Nebraska Law Review

Volume 100 | Issue 3 Article 5

6-2022

When the Math Matters: Improving Statistical Advocacy in Gerrymandering Litigation

Robin L. Juni
George Washington University Law School

Robin L. Juni*

When the Math Matters: Improving Statistical Advocacy in Gerrymandering Litigation

ABSTRACT

Lawyers regularly joke about their supposed inability to address mathematical issues. However, mathematical concepts are sometimes at the core of a legal dispute, and lawyers do a disservice to their clients if they are not able to engage in effective advocacy in these contexts. This Article discusses Gill v. Whitford*, a gerrymandering dispute involving an important mathematical idea—the core statistical concept of regression analysis, particularly multivariable regression analysis—which Chief Justice Roberts referred to in* Whitford *as "sociological gobbledygook." In fact, the mathematical analysis has crucial implications and connections to the legal issues. This Article explains the statistics behind that analysis, connects them to the issues before the Court, and demonstrates that these concepts are employed in many types of litigation. The Article concludes that, even facing a skeptical court, lawyers who can understand and explain mathematical concepts in a way that resonates fully with the legal issues are best positioned to advocate for their clients. With new census results released in late 2021 and the opportunities for redistricting that will subsequently arise, mathematical understanding in the context of gerrymandering litigation is more critical than ever.*

TABLE OF CONTENTS

* Visiting Associate Professor, Fundamentals of Lawyering Program, the George Washington University Law School. J.D. Harvard Law School; B.A. Hamline University. The author thanks Brad Bush, GW Law Class of 2022, for superlative research assistance.

I. INTRODUCTION: MATH CAN MATTER

Lawyers are famously math averse, affirmatively disengaging when technical concepts come into play and regularly joking about innumeracy as a core characteristic of most in the profession.[1] Noting this phenomenon, scholars have sought to demonstrate areas in which mathematical reasoning is critical to various legal disputes and to explain how numerical skills might be improved among lawyers and law students.[2]

This Article seeks to build on that scholarship, taking one example from a legal dispute involving an important mathematical idea—the core statistical concept of regression analysis, particularly multivariable regression analysis—then explaining the underlying math involved and identifying ways in which counsel can better advocate for their clients in similar litigation.[3] The premise of this effort is that in many cases, mathematical concepts are critical to the result reached by a court or regulatory agency, and lawyers who cannot effectively engage with these concepts may disadvantage their clients. In short,

1. *See* Lisa Milot, *Illuminating Innumeracy*, 63 CASE W. RSRV. L. REV. 769 (2013) (leading with a quotation from then-First Lady Michelle Obama in remarks to the National Science Foundation: "I know for me, I'm a lawyer because I was bad at [science and math]. All lawyers in the room, you know it's true. We can't add and subtract, so we argue." (alteration in original) (quoting Michelle Obama, Remarks by the First Lady at the National Science Foundation Family-Friendly Policy Rollout (Sept. 26, 2011))); *see also* Edward K. Cheng, *Fighting Legal Innumeracy*, 17 GREEN BAG 2D 271, 271 (Spring 2014) (crediting mathematician John Allen Paulos with the observation that "while readers frequently condemn grammatical errors, wild mathematical ones often pass undetected" (quoting JOHN ALLEN PAULOS, INNUMERACY: MATHEMATICAL ILLITERACY AND ITS CONSEQUENCES 3–4 (2001))).

2. *See, e.g.,* Arden Rowell & Jessica Bregant, *Numeracy and Legal Decision Making*, 46 ARIZ. ST. L.J. 191 (2014); Theresa A. Gabaldon, *Strength in Numbers: Teaching Numeracy in the Context of Business Associations*, 59 ST. LOUIS U. L.J. 701 (2015).

3. Multivariable regression analysis has been recognized as having key applications in employment discrimination, antitrust, capital punishment, and patent damages matters, as well as voting rights litigation. *See* Franklin M. Fisher, *Multiple Regression in Legal Proceedings*, 80 COLUM. L. REV. 702 (1980).

the math matters. This Article tackles the issue by explaining how the courts addressed the statistical analysis in one piece of important litigation in which math was critical. Backing up, the Article lays out the relevant mathematical concepts before returning to their application in the litigation. Finally, the conclusion provides advice on how to move forward. The goal is an understandable roadmap to the necessary integration of math and law, such that a practitioner involved in similar litigation will be a better prepared advocate. With new census results recently released setting up a new round of redistricting efforts,[4] lawyers' ability to effectively engage with these issues is critical.

This analysis is structured around *Gill v. Whitford*, in which Wisconsin voters alleged the state legislative redistricting plan unconstitutionally diluted their votes through partisan gerrymandering.[5] The Supreme Court found the plaintiffs had not shown sufficiently concrete harm to establish Article III standing and thus remanded to the district court.[6] Though *Whitford* ultimately was dismissed after a subsequent Supreme Court holding that political gerrymandering cases are entirely non-justiciable,[7] during briefing and oral argument the Court engaged on the merits—including whether the "efficiency gap," the mathematical model used by the plaintiffs, effectively demonstrated improper gerrymandering.[8] During oral argument, Chief Justice Roberts, addressing appellees' counsel, Paul Smith, offered a critique of this evidence:

> Mr. Smith, I'm going to follow an example of one of my colleagues and lay out for you as concisely as I can what—what is the main problem for me and give you an opportunity to address it.

---

4. *2020 Census Results*, U.S. CENSUS BUREAU, https://www.census.gov/programs-surveys/decennial-census/decade/2020/2020-census-results.html [https://perma.cc/3YGN-DGMY] (last visited Sept. 23, 2021) ("On August 12, we released the redistricting data to the states and the public. States may use these data in redrawing congressional, legislative, and local district boundaries.").

5. Whitford v. Gill (*Whitford IV*),138 S. Ct. 1916 (2018).

6. *Id.* at 1934.

7. *Id.* at 1933. In July 2019, the district court dismissed *Whitford* without prejudice "in light of *Rucho v. Common Cause*, in which the Supreme Court held that 'partisan gerrymandering claims present political questions beyond the reach of the federal courts.'" Whitford v. Gill, No. 15-cv-421, 2019 U.S. Dist. LEXIS 111625 at *3 (W.D. Wis. July 2, 2019) (quoting Rucho v. Common Cause, 139 S. Ct. 2484, 2506 (2019)).

8. A mathematical model is a "mathematical framework representing variables and their interrelationships to describe observed phenomena or predict future events." PIETER EYKHOFF, SYSTEM IDENTIFICATION: PARAMETER AND STATE ESTIMATION (1974); *see also* Teegwendé V. Porgo et al., *The Use of Mathematical Modeling Studies for Evidence Synthesis and Guideline Development: A Glossary*, 10 RSCH. SYNTHESIS METHODS 125 (2018) (describing and defining a variety of mathematical models).

> I would think if these—if the claim is allowed to proceed, there will naturally be a lot of these claims raised around the country. Politics is a very important driving force and those claims will be raised.
>
> And every one of them will come here for a decision on the merits. These cases are not within our discretionary jurisdiction. They're the mandatory jurisdiction. We will have to decide in every case whether the Democrats win or the Republicans win. So it's going to be a problem here across the board.
>
> And if you're the intelligent man on the street and the Court issues a decision, and let's say, okay, the Democrats win, and that person will say: "Well, why did the Democrats win?" And the answer is going to be because EG was greater than 7 percent, where EG is the sigma of party X wasted votes minus the sigma of party Y wasted votes over the sigma of party X votes plus party Y votes.
>
> And the intelligent man on the street is going to say that's a bunch of baloney. It must be because the Supreme Court preferred the Democrats over the Republicans. And that's going to come out one case after another as these cases are brought in every state.
>
> And that is going to cause very serious harm to the status and integrity of the decisions of this Court in the eyes of the country.
>
> . . . .
>
> It is just not, it seems, a palatable answer to say the ruling was based on the fact that EG was greater than 7 percent. That doesn't sound like language in the Constitution.
>
> . . . .
>
> . . . [T]he whole point is you're taking these issues away from democracy and you're throwing them into the courts pursuant to, and it may be simply my educational background, but I can only describe as sociological gobbledygook.[9]

While his disclaimer attracted some media attention,[10] Chief Justice Roberts is far from the first justice to raise concerns about the ability of the Court to address scientific and mathematical issues.[11] The rise

9. Transcript of Oral Argument at 37–40, *Whitford IV*, 138 S. Ct. 1916 (No. 16-1161), 2017 WL 4517131, at *37–40.

10. Garrett Epps, *The Supreme Court's Choice on Partisan Gerrymandering*, ATLANTIC (Mar. 28, 2018), https://www.theatlantic.com/politics/archive/2018/03/the-supreme-courts-choice-on-partisan-gerrymandering/556661/ [https://perma.cc/8ALK-ETCU]; Oliver Roeder, *The Supreme Court Is Allergic to Math* (Oct. 17, 2017, 6:00 AM), https://fivethirtyeight.com/features/the-supreme-court-is-allergic-to-math/ [https://perma.cc/QW8J-73MP]; Philip Rocco, *Justice Roberts Said Political Science Is 'Sociological Gobbledygook.' Here's Why He Said It and Why He's Mistaken*, WASH. POST (Oct. 4, 2017), https://www.washingtonpost.com/news/monkey-cage/wp/2017/10/04/justice-roberts-said-political-science-is-sociological-gobbledygook-heres-why-he-said-it-and-why-hes-mistaken/ [https://perma.cc/3JYE-94FP].

11. *See, e.g.*, Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 599 (1993) (Rehnquist, C.J., concurring in part and dissenting in part) ("The various briefs filed in this case . . . deal with definitions of scientific knowledge, scientific method, scientific validity, and peer review—in short, matters far afield from the expertise of judges."); Marconi Wireless Tel. Co. v. United States, 320 U.S. 1, 60–61 (1943) (Frankfurter, J., dissenting in part) ("It is an old observation that the training of Anglo-American judges ill fits them to discharge the duties cast upon them by patent legislation.").

of ever-more sophisticated data analytics, however, makes the second decade of the twenty-first century a particularly inopportune time to continue the "lawyers can't do math" trope. When a court dismisses a key argument as inscrutable nonsense, it implicates a breakdown of advocacy as much as the limitations of the judiciary. The remainder of this Article seeks to unpack the issue before the Court in *Whitford* and to help figure out how to do it better.[12]

II. *GILL V. WHITFORD*

A. Case Background

The rise of the Republican Tea Party after the election of former president Barack Obama led to the "Tea Party Wave" of 2010—Republicans swept into power in many federal and state government positions, in many places for the first time in years.[13] This change in the political dynamic coincided with the 2010 census[14] and an opportunity for political redistricting based on new census numbers.[15]

12. At the outset, it is fair to recognize that the Supreme Court's reluctance to hear partisan gerrymandering cases might have carried the day no matter what arguments were made. *See* Roeder, *supra* note 10 ("[M]aybe this allergy to statistical evidence is really a smoke screen—a convenient way to make a decision based on ideology while couching it in terms of practicality." As University of Chicago law professor Daniel Hemel states: "I don't put much stock in the claim that the Supreme Court is afraid of adjudicating partisan gerrymanders because it's afraid of math . . . [Roberts] is very smart and so are the judges who would be adjudicating partisan gerrymandering claims—I'm sure he and they could wrap their minds around the math. The 'gobbledygook' argument seems to be masking whatever his real objection might be." (alteration in original)). However, even a skeptical court is less likely to dismiss mathematical arguments out of hand when those arguments are clearly made and directly tied to the legal claims at issue.
13. *See* Vanessa Williamson, Theda Skoepol & John Coggin, *The Tea Party and the Remaking of Republican Conservatism*, 9 PERSPS. ON POL. 25 (2011); Alan I. Abramowitz, *Right Turn: The 2010 Midterm Elections*, *in* PENDULUM SWING 55 (Larry J. Sabato ed., 2012).
14. *History: Agency History*, U.S. CENSUS BUREAU, https://www.census.gov/history/www/census_then_now/ [https://perma.cc/RZE8-Z5JG] (last visited Sept. 23, 2021); *see also* SARAH J. ECKMAN, CONG. RSCH. SERV., IN11360, APPORTIONMENT AND REDISTRICTING FOLLOWING THE 2020 CENSUS 1 (2021), https://crsreports.congress.gov/product/pdf/IN/IN11360 [https://perma.cc/U3PG-AN8D] (describing problems with 2020 census and explaining that "April 1 of any year ending in '0' marks the official decennial census date," although the Census Bureau begins certain population counts and outreach earlier. "A count known as the *apportionment population*, which reflects the total resident population in each state," including minors and noncitizens, in addition to Armed Forces and federal civilian personnel and their dependents living abroad, "is typically used to distribute House seats. Within nine months of the decennial census date (December 31 of the year ending in '0'), the Secretary of Commerce is to report the apportionment population to the President.").
15. U.S. CONST. art. 1, § 2, cl. 3. Although state legislatures enjoy substantial discretion in the process, constitutional guidelines for district creation require each dis-

In Wisconsin, Republicans captured a majority of state legislative seats in 2010. The Republican-led legislature then enacted a new districting plan in 2011, Wisconsin Act 43, based on the 2010 census.[16] The next year, the American people—and Wisconsin citizens—voted in the 2012 election. In the 2012 presidential race, Barack Obama won the state with 53.5% of the statewide vote, receiving Wisconsin's ten electoral votes.[17] In elections for the state legislature, however, Republicans won sixty out of ninety-nine Wisconsin Assembly districts, a result that engendered concern among Democrats.[18]

William Whitford and other Wisconsin citizens ultimately filed a complaint in federal court alleging that they were redistricted based on their affiliation to a specific political party—Whitford and the other plaintiffs registered as Democrats. They claimed that Republicans analyzed the geographic location and distribution of Democratic voters and shifted the lines of political districts to dilute Democratic Party votes throughout the state.[19] As explained by the district court, "the complaint alleges that Act 43 purposely distributed the predicted Re-

---

trict to be as equal in population to all other districts in the state "as practicable," while the boundaries and parameters of the districts should adhere to traditional districting principles like compactness and contiguity. *See* Growe v. Emison, 507 U.S. 25, 34 (1993); Larios v. Cox, 300 F. Supp. 2d 1320, 1330 (N.D. Ga. 2004); Justin Levitt, *Where Are the Lines Drawn?*, ALL ABOUT REDISTRICTING, https://redistricting.lls.edu/redistricting-101/where-are-the-lines-drawn/ [https://perma.cc/A247-PBK8] (last visited Sept. 23, 2021).

16. Joe Forward, *Redistricting: Litigation Common, Current Cycle Unique in Wisconsin*, STATE BAR WIS. (Mar. 7, 2012), https://www.wisbar.org/NewsPublications/RotundaReport/Pages/Article.aspx?ArticleID=7996#bio [https://perma.cc/Q4X2-NP6J].

17. Brendan Fischer, *Wisconsin's "Shameful" Gerrymander of 2012*, PRWATCH (Feb. 4, 2013, 8:15 AM), https://www.prwatch.org/NEWS/2013/02/11968/WISCONSINS-SHAMEFUL-GERRYMANDER-2012 [https://perma.cc/X9EP-XW7Y].

18. *Id.* Indeed, the plaintiffs' expert for the ensuing litigation characterized this result as absurd. Expert Report of Kenneth Mayer at 38–39, Whitford v. Nichol (*Whitford I*), 151 F. Supp. 3d 918 (W.D. Wis. 2015) (No. 3:15-cv-421), 2015 WL 10091018, at *38–39 ("Any discussion of Act 43 must begin with the basic fact that in 2012 Republicans achieved a 60–39 majority in the Assembly in an election in which the Democratic Party achieved 53.5% of the statewide two-party presidential vote. The imbalance between the Republican Party's statewide vote margin at the top of the ticket (46.5%) and its Assembly majority (60.6%) turns the very notion of partisan symmetry on its head. That standard . . . 'requires that the number of seats one party would receive if it garnered a particular percentage of the vote be identical to the number of seats the other party would receive if it had received the same percentage of the vote' . . . . Here, it means that Democrats would have had to obtain 60 Assembly seats with 46.5% of the vote, an absurd proposition that requires a party's legislative seat share to go *up* as its share of the vote goes *down*.").

19. Complaint at 9, *Whitford I*, 151 F. Supp. 3d 918 (No. 3:15-cv-421) 2015 WL 4651084, at *9. Like other gerrymandering cases, the case was heard by a three-judge panel and proceeded to direct review in the Supreme Court, pursuant to 28 U.S.C. § 2284. That section provides:

publican vote share with greater efficiency so that it translated into a greater number of seats, while purposely distributing the Democratic vote share with less efficiency so that it would translate into fewer seats."[20]

Specifically, the complaint alleged that all plaintiffs were affected by the general electoral map drawn by Republicans in 2011, and four plaintiffs further alleged they were directly affected by the redistricting map—Mary Lynn Donohue in the twenty-sixth district, Wendy Sue Johnson in the ninety-first district, Janet Mitchell in the sixty-sixth district, and Jerome Wallace in the twenty-third district. All four of these plaintiffs alleged their vote was directly diluted, either by the "cracking" or "packing" of their district.[21]

As in many gerrymandering cases,[22] the plaintiffs asserted that the Wisconsin electoral districts at issue were so gerrymandered that

> (a) A district court of three judges shall be convened . . . when an action is filed challenging the constitutionality of the apportionment of congressional districts or the apportionment of any statewide legislative body.
> (b) In any action required to be heard and determined by a district court of three judges under subsection (a) of this section, the composition and procedure of the court shall be as follows:
> (1) Upon the filing of a request for three judges, the judge to whom the request is presented shall, unless he determines that three judges are not required, immediately notify the chief judge of the circuit, who shall designate two other judges, at least one of whom shall be a circuit judge. The judges so designated, and the judge to whom the request was presented, shall serve as members of the court to hear and determine the action or proceeding.

28 U.S.C. § 2284(a), (b)(1). The panel in *Whitford v. Gill (Whitford III)*, 218 F. Supp. 3d 837 (W.D. Wis. 2016), was comprised of Seventh Circuit Judge Ripple and District Judges Crabb and Griesbach. A party is permitted direct appeal to the Supreme Court pursuant to 28 U.S.C. § 1253, which authorizes review of "an order granting or denying, after notice and hearing, an interlocutory or permanent injunction in any civil action, suit or proceeding required by any Act of Congress to be heard and determined by a district court of three judges."

20. *See Whitford III*, 218 F. Supp. 3d at 854.
21. Complaint, *supra* note 19, at 9.
22. The practice of gerrymandering, of course, has deep historical roots, with the name itself deriving from an 1812 Massachusetts election. *See* THE VOLUME LIBRARY: AN ENCYCLOPEDIA OF PRACTICAL AND CULTURAL INFORMATION 1407 (1939) (defining gerrymander as "the process of redistricting a state or other political division so that one party gets an unfair advantage over another" and further explaining that "[i]n 1812 when Elbridge Gerry was governor of Massachusetts, the election districts were redistributed to the advantage of the Democratic-Republican party, and one district (a part of Essex County) had boundaries so eccentric that on the map the outline looked like a dragon. One politician compared it to a salamander. 'No, a Gerrymander,' another said. So the word (originally pronounced with a hard *g,* but now with a *j* sound for the initial letter) has become a general term for unfair districting. (The history experts say that Governor Gerry was not to blame; the legislature made the new election districts.)"). Gerry later served from March 1813 until his death in November 1814 as the fifth vice president of the United States under James Madison. Senate Hist. Off., *Elbridge Gerry, 5th Vice President (1813–1814)*, U.S. SENATE, https://www.senate.gov/

they adversely affected the plaintiffs' rights to free speech, free association, and equal protection.[23] The plaintiffs claimed they were targeted for gerrymandering because of their affiliation with the Democratic Party of Wisconsin and their past voting history as Democrats.[24] The complaint averred: "Wisconsin Act 43 violates the First and Fourteenth Amendment because it intentionally uses [plaintiffs'] partisan affiliation to affect the weight of their votes."[25] Thus, the plaintiffs alleged, Wisconsin Act 43 infringed on their First Amendment rights to express a political opinion and to affiliate with the party of their choice, and their Fourteenth Amendment guarantee of equal treatment as citizens, regardless of party affiliation or voting characteristics.[26]

The district court pointed out that the Supreme Court had not precluded relief for a partisan gerrymandering claim, though the Court had not heard the merits of such a case for lack of an appropriate standard of review. However, the plaintiffs proposed a three-part test for establishing an unconstitutional partisan gerrymander, and the district court found it judicially manageable enough to survive the defendants' motion to dismiss.[27] First, the plaintiffs must prove "that the defendants intended to discriminate against an 'identifiable political group' of which the plaintiffs are a member."[28] Second, the plaintiffs had to demonstrate discriminatory effect of the electoral map, which was accomplished in *Whitford* with efficiency gap statistical

about/officers-staff/vice-president/gerry-elbridge.htm [https://perma.cc/CT9B-SANL] (last visited Sept. 23, 2021).

23. *See* Gaffney v. Cummings, 412 U.S. 735 (1973); Davis v. Bandemer, 478 U.S. 109 (1986); Vieth v. Jubelirer, 541 U.S. 267 (2004); League of United Latin Am. Citizens v. Perry, 548 U.S. 399 (2006). In particular, districts must be approximately equal in population, so as not to violate the "one-person, one-vote" principle of the Equal Protection Clause of the Fourteenth Amendment. *See* Reynolds v. Sims, 377 U.S. 533, 568 (1964). Moreover, a voting plan must comply with Section 2 of the Voting Rights Act of 1965, which preserves the voting power of minority groups. 52 U.S.C. § 10301; Thornburg v. Gingles, 478 U.S. 30, 47 (1986).
24. Complaint, *supra* note 19 at 9.
25. *Id.* at 24–28.
26. *Id*.
27. *Whitford I*, 151 F. Supp. 3d 918, 928, 931 (W.D. Wis. 2015); Recent Case, Whitford v. Gill*, No. 15-cv-421-bbc, 2016 WL 6837229 (W.D. Wis. Nov. 21, 2016)*, 130 HARV. L. REV. 1954 (2017).
28. *Whitford I*, 151 F. Supp. 3d at 927. The plaintiffs presented numerous allegations on this issue, including that the Republican state legislators formed a plan to dilute Democratic voters, hired mathematicians to create an electoral map with the intent to dilute Democratic votes, and asked Republican legislators to sign non-disclosure agreements. The complaint specifically names several people who allegedly orchestrated the creation of the new, intentionally discriminatory plan. The complaint alleges that the Republican legislators not only intended to create an electoral map that favored Republicans but also expressly intended to create a map with a twelve percent efficiency gap favoring Republicans. Complaint, *supra* note 19 at 9.

analysis. If the plaintiffs successfully established intent and discriminatory effect, the burden shifted to the defendants for step three, to controvert the claim by showing the electoral map bias was "the necessary result of either a legitimate state policy or the state's underlying political geography."[29]

B. Efficiency Gap

The *Whitford* complaint presents the efficiency gap as a mathematical tool to determine if political districting is unconstitutional. Generally, the political gerrymandering claim avers that ignoring natural geographic lines, votes are either: (1) "cracked" between districts, where a geographic block of opposite-party voters is split to dilute the votes across districts and fall short of a majority in each one; (2) "packed," where opposite-party voters are consolidated into a small number of districts where they elect their preferred candidates but waste many votes in landslide victories, votes that may have made other district elections competitive, had they not been packed; or (3) both.[30] One party thus is able to maximize the total number of elected representatives without winning any more voters from one election to the next but rather by moving voters from one jurisdiction to another.[31] The number of votes needed—and therefore the amount of cracking or packing required—can be calculated by reviewing voting

29. *Whitford I*, 151 F. Supp. 3d at 928. Such "underlying political geography" might occur if, for example, "political groups that tend to cluster (as is the case with Democratic voters in cities) [c]ould be systematically affected by what might be called a 'natural' packing effect." Vieth v. Jubelirer, 541 U.S. 267, 290 (2004). *But see* ANTHONY J. MCCANN, ET AL., GERRYMANDERING IN AMERICA 135 (2016) ("[G]eographic and demographic constraints (such as the urban concentration of Democratic voters, the requirement to draw majority–minority districts, and the geographic sorting of voters) cannot account for the level of partisan bias we observe and certainly cannot account for the increase in bias we observe between the 2000 and 2010 districting rounds.").

30. Nicholas O. Stephanopoulos & Eric M. McGhee, *Partisan Gerrymandering and the Efficiency Gap*, 82 U. CHI. L. REV. 831, 850–51 (2015) ("Our analysis begins with the premise that the goal of a partisan gerrymander is to win as many *seats* as possible given a certain number of *votes*. To accomplish this aim, a party must ensure that its votes translate into seats more 'efficiently' than do those of its opponent. In the plurality-rule, single-member district (SMD) elections that are almost universal in American politics, 'inefficient' votes are those that do not directly contribute to victory. Thus, any vote for a losing candidate is wasted by definition, but so too is any vote beyond the 50 percent threshold needed (in a two-candidate race) to win a seat. If these supporters could be moved through redistricting to a different seat, they could help the party claim that seat as well without changing the outcome in the seat from which they were moved." (footnote omitted)).

31. *Id.* at 850 ("A gerrymander is simply a district plan that results in one party wasting many more votes than its adversary. And the efficiency gap indicates the magnitude of the divergence between the parties' respective wasted votes. It aggregates all of a plan's cracking and packing choices into a single number.").

data across multiple elections, then determining the necessary minimum threshold for victory.[32] Votes affected by cracking and packing are considered "diluted," because they lack the power to which each is otherwise entitled,[33] while "wasted" votes are those in excess of what was needed to win a given district or all votes in districts that are not won. The efficiency gap is calculated by dividing the difference in wasted votes by the total number of votes cast, quantifying the effects of cracking and packing. In other words, the purpose of cracking and packing is not just to dilute and waste votes for the opposing party, but to maximize the weight and importance of each vote for the preferred candidate, so the gulf between the parties' wasted votes will be significant where districts have been engineered for precisely that result. Once the discriminatory effect is quantified as the efficiency gap, it can be compared to the efficiency gaps of previous district maps or of similarly situated districts elsewhere.

Building on the cracking and packing concept, the *Whitford* approach explains the efficiency gap argument in terms of earlier Supreme Court jurisprudence regarding partisan bias.[34] The crux of those earlier arguments is the premise that electoral systems should treat similarly-situated political parties equally, such that in theory, partisan districts undercutting that premise could be unconstitutional.[35] However, the Court has articulated a critical concern regarding the partisan bias approach, finding that "[t]he existence or degree of [bias] may in large part depend on conjecture about where possible vote-switchers will reside. . . . [W]e are wary of adopting a constitutional standard that invalidates a map based on unfair results that would occur in a hypothetical state of affairs."[36]

---

32. *Id.*

33. *See, e.g.*, Thornburg v. Gingles, 478 U.S. 30, 47 (1986) (discussing predicates for vote dilution in the context of racial gerrymandering claims); *see also* Benjamin Plener Cover, Rucho *for Minimalists*, 71 MERCER L. REV. 695, 697 (2020) ("[T]he Court conceptualized partisan gerrymandering as allocative vote dilution, whereby a mapmaker reallocates votes towards a favored party by subjecting disfavored voters to packing and cracking techniques that diminish the efficacy of their votes.").

34. *See* Eric McGhee, *Measuring Partisan Bias in Single-Member District Electoral Systems*, 39 LEGIS. STUD. Q. 55, 68 (2014) (defining "partisan bias" as the difference between the share of the seats that a party would win if the parties tied in a statewide election, and the actual share received).

35. *See* Nicholas O. Stephanopoulos & Eric M. McGhee, *The Measure of a Metric: The Debate over Quantifying Partisan Gerrymandering*, 70 STAN L. REV. 1503, 1505 (2018) ("An analyst would estimate the seat shares the major parties would win in a state if (hypothetically) they each received the same vote share. The greater the divergence between the parties' seat shares for the same (counterfactual) vote share, the larger a district plan's *partisan bias*, and the more gerrymandered the plan.").

36. League of United Latin Am. Citizens v. Perry, 548 U.S. 399, 420 (2006).

Importantly, Justice Kennedy's concurring opinion in *Vieth v. Jubelirer* stated that though the plaintiffs in that case had not presented the Court with a manageable standard to adequately assess the effects of political gerrymandering, it was possible that another plaintiff may present a justiciable claim in the future.[37] For Justice Kennedy, the key was giving a "limited and precise rationale" that established a constitutional violation, stating that "if workable standards do emerge . . . courts should be prepared to order relief."[38] The *Whitford* plaintiffs stepped into that breach, presenting the efficiency gap as a metric for partisan bias in gerrymandering cases that avoided the flaws inherent in previously proposed standards, flaws such as requiring that courts evaluate hypothetical outcomes of a future vote—an approach the Court discouraged in *League of United Latin American Citizens v. Perry*.[39] Instead, *Whitford* sought to present accurate and measurable discriminatory effects of gerrymandering by creating a district plan that minimized the efficiency gap and then comparing that plan to the actual map used in Wisconsin Act 43.

Specifically, the *Whitford* plaintiffs presented two empirical studies, one conducted by Kenneth Mayer, professor of political science at the University of Wisconsin-Madison, and one by Simon Jackman, then a professor of political science at Stanford.[40] Citing the Mayer report, the complaint breaks down the numbers and percentages of wasted votes in specific counties in Wisconsin and the percentile calculation of partisanship bias based on the efficiency gap metric.[41]

The plaintiffs detailed Wisconsin's efficiency gap data under Act 43, alleging that the Wisconsin plan had the twenty-eighth worst efficiency gap score—meaning it has an extremely high level of partisan gerrymandering—out of nearly 800 total plans surveyed. The complaint described six Wisconsin districts gerrymandered through cracking or packing and alleged seven more were otherwise adversely

37. Vieth v. Jubelirer, 541 U.S. 267, 306 (2004) (Kennedy, J., concurring). Of course, the Court's decision in *Rucho v. Common Cause*, 139 S. Ct. 2484 (2019), finding that political gerrymandering cases are nonjusticiable, has undercut the approach, though it still is relevant in racial gerrymandering cases and perhaps in future political gerrymandering litigation.

38. *Vieth*, 541 U.S. at 306, 317.

39. *League of United Latin Am. Citizens*, 548 U.S. at 420.

40. Expert Report of Kenneth Mayer, *supra* note 18, at 1; Expert Report of Simon Jackman at 1, *Whitford I*, 151 F. Supp. 3d 918 (W.D. Wis. 2015) (No. 3:15-cv-421), 2015 WL 10091020, at *1.

41. Complaint, *supra* note 19, at 18. Professor Jackman's expert report explains:

> The efficiency gap (*EG*) is a relative, wasted vote measure, the ratio of one party's wasted vote rate to the other party's wasted vote rate. *EG* can be computed directly from a given election's results, without recourse to extensive statistical modeling or assumptions about counterfactual or hypothetical election outcomes, unlike other extant measures of the fairness of an electoral system (e.g., partisan bias).

Expert Report of Simon Jackman, *supra* note 40 at 3.

affected by redistricting efforts, morphing from not heavily partisan to solidly Republican in 2012.[42]

Ruling on a motion to dismiss, the District Court for the Western District of Wisconsin allowed the case to proceed, finding that electoral mapping must give equal weight to all votes unless the discrepancy is based on permissible, legislative concerns.[43] The court discussed the efficiency gap methodology, explaining that the calculation is performed by creating statewide totals for each party of (1) the number of votes cast for the losing candidates in district races (as a measure of cracked voters), along with (2) the number of votes cast for the winning candidates in excess of the fifty-percent-plus-one votes necessary to secure the candidate's victory (as a measure of packed voters). The resulting figure is the total number of wasted votes for each party.[44] As explained by the district court in its merits ruling, the efficiency gap then is calculated by taking the difference in the wasted votes for each party, divided by the overall votes cast.[45] Under the methodology, the higher the efficiency gap, the greater the number of seats won because of the opposing party's wasted votes.[46] The plaintiffs' experts testified that an efficiency gap model showing a threshold number greater than seven percent[47] in the first year indicates the redistricting plan likely will continue to favor one party for the life of the plan.[48] Thus, the plaintiffs contended, because the

42. Complaint, *supra* note 19, at 19–22.

43. *See* Whitford v. Nichol (*Whitford II*), 180 F. Supp. 3d 583, 587 (W.D. Wis. 2016).

44. *Id.* at 588–89.

45. *Whitford III*, 218 F. Supp. 3d 837, 900 (W.D. Wis. 2016).

46. *See* Stephanopoulos & McGhee, *supra* note 30, at 834.

47. *Whitford III*, 218 F. Supp. 3d at 905–06.

48. Chief Justice Roberts focused on the seven percent number at oral argument, making the entire enterprise seem arbitrary or nonsensical. Transcript of Oral Argument, *supra* note 9 at 37–40 ("It is just not, it seems, a palatable answer to say the ruling was based on the fact that EG was greater than 7 percent. That doesn't sound like language in the Constitution."). A dissent in the district court likewise confuses relevant numbers. Posing a hypothetical district where Party A wins 5,200 votes and Party B wins 4,800 votes, Judge Griesbach claims that Party A wasted only 199 votes while party B wasted all 4,800—producing, in his words, "an eye-popping efficiency gap of 46%!" *Whitford III*, 218 F. Supp. 3d at 956 (Griesbach, J., dissenting). He asserts that the idea of the efficiency gap is founded on circular logic—that as Republicans win more close elections—which, as in Judge Griesbach's hypothetical, generate a large number of wasted Democratic votes and a small number of wasted Republican votes—the efficiency gap would indicate that the election was the result of gerrymandering. He crucially misunderstands that the metric is to be applied statewide, as some districts will be cracked and others packed, and that its usefulness is as a comparative tool to provide evidence of intentional gerrymandering, not a prescription for proportional representation. *Id.*

model forecast a twelve percent efficiency gap for Wisconsin's redistricting plan, it should be found presumptively unconstitutional.[49]

C. Supreme Court Guidance

Although the Supreme Court ultimately remanded the case to the lower court on standing grounds,[50] the *Whitford* decision provides guidance for potential future plaintiffs seeking to litigate similar claims. Most fundamentally, the majority opinion—authored by Chief Justice Roberts—explains its concept of individualized harm in gerrymandering cases:

> We have long recognized that a person's right to vote is "individual and personal in nature." . . . [T]he plaintiffs claim a constitutional right not to be placed in legislative districts deliberately designed to "waste" their votes in

---

49. Complaint, *supra* note 19, at 26; Expert Report of Simon Jackman, *supra* note 40, at 5. The district court ultimately did not accept that argument, explaining that:

> [I]n any event, that we have not determined that a particular measure of EG establishes presumptive unconstitutionality, which itself diminishes all of the defendants' policy-based arguments. Instead, we acknowledge that the expert opinions in this case have persuaded us that, on the facts before us, the EG is corroborative evidence of an aggressive partisan gerrymander that was both intended *and likely* to persist for the life of the plan.

Whitford III, 218 F. Supp. 3d at 910.

50. Others have sought to explain why the Court addressed the case on the merits at all, especially in light of the institutional legitimacy concerns articulated by Chief Justice Roberts at oral argument that "there will naturally be a lot of these claims raised around the country . . . [a]nd every one of them will come here for a decision on the merits." Transcript of Oral Argument, *supra* note 9, at 37. Professors Joshua A. Douglas and Michael E. Solimine explain why the Court engaged with the substantive issues despite concerns about appearing political:

> The answer is that the cases did not come to the Court through its normal discretionary certiorari jurisdiction. Instead, both *Gill* and *Benisek* involved direct appeals to the Supreme Court from three-judge district courts. The Court had no choice but to hear the cases. Chief Justice Roberts's concern was that if the Court recognized a claim for partisan gerrymandering, then *all* of these cases would reach the Supreme Court on mandatory direct appeal, not via the Court's discretionary certiorari jurisdiction where the Court could avoid ruling on them.
>
> Under federal law, when three-judge district courts hear cases such as *Gill* and *Benisek*, a losing party may appeal directly to the Supreme Court, skipping the court of appeals. Thus, in *Gill* and *Benisek*, because the losing parties appealed the decisions, the Court had to do *something* with respect to the cases. Unlike a denial of certiorari, refusing to hear them would have been tantamount to a summary affirmance on the merits, which lower courts likely would have considered to have at least some precedential value. Instead, the Court took a middle ground in *Gill* and *Benisek,* hearing them on the merits but providing little guidance on the doctrine of partisan gerrymandering beyond saying in *Gill* that plaintiffs must demonstrate individual harm from a map in a specific district.

Joshua A. Douglas & Michael E. Solimine, *Precedent, Three-Judge District Courts, and the Law of Democracy,* 107 GEO. L.J. 413, 415–16 (2019) (footnotes omitted).

> elections where their chosen candidates will win in landslides (packing) or are destined to lose by closer margins (cracking).
>
> To the extent the plaintiffs' alleged harm is the dilution of their votes, that injury is district specific.[51]

Because "a plaintiff's remedy must be 'limited to the inadequacy that produced [his] injury in fact.' . . . the remedy that is proper and sufficient lies in the revision of the boundaries of the individual's own district."[52] The majority opinion concludes: "On the facts of this case, the plaintiffs may not rely on 'the kind of undifferentiated, generalized grievance about the conduct of government that we have refused to countenance in the past.'"[53] This conception of the case in the majority opinion allowed the Court to dismiss the relevance of the mathematical analysis.[54]

---

51. *Whitford IV*, 138 S. Ct. 1916, 1929–30 (2018) (citations omitted). The Court further explained:

> [Lead plaintiff Whitford] testified that he lives in Madison in the 76th Assembly District, and acknowledged on cross-examination that this is, under any plausible circumstances, a heavily Democratic district. Under Act 43, the Democratic share of the Assembly vote in Whitford's district is 81.9%; under the plaintiffs' ideal map—their Demonstration Plan—the projected Democratic share of the Assembly vote in Whitford's district would be 82%. Whitford therefore conceded that Act 43 had not "affected [his] ability to vote for and elect a Democrat in [his] district." Whitford testified that he had nevertheless suffered a harm "relate[d] to [his] ability to engage in campaign activity to achieve a majority in the Assembly and the Senate." As he explained, "[t]he only practical way to accomplish my policy objectives is to get a majority of the Democrats in the Assembly and the Senate ideally in order to get the legislative product I prefer."

Id. at 1924–25 (alterations in original) (citations omitted).

52. *Id.* at 1930 (quoting Lewis v. Casey, 518 U.S. 343, 357 (1996)) ("For similar reasons, we have held that a plaintiff who alleges that he is the object of a racial gerrymander—a drawing of district lines on the basis of race—has standing to assert only that his own district has been so gerrymandered.").
53. *Id.* (quoting Lance v. Coffman, 549 U.S. 437, 442 (2007)).
54. *See id.* ("The plaintiffs asserted in their complaint that the 'efficiency gap captures in a single number all of a district plan's cracking and packing.' That number is calculated by subtracting the statewide sum of one party's wasted votes from the statewide sum of the other party's wasted votes and dividing the result by the statewide sum of all votes cast, where 'wasted votes' are defined as all votes cast for a losing candidate and all votes cast for a winning candidate beyond the 50% plus one that ensures victory. The larger the number produced by that calculation, the greater the asymmetry between the parties in their efficiency in converting votes into legislative seats. Though they take no firm position on the matter, the plaintiffs have suggested that an efficiency gap in the range of 7% to 10% should trigger constitutional scrutiny. . . . We need not doubt the plaintiffs' math. The difficulty for standing purposes is that these calculations are an average measure. They do not address the effect that a gerrymander has on the votes of particular citizens. Partisan-asymmetry metrics such as the efficiency gap measure something else entirely: the effect that a gerrymander has on the fortunes of political parties. . . . That shortcoming confirms the fundamental problem with the plaintiffs' case as presented on this record. It is a case about group political interests, not individual legal rights. But this Court is not responsible

A concurring opinion, authored by Justice Kagan and joined by Justices Ginsburg, Breyer, and Sotomayor, agrees with this standing analysis but provides further guidance for future plaintiffs making a claims of individualized vote dilution and of associational harm, both of which might be founded on the types of statistical methodologies presented by the *Whitford* plaintiffs.[55]

More specifically, Justice Kagan explains that:

> To have standing to bring a partisan gerrymandering claim based on vote dilution, then, a plaintiff must prove that the value of her own vote has been "contract[ed]." . . . Consider the perfect form of each variety. When a voter resides in a packed district, her preferred candidate will win no matter what; when a voter lives in a cracked district, her chosen candidate stands no chance of prevailing. But either way, such a citizen's vote carries less weight—has less consequence—than it would under a neutrally drawn map. So when she shows that her district has been packed or cracked, she proves, as she must to establish standing, that she is "among the injured."[56]

Importantly for our purposes, Justice Kagan elucidates: "Among other ways of proving packing or cracking, a plaintiff could produce an alternative map (or set of alternative maps)—comparably consistent with traditional districting principles—under which her vote would carry more weight."[57] In other words, Justice Kagan would look for precisely the kind of mapping and statistical analyses presented in *Whitford*.[58]

---

for vindicating generalized partisan preferences. The Court's constitutionally prescribed role is to vindicate the individual rights of the people appearing before it." (citations omitted)).

55. *Id.* at 1934 (Kagan, J., concurring) ("I write to address in more detail what kind of evidence the present plaintiffs (or any additional ones) must offer to support that allegation. And I write to make some observations about what would happen if they succeed in proving standing—that is, about how their vote dilution case could then proceed on the merits.").

56. *Id.* at 1935–36 (citations omitted).

57. *Id.* at 1936. Justice Kagan further explains that:

> Here, the Court is right that the plaintiffs have so far failed to make such a showing. William Whitford was the only plaintiff to testify at trial about the alleged gerrymander's effects. He expressly acknowledged that his district would be materially identical under any conceivable map, whether or not drawn to achieve partisan advantage. That means Wisconsin's plan could not have diluted Whitford's own vote. So whatever other claims he might have, Whitford is not "among the injured" in a vote dilution challenge. Four other plaintiffs differed from Whitford by alleging in the complaint that they lived in packed or cracked districts. But for whatever reason, they failed to back up those allegations with evidence as the suit proceeded. So they too did not show the injury—a less valuable vote—central to their vote dilution theory.

Id. (citations omitted) (quoting Lujan v. Defenders of Wildlife, 504 U.S. 555, 563 (1992)).

58. *But see* Rucho v. Common Cause, 139 S. Ct. 2484, 2503–04 (2019) ("Even the most sophisticated districting maps cannot reliably account for some of the reasons voters prefer one candidate over another, or why their preferences may change. Voters elect individual candidates in individual districts, and their selections depend on the issues that matter to them, the quality of the candidates, the

Justice Kagan goes on to address a claim based on infringement of the First Amendment right of association—which would not hinge on a showing of vote dilution based on cracking or packing—finding that "the plaintiffs' evidence of partisan asymmetry well fits a suit alleging associational injury."[59] She concludes that although the First Amendment claim was not fleshed out enough to avoid dismissal for lack of standing, since the statistical proof of partisan asymmetry best measures "a gerrymander's effect 'on the fortunes of political parties' and those associated with them," the evidence "could easily have supported the alternative theory of associational harm."[60]

Thus, while *Rucho*[61] has abrogated judicial limits on partisan gerrymandering cases for the moment, racial gerrymandering and other cases that raise similar analyses may yet find a salutary reception at the Court. We now turn to unpacking those statistical analyses in *Whitford*.

III. THE MATH

The efficiency gap study presented in the complaint—the essential piece of mathematical analysis that underpins the plaintiffs' arguments—was conducted and authored by Kenneth Mayer, a professor of political science at the University of Wisconsin-Madison.[62]

Professor Mayer based his mathematical analysis on the efficiency gap—a term for the number of wasted voters of one party compared to another.[63] The efficiency gap compares how effectively the parties translate votes into seats—when a party has a high efficiency gap, the percentage of seats won is significantly higher than the percentage of

---

tone of the candidates' campaigns, the performance of an incumbent, national events or local issues that drive voter turnout, and other considerations. Many voters split their tickets. Others never register with a political party, and vote for candidates from both major parties at different points during their lifetimes. For all of those reasons, asking judges to predict how a particular districting map will perform in future elections risks basing constitutional holdings on unstable ground outside judicial expertise."). As in *Whitford*, Chief Justice Roberts, who authored the majority opinion, continues in *Rucho* to question the ability of the judiciary to cope with the mathematical underpinnings of partisan gerrymandering litigation, despite broad judicial acceptance of the necessary multivariable regression techniques, as discussed in Part III of this Article.

59. *Whitford IV*, 138 S. Ct. at 1937–39 (Kagan, J., concurring).

60. *Id*. at 1939.

61. *Rucho*, 139 S. Ct. 2484.

62. Expert Report of Kenneth Mayer, *supra* note 18, at 1.

63. *See* Stephanopoulos & McGhee, *supra* note 30, at 831.

the total votes received.[64] An efficiency gap of zero would mean that the parties had the same number of wasted votes.[65]

Professor Mayer analyzed the 2012 Presidential vote in Wisconsin to determine the efficiency gap of the Act 43 map. He explained that:

> Instead of estimating future election results for an existing or proposed hypothetical plan, [his] goal was to determine whether it was possible to draw a district plan following the 2010 Census that minimized the efficiency gap while maintaining strict fidelity to the federal and state constitutional requirements of population equality, contiguity, compactness, respect for political subdivisions, and compliance with the Voting Rights Act.[66]

He used reliable statistical methodologies to draw a hypothetical map and determined the efficiency gap for the hypothetical map in the 2012 election.[67] In doing so, he demonstrated that the Act 43 electoral map was politically skewed and how applied mathematics can inform political decision-making. This section describes the structure of his analysis.

A. Overview

To create a new, hypothetical electoral map, Professor Mayer moved wards into different districts—or, to think about it another

64. *Id.* at 853 ("Beyond its positive and normative appeal, the efficiency gap has a number of useful properties that warrant discussion. First, under circumstances that are very common in US [sic] elections, it is unnecessary to sum the wasted votes in each individual district—a process that can be somewhat cumbersome. Instead, if we assume that all districts are equal in population (which is constitutionally required), and that there are only two parties (which is typical in [single-member district] systems), then the computation reduces through simple algebra to something quite straightforward: Efficiency Gap = Seat Margin – (2 × Vote Margin)[.] In this formula, 'Seat Margin' is the share of all seats held by a party, minus 50 percent. 'Vote Margin' is the same for votes: the share received by a party, minus 50 percent. A party has an electoral advantage when the efficiency gap is positive, and a disadvantage when it is negative. When the number is equal to zero, there is no efficiency gap and so no partisan benefit derived from redistricting." (footnotes omitted)).

65. Though an efficiency gap is unlikely to ever be precisely zero, the measure delineates negative numbers as reflecting a higher number of wasted votes for Democrats, and positive numbers reflect a higher number of wasted votes for Republicans. *See id.* at 849; Expert Report of Simon Jackman, *supra* note 40, at 3 ("The efficiency gap is an 'excess seats' measure, reflecting the nature of a partisan gerrymander. An efficiency gap in favor of one party sees it wasting fewer votes than its opponent, thus translating its votes across the jurisdiction into seats more efficiently than its opponent. This results in the party winning more seats than expected, given its vote share (V) and if wasted vote rates were the same between the parties. $EG = 0$ corresponds to no efficiency gap between the parties, or no partisan difference in wasted vote rates. In this analysis (but without loss of generality) EG is normed such that negative EG values indicate higher wasted vote rates for Democrats relative to Republicans, and $EG > 0$ the converse.").

66. Expert Report of Kenneth Mayer, *supra* note 18, at 6.

67. *Id.*

way, drew the lines of the Congressional districts to include wards in new assembly districts. One benefit to working at the ward level when analyzing voting behavior is that each ward contains a small number of people (and therefore, voters), allowing greater precision when drawing the new electoral map. Since Mayer worked with smaller base units of voters, he was able to more accurately place voters into hypothetical electoral boundaries.[68]

Wisconsin has ninety-nine assembly districts, each of which elects a state representative. Each district has between 24 and 135 wards and wards each have an average of 869 people.[69] Professor Mayer created a regression model for assembly votes based on the 2012 voting data. However, Mayer could not simply count the number of votes in the new district based on the votes cast in each ward. Crucially, some assembly districts had uncontested elections, which would skew the statistical data.[70] In many cases, an assembly district with an uncontested race might indicate a heavily partisan district—but also might indicate that the ward was involved in cracking or packing. To avoid this issue, Professor Mayer created a regression model based on the 2012 votes to predict *how people would have voted in 2012 under a different electoral map.*

B. Simple Linear Regression

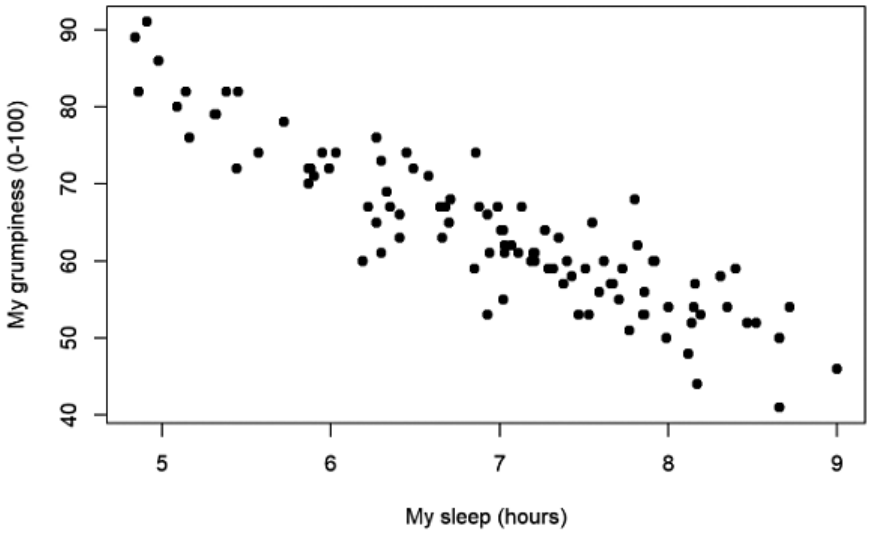
In statistics, a linear regression model asks predictive questions about future behavior based on characteristics, which researchers describe as variables. For example, a linear model between two variables, the effect of hours of sleep on "level of grumpiness," might appear as follows:[71]

68. *Id.* at 8–9. Although statistical analysis is typically built on large numbers, the smaller the base unit of the analysis, the more accurate the calculations. PHILIP H. POLLOCK III, THE ESSENTIALS OF POLITICAL ANALYSIS 27 (5th ed. 2016).

69. *See generally* Legis. Tech. Servs. Bureau, *Wisconsin Legislative Map Library*, WIS. STATE LEGISLATURE, https://legis.wisconsin.gov/ltsb/gis/maps/ [https://perma.cc/3AU3-8WNU] (last visited Sept. 23, 2021).

70. Expert Report of Kenneth Mayer, *supra* note 18, at 5.

71. DANIELLE NAVARRO, *Linear Regression, in* LEARNING STATISTICS WITH R: A TUTORIAL FOR PSYCHOLOGY STUDENTS AND OTHER BEGINNERS chapter 15 (Emily Kothe trans., 2019), https://learningstatisticswithr.com/book/regression.html [https://perma.cc/ZPU4-XMBX].



In this experiment, the researcher noted how many hours they slept and then also took note of their perceived level of grumpiness on a scale of zero to 100.[72] The researcher collected these data points over time, and then documented them in the above graph. The *regression line* is the line that best fits the data points. The researcher is looking for the line that runs through each of the data points, minimizing the distance between the line and the data point for all the data points.[73] This line is the most common way of using the collected data to predict future grumpiness levels based on hours of sleep.

The equation, $y = a + b(x)$, expresses the regression model, where x is the independent variable (number of hours of sleep), y is the dependent variable (grumpiness), and b is the slope of the line. In the above equation, the slope of the line will be negative—as the number of hours of sleep increase, grumpiness decreases. The letter a is the y-intercept, representing the level of grumpiness if the person were to get no sleep at all. So given this equation, a person can plug in any

72. *Id.* at fig. 15.1.

73. More specifically, the least squares regression line is one that minimizes the sum of the squared vertical distances from each individual data point to the line. SAMPRIT CHATTERGEE & ALI S. HADI, REGRESSION ANALYSIS BY EXAMPLE 33 (5th ed. 2012). These vertical distances represent "errors" or "residuals" in the predicted (or "fitted") verse actual data points. *Id.* The least squares regression line, then, is the line that minimizes the sum of these squared errors. *Id.* at 34. By design, the vertical distances sum to zero which means the sum of the vertical distances above the line equal the sum of the vertical distances below the line. *Id.*

number for x, the number of hours of sleep, to predict the level of grumpiness.[74]

Variance is a mathematical term for the spread in a data set; higher levels of variance make the data prediction less accurate.[75] Some variables—like sleep and grumpiness—may be strongly correlated, which means that one variable can explain the other. Some variables are not related at all. A linear regression line is most accurate when looking at data between two variables that are strongly related to one another—strong correlation and low variance means the predictive power of the data is higher.[76]

Variance is often calculated using a least squares regression, which minimizes the distance between each point of a scatterplot and the generated line.[77] Given any set of data, there exists one line that minimizes this distance. The generated least squares regression line thus has the lowest level of variance possible given the data points, and in this way, it is the *best fit* for the data.[78]

---

74. Note that the relationship is associational and not necessary causal. It might be that the lack of sleep is caused by other factors (e.g. pressure at work), and these are direct causes of the grumpiness, not lack of sleep. Regression is not, in and of itself, able to identify causal relationships. *See* D.H. Kaye, *The Dynamics of* Daubert*: Methodology, Conclusions, and Fit in Statistical and Econometric Studies*, 87 VA. L. REV. 1933, 1994 (2001) (explaining linear regression equation in the context of market share litigation).

75. *Trend (Regression) Analysis*, *in* ENVIRONMENTAL SCIENCE DESKBOOK § 1:24 (James W. Conrad, Jr. ed., 2020) Westlaw ("A line can be fit through any series of points. The important question is whether the slope and direction of the line illustrate some meaningful correlation between the two variables, or whether the line is merely a function of chance. . . . The evaluation involves testing how close the measured data values equal the values that would be predicted by the regression line. These residuals (differences) are expected to be normally distributed, with a mean of zero and a constant variance. . . . [T]he null hypothesis is assumed (typically, that the slope (m) is zero, which would mean that the dependent variable is not changing as the independent variable changes). One then asks how likely it would be that the observed variance would occur if the dependent variable was not changing in some relationship with the independent variable. A t-distribution is generally used to answer this question, because it yields the interval that would contain the variance at a given level of confidence. Typically, a 95-percent confidence level is used; the higher the level used, the smaller the variance must be for the slope to be significant.").

76. *Id.*

77. *Id*. *See generally* POLLOCK III, *supra* note 68, at 183–241.

78. *See* 1 DAVID L. FAIGMAN ET AL., *The Basics of Multiple Regression*, *in* MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY app. 6A (2020) ("A *regression line* is the best-fitting straight line through a set of points in a scatterplot. . . . The regression line typically is estimated using the standard method of *least-squares*, where the values of a and b are calculated so that the sum of the squared deviations of the points from the line are minimized. In this way, positive deviations and negative deviations of equal size are counted equally, and large deviations are counted more [than] small deviations."). *See generally* POLLOCK III, *supra* note 68, at 183–241.

C. Multivariable Regression

In research like Professor Mayer's, scientists analyze how a dependent variable relates to many different interrelated independent variables. When trying to find a regression equation to determine voting patterns, scientists must consider how votes, the dependent variable, relate to the many different variables that play a part in determining how those votes are cast.[79]

A multivariable regression[80] looks at how those independent variables correlate to the dependent variable. Statistical software can examine many data points, even with many different characteristics, and produce a predictive equation that considers all the variables. The software will generate an equation like the one utilized in simple linear regression, but with more independent variables: $y = a + b_i(x_i)$. In

79. Multivariable regression is well accepted as a tool in judicial decisions. *See* Daniel L. Rubinfeld, *Reference Guide on Multiple Regression*, *in* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 305–06 (3d ed. 2011), https://www.nap.edu/read/13163/chapter/8 [http://perma.cc/2JY7-SGM7] ("Multiple regression analysis is a statistical tool for understanding the relationship between two or more variables. Multiple regression involves a variable to be explained—called the *dependent variable*—and additional *explanatory variables* that are thought to produce or be associated with changes in the dependent variable. For example, a multiple regression analysis might estimate the effect of the number of years of work on salary. Salary would be the dependent variable to be explained; years of experience would be the explanatory variable. Multiple regression analysis is sometimes well suited to the analysis of data about competing theories in which there are several possible explanations for the relationship among a number of explanatory variables. Multiple regression typically uses a single dependent variable and several explanatory variables to assess the statistical data pertinent to these theories. In a case alleging sex discrimination in salaries, for example, a multiple regression analysis would examine not only sex, but also other explanatory variables of interest, such as education and experience. The employer–defendant might use multiple regression to argue that salary is a function of the employee's education and experience, and the employee–plaintiff might argue that salary is also a function of the individual's sex. Multiple regression also may be useful (1) in determining whether a particular effect is present; (2) in measuring the magnitude of a particular effect; and (3) in forecasting what a particular effect would be, but for an intervening event." (footnotes omitted)).

80. Multivariable regression is sometimes inappropriately described as multivariate regression; however, the latter is a more complicated field of statistical analysis. *See* Bertha Hidalgo & Melody Goodman, *Multivariate or Multivariable Regression?*, 103 AM. J. PUB. HEALTH 39 (2013) (explaining that either a single or multivariable regression model "has a single outcome variable and [one] or more independent or predictor variables. Multivariate, by contrast, refers to the modeling of data that are often derived from longitudinal studies, wherein an outcome is measured for the same individual at multiple time points (repeated measures), or the modeling of nested/clustered data, wherein there are multiple individuals in each cluster").

the multivariable equation, *n* represents the variable, and each variable will have a coefficient multiplied by the value of that variable.[81]

Statistical software will create the most accurate equation possible given the diversity of the input and report whether each variable coefficient is statistically significant. If it is, then the software has concluded that coefficient may be an important predictor in determining the outcome, based on the data given.[82]

For example, statistical software would analyze the sleep and grumpiness data above and determine the equation $y = 125.6 + (-8.94)x$ to describe that relationship.[83] In real terms, a person might have a grumpiness level of 125.6 with no sleep. But each hour of sleep decreases that level by 8.94 points—so with eight hours of sleep, the same person would have a grumpiness level of 54.08. Now imagine that we take other variables into account to determine grumpiness level—like stress at work. In a multivariable analysis, the software will create a precise equation for all variables and at the same time, will determine whether each coefficient is accurately predictive.

In other words, the software will tell you if the (–8.94) above is sufficiently different from zero to indicate the likelihood that the relationship between sleep and grumpiness is real. This evaluation of the difference from zero or a null effect is called statistical significance—if the coefficient is sufficiently different from zero, it means that the coefficient may be predictive based on the analyzed data.[84]

---

81. *See id.* 39–40. Choice of both the independent and dependent variables is critical to reliability of the regression results. *See, e.g.*, 1 FAIGMAN ET AL., *supra* note 78, § 6:17 ("It is essential in multiple regression analysis that the explanatory variable of interest [the dependent variable] not be correlated perfectly with one or more of the other explanatory variables [the independent variables]. If there were perfect correlation between two variables, the expert could not separate out the effect of the variable of interest on the dependent variable from the effect of the other variable. Suppose, for example, that in a sex discrimination suit a particular form of job experience is determined to be a valid source of high wages. If all men had the requisite job experience and all women did not, it would be impossible to tell whether wage differentials between men and women were due to sex discrimination or differences in experience.").

82. *See* POLLOCK III, *supra* note 68, at 156–79.

83. NAVARRO, *supra* note 71.

84. NAVARRO, *supra* note 71; *see also* RAMONA L. PAETZOLD & JASON R. BENT, THE STATISTICS OF DISCRIMINATION: USING STATISTICAL EVIDENCE IN EMPLOYMENT DISCRIMINATION CASES § 6:7 (2020) ("The individual regression coefficients are weights attached to the predictor variables; for example, the coefficient of the sex variable in our example (-$321) is the weight explicitly attached to being female for the community college's compensation model. Assuming that individual coefficients can be interpreted without error, it is of interest whether this coefficient is significantly different from zero. If the coefficient is not significantly different from zero, then sex can be viewed as playing no role in the community college's compensation system. If the coefficient is significantly different from zero, then a community college teacher's sex would appear to be associated with that teacher's

Thus, based on the statistical software's information, we can determine if the data analyzed is adequate to predict future behavior. This approach is the basic premise of Mayer's analysis. He compiled voting data based on many different characteristics, then used statistical software to predict voting behavior—not in future elections, but in a hypothetical 2012 election where the district boundaries were drawn differently.

D. A Hypothetical Voter

Think about a hypothetical voter living in hypothetical Wisconsin assembly district 1. That district may have many characteristics: wealthy, diverse, or politically polarized, for example. Perhaps the district is so heavily partisan that most races go uncontested. Such factors may affect how our hypothetical voter collectively votes at every political level—local, state, and national. In 2012, such a Wisconsin voter cast ballots for representatives in the city, county, assembly district, United States House of Representatives, United States Senate, and Presidency. For each race, a certain number of people voted for each candidate on the ballot. These are *specific vote totals.*

Now, imagine that assembly district 1 included more or fewer neighborhoods. Imagine how different the assembly district might look if it included other voters. Our hypothetical voter might no longer be voting with people who are wealthy, diverse, or belonging to the same party. These changes might—and almost certainly would—affect the specific vote totals in assembly district 1.

To understand Mayer's equation, the hypothetical voter is an effective way to conceptualize the relationship between voter characteristics and vote totals. In his model, the dependent variable is the specific vote totals—the number of votes received—for each 2012 assembly candidate. This number is dependent on the other variables, specifically, the characteristics of voters in each ward that makes up the assembly district. These characteristics are *independent variables*. Mayer analyzed eight independent variables, categorized in four types: district demographics, underlying partisanship, incumbency, and fixed geographic effects:[85]

---

compensation level, so that males tend to earn, on average, $321 more than females." (footnotes omitted)).

85. These variables were chosen because they have been shown to reliably affect how votes are cast. Expert Report of Kenneth Mayer, *supra* note 18, at 11. For experts making choices about these variables, lawyers often can add value, questioning the basis for each choice of input to the model, and ensuring that each can be factually supported and withstand cross-examination. *See, e.g.*, Spaulding v. Univ. of Wash., 740 F.2d 686, 703 (9th Cir. 1984) ("Thus, the weight of statistical proof relies implicitly on 'the existence of proper supportive facts and the absence of variables which would undermine the reasonableness of the inference of discrimination.'" (quoting White v. City of San Diego, 605 F.2d 455, 460 (9th Cir.

(1) The eligible number of voters in a ward, as measured by the 2010 census;
(2) The voting eligible Black population in a ward;
(3) The voting eligible Hispanic population in a ward;[86]
(4) The Number of votes cast for Barack Obama [the Democrat] in the 2012 Presidential election in a ward;
(5) The Number of Votes cast for Mitt Romney [the Republican] in the 2012 Presidential election in a ward;[87]

1979))); Bullington v. United Air Lines, Inc., 186 F.3d 1301, 1313 n.8 (10th Cir. 1999) (explaining that an analysis that fails to account for "important variations" might create "statistical anomalies that are not really probative"); Gallaway v. Empire Fire & Marine Ins., No. 03-113, 2007 WL 1199502 (W.D. La. Apr. 20, 2007), *aff'd*, Medlin v. Newman, 255 F. App'x 892 (5th Cir. 2007) (providing an example of a draconian sanction rendered after inappropriate modeling). In *Gallaway*, the plaintiffs alleged that they had been "engulfed" by a "cloud" of hydrochloric acid (HCl) fumes after a truck carrying liquid HCl was involved in an accident that caused acid to leak from the truck. *Id.* at *2. Although no one disputed that some level of fumes was present at the accident, absent defensive modeling efforts, the plaintiffs' testimony of their exposure might have gone unchallenged. *Id.* at *2–3. However, an air dispersion model used by emergency responders to predict chemical dispersion after an accidental release showed that the plaintiffs "could not have been exposed to harmful levels of HCl such that would have caused their alleged chronic symptoms." *Id.* at *3. The court was persuaded that the defense expert who presented the model had reviewed available factual information about the event, including meteorological records and emergency response reports, and had made appropriate inputs to the model regarding—among other things—the amount of chemical spilled, the location of the chemical source, the duration of the spill, and the position of the plaintiffs relative to that source. *Id.* Because the plaintiffs did not present any evidence to counter the defendants' model, the court granted summary judgment to the defendants and dismissed all of the plaintiffs' claims. *Id.; see also* Thomas C. Redman, *If Your Data Is Bad, Your Machine Learning Tools Are Useless*, HARV. BUS. REV. (Apr. 2, 2018), https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless [https://perma.cc/94V6-MWTL] (setting forth five steps to ensure for a well-executed quality program for analyzing data).

86. More specifically, Mayer explains that:

> Total VEP [Voting Eligible Population] constitutes a baseline of the size of the voting population, reflecting the fact that the number of votes will be a function of total population. Black and Hispanic VEP are additional controls that reflect the partisan tendencies of key subpopulations as well as turnout likelihood. Traditionally, both African American and Hispanic populations vote at lower rates tha[n] whites, although in 2012 African American turnout was comparable to white turnout. Hispanic populations vote at lower rates than other demographic groups, in part because of a higher noncitizen population, but also because of socioeconomic factors known to reduce turnout.

Expert Report of Kenneth Mayer, *supra* note 18, at 12–13.

87. Presidential voting is an important variable because:

> The presidential vote is widely used as an exogenous measure of district level partisanship, and it correlates very strongly with other more complex measures of partisan strength.
>
> The presidential vote is, not surprisingly, an extremely strong predictor of the legislative vote. If we know how many votes were cast for the Republican presidential candidate in a ward we will have a very good idea, subject to some conditions, of how many votes will be cast for the Republican candidate in the legislative election in that ward. While not

(6)Whether the Assembly election in the ward has a Democratic incumbent;
(7) Whether the Assembly election in the ward has a Republican incumbent;[88]
(8) A set of geographic effects for each county.[89]

---

everyone who votes for the Republican presidential candidate will vote for the Republican state legislative candidate, nearly all will, and we can precisely quantify the nature of that relationship.

. . . .

An important property of the presidential vote as an independent variable in this model is that it can be treated as exogenous to (i.e., not caused by) the legislative vote. Exogeneity can be described in two ways. The first is in causal terms. Most voters will vote for the same party for the president and state Assembly. . . . These voters are consistent because they are Democrats or Republicans, and partisanship is the factor that explains both vote choices. Other voters will make their Assembly choice based on their presidential vote, because they use party labels as a cue when voting in a down-ticket race. "[P]arties are generally known by the presidential candidates they nominate, and candidates for state legislative races are a good deal less well known to voters than the congressional candidates who ride presidential coattails." Few voters, if any at all, will decide on an Assembly candidate first and then vote for president on the basis of their Assembly vote preference. The causal arrow runs from the presidential vote to the Assembly vote, not from the Assembly vote to the presidential vote. This is why we speak of presidential coattails affecting legislative races, and not the other way around.

The second reason why the presidential vote is exogenous to the Assembly vote is that it is not affected by local district-level conditions such as incumbency, spending, or candidate quality. The broader factors that influence the presidential vote, and the presidential candidates themselves, are the same in every Assembly district. The presidential vote is affected by underlying partisanship, national conditions and the characteristics of the presidential candidates, factors that are constant whether that vote is aggregated at the state, district, or ward levels.

Id. at 13, 16–17 (citations omitted) (quoting James E. Campbell, *Presidential Coattails and Midterm Losses in State Legislative Elections* 80 AM. POL. SCI. REV. 45, 46 (1986)).

88. Professor Mayer stated:

> The incumbency advantage is perhaps the most well-known feature of contemporary legislative elections. Legislative incumbents rarely lose, and usually win by large margins. All other things being equal, an incumbent will get more votes than a non-incumbent. The causes of this advantage are less important in this context than its magnitude. The model takes into account the incumbency advantage by noting whether an incumbent is running in an Assembly district.
>
> Incumbency effects are measured with a dummy variable equal to [one] when a candidate is an incumbent, and [zero] otherwise, multiplied by the ward voting eligible population to create an interactive variable that accounts for differences in size from one ward to the next. Since the dependent variable is an actual vote count, the value of incumbency—in terms of how many additional votes incumbents receive—will vary with the number of voters who reside in a ward.

Id. at 17–18 (footnotes omitted) (citation omitted).

89. Professor Mayer explained:

> The last set of variables estimate the effect that county geography has on the Assembly vote. Some counties in Wisconsin are heavily Republican and some heavily Democratic. It is possible that a voters' county of resi-

Professor Mayer calculated the assembly vote total for each ward twice—once for the Democratic vote total and once for the Republican vote total and explained the assembly vote as a function of these eight variables.

Importantly, however, Mayer did not just look at the vote totals in each ward, redraw the political districts, and tally up the number of votes in each ward. To better understand why that is, imagine this rectangle represents voting districts. District 1 includes four wards: A–D. District 2 includes two wards: E and F.

Assembly District 1 – 750 Votes				**Assembly District 2 – 400 Votes**	
Ward A: 100 Votes	Ward B: 100 Votes	Ward C: 200 Votes	Ward D: 350 Votes	Ward E: 200 Votes	Ward F: 200 Votes

Now assume ward D is moved into assembly district 2. It might be tempting simply to shift the 350 votes in ward D out of district 1 and into district 2. However, this choice would be inappropriate, because it would ignore the fact that voting totals are a product of the characteristics of *that district*—and thus require a more fulsome analysis. While the number of votes cast will likely remain relatively unchanged if ward D is pushed into district 2, the multivariable regression equation ensures that any differences in the independent variables prevalent in ward D are considered.

The beauty of models is that they can be tested. In his analysis, Professor Mayer first used the efficiency gap model he created to predict the 2012 vote with the Act 43 districting plan; in other words, he compared the model's results with the actual results. The model was extremely accurate, correctly predicting the winner in seventy of seventy-two districts. In the two incorrect districts, the Republican candidates received 51.9% and 49.7% of the vote, with an average absolute error in the vote margin of 1.49%.[90] Importantly, by confirming the

dence could have an effect on the vote choice, whether because of sorting, socialization or assimilation, or other unobserved effects.

Id. at 18 (citations omitted).

90. *Id.* at 24–25.

> As important as the prediction of actual district vote totals is the model's ability to accurately identify the winner, as the efficiency gap calculation is sensitive to the party of the winners and losers. . . .
>
> . . . .
>
> These results, which compare predicted election results to the actual election results, demonstrate that the model is very accurate. A harder test involves the accuracy of predictions using data not in the sample—that is, applying the model to data and election results that are different from the data used to estimate the model. . . .
>
> . . . [The figure showing] actual versus predicted vote totals . . . are not grouped as tightly . . . but still show a very high degree of accuracy.

accuracy of his model, Mayer validated his choices and calculations of independent variables. The next step in the process: test the application of the model with a different districting plan, as if Act 43 were not enacted.

E. A Demonstration District Plan

Professor Mayer created a different districting plan for the 2012 election—one he designed to work within constitutional requirements to draw a non-partisan map.[91] To do this, he evaluated each census block within a specific ward. Wisconsin census blocks have between zero and 2,988 eligible voters, with an average of seventeen,[92] and each ward is comprised of approximately forty census blocks. Mayer determined the census block-level voting characteristics, using the same eight independent variables as his earlier analysis, and entered the data into Maptitude, a commercial redistricting software.[93] Within the state and federal constitutional parameters of "equal population, continuity, compactness, and respect for political subdivisions," Mayer created a redistricting plan that minimized the efficiency gap and maximized competitive districts.[94]

Professor Mayer then analyzed the voting data to create an equation to predict the outcome of the 2012 vote under different electoral maps. He did not create the equation to predict future votes; rather, he created an equation to predict how the vote would have occurred if the Wisconsin Legislature never enacted Act 43. He found that the demonstration plan resulted in an efficiency gap of just 2.2%, more than five times lower than the Act 43 efficiency gap of 11.69%.[95] Professor Mayer concluded that:

> The model does an excellent job accurately forecasting vote totals and election results, and provides a solid foundation for estimating hypothetical vote totals in an alternative district plan.

Id. at 23–28 (footnotes omitted)).

91. *Id.* at 6. Constitutional guidelines for district creation require each district to be as equal in population to all other districts in the state "as practicable," while the boundaries and parameters of the districts should adhere to traditional districting principles like compactness and contiguity. *See* U.S. CONST. art. I, § 2.
92. Expert Report of Kenneth Mayer, *supra* note 18, at 31.
93. *Id.* Specifically, Mayer used Maptitude for Redistricting 2013, Build 2060. *Id.*
94. *Id.* at 34.
95. *Id.* at 45 (summarizing data as "show[ing] that the Demonstration Plan results in 741,984 wasted Democratic votes . . . obtained by adding the number of lost Democratic votes cast for losing candidates . . . and the number of surplus Democratic votes cast for winners above what was necessary to win . . . . The same calculation for Republicans . . . results in 689,570 wasted Republican votes. The difference between these two numbers, 781,984 – 689,570 = 62,414 net wasted Democratic votes. Dividing 62,414 by the predicted total number of votes 2,843,108, produces the baseline efficiency gap for [the] plan, 0.0220, or 2.20%. Table 8 . . . shows the same calculation for Act 43 districts, using estimated partisan vote totals with incumbent advantages removed. Act 43 resulted in a total of 332,552 net wasted

> The results demonstrate that Act 43 was an egregious gerrymander, packing Democratic voters into a small number of districts and distributing Republican voters efficiently in a large number of districts in which they constituted safe majorities. . . . [A]reas of Democratic strength large enough to constitute majorities in single districts were unnecessarily split and then combined with larger Republican populations to create additional Republican districts and eliminate Democratic districts. . . . This packing and cracking was so successful that Republicans won 61% of Assembly seats in 2012, while obtaining only 46.5% of the statewide presidential vote.[96]

F. Permanence of the Efficiency Gap

We return now to the effect of the efficiency gap present in the Wisconsin data and the seven percent threshold that so vexed Chief Justice Roberts. Simon Jackman, then a professor of political science at Stanford University, built on Mayer's work to further explain the efficiency gap, comparing the Wisconsin districting plan to other legislative maps to show it was gerrymandered.[97] In his analysis, he compares the Wisconsin districts against legislative maps in other states across the United States to show that the Republican version of the Wisconsin map (Act 43) was an outlier.[98] He takes the argument a step further by calculating the efficiency gap of legislative districting plans across multiple elections to show that a legislative districting plan like the one in Wisconsin can have a sustained effect of disproportionately allocating legislative seats. Therefore, the gerrymandering could not be explained away based on the vagaries of one election.[99]

Professor Jackman likewise analyzed the level of wasted votes, so critical to the efficiency gap approach, concluding that: "To be sure, both parties are wasting votes. But partisan advantage ensues when one part is wasting fewer votes than the other, or, equivalently, more efficiently translating votes into seats."[100] While he boils down the efficiency gap as a measure of excess seats, his analysis takes on some complexity when he describes the "seats-vote" curve, which is his way of saying the proportion of legislative seats to votes in an election.[101] Jackman points out that seat totals are not directly proportional even

---

Democratic votes. The efficiency gap of Act 43 is 11.69%, more than five times larger than the Demonstration Plan.").

96. *Id.* at 54.

97. Expert Report of Simon Jackman, *supra* note 40, at 1. Since April 2016, Professor Jackman has been a professor of political science and chief executive officer of the United States Studies Centre at the University of Sydney in his native Australia. *Professor Simon Jackman*, U.S. STUD. CTR., https://www.ussc.edu.au/people/simon-jackman [https://perma.cc/ZZT6-M84M] (last visited Sept. 23, 2021).

98. Expert Report of Simon Jackman, *supra* note 40 at 69.

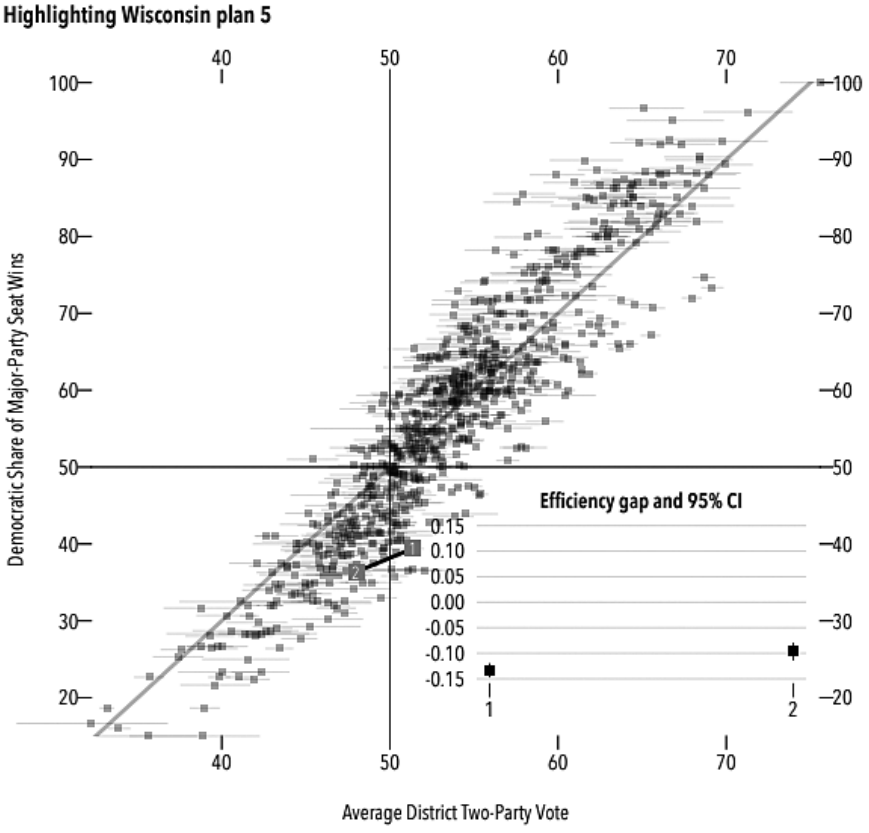
99. *Id.* at 19.

100. *Id.* at 15.

101. *Id.*

with a legislative map with no efficiency gap—a party will not win fifteen percent of seats in a legislative body by winning fifteen percent of the statewide vote. Rather, Professor Jackson points out that—assuming a two-candidate election—a party must win fifty-one percent of the vote in *a district* to win a seat in the legislature.[102]

The crux of his analysis examines numerous districting plans in various states over a number of years, comparing the extent to which their efficiency gaps diverge from zero.[103] More specifically, Professor Jackman calculated the efficiency gap in state legislative elections from 1972 to 2014, examining data from 83,269 district-level state legislative contests from 786 elections in forty-one states.[104] He then compared the Wisconsin efficiency gap against these measures to show the effect of the Act 43 plan.



102. *Id.*
103. *See* POLLOCK III, *supra* note 68, at 133.
104. Expert Report of Simon Jackman, *supra* note 40, at 15.

On the graph of all efficiency gaps in the elections in the comparison, Wisconsin's 2012 and 2014 are numbered squares 2 and 1, respectively.[105] The diagonal line represents an efficiency gap of zero (equal amount of wasted Democratic and Republican votes).[106] He concluded:

> The historical analysis reported above supports the proposition that Wisconsin's *EG* scores are likely to endure over the course of the plan. Few states ever record *EG* scores as large as those observed in Wisconsin; indeed, there is virtually no precedent for the lop-sided, two election sequence of *EG* scores generated in Wisconsin in 2012 and 2014 in the data [he] analyze[d] here (1972–2014). The closest historical analogs suggest that a districting plan that generates an opening, two-election sequence of *EG* scores like those from Wisconsin will continue to do so, generating seat shares for Democrats that are well below those we would expect from a neutral plan.[107]

Professor Jackman also discusses pinpointing the efficiency gap for a districting plan that will likely ensure the district remains uncompetitive until the next ten-year census.[108] His analysis attempts to find the *threshold* number for the efficiency gap—how many votes need to be packed into a district to ensure that the party continues to win elections in the district, thus making the district uncompetitive for the foreseeable future. He concluded, based on the data, that "an efficiency gap above 7% in any districting plan's first election year will continue to favor that party for the life of the plan."[109] Professor Jack-

105. *Id.* at 15.
106. *Id.* at 9, 70 fig. 34.
107. *Id.* at 71.
108. Professor Jackman was cautious in reaching this result:

> We do need to be careful when making these kinds of *relative* assessments about the magnitude of the efficiency gap. If pro-Republican gerrymandering is widespread, then it will be less unusual to see a large, negative *EG* estimate, at least contemporaneously; in fact this appears to the case in the post-2010 set of elections, where the longer-term distinctiveness of the Wisconsin numbers is matched and in some cases exceeded by other states also recording unusually large, negative *EG* estimates (e.g., Florida, Michigan, Virginia and North Carolina). This speaks to the utility of the longer-term, historical analysis in other research and in this report. It is important to remember that *EG = 0* corresponds to a partisan symmetry in wasted vote rates; we should be wary of arguments that would lead us to tolerate small to moderate levels of the efficiency gap because they appear to be the norm in some period of time, or in some set of jurisdictions.

   Id. at 65 (alterations in original) (citing Stephanopoulos & McGhee, *supra* note 30).
109. *Whitford III*, 218 F. Supp. 3d 837, 905 (W.D. Wis. 2016). As explained by the district court:

> First, Professor Jackman compared districting plans across a wide variety of states, and determined that over 95% of plans with an EG of at least 7% will never have an EG that favors the opposite party. Second, Professor Jackman conducted a "swing analysis" of all redistricting plans since 2010 and determined that nearly all plans that resulted in a 7% efficiency gap favoring one party in the first election year will retain an efficiency gap that favors that same party, even when one adjusts a party's statewide vote share by five points.

man estimated Wisconsin's efficiency gap was –.13 (thirteen percent more favorable to Republican candidates) in 2012 and –.10 (ten percent more favorable to Republican candidates) in 2014.[110]

G. What Does It All Mean?

It is important to understand the purpose of these analyses to understand what they tell us. Professor Mayer created a mathematical equation to accurately predict the voting totals at the ward level across the state in 2012, which allowed him to draw up different electoral maps to simulate the 2012 Presidential vote using those new maps and thereby show the Act 43 electoral map was *purposefully* gerrymandered.[111] By creating an electoral map where each party won seats proportional to their percentage of the overall vote totals—that districts could be drawn so there was no efficiency gap at all (or only a small one)—Mayer was able to show that the decision to select a map with such a high efficiency gap, one that could (and did) benefit the party drawing the map, was deliberate. In short, the presence of a large efficiency gap is evidence of a gerrymandered map.

Similarly, Professor Jackman evaluated efficiency gaps over a broad set of electoral data spanning over forty years. After comparing Wisconsin's election results under Act 43 to this expansive dataset, he concluded that there was "no precedent" for the Act 43 election results, and thus the Act 43 map likely was gerrymandered. Jackman also used the data comparison to determine that any map with an efficiency gap of more than seven percent—when one political party is utilizing its votes at least seven percent more efficiently than the other—is both gerrymandered and likely to dictate outcomes until electoral maps are redrawn, presumably no sooner than the next cen-

---

. . . .

> . . . Professor Jackman observed that, even with a 5% swing in the Democrats' favor, the EG would not drop below 7%.

Id. (footnotes omitted). Professor Jackman himself stated:

> This analysis points to a benchmark of about –.06 or –.07 as the actionable threshold given a first election with $EG < 0$ (Democratic disadvantage) or .08 or .09 when we observe $EG > 0$ in the first election under a redistricting plan (Democratic advantage); the asymmetry here reflects the fact that districting plans evincing apparent Democratic advantages are not as durable or as common (in recent decades) as plans presenting evidence of pro-Republican gerrymanders. At these proposed benchmarks the overall confidence rates are estimated to be 95%, with this confidence rate corresponding to a benchmark used widely in statistical decision-making in many fields of science.

Expert Report of Simon Jackman, *supra* note 40, at 66.

110. Expert Report of Simon Jackman, *supra* note 40, at 69.

111. Expert Report of Kenneth Mayer, *supra* note 18, at 38.

sus. His findings illustrate that the 2012 result in Wisconsin was not a single-election anomaly.[112]

Thoroughly unpacked and evaluated, these conclusions seem understandable and relatable to the Wisconsin gerrymandering claim. But law is about advocacy, and lawyers and law students are not always familiar with the necessary mathematical concepts. These concepts can provide powerful evidence and can advance legal arguments, but often remain underutilized or neglected, highlighting the need for increased mathematical awareness among legal professionals.

IV. MATH AND LEGAL ADVOCACY

Math is a tool that can advance legal argument—but lawyers and legal professionals must keep abreast of math, statistics, and applied statistical analysis. *Gill v. Whitford* provides one important example of mathematical application in litigation, and some guidance on how to move forward in effective legal advocacy using mathematical concepts.

A. Get Familiar with Math

A lawyer can effectively use math in advocacy—or competently counter opposing counsel's use—only if they thoroughly understand the mathematical concepts at hand. In some cases, the math is relatively simple; in other cases, more complex. In *Whitford*, the mathematical analysis is not incredibly complex. A multivariable regression equation is relatively simple to create and run with simple statistics software. That said, the statistical concepts underlying the argument are important and understanding the fundamental principles (a regression line describing hours of sleep and level of grumpiness, for example) may mean not giving up when understanding does not come

112. Though not directly relevant to his Wisconsin findings, Professor Jackman further observes:

> It is also interesting to note that the estimate of the 75th percentile of the distribution of *EG* magnitudes jumps markedly after 2010, suggesting that districting plans enacted after the 2010 census are systematically more gerrymandered than in previous decades. Of the almost 800 *EG* estimates in the analysis, spanning 42 years of elections, the largest, negative estimates (an efficiency gap disadvantaging Democrats) are more likely to be recorded in the short series of elections after 2010. These include Alabama in 2014 (–.18), Florida in 2012 (–.16), Virginia in 2013 (–.16), North Carolina in 2012 (–.15) and Michigan in 2012 (–.14); these five elections are among the 10 least favorable to Democrats we observe in the entire set of elections. Among the 10 most pro-Democratic *EG* scores, *none* were recorded after 2000. The most favorable election to Democrats in terms of *EG* since 2010 is the 2014 election in Rhode Island (*EG* = .12), which is only the 20th largest (pro-Democratic) *EG* in the entire analysis.

Expert Report of Simon Jackman, *supra* note 40, at 44.

easily. Most importantly, lawyers should understand enough about math to identify issues and recognize when they do not understand a concept, then work even harder to figure out how it relates to the legal claim.

In most cases that involve significant mathematics, plaintiffs and defendants will hire mathematical experts, as each side did in *Whitford*. An expert's role can include both teaching the lawyers about the concepts and also explaining them to the judge or jury. But experts are not lawyers; lawyers understand the legal claims and the standards of proof for those claims. Experts can help, but they cannot translate the entire analysis into those claims; that is the lawyer's job.[113]

So, research the math. Ask your experts for their opinions and ask them to teach you. Get familiar enough with the concepts to talk about them with others and test how well you truly know the material. These might seem like basic tips and tricks, but the goal of your advocacy is to take the math and use it to your advantage. To do that, you must understand the concepts and understand how to talk about them.

B. Frame Math Within the Legal Issue

The next step is to evaluate how the math factors into the legal analysis. The difficulties of the *Whitford* plaintiffs provide an illustration. As explained in Justice Kagan's concurring opinion, certain plaintiffs (though not Professor Whitford himself, based on the heavily Democratic Party leanings in the city of Madison) could have shown individualized harm because their votes were diluted or because their First Amendment right of association was infringed, but they did not present sufficient facts to do so.[114] As Justice Kagan suggested, "[a]mong other ways of proving packing or cracking, a plaintiff could produce an alternative map (or set of alternative maps)—comparably

113. Fisher, *supra* note 3, 735–36 ("Multiple regression analysis can play a vital role in legal proceedings. Used properly, it is an accurate and reliable method of determining the relationships between two or more variables, and it can be a valuable tool for resolving factual disputes. In order for this to happen, however, multiple regression must be better understood by the legal community; in particular, there must be an understanding of both the potential and the limits of the technique. It is not necessary that lawyers understand the mechanics of multiple regression in terms of what goes on inside the computer. It is necessary, however, that they understand the regression model and the assumptions being used in any given regression study, how the results of the regression bear on the hypothesis to be tested, and how the results distinguish this particular hypothesis from other hypotheses. The expert constructing the analysis should be able to explain all of this to the attorney who employs him, and an expert who cannot explain such things is likely to fall apart on cross-examination.").

114. *Whitford IV*, 138 S. Ct. 1916, 1936, 1939 (2018) (Kagan, J., concurring); *see supra* notes 55–58 and accompanying text.

consistent with traditional districting principles—under which her vote would carry more weight."[115] With respect to the claim for associational harm, Justice Kagan found that "the plaintiffs' evidence of partisan asymmetry well fits" such an allegation.[116] Though argumentation hindsight is twenty/twenty, Justice Kagan's points highlight how crucial it is for lawyers to think critically about the mathematical concepts and precisely how they support the legal theory being advanced.

In *Whitford*, the plaintiffs built their case around mathematical analyses yet did not use the math to meet necessary legal requirements for standing (or for associational harm). A legal advocate must understand what the math means to effectively use the math to make a legal point. Evaluation of cracking and packing as mathematical functions explained political theory in a way that certainly resonated with Professors Mayer and Jackman and their colleagues, but the math (apparently) did not sufficiently convey to the Court how the plaintiffs were affected, nor were proper factual predicates laid to support those mathematical conclusions.

C. Connect the Dots: Define and Dissect

The key to using math in legal advocacy is to connect the dots for your audience. You cannot assume that the court or the public understands what you are saying to them in the same way you understand it, and thus you must become "fluent" in math.

First, define the terms. Most mathematics involve terms of art, just like the law. Regression means drawing a line that best fits the data, and multivariable regression means drawing a line that best fits a lot of data with many variables. Variance describes, collectively, how far data points stray from the average of that line. It's far more complicated than that, of course, but starting at first principles allows the lawyer to integrate the statistical discussion into the legal arguments. Second, dissect the expert analysis. Take on the big pieces first—here, understanding regression and multivariable regression—then dissect down. Understand how each point wound up where it did on the graph and what assumptions were behind the analysis that put them there. Sometimes baseless assumptions sneak into expert analysis: correct yours and exploit those of your opponent.[117] Once fully dissected, the expert analysis can be put back together to support the legal discussion. The connective tissue between math and law is where cases are won or lost.

---

115. *Id.* at 1936 (Kagan, J., concurring).
116. *Id*. at 1937.
117. *See supra* note 85 (describing legal penalties for failure to support model assumptions with facts introduced in court).

V. CONCLUSION

This Article has sought to explain one important case involving statistical analysis to show how the claim of "sociological gobbledygook" was unfounded and to provide a roadmap for future lawyers presenting similar claims. The math mattered in *Whitford* and will matter again in future gerrymandering cases.