shavell, W. Kip Viscusi & David Cope (eds.), *Analytical Methods for Lawyers* (2d ed. 2010). Additionally, many law schools employ a resident “methodologist” in the law school to interact with faculty and students regarding quantitative methodologies. See Matthew Spitzer, Evaluating Valuing Empiricism (at Law Schools), 53 J. Leg. Educ. 328, 331 (2003).

become judges, we are cautiously optimistic about the long-term benefits of this training.

Increasing the statistical training of judges and law students has multiple potential beneficial effects. Most apparent, this training will improve the ability of judges and future clerks to analyze and utilize statistical lines of reasoning in judicial opinions. This benefit will also dovetail with our third recommendation below regarding increased judicial discrimination of statistical arguments because it will provide the judges and their clerks the increased ability to directly engage with potential lines of statistical reasoning, rather than solely relying on the expertise of witnesses at trial or amici briefs during litigation. Finally, increasing the training for judges will allow these knowledgeable judges to become better at critically engaging with statistical reasoning when other judges use poor statistical reasoning in their opinions in both previous precedential opinions and in other opinions in the case before the knowledgeable judge. Importantly, increased statistical sophistication will mitigate the perpetuation of bad judicial precedent,¹⁴⁸ such as what occurred with the negative effect fallacy, and help judges to distinguish a current case from a previous line of precedent.¹⁴⁹

The extent to which increased training will avoid problems such as the negative effect fallacy partly depends on the reasons that these kinds of mistakes are made. If judges unwittingly utilize a fallacy, perhaps because they have seen it in previous cases and assume the logic is valid, then increased education is an effective solution. However, if judges knowingly make incorrect assertions, perhaps because they are ideologically motivated to arrive at a particular decision, then the effect of education would be smaller. Improving the scientific and statistical literacy of judges would still be beneficial in a world of ideological and insincere judges. When other judges and legal actors can more easily identify and refute fallacious reasoning, its use, in equilibrium, will become less prevalent.

A second small-scale institutional proposal is that the federal court system hire staff members with expertise in statistical and scientific inquiry. Commentators have previously suggested somewhat similar reforms using the federal court infrastructure to increase scientific knowledge, such as using magistrate or other types of judges with a scientific background,¹⁵⁰

¹⁴⁸This is similar in concept to, but narrower in scope than, the proposed constitutional anti-canon, or cases that are important but normatively disproved. For discussions of the constitutional anti-canon, see generally Jamal Greene, *The Anticanon*, 125 *Harv. L. Rev.* 379 (2011); Richard A. Primus, *Canon, Anti-Canon, and Judicial Dissent*, 48 *Duke L.J.* 243 (1999). Analogously, we argue that increased statistical training among the judiciary can allow the judiciary to better create lines of precedential reasoning that are important, but statistically disproved, such as arguments that utilize the negative effect fallacy.

¹⁴⁹On the importance of the ability of judges to distinguish precedent, see Richard A. Posner, *How Judges Think* 184 (2010); Michael W. McConnell, *The Importance of Humility in Judicial Review: A Comment on Ronald Dworkin's Moral Reading of the Constitution*, 65 *Fordham L. Rev.* 1269, 1289 (1997); Alex Kozinski, *What I Ate for Breakfast and Other Mysteries of Judicial Decision Making*, 26 *Loy. L.A. L. Rev.* 993, 997 (1993).

¹⁵⁰Edward V. Di Lello, *Note, Fighting Fire with Firefighters: A Proposal for Expert Judges at the Trial Level*, 93 *Colum. L. Rev.* 473 (1993); John W. Osborne, *Judicial/Technical Assessment of Novel Scientific Evidence*, 1990 *U. Ill. L. Rev.* 497, 540–43 (1990).

using litigant agreed-upon technical specialists,¹⁵¹ and creating special science courts.¹⁵² While some of these proposed changes may improve the ability for federal district courts to engage with empirical information, such changes are unlikely to take place due to the difficulties associated with such large changes to the structure of the federal court system.¹⁵³ Instead, federal courts should create positions for statistical experts. While expert witnesses partially fill this role, their incentive to support the litigant that hired them may make their testimony less trustworthy.¹⁵⁴ Court-staffed experts could serve as an impartial resource in cases involving complex statistical or empirical questions.

Our final proposal is that courts alter their norms and standards regarding the consideration of statistical evidence. Even with increased statistical training and resources, judges are still free to ignore statistical evidence. As Cho and Yoon bleakly put the situation, “it appears as though judges will continue to evaluate scientific evidence as he or she sees fit, subject to an abuse of discretion standard. The current system gives trial judges considerable discretion to be ambitious or lazy regarding scientific evidence.”¹⁵⁵ If this is correct, a meaningful improvement in judicial evaluation of statistical evidence may require a fundamental change in the way that judges think about and incorporate this evidence into their decisions.

While the general issue of the distinction between law and fact has been discussed by commentators for over a century,¹⁵⁶ it was only beginning in the 1950s that legal scholars started to notice that empirical reasoning was becoming embedded within legal decisions and therefore becoming a basis for legal precedent.¹⁵⁷ Walker and Monahan’s

¹⁵¹Ellen R. Jordan, *Specialized Courts: A Choice*, 76 *Nw. L. Rev.* 745, 745 (1981).

¹⁵²James A. Martin, *The Proposed “Science Court,”* 75 *Mich. L. Rev.* 1058, 1058 (1977).

¹⁵³This being said, individual judges have the option to hire law clerks with technical expertise. However, we doubt there are enough law school graduates with such a background to substantially alter the functioning of the federal court system.

¹⁵⁴For discussions of the partisanship and bias of expert witnesses, see Jennifer Mnookin, *Expert Evidence, Partisanship and Epistemic Competence*, 73 *Brook. L. Rev.* 587, 588–92 (2008); David E. Bernstein, *Expert Witnesses, Adversarial Bias, and the (Partial) Failure of the Daubert Revolution*, 93 *Iowa L. Rev.* 101, 104–09 (2008); Roscoe Pound, *The Causes of Popular Dissatisfaction with the Administration of Justice*, 14 *Am. Law.* 445, 448 (1906).

¹⁵⁵Cho & Yoon, *supra* note 135, at 263.

¹⁵⁶For some canonical pieces on the distinction between law and fact, see generally Ronald J. Allen & Michael S. Pardo, *The Myth of the Law-Fact Distinction*, 97 *Nw. U. L. Rev.* 1769 (2002); Clarence Morris, *Law and Fact*, 55 *Harv. L. Rev.* 1303 (1942); Nathan Isaacs, *The Law and the Facts*, 22 *Colum. L. Rev.* 1 (1922). See also Edward Coke, *Institutes of the Laws of England, or, A Commentary on Littleton* 460 (John Henry Thomas ed. 1818).

¹⁵⁷For an early discussion on the difficulty of the distinction between law and fact created by judicial reliance on empirical information, see generally Harold L. Korn, *Law, Fact, and Science in the Courts*, 66 *Colum. L. Rev.* 1080 (1966). Korn dutifully cataloged the possible sources of error for courts that engage with scientific evidence, the various ways in which scientific facts may be integrated within the law, and the important distinctions between legal adjudication and factual information that make their marriage a potentially difficult one to cement. *Id.* at 1081–1106. However, Korn provided little in the way of a prescriptive framework for how legal agents should view empirical information—only offering the vague possibility that such information be viewed as “precedents of less than conclusive force.” *Id.* at 1106.

pioneering work in the 1980s then began the process of attempting to form a systematic framework of how courts should approach using empirical information in the creation of legal rules.¹⁵⁸ Within this broad topic, we would like to make a narrow point: *when an empirical question is relevant for a particular decision, judges should fairly consider the empirical evidence on that question and accept or reject it based on its own merits.* A judge might reject a piece of empirical evidence because it is irrelevant for the case at hand, because the evidence is inconclusive, or because the assumptions of the researcher are unsound, but the evidence should not be dismissed out of hand. We believe these recommendations should apply both to judges who are first considering new statistical evidence and also to judges who are reconsidering the same kind of statistical evidence evaluated in previous cases.

Our discussion of the negative effect fallacy suggests that even legal precedent does not provide compelling grounds upon which to dismiss empirical evidence. Allison Orr Larsen documents the way in which the findings of the Supreme Court are often accepted uncritically as “factual precedent” by lower courts.¹⁵⁹ Instead of engaging on their own with empirical facts, lower federal courts often given precedential value to findings of fact made by the Supreme Court.¹⁶⁰ As demonstrated by the case of the negative effect fallacy, the same problem can arise with flawed logical arguments. Judges repeatedly utilized the negative effect fallacy across cases and legal domains, thereby ignoring potentially valuable empirical evidence. Once a piece of reasoning enters the jurisprudence, any judge can take that piece of reasoning—flawed or not—isolate it from its surrounding context in the original opinion, and apply it in a different case.¹⁶¹

When an empirical claim is relevant, judges should engage with the empirical evidence as opposed to dismissing the evidence or embracing precedent without

¹⁵⁸See generally Laurens Walker & John Monahan, *Scientific Methodology as Legal Precedent*, 76 Cal. L. Rev. 877 (1988); Laurens Walker & John Monahan, *Social Frameworks: A New Use of Social Science in Law*, 73 Va. L. Rev. 559 (1987); Laurens Walker & John Monahan, *Social Authority: Obtaining, Evaluating, and Establishing Social Science in Law*, 134 U. Pa. L. Rev. 477 (1986). Walker and Monahan argued it was a mistake for courts to view empirical information as simply a legislative or adjudicative fact, especially when such empirical information is involved in the creation of legal rules. Laurens Walker & John Monahan, *Social Authority: Obtaining, Evaluating, and Establishing Social Science in Law*, 134 U. Pa. L. Rev. 477, 479–85 (1986). Instead, Walker and Monahan put forth the idea that empirical information should be viewed as more analogous to “law,” and as a result should be treated by courts as similar to legal precedent. *Id.* at 490–91. Such empirical information that had precedential value should be labeled “social authority.” *Id.* at 488. In later writings, Walker and Monahan proposed that empirical methodologies should be given legal precedential effect, but that specific results from empirical methodologies should not be given precedential effect. Laurens Walker & John Monahan, *Scientific Methodology as Legal Precedent*, 76 Cal. L. Rev. 877, 879 (1988).

¹⁵⁹Alison Orr Larsen, *Factual Precedents*, 162 U. Pa. L. Rev. 59, 62 (2013).

¹⁶⁰Larsen identified five categories of factual precedents: (1) facts imported from one context to another, (2) facts that supplement the record for a strategic purpose, (3) facts that answer residual questions following landmark decisions, (4) historical factual precedents, and (5) facts that form the premise of a new legal rule. *Id.* at 79–97 (discussing various examples of factual precedent in Supreme Court jurisprudence).

¹⁶¹See Molly W. Lien, *Technocentrism and the Soul of the Common Law Lawyer*, 48 Am. U. L. Rev. 85, 130–31 (1998).

discussion. In many cases that cite the negative effect fallacy, judges typically quoted the fallacy from *Elkins* and quickly moved onto the next line in their argument. These opinions did not consider the flawed logic of the fallacy, nor is there any apparent humility in regard to the quality of evidence from a different discipline in which most judges have little expertise. Perhaps even more disquieting, judges frequently used the fallacy in order to reject wholesale the application of statistics to the legal question at issue in the case, which allowed them to rhetorically fall back to the more familiar terrain of legal reasoning through more free-form value judgments.

The active engagement with statistical evidence serves as an important signaling mechanism both to other judges and to other members of the legal community.¹⁶² When judges engage with evidence, they should explain the extent to which they find the evidence compelling and the extent to which the evidence influences their decision. As discussed above, in addition to increasing transparency, this practice will assist judges in future cases as they engage with the same evidence and incorporate new evidence on the same question as it becomes available.

Underlying our discussion is the belief that statistical evidence should be considered persuasive only to the extent that the analysis is persuasive to those expert in the discipline¹⁶³ and effectively communicated to those outside the discipline. A critical and skeptical eye toward empirical evidence by lower courts is especially warranted considering the uneasy and elusive distinction in practice between dicta and holding because the use of empirics in the former could easily be used to bind future courts by its use in the latter. As noted by Judge Pierre Leval, “[t]he distinction between dictum and holding is more and more frequently disregarded.”¹⁶⁴ The case study of the negative effect fallacy only further demonstrates that statistical claims utilized by the Supreme Court in dicta in one opinion can quickly jump into another jurisprudential domain.¹⁶⁵ Nonetheless, to the extent that empirical evidence is incorporated into legal decisions or ignored, we believe these decisions should be made on the grounds of legal relevance and the internal credibility of the results. A shift of norms and standards in this direction would make it more difficult for judges to ignore relevant evidence and would potentially improve the quality of judicial decision making.

¹⁶²For discussions regarding the signaling mechanisms involved in judicial opinions and precedence formation, see Robert J. Hume, *The Impact of Judicial Opinion Language on the Transmission of Federal Circuit Court Precedent*, 43 *Law & Soc. Rev.* 127, 131–34 (2009); Ethan Bueno De Mesquita & Matthew Stephenson, *Informative Precedent and Intrajudicial Communication*, 96 *Am. Pol. Sci. Rev.* 755, 765 (2002); William N. Eskridge & Philip P. Frickey, *Law as Equilibrium*, 108 *Harv. L. Rev.* 26, 39–42 (1994).

¹⁶³See Larsen, *supra* note 159, at 107.

¹⁶⁴Pierre N. Leval, *Judging Under the Constitution: Dicta About Dicta*, 81 *N.Y.U. L. Rev.* 1249, 1250 (2006). See also Judith M. Stinson, *Why Dicta Becomes Holding and Why it Matters*, 76 *Brook. L. Rev.* 219, 260 (2010).

¹⁶⁵See Larsen, *supra* note 159, at 74–76.

VI. CONCLUSION

In this article, we identify the negative effect fallacy and discuss its proliferation into numerous legal domains at all levels of the federal courts since 1960. Since its creation by Judge Stewart in the exclusionary rule case *Elkins v. United States*, the fallacy has spread to numerous legal domains in the Supreme Court and in lower courts, including free speech, voting rights, and campaign finance. Our analysis shows that the negative effect fallacy is typically used to dismiss actual or potential quantitative information that is highly relevant to the case at hand. We suspect a combination of statistical ignorance, reasoning from precedent, and motivated reasoning is largely responsible for the vitality of the fallacy.

We hope to bring the negative effect fallacy to light and prevent its further use. Our discussion highlights the challenges associated with interpreting statistical evidence in federal courts, and it produces several concrete suggestions for preventing the use of the negative effect fallacy and similar errors in future cases. As our case study of the negative effect fallacy shows, precedent alone is an insufficient justification for dismissing statistical evidence. Instead, we recommend that judges take a critical stance toward statistical evidence and actively evaluate the merits of statistical arguments on their own terms.