

**Trial by Internet:
A Randomized Field Experiment on Wikipedia's
Influence on Judges' Legal Reasoning¹**

Preprint

Forthcoming in Kevin Tobia (ed.) *The Cambridge Handbook of Experimental Jurisprudence*. New York: Cambridge University Press.

Neil C. Thompson
MIT

Brian Flanagan
Maynooth University

Edana Richardson
Maynooth University

Brian McKenzie
Maynooth University

Xueyun Luo
Cornell University

¹ The authors would like to thank Jana Khalil for her research assistance and MIT and Maynooth University for research funding. The authors would also like to thank Michael Doherty and Maria Murphy for their valuable contributions, Richard Albert and Felipe Jiménez for their insightful comments, and the students of Maynooth University's School of Law and Criminology who assisted with Wikipedia article drafting. Correspondence about the chapter should be sent to neil_t@mit.edu.

TABLE OF CONTENTS

Introduction

- I. A New Empirical Method for the Field**
 - a. The Field Experiment**
 - b. The Empirical Turn**
 - c. Limitations of Existing Behavioral Methods**
- II. Legal Reasoning: Theory and Evidence**
 - a. Theoretical Foundation**
 - b. Evidential Foundation**
 - c. Normative Complications**
- III. It's a Wiki World**
 - a. Knowledge on Tap**
 - b. Expertise Optional**
- IV. Theory to Data**
 - a. Hypothesis**
 - b. Causal Mechanism**
- V. Experimental Design**
 - a. Choosing the Sample**
 - b. Building the Control and Treatment Groups**
 - c. Writing the Wikipedia Articles**
 - d. Integrating the New Articles**
- VI. Results**
 - a. Citation Behavior**
 - b. Argument Structure**
 - c. Discussion**
- VII. Conclusion**

**Trial by Internet:
A Randomized Field Experiment on Wikipedia's Influence on Judges' Legal Reasoning**

In the common law tradition, legal decisions are supposed to be grounded in both statute and precedent, with legal training guiding practitioners on the most important and relevant touchstones. But actors in the legal system are also human, with the failings and foibles seen throughout society. This may lead them to take methodological shortcuts, even to relying on unknown internet users for determinations of a legal source's relevance.

In this chapter, we investigate the influence on legal judgments of a pervasive, but unauthoritative source of legal knowledge: Wikipedia. Using the first randomized field experiment ever undertaken in this area—the gold standard for identifying causal effects—we show that Wikipedia shapes judicial behavior. Wikipedia articles on decided cases, written by law students, guide both the decisions that judges cite as precedents and the textual content of their written opinions. The information and legal analysis offered on Wikipedia led judges to cite the relevant legal cases more often and to talk about them in ways comparable to how the Wikipedia authors had framed them.

Collectively, our study provides clear empirical evidence of a new form of influence on judges' application of the law—easily accessible, user-generated online content. Because such content is not authoritative, our analysis reveals a policy-gap: if easily-accessible analysis of legal questions is already being relied on, it behooves the legal community to accelerate efforts to ensure that such analysis is both comprehensive and expert.

INTRODUCTION

To avoid an arbitrary discretion in the courts, it is indispensable that they should be bound down by strict rules and precedents...and it will readily be conceived from the variety of controversies which grow out of the folly and wickedness of mankind, that the records of those precedents must unavoidably swell to a very considerable bulk, and must demand long and laborious study to acquire a competent knowledge of them. Hence it is that there can be but few men in the society, who will have sufficient skill in the laws to qualify them for the stations of judges.

- Alexander Hamilton²

The adjudicatory challenge identified by Hamilton is a real one: “[j]udges everywhere face crowded dockets and enormous time pressures...the busy judge cannot always have the luxury of constantly revisiting their approach to each and every case.”³ In the face of these pressures, judges may turn to shortcuts and heuristics to stay abreast of the ever-growing body of case law. This chapter shows that Wikipedia is one of the shortcuts being used.

² THE FEDERALIST No. 78, 418 (A. Hamilton) (J.R. Pole ed. 2005).

³ Jeffrey Rachlinski and Andrew Wistrich, *Judging the Judiciary by the Numbers: Empirical Research on Judges*, 13 ANNUAL REV. OF L. & SOCIAL SCIENCE 203, 223 (2017). See also, Holger Spamann and Lars Klohn, *Justice is less Blind, and less Legalistic, than We Thought: Evidence from an Experiment with Real Judges*, 45 J. LEGAL Stud. 255, 274 (2016) noting “the severe constraints facing many congested courts for many decisions.”

An explanation of why judges dispose of litigation as they do is critical to our understanding of society. A scholarly project has blossomed in the last hundred years that looks beyond the pages of the law reports to draw evidence of judicial motivations from a wide range of external sources, such as intellectual history, laboratory experiments, and observational studies. While this project has supplied a much richer understanding, the strongest form of evidence for what causes judicial decisions—that from randomized field experiments—has been absent. In this research, we correct that absence. The result is the opening of a new, more empirically rigorous window of knowledge into the practice of adjudication.

At the heart of our analysis is the question of how adjudication should happen. When a litigant asks a court to vindicate their legal rights, they are entitled to assume that the judge, relying on her professional training and experience,⁴ will resolve their claim by reaching an expert determination of whose side the law is on. As a purely descriptive matter, most scholars agree that, for most prospective litigation, this is not an unreasonable assumption to make. The uncertainty that most scholars have about the nature of judicial decision-making instead concerns the relatively rare cases that find their way to the highest appellate courts, where the guidance provided by legal sources may be opaque.⁵

The consensus about the adjudication of typical cases rests on a silent assumption: that judges, when establishing what the law says, forgo the convenience of readily accessible but potentially unreliable sources of information. If, conversely, unreliable sources are used, then, even in routine cases, the law might not be determinative. Such a scenario would, in turn, raise worries that the practice of adjudication might compromise both the predictability of litigation and the ideal of the rule of law alike. Here we test for the first time whether routine questions of legal rights can, if litigated, be counted on to receive an expert answer. We find that, in some cases, the contemporary substitute for Hamilton's "long and laborious study" is to consult Wikipedia.

In the early years of the internet, the technology's evident potential to educate and inform led to its characterization as an "information superhighway". Today, Wikipedia, the collectively-written, online encyclopedia, with its 6,375,621 articles and counting in English, is the apotheosis of this vision.⁶ Despite Wikipedia's many strengths as a means of making knowledge available to the world, its fundamental feature of collective self-creation can also make it unreliable: specialized or obscure topics often reflect the perspective of one or two contributors. For citizens and legal professionals alike, the use of Wikipedia as a source of guidance as to what the law says therefore presents a challenge. The first step in establishing the extent of this challenge is to discover whether judges do in fact rely on Wikipedia. We do so by investigating the practices of judges in Ireland—and generalizing the trends that we find there.

As a former British colony, Ireland traces the origins of its modern legal system to the English common law tradition, through which many of its legal rules have been articulated, shaped, and developed by judges.⁷ As a common law jurisdiction, the Irish legal system

⁴ In this article, we refer to a female hypothetical judge simply for consistency and readability rather than as indicative of those affected by the experiment.

⁵ See Section II below.

⁶ Wikipedia Main Page, https://en.wikipedia.org/wiki/Main_Page, accessed September 14, 2021.

⁷ *McCaffrey v. Central Bank of Ireland* [2017] IEHC 546 (Ir.) at [95] per Noonan J, "[u]nlike the civil law systems that [exist in] most European jurisdictions, [Ireland's] is a common law system shared with countries

shares a key similarity with other national legal systems such as the UK and the US: it operates within a hierarchical court structure in which decisions of higher courts subsequently bind lower courts, and in which judges cite earlier cases as determinative of the applicable legal principle(s). Just as this feature has long grounded comparative analyses of the legal doctrines of Irish and other common law systems, it also anchors comparative empirical scrutiny of judicial behavior, such that a pattern in how Irish judges perform core functions may be expected to emerge in the performance of such functions in other common law jurisdictions. Against this backdrop of “likeness”, the Irish legal system and, within it, Irish case law, offers an important advantage—unlike, say, the U.S. Supreme Court, about whose decisions a multitude of Wikipedia articles already exist—the decisions of the Irish Supreme Court (or of any other Irish court) have attracted hardly any such coverage. This feature allows us to manipulate the inclusion on Wikipedia of the case law to be applied by one common law judiciary to determine the website’s possible influence on judicial legal reasoning more generally.

To investigate Wikipedia’s influence on legal reasoning, we begin by describing a style of judging that might be receptive to its use, namely, moot court adjudication. We then develop the hypothesis that, in its reliance on Wikipedia, real-world judging exhibits the moot court style. We then distinguish the sorts of evidence that would indicate alternative mechanisms for Wikipedia’s influence on judges’ citations of previous cases and on the textual content of their judgments, i.e., whether it operates indirectly through the filings of the parties or directly through the research of the judge herself (or her clerks). Designed to uncover such evidence, our experiment reveals that judges’ application of the law is now influenced by the same internet forces that shape other professional domains.

Approved by the respective ethics boards of MIT and Maynooth University, the experiment amounted to a friendly stress-test of the potential vulnerability of judicial legal reasoning to the limitations of reliance on Wikipedia, notably, its ad hoc topic coverage and unknown author/editorship. The experiment featured Wikipedia entries authored by faculty and by law students under faculty supervision, who each had access, through their university library, to all the relevant primary and secondary legal materials available to judges and their clerks. This assurance of accuracy and of informed analysis in the content of the entries—though short of that offered by a specialist textbook—indicates that judges or lawyers would be unlikely to be misled by what they might read. However, as the authorship of Wikipedia articles is opaque, this fact would not be known to any legal professional when using them. From the users’ perspective, there was no particular reason to imagine that the creators of the relevant entry had any legal expertise—or even that they lacked an ulterior agenda.

This chapter is structured as follows: Section I establishes the importance and novelty of using a randomized field experiment to investigate the causes of judicial behavior; Section II outlines the contemporary understanding of judicial legal reasoning; Section III situates Wikipedia within the ecology of information sources; Section IV describes the hypothesis; and, finally, Sections V and VI detail the experiment’s design and results, respectively.

I. A NEW EMPIRICAL METHOD

such as the United Kingdom, the United States, Canada and Australia. All have their roots in the common law of England.”

To our knowledge, this research reports the first randomized field experiment that investigates the influence of any non-procedural factor – legal, personal, or ideological – on judicial behavior. It thereby establishes a new frontier in the empirical study of law.

a. The Field Experiment

Randomized field experiments are recognized as the gold standard for evidence of causal effects: “no field can claim to be evidence-based without a central role for the RCT [randomized control trial] as a means of accumulating knowledge”.⁸ Such studies achieve two key social scientific desiderata. First, they are *ecologically valid*, meaning that they are conducted in-situ using real-world decision-making—in our case, judges are conducting their usual adjudicatory role unaware of the presence of an experiment. This improves the likelihood that the experiment’s findings will generalize outside the experiment and contrasts with, for example, laboratory studies where other effects can bias outcomes artificially. Such effects can include the desire to please or look good to the experimenters (“social desirability bias”)⁹ and “Hawthorne” effects where the knowledge of monitoring changes behavior.¹⁰ Hence, the ecological validity of field experiments makes them more informative of real-world behavior and, by extension, of the impact of potential policy interventions.

The second, more important advantage of randomized field experiments is that they can distinguish causation from correlation. The ability to prove causal relationships derives from the combination of two characteristics. The first is having a *control group*, that is, a group unaffected by the intervention (in our case, publication of a Wikipedia article on the topic) that can be used as a counterfactual to estimate the size of causal effects. The second is *randomization*, that is, random assignment into the control and intervention groups. With sufficient data and a sound experimental design, the experiment can reduce the probability of being misled by correlation or noise to whatever arbitrarily small value is desired.¹¹

Field experimentation in law has to date been limited in both depth and scope. Research into effective procedural interventions on legal outcomes has benefited from the method, establishing the causal effect of legal representation,¹² of the provision of information on bail applicants’ likelihood of absconding,¹³ and of the volume of judicial hearings.¹⁴ Inquiry into the nature of judicial motivation has not, however, generated evidence of a comparable quality. Given the established use of randomized control trials in other social scientific

⁸ James Greiner and Andrea Matthews, *Randomized Control Trials in the United States Legal Profession*, 12 ANNUAL REV. OF L. & SOCIAL SCIENCE 295, 296 (2016); Donald Green and Dane Thorley, *Field Experimentation and the Study of L. and Policy*, 10 ANNUAL REV. OF L. & SOCIAL SCIENCE 53, 59 (2014); Donald Rubin, *The Design Versus the Analysis of Observational Studies for Causal Effects: Parallels with the Design of Randomized Trials*, 26(1) STATISTICS IN MEDICINE 20, 20 (2007).

⁹ Ivar Krumpal, *Determinants of social desirability bias in sensitive surveys: a literature review*, 47 QUAL QUANT 2025–2047 (2013).

¹⁰ C.W.M. Hart, *The Hawthorne experiments*, 9(2) THE CANADIAN JOURNAL OF ECONOMICS AND POLITICAL SCIENCE 150-163 (1943).

¹¹ John A. List, *Why economists should conduct field experiments and 14 tips for pulling one off*, 25(3) JOURNAL OF ECONOMIC PERSPECTIVES 3-16 (2011).

¹² D.J. Greiner et. al., *The limits of unbundled legal assistance: a randomized study in a Massachusetts district court and prospects for the future*, 126(4) HARV. L. REV. 901 (2013).

¹³ J.S. Goldkamp and M.R. Godfredson, *Judicial Guidelines for Bail: The Philadelphia Experiment*. Washington, DC: US Dep. Justice, Natl. Inst. Justice (1984).

¹⁴ D.B. Marlowe et. al., *Are judicial status hearings a key component of drug court? During-treatment data from a randomized trial*, 30(2) CRIM. JUSTICE BEHAV. 141 (2003).

domains, including the cognate field of criminology,¹⁵ the absence of an equivalent knowledge base on the critical question of why judges decide as they do is striking.

b. The Empirical Turn

Following critiques of the traditional focus on legal doctrine by “legal realists” such as Karl Llewellyn,¹⁶ scholars have pursued a systematic empirical alternative to the interpretation of official opinions, unofficial writings and biographies.¹⁷ As part of this decades-long effort to quantifiably identify judicial motivation, some have tried to uncover judicial approaches to legal reasoning by asking judges directly in surveys.¹⁸ However, as introspection is thought to provide unreliable access to cognitive processes,¹⁹ the ability of judicial survey participants to accurately report their decision-making is open to question:

[A]sking someone to identify his or her motive is one of the worst methods of measuring motive. People often do not know, or cannot articulate, why they act as they do... This is obvious from the example of asking justices about how they reach decisions....²⁰

Alternatively, surveys have investigated judicial motivation, not by asking judges explicitly, but by considering their responses to hypothetical legal cases directly.²¹ Some of these studies have incorporated randomization and controls to establish how, for example, judges' stereotypes and prejudices determine their verdicts on the presented legal vignettes, thereby allaying concerns about the self-reporting of motivation. But a related difficulty emerges. To infer that such factors actually influence judicial behavior, we must assume that judges' behavior in contrived, artificial contexts will mimic their actual, real-world exercise of judicial office. There will be no guarantee of correlation with behavior in naturalistic settings, and little evidence with which to assess its likelihood.

These problems facing the use of survey methods both stem from a concern about the ecological validity of survey data: that study participants' responses, being removed in

¹⁵ D.P. Farrington et. al., *Randomized experiments in criminology: What have we learned in the last two decades?*, 1 J. EXP. CRIMINAL 9 (2005).

¹⁶ K.N. Llewellyn, *THE BRAMBLE BUSH: ON OUR LAW AND ITS STUDY* (1930, Columbia University, New York).

¹⁷ Jeffrey Rachlinski and Andrew Wistrich, *Judging the judiciary by the numbers: empirical research on judges*, 13 ANNU. REV. L. SOC. SCI. 203 (2017).

¹⁸ Howard Becker, *Approach to the Study of Social Change*, 2(3) THE SOCIOLOGICAL QUARTERLY 155 (1961); D.E. KLEIN, *MAKING LAW IN THE UNITED STATES COURTS OF APPEALS* (Cambridge Univ. Press 2002); Brian Flanagan and Sinéad Ahern, *Judicial Decision-Making and Transnational Law: A Survey of Common Law Supreme Court Judges*, 60(11) THE INT'L & COMPARATIVE L. QUARTERLY (2011).

¹⁹ R.E. Nisbett and T.D. Wilson, *The halo effect: Evidence for unconscious alteration of judgments*, 35(4) J. OF PERSONALITY & SOCIAL PSYCHOLOGY 250 (1977); W. Rahn et. al., *Rationalization and Derivation Processes in Survey Studies of Political Candidate Evaluation*, 38 AM. J. POLIT. SCI. 582 (1994).

²⁰ Lee Epstein and Gary King, *The Rules of Inference*, 69(1) UNIVERSITY OF CHICAGO L. REV. 93 (2002); JEFFREY SEGAL AND HAROLD SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL* (Cambridge Univ. Press 1994) 16.

²¹ C. Guthrie et. al., *Blinking on the bench: how judges decide cases*, 93 CORNELL L. REV. (2007); J.J. Rachlinski et. al., *Does unconscious bias affect trial judges?*, 84 NOTRE DAME L. REV. 1195 (2009); Dan Kahan et. al., *'Ideology' or 'Situation Sense'? An Experimental Investigation of Motivated Reasoning and Professional Judgment*, 116 U. PA. L. REV. 349 (2016); M. Kneer & S. Bourgeois-Gironde, *Mens rea ascription, expertise and outcome effects: Professional judges surveyed*, 169 COGNITION 139 (2017).

different ways from “the complexity and dynamics of stimuli and behaviors in real-life”,²² will fail to reflect what happens in complex, natural settings. By contrast, in studies that measure influences on judicial behavior directly—in the field—what you see is what you get: we know to expect the behavior to occur in real situations because those are precisely the situations in which it has been shown to occur.

Starting with C. Herman Pritchett's study of the relationship between the voting record and partisan affiliations of the members of the “Roosevelt Court”,²³ behavioral research has long generated insight into the connection between the work of the judiciary and underlying social structures, e.g., judicial ideology and case votes,²⁴ and race and sentencing severity.²⁵ Moreover, unlike the survey method, in establishing correlations between characteristics and actual judicial behaviors, this research leaves no room for doubt that, given the relevant characteristic, the behavior in question will be more likely to materialize.

c. Limitations of Existing Behavioral Methods

There is a key limitation to nearly all observational behavioral inquiry. Without the random introduction of the posited influence into a subset of otherwise equivalent opportunities for judicial behavior (i.e., actions in cases) there is no way to tell whether the correlation between characteristic and action arises because the former causes the latter or because both are caused by the presence of some unconsidered characteristic.²⁶ Because it provides no assurance that “control and treatment groups will be similar even with regard to attributes that are unobservable to the researcher”,²⁷ the bulk of behavioral inquiry to date cannot strictly adduce evidence for any causal conclusion, e.g., that “Rehnquist votes the way he does *because* he is extremely conservative... [and] Marshall voted the way he did *because* he is extremely liberal”.²⁸

²² Saurabh Sonkusare et. al., *Naturalistic Stimuli in Neuroscience: Critically Acclaimed*, 23(8) TRENDS IN COGNITIVE SCIENCES 699, 699 (2019).

²³ CHARLES HERMAN PRITCHETT, *THE ROOSEVELT COURT* (New York: The Macmillan Company 1948).

²⁴ JEFFREY SEGAL AND HAROLD SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED* (Cambridge Univ. Press 2002).

²⁵ C.A. Albonetti, *An integration of theories to explain judicial discretion*, 38(2) SOCIAL PROBLEMS 247 (1991).

²⁶ “[I]f the[] proxies for [ideological] judicial decision making are correlated with unobserved factors, these studies may suggest the presence of ideological decision making where none exists” Matthew Hall, *Randomness Reconsidered: Modeling Random Judicial Assignment in the U.S. Courts of Appeals* JOURNAL OF EMPIRICAL LEGAL STUDIES 7(3):574–589, 574 (2011).

²⁷ Michael Abramowicz et. al., *Randomizing Law*, 159 U. PA. L. REV. 929, 936 (2011).

²⁸ JEFFREY SEGAL AND HAROLD SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL* 32-33 (New York: Cambridge University Press, 1993) (emphasis added). Exploiting the ostensible randomization of judges' case assignments, a few studies have reported *natural* experiments on the influence on a judge's behavior of her partisan affiliation (e.g., Orley Ashenfelter et. al., *Politics and the judiciary: the influence of judicial background on case outcomes*, 24(2) J. LEGAL STUD. 257–281 (1995); CASS SUNSTEIN ET. AL., *ARE JUDGES POLITICAL? AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY* (Brookings Institute Press 2006); Chelsea Liu, *Judge Political Affiliation and Impacts of Corporate Environmental Litigation*, 64(4) J. OF CORPORATE FINANCE 101670 (2020)) and of a defendant's race (e.g., Crystal Cohen and Alma Yang, *Judicial Politics and Sentencing Decisions*, 11(1) AMERICAN ECONOMIC J.: ECONOMIC POLICY 160 (2019); Oren Gazal-Ayal and Raanan Sulitzeanu-Kenan, *Let My People Go: Ethnic In-Group Bias in Judicial Decisions—Evidence from a Randomized Natural Experiment*, 7(3) J. OF EMPIRICAL LEGAL STUDIES 403 (2010); D.S. Abrams et. al., *Do judges vary in their treatment of race?*, 41(2) J. LEGAL STUD. 347–83 (2012)). There are indications, however, of implicit practices of non-random case assignment (e.g., Clark Hildabrand, 2019. *The Curiously Nonrandom Assignment of Sixth Circuit Senior Judges*, KENTUCKY L. J. 108; Adam Chilton and Marin Levy (2015) *Challenging the*

The obvious impediment to conducting randomized field experiments on judges is the difficulty for the researcher in randomly stipulating that certain acts of adjudication should be performed under a particular condition. Certainly, the researcher cannot randomly assign judges to cases; only a court's chief judge has assignment authority, and they could scarcely use it to experiment with citizens' efforts to vindicate their legal rights:

[O]f course... executing an experiment of this sort is nearly as impossible as rerunning history. As a result, judicial specialists, again as included, must work with observational data, which substantially complicate the inferential task.²⁹

Equally, it would be impossible to select a specific subset only of a set of equivalent legal cases to be resolved according to a particular legal authority, i.e., regulation or precedent.³⁰ The key methodological innovation of this research, therefore, is to simplify the “inferential task” by randomly stipulating certain court decisions to be made under the condition of interest, namely, the availability of a particular source of inexpert information about a possible legal authority, viz., a precedent. In this way, we exclude any effect of both observed and *unobserved* attributes on our estimate of the judicial application of that authority in an expert or inexpert fashion.

By designing a research project that features a randomized control trial, we avoid the concern that an overlooked variable is what is truly making the difference. The result is a study of judicial motivation that uniquely satisfies the social scientific maxim: “no causation without manipulation”.³¹ By also matching the ecological validity of existing behavioral studies, we establish a rigorous basis for interpreting judicial behavior as having been *caused* by a particular factor.

II. LEGAL REASONING: THEORY AND EVIDENCE

It has been observed that, “in most contemporary legal systems, there is a requirement—formal or informal—for courts, administrative agencies, and other public institutions to provide reasons for their decisions.”³² The theory of the nature of legal

Randomness of Panel Assignment in the Federal Courts of Appeals. Duke Law School Public Law & Legal Theory Series No. 2015-1. Univ. of Chicago, Public Law Working Paper No. 529). It has been observed, accordingly, that ‘randomization of cases in [US] federal courts is very far from perfect, and scholars should be careful in claiming randomization for purposes of causal identification’ (Adam Bonica and Maya Sen, *Estimating Judicial Ideology*, 35(1) J. OF ECONOMIC PERSPECTIVES 97, 107 (2021); similarly, C.L. Boyd et al, *Untangling the causal effects of sex on judging*, 54(2) AM. J. POLITICAL SCI. 389, 394 (2010)). In any event, a potentially significant, inherent limitation of natural experiments is that, unlike a field experiment, which is designed to specification, controls cannot be introduced to ensure that the subjects randomly assigned to both treatment and control populations are otherwise as similar as possible. This study’s employment of stratification (described below) provides an additional assurance of comparability.

²⁹ C.L. Boyd et. al., *Untangling the causal effects of sex on judging*, 54(2) AM. J. POLITICAL SCI. 389, 395 (2010).

³⁰ E Braman, *Reasoning on the Threshold: Testing the Separability of Preferences in Legal Decision Making*, 68(2) J. OF POLITICS 308, 311 (2006) “Clearly this would be impossible in the real world.”

³¹ Donald Rubin, *Bayesian inference for causality: The role of randomization*, PROCEEDINGS OF THE SOCIAL STATISTICS SECTION OF THE AMERICAN STATISTICAL ASSOCIATION 233, 238 (1975).

³² J. Cohen, *Procedure and Substance in Deliberative Democracy*, in Bohman, J. & Rehg, W. (eds.) DELIBERATIVE DEMOCRACY: ESSAYS ON REASON AND POLITICS, 407, 412–13 (1997). In *Delaney v Judge Donnchadh O Buachalla and anor* [2011] IEHC 138 (Ir.) [33]-[34], McMahon J. noted that ‘[c]onfidence in the judicial process is based on the assumption that decisions are based on rational foundations and are not

reasoning is fraught with controversy, while the voluminous empirical evidence on the practice of adjudication lacks a unifying theme. Nevertheless, a core of agreement on how the law both is and ought to be applied remains, namely, that judges resolve most legal issues by expertly interpreting the relevant legal authorities - often with the help of their clerks.

a. Theoretical Foundation

Closely associated with the Western, common law tradition of legal adjudication, a precedent is an independent reason for reaching the same decision in analogous cases:

When we make a decision on the basis of precedent, we consider significant the fact that our current predicament has been addressed before, but we will not necessarily value a precedent for what it teaches us.³³

A precedent “represent[s] a decision on the balance of reasons in the individual case before the court that later courts are required to treat as correctly decided”.³⁴ Accordingly, whether a prior case serves as a precedent for a particular matter depends on the original court’s reasons for its decision, as extracted from judges’ written opinions. To establish the relevance of previous decisions, a judge might study those decisions herself, or identify an expert secondary source of information from which to draw, e.g., an academic treatise or the report of a judicial clerk. Alternatively, she might google it.

Suppose A sues B in federal district court. A argues that B is liable for breach of contract; B acknowledges A’s account of the facts but maintains that they gave rise to no contract between them. The assigned judge, conscious of the heavy work that she has already delegated to her clerks, decides to conduct her own research. On reviewing the parties’ submissions, the judge forms the preliminary view that a contract has not truly been formed and that she should give judgment for the defendant. For the purpose of writing her official opinion, the judge googles some previous decisions cited in B’s brief that seem similar to the instant case. On confirming their similarity by reading the relevant case summaries on Wikipedia, the judge paraphrases some of the text of the Wikipedia entries in her draft opinion to complete her analysis. The judge then enters her judgment and publishes her opinion. What has just happened?

The issue of the nature of law and legal reasoning provokes a series of overlapping debates about whether morality is intrinsic to legality; whether judicial disagreement in hard cases is genuine; and what it is that particular legal sources (such as legislation or precedent) truly consist in. In pursuing these questions, scholars continue to deploy both the traditional,

arbitrarily arrived at. Moreover, public confidence is best secured when the reasons for the decision are explained and furnished’.

³³ NEIL DUXBURY, *THE NATURE AND AUTHORITY OF PRECEDENT* 2 (Cambridge Univ. Press 2008); Frederick Schauer, *Precedent*, 39 STAN. L. REV. 571, 575-576 (1987).

³⁴ Grant Lamond, *Precedent and Analogy in Legal Reasoning*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Summer 2020 Edition), <https://plato.stanford.edu/archives/sum2020/entries/legal-reas-prec/>, accessed September 14, 2021; H.L.A. HART, *THE CONCEPT OF LAW* 135-135 (Oxford Uni. Press 1961); RONALD DWORKIN, *LAW’S EMPIRE* 24-25 (Harvard Univ. Press 1986).

analytic method,³⁵ and, increasingly, the tools of cognitive science.³⁶ Although views are not uniform,³⁷ the organizing assumption of much of this research is that law is a distinct realm of social order, the application of which involves distinctive sorts of practical reasoning. On this “legal” model of adjudication, judges’ application of the law suffices to determine the outcome of most litigation, such that “judges... seek[ing] to capture and be faithful to the content of the law... always seem to be able to decide cases by interpreting the law”.³⁸ In this context, the doctrine of precedent (or *stare decisis*) is seen within the common law tradition as a source of consistency and predictability.³⁹

Notice that conformity with the legal model means determining the applicable law *tout court*—not merely deciding which party has hired the better lawyer. Crucially, a judge cannot claim to capture the content of the law if she defers to analysis of unknown origin, that is, to analysis of which she can determine neither author nor editor.

b. Evidential Foundation

Both doctrinal and empirical investigations of law’s influence on judicial decision-making have so far tended to support the descriptive accuracy of the legal model. While there is ample evidence that, at the top of a judicial hierarchy, judges often vote as a legislator might,⁴⁰ the interpretation of this finding is complicated by the overlap in the behavior of a mere “politician in robes” and that of a judge who adheres to a moralistic conception of legal judgment. On a moralistic account of the nature of law,⁴¹ a judge’s ideological concerns are in fact part and parcel of her legal reasoning:

³⁵ Hasan Dindjer, *The New Legal Anti-Positivism* LEGAL THEORY 26(3) 181 - 213 (2020); SCOTT SHAPIRO, LEGALITY (Belknap Press 2011); David Plunkett and Timothy Sundell, *Disagreement and the Semantics of Normative and Evaluative Terms*, 13(23) PHILOSOPHERS’ IMPRINT 1 (2013).

³⁶ J. Turri and P. Blouw, *Excuse validation: a study in rule-breaking* 172(3) PHILOSOPHICAL STUDIES, 615–634 (2015); Kevin Tobia, *Testing Ordinary Meaning*, 134 HARV. L. REV. 726 (2020); Brian Flanagan and Ivar Hannikainen, *The Folk Concept of Law: Law is Intrinsically Moral*, AUSTRALASIAN J. OF PHILOSOPHY 1-15 (2020).

³⁷ S. FISH, DOING WHAT COMES NATURALLY: CHANGE, RHETORIC, AND THE PRACTICE OF THE THEORY IN LITERARY AND LEGAL STUDIES (Duke Univ. Press 1989); M. Tushnet, *Defending the Indeterminacy Thesis*, in B. Bix (ed.), ANALYZING LAW: NEW ESSAYS IN LEGAL THEORY (Oxford University Press 1998) 223–238; D. Patterson, *Methodology and Theoretical Disagreement*, in U. Neergaard, R. Nielsen, & L. Roseberry (eds.), EUROPEAN LEGAL METHOD: PARADOXES AND REVITALISATION. (DJØF Publishing 2011) 227–241.

³⁸ Julie Dickson, *Interpretation and Coherence in Legal Reasoning*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Summer 2020 Edition), <https://plato.stanford.edu/archives/sum2020/entries/legal-reas-interpret/>, accessed September 14, 2021.

³⁹ See *McCaffrey v. Central Bank of Ireland* [2017] IEHC 546 (Ir.) at [95] per Noonan J. ‘Common law is sometimes defined as judge-made law or the law of judicial precedent. Its origins are ancient. *Stare decisis* is at its core. Students of law and lawyers alike study decided cases to learn the law. The common law evolves to mirror societal changes but it does so slowly. Lawyers speak in terms of the law being settled by virtue of long standing and long followed authorities.’

⁴⁰ GUNNAR GRENSTAD ET. AL., POLICY MAKING IN AN INDEPENDENT JUDICIARY: THE NORWEGIAN SUPREME COURT (ECPR Press 2015); Christoph Hönnige, *The Electoral Connection: How the Pivotal Judge Affects Oppositional Success at European Constitutional Courts*, 32(5) WEST EUROPEAN POLITICS 963 (2009); J.A. Segal, *Judicial Behaviour* in THE OXFORD HANDBOOK OF LAW AND POLITICS, KE Whittington, RD Kelemen, GA Caldeira (eds) (Oxford Univ. Press 2008) 19–35.

⁴¹ LON FULLER, MORALITY OF LAW (Yale University Press 1969); RONALD DWORKIN, LAW’S EMPIRE (Harvard Univ. Press 1986); JOHN FINNIS, NATURAL LAW AND NATURAL RIGHTS (Clarendon Press 1980).

It is no surprise, or occasion for ridicule or suspicion, that a constitutional theory [and consequent judicial votes] reflects a moral stance. It would be an occasion for surprise – and ridicule – if it did not.⁴²

Equally, evidence of the operation of a variety of personal factors sits alongside positive indications that legal sources are indeed influential,⁴³ especially in lower court decision-making. Some personal motivations appear to be consistent with respect for the role of judicial expertise in general, e.g., refraining from issuing a dissent to avoid burdening colleagues with writing lengthier opinions⁴⁴ or writing higher quality opinions to increase one's chances of promotion.⁴⁵ Other motivations allow at least for the application of law to facts itself to be guided by expertise, e.g., the impact of the desire for leisure time on trial management,⁴⁶ of the possible effect of the proximity of mealtimes on parole decisions,⁴⁷ or of the influence of the prospect of promotion on sentencing severity.⁴⁸

Significantly, in courts of first instance (where the vast majority of cases are resolved) no evidence has so far cast serious doubt on judges' general adherence to the legal model of adjudication.⁴⁹ It is widely accepted, thus, that, in disposing of most litigation, judges are led by their "simple desire to 'follow the law'".⁵⁰ Now recall the story of the Wikipedia-using judge told at the outset of this section. The task that the judge purported to perform at the litigants' behest is the core adjudicative function of applying law to facts—not any of the many ancillary judicial behaviors, such as ordering discovery, admitting evidence, setting bail, or sentencing a convict. On the standard account of adjudication in courts of first instance, the depicted judge's failure to perform this function ought to be a bug or glitch

⁴² RONALD DWORKIN, *FREEDOM'S LAW: THE MORAL READING OF THE AMERICAN CONSTITUTION* (Oxford Univ. Press 2008) 37; similarly, Howard Gillman, *What's Law Got to Do with It? Judicial Behaviorists Test the "Legal Model" of Judicial Decision Making*, 26 L. & SOCIAL INQUIRY 465, 493 (2001); Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies That Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 DUKE L.J. 1895, 1945-46 (2008).

⁴³ Brandon Bartels and Andrew O'Geen, *The Nature of Legal Change on the U.S. Supreme Court: Jurisprudential Regimes Theory and Its Alternatives*, 59(4) AMERICAN J. OF POLITICAL SCIENCE 880 (2015).

⁴⁴ Lee Epstein, William M. Landes and Richard A. Posner RA, *Why (and when) judges dissent*, 3 JOURNAL OF LEGAL ANALYSIS 101-137 (2011).

⁴⁵ Eric Posner, Stephen J. Choi & G. Mitu Gulati, *What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals*, 28 JOURNAL OF LAW, ECONOMICS AND ORGANIZATION 518 (2012).

⁴⁶ LEE EPSTEIN, WILLIAM M. LANDES AND RICHARD A. POSNER, *THE BEHAVIOR OF FEDERAL JUDGES* (Harvard Univ. Press 2013).

⁴⁷ Shai Danziger et. al., *Extraneous Factors in Judicial Decisions*, 108(17) PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA (2011).

⁴⁸ Mark A. Cohen, *The motives of judges: empirical evidence from anti-trust sentencing* 12(1) INTERNATIONAL REVIEW OF LAW AND ECONOMICS 13-30 (1992).

⁴⁹ Ryan Hübert and Ryan Copus, *Political Appointments and Outcomes in Federal District Courts*, J. OF POLITICS (forthcoming) present evidence of ideological influence on the disposition of litigation in the relatively politically charged field of civil rights by US federal district courts of the 9th Circuit. But there is also prominent evidence of law's influence, see e.g., Michael Bailey and Forrest Maltzman, *Does Legal Doctrine Matter? Unpacking Law and Policy Preferences on the U.S. Supreme Court*, 102(3) AMERICAN POLITICAL SCIENCE REVIEW 369-384 (August 2008); Chad Westerland, *Strategic Defiance and Compliance in the U.S. Courts of Appeals*, 54(4) AMERICAN J. OF POLITICAL SCIENCE 891 (2010); D.R. Pinello, *Linking party to judicial ideology in American courts*, 20 JUSTICE SYSTEM JOURNAL 219 (1999).

⁵⁰ Lee Epstein and Jack Knight, *Reconsidering Judicial Preferences*, 16 ANNUAL REV. OF POLITICAL SCIENCE 11, 25 (2013); similarly, Jeffrey Rachlinski and Andrew Wistrich, *Judging the judiciary by the numbers: empirical research on judges*, 13 ANNU. REV. L. SOC. SCI. 203, 206 (2017).

rather than a systematic feature of judicial practice. By showing the opposite to be true, this research unsettles any complacency about law's influence even in routine contexts.

c. Normative Complications

The broad scope of the descriptive investigation into whether court decisions are guided by law is commensurate with the question's normative importance: "[t]he stakes of the debate over legal... constraints are high because they get to foundational questions... about [a court's] core functions".⁵¹ Deviations from the legal model of adjudication are recognized to compromise the normative ideals of the rule of law, notably, the principle that justice is administered in accordance with law. It is "... an important part of the Rule of Law that there be a competent profession available to offer... advice [as to what the law requires] and that the law must be such as to make it possible for professionals at least to get a reliable picture of what the law at any given time requires".⁵² Violation of this principle threatens to introduce a cascade of legal pathologies. Recall again the story of our judge.

The story's normatively notable feature is that the judge, in failing to satisfy an elementary standard of legal professionalism, defies the public's expectation that "judges [should] be deliberative and... decide... on the record".⁵³ The court cannot certify that, in its professional judgment, the outcome rests on all relevant legal considerations, and the parties' entitlement to the disposition of their litigation by means of judicial expertise is undercut. Moreover, the judge's failure to consult appropriate materials might lead her to decide otherwise than if she had taken due care and considered all relevant precedents. The mere possibility of a perverse outcome undermines the predictability of the legal system, and, if it should result, would, through its own precedential force, generate a new source of law of seemingly arbitrary character.

Equally, this potential for a perverse outcome would create unsalutary incentives. A litigant's deliberate authorship or revision of a Wikipedia entry on a particular court decision might present that decision as a useful (or, alternatively, irrelevant) precedent. The judge, on consulting the entry, might then be led to effectively adjudicate a party's proposed reading of a legal authority by reference to that party's very own view. From a cynical litigant's perspective, the anonymity associated with such an effort might make it an attractive method of helping to achieve their preferred legal outcome.⁵⁴

The findings that we report below do not directly challenge the assumption that legal materials are always ascribed binding authority over the disposition of litigation, i.e., that "[judges] believe that their decisions always represent the state of the law at a time just prior

⁵¹ Brandon Bartels and J. O'Geen, *The Nature of Legal Change on the U.S. Supreme Court: Jurisprudential Regimes Theory and Its Alternatives*, 59(4) AMERICAN J. OF POLITICAL SCIENCE 880, 881 (2015).

⁵² Jeremy Waldron, *The Rule of Law*, THE STANFORD ENCYCLOPEDIA OF PHILOSOPHY (Summer 2020 Edition), <https://plato.stanford.edu/archives/sum2020/entries/rule-of-law/>, accessed September 14, 2021.

⁵³ Jeffrey Rachlinski and Andrew Wistrich, *Judging the Judiciary by the Numbers: Empirical Research on Judges*, 13 ANNUAL REV. OF L. & SOCIAL SCIENCE 203, 222 (2017); see also, MATTHEW P. HITT., *INCONSISTENCY AND INDECISION IN THE UNITED STATES SUPREME COURT* (University of Michigan Press 2019) at 82, noting that "[t]he rule of law in the United States requires that citizens receive predictable and nonarbitrary treatment when they seek relief in the judicial system."

⁵⁴ Presumably, the incentives to mischaracterize cases by editing their Wikipedia entries or by or authoring new entries will apply equally to foreign powers or foreign nationals who have potential interests in litigation before the courts. Evidently, the significant cyber resources at the disposal of adversarial state actors would significantly facilitate such interference.

to their decision".⁵⁵ The paper's challenge, rather, is to the rationality of the process by which, in specific cases, particular legal materials are actually accorded such authority. In the broadest range of cases, judges may well seek an outcome exclusively in the legal materials; whether a judge will apply her *expertise*—or that of her clerk—to discern what these materials in fact imply is another question entirely.

III. IT'S A WIKI WORLD

In recent decades, the falling cost of digital production technologies (e.g., video recording devices) and the ability to distribute material via the internet has led to an efflorescence in digital media.⁵⁶ For example, more than 500 hours of video content is uploaded to YouTube every minute.⁵⁷ There are strong theoretical reasons to believe that this proliferation is not just low-quality "noise" clogging up searches, but that it contains nuggets of high-quality content that would not have otherwise been produced and that these can substantially improve public welfare; great ideas that did not seem promising beforehand can be re-evaluated once they are actually published.⁵⁸ An important feature of this democratization of digital production is the creation of user-generated content (UGC), such as blogs, Instagram feeds, homemade podcasts or YouTube videos.

a. Knowledge on Tap

The Organization for Economic Cooperation and Development defines UGC as "(i) content made publicly available over the Internet, ii) which reflects a 'certain amount of creative effort', and iii) which is 'created outside of professional routines and practices'."⁵⁹ Because UGC carries no guarantees about the creator or the process of creation, there have always been debates about the quality of information provided on these platforms.⁶⁰ Of particular concern are questions about insufficient expertise and contributor bias.⁶¹ Notwithstanding these weaknesses, UGC clearly carries important information. For example, it can be used to predict stock market outcomes⁶² and blog mentions were a stronger predictor of 2008 US presidential polls than other forms of media.⁶³ UGC has also been shown to

⁵⁵ J. Raz, *Two Views of the Nature of the Theory of Law: A Partial Comparison*, in J. Coleman (ed), *HART'S POSTSCRIPT: ESSAYS ON THE POSTSCRIPT TO THE CONCEPT OF LAW* (Oxford University Press, 2001) 1, 34.

⁵⁶ Anderson, S.P., Waldfogel, J. and Stromberg, D. eds., 2016. *Handbook of Media Economics, vol 1A*. Elsevier.

⁵⁷ STATISTA, *Hours of video uploaded to YouTube every minute as of May 2019*, (May 2019), <https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/>, accessed September 14, 2021.

⁵⁸ Joel Waldfogel, *The Random Long Tail and the Golden Age of Television*, 17 *INNOVATION POLICY & THE ECONOMY* (2017).

⁵⁹ Organisation for Economic Co-operation and Development, *Working Party on the Information Economy, PARTICIPATIVE WEB: USER-CREATED CONTENT (DSTI/ICCP/IE(2006)7/FINAL)* (April 12, 2007) <https://www.oecd.org/digital/ieconomy/38393115.pdf>, accessed September 14, 2021.

⁶⁰ Roman Lukyaneko et. al., *The IQ of the crowd: understanding and improving information quality in structured user-generated content*, 25 *INFORMATION SYSTEMS RESEARCH* 669.

⁶¹ David Coleman et. al., *Volunteered geographic information: The nature and motivation of producers* 4(4) *INTERNATIONAL JOURNAL OF SPATIAL DATA INFRASTRUCTURES RESEARCH* 332-358 (2009).

⁶² Shawndra Hill and Noah Ready-Campbell, *Expert Stock Picker: The Wisdom of (Experts in) Crowds*, 15(3) *INT'L J. OF ELECTRONIC COMMERCE* 73 (2011).

⁶³ Sunil Wattal et. al., *Web 2.0 and Politics: The 2008 U.S. Presidential Election and an E-Politics Research Agenda*, 34(4) *MANAGEMENT INFORMATION SYSTEMS Q.* 669 (2010).

impact decision-making by venture capitalists, as well as purchasing behavior by consumers.⁶⁴

Wikipedia is one of the world's largest UGC platforms. It gets 13.6 billion visits per month, making it the fourth most-visited website in the world.⁶⁵ Wikipedia content is also used by Google and other search engines as part of their results, which bolsters Wikipedia readership beyond even these lofty figures.⁶⁶ Wikipedia covers technical topics with great breadth. An analysis in the field of chemistry showed that nearly 90% of university undergraduate topics and 50% of graduate topics are covered by Wikipedia articles, and that Wikipedia is either the largest or second largest source of review-like articles in the world—only the academic literature itself may have more.⁶⁷ The academic articles and monographs cited on Wikipedia tend to be highly cited,⁶⁸ suggesting they may be viewed favorably by the academic community. Given this breadth of knowledge and connection to important references, it may not be surprising that a survey of Spanish academics found that 38.1% of faculty consult Wikipedia articles from their own discipline “frequently” or “very frequently”⁶⁹ and that many use Wikipedia articles as a stepping stone to the sources they reference.⁷⁰

Wikipedia has many of the advantages associated with UGC. First, Wikipedia is easily and freely accessible. Second, Wikipedia acts as an aggregator of information,⁷¹ and, as far back as 2006, researchers have credited aspects of Wikipedia's coverage as being largely accurate and credible.⁷² Third, consistent with the principle that “a greater number of contributors to an article makes an article more neutral” it has been shown that Wikipedia can be good at minimizing bias.⁷³

b. Expertise Optional

⁶⁴ Rohit Aggarwal and Harpreet Singh, *Differential Influence of Blogs Across Different Stages of Decision Making: The Case of Venture Capitalists*, 37(4) *MIS QUARTERLY* (2013) 1093–1112. See also, Khim-Yong Goh et. al., *Social Media Brand Community and Consumer Behavior: Quantifying the Relative Impact of User-and Marketer-Generated Content*, 24(1) *INFO. SYSTEMS RESEARCH* 88 (2013).

⁶⁵ STATISTA, *Most popular websites worldwide as of June 2021, by total visits*, (July 2021), <https://www.statista.com/statistics/1201880/most-visited-websites-worldwide/>, accessed September 14, 2021.

⁶⁶ Conor McMahon, Isaac Johnson and Brent Hecht, *The Substantial Interdependence of Wikipedia and Google: A Case Study on the Relationship Between Peer Production Communities and Information Technologies*, *PROCEEDINGS OF THE ELEVENTH INTERNATIONAL AAAI CONFERENCE ON WEB AND SOCIAL MEDIA (ICWSM 2017)* 142, 148-149 (2017).

⁶⁷ Neil Thompson and Douglas Hanley, *Science is Shaped by Wikipedia: Evidence from a randomized control trial*, MIT Sloan Research Paper No. 5238-17 (February 13, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3039505, accessed September 14, 2021.

⁶⁸ Perry Evans and Michael Krauthammer, *Exploring the Use of Social Media to Measure Journal Article Impact*, *AMIA ANNU. SYMP. PROC.* 374, 380 (2011); Kayvan Kousha and Mike Thelwall, *Are Wikipedia Citations Important Evidence of the Impact of Scholarly Articles and Books?*, 68 *J. OF THE ASSOC. FOR INFORMATION SCIENCE AND TECHNOLOGY* 762, 774-775.

⁶⁹ E. Aibar, J. Lladós-Masllorens, A. Meseguer-Artola, J. Minguillón, and M. Lerga, *Wikipedia at University: What Faculty Think and Do About It*, 33 *THE ELECTRONIC LIBRARY* 668, 675 (2015).

⁷⁰ T. Piccardi, M. Redi, G. Colavizza, and R. West, *On the Value of Wikipedia as a Gateway to the Web*, *PROCEEDINGS OF THE WEB CONFERENCE 2021 (WWW '21)* 249, 255-256 (2021).

⁷¹ Sean Xin Xu and Xiaoquan (Michael) Zhang, *Impact of Wikipedia on Market Information Environment: Evidence on Management Disclosure and Investor Reaction* *MIS QUARTERLY*, 37(4) 1043–1068 (2013).

⁷² T. Chesney, *An Empirical Examination of Wikipedia's Credibility*, *FIRST MONDAY* (2006).

⁷³ Shane Greenstein et. al., *Ideological Segregation among Online Collaborators: Evidence from Wikipedians*, N.B.E.R. Working Papers 22744 (2016). Greenstein et al present the stated principle as a narrow interpretation of Linus' Law of software development, which holds that “with enough eyeballs, all bugs are shallow” (coined by ERIC S. RAYMOND, *THE CATHEDRAL AND THE BAZAAR* (O'Reilly Media 1999)).

Wikipedia also has distinct weaknesses. Articles whose content is more “peripheral” have been found to be of lower quality.⁷⁴ Equally, there is evidence that Wikipedia articles discuss contested knowledge (that is, “subjective, unverifiable, or controversial information”) in ways that are more in line with America’s Democratic Party voters than its Republican Party voters.⁷⁵ Moreover, it appears that these biases in Wikipedia content are larger than those in comparable expert-based articles in Encyclopaedia Britannica.⁷⁶ Thus, the evidence about the accuracy and unbiasedness of Wikipedia’s content is mixed. At its best, Wikipedia is voluminous, accurate, and unbiased. At its worst, it is none of these.

Two previous randomized field experiments on Wikipedia prove the causal role that Wikipedia can play in shaping knowledge and behavior. The first of these studies found that Wikipedia shapes academic science, showing that adding articles on a chemistry topic changes how the topic is discussed in the scientific literature.⁷⁷ It further shows that scientific articles added as references to Wikipedia get more academic citations as a result. The second study tested the effects of Wikipedia on where tourists choose to visit.⁷⁸ It found that adding more information on tourist sites to a city’s Wikipedia page leads to more overnight tourism in that city. Given these results, there is no doubt that Wikipedia is shaping public and specialist knowledge. But could it also affect the practice of adjudication?

In keeping with its currency, collaborative design, breadth, and internationalism,⁷⁹ Wikipedia is increasingly cited as a general source of information in legal scholarship and court judgments alike.⁸⁰ Citations to Wikipedia in US judicial opinions first appeared in 2004.⁸¹ By May 2010, there were at least 117 US state and 326 federal cases featuring judicial opinions that cited Wikipedia.⁸² As of September 2021, there were 6,901 references to Wikipedia in documents indexed in US Westlaw’s JLR (“Journals and Law Reviews”) Database and 1,627 such references in the All-briefs Database.

In respect of certain jurisdictions, Wikipedia also carries substantial information about individual legal cases. 3,315 US Supreme Court decisions had Wikipedia pages as of September 2021. In contrast, the coverage of cases in other jurisdictions is often poorer.

⁷⁴ Gerald Kane and Sam Ransbotham, *Content as Community Regulator: The Recursive Relationship Between Consumption and Contribution in Open Collaboration Communities*, 27(5) ORGANIZATION SCIENCE 1258 (2016). Technically, “peripheral (low centrality)” means a lack of graph centrality where Wikipedia pages are nodes on the graph and edges are the links between them. Quality was measured on a seven-point scale from lowest to highest quality evaluated by the Medicine WikiProject. They also use additional measurements of quality, including agreement with experts, such as medical students. See further, Linton C. Freeman, *Centrality in social networks conceptual clarification*, 1(3) SOCIAL NETWORKS 215-239 (1978-1979).

⁷⁵ Shane Greenstein et. al., *Do Experts or Crowd-Based Models Produce More Bias? Evidence from Encyclopædia Britannica and Wikipedia*, 42(3) MIS QUARTERLY 945–959 (2018).

⁷⁶ Id.

⁷⁷ Neil Thompson and Douglas Hanley, *Science is Shaped by Wikipedia: Evidence from a randomized control trial*, MIT Sloan Research Paper No. 5238-17 (February 13, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3039505, accessed September 14, 2021.

⁷⁸ Marit Hinnosaar et. al., *Wikipedia matters*, J. OF ECONOMICS AND MANAGEMENT STRATEGY (2019).

⁷⁹ R. Anderson, M.T. González and S. Lee, *Toward a new student insurgency: A critical epistolary*, 94(6) CALIFORNIA L. REV. 1879 (2006); D.R. Irwin, *Freedom of thought: The First Amendment and the scientific method*, WIS. L. REV. 1479 (2005); J. Knouse, *From identity politics to ideology politics*, UTAH L. REV. 749 (2009); B. Meyer, *Defense and illustration of Wikipedia*, EIFFEL WORLD COLUMN (2006).

⁸⁰ Wikipedia, *Wikipedia as a court source*, https://en.wikipedia.org/wiki/Wikipedia:Wikipedia_as_a_court_source, accessed September 14, 2021.

⁸¹ L.F. Peoples, *The citation of Wikipedia in judicial opinions*, 12 YALE JL & TECH 1 (2009).

⁸² J.L. Gerken, *How Courts Use Wikipedia Developments*, 11 J. APP PRAC. & PROCESS 191-227, 191 (2010).

Notably, as of the start of our experiment in early 2019, there were only nine Irish Supreme Court decisions that had associated Wikipedia articles.

IV. THEORY TO DATA

The two basic empirical questions about Wikipedia's influence, if any, on the application of the law are a) what to recognize as evidence of such influence, and b) what such evidence indicates about how Wikipedia is influencing adjudication.

a. Hypothesis

Our discussion of the theory of legal reasoning prompts two competing hypotheses about the role of Wikipedia in judges' behavior. As noted above in Section II, most contemporary theorists endorse the legal model as a descriptive matter, supposing that judges resolve most cases by reference to what the relevant legal materials say. This generates Hypothesis 1:

Legal Model - Wikipedia does not replace judges' expert application of the law.

Wikipedia case summaries are generated on an ad hoc basis. Likewise, they may be authored and edited by non-specialists who may have no training whatsoever in law, let alone any training in the relevant jurisdiction. Judges cannot safely determine either the authors or the editors of Wikipedia case analyses. In deferring to analyses of unknown origin and/or arbitrary coverage, the judge is abdicating her personal responsibility to apply the law. To allow a case's consideration as a precedent to depend either on the chance event of its inclusion on Wikipedia or on the interpretive decisions of Wikipedia author/editors would compromise the deliberative character of her decision. That, in turn, would undercut parties' entitlement to the disposition of their litigation by means of judicial expertise. So, too, would the judge's failure to independently analyze the texts of the decisions cited in parties' legal submissions before determining their significance. Any practice of deciding cases at an inexpert remove from the texts of the relevant precedents would also make their resolution harder to predict, thereby undercutting a principal motivation for courts' adherence to the doctrine of precedent, namely, certainty in the law (on which, see Section II, above).

If the legal model is accurate, then the judge *will not* rely on Wikipedia to help establish a previous case's relevance. Conversely, the breadth of Wikipedia's influence on even highly specialized research scientists (see Section III above) suggests that judges might instead prefer the convenience of googling a case's name, clicking on the closest, most prominent link, and perhaps paraphrasing the information found there. This suggestion generates Hypothesis 2:

Moot Court Model - Wikipedia does replace judges' expert application of the law.

Taken to its fullest expression, this model suggests that judges treat legal reasoning as a judge might treat the task of identifying the best legal team in a student mooted competition. Attention is duly paid to the parties' filings and oral arguments but the actual veracity of their respective descriptions of the legal materials is of secondary importance. Judges might just rely on the descriptions in the parties' filings themselves; or they might supplement this information by referring to all or some discussions of these cases in the filings submitted as

part of previous editions of the competition. On being required to write up a mock judgment, they may simply echo the analysis and/or case quotations from the filings themselves:

[T]here is nothing unexpected about a judge using materials that happen to be at hand; the only interesting question may be: How did that particular tool happen to be at hand now, when it seems not to have been available earlier?⁸³

If the moot court model of judicial behavior is accurate, then the judge *will* use a source such as Wikipedia to help establish a previous case's relevance. In view of the reliance by scientists on the convenience of Wikipedia in making research choices, we predicted that at least sometimes judges (or their clerks) would rely on it similarly for the purpose of justifying their decisions, i.e., that they would behave according to the moot court model. Considering their heavier caseloads, we anticipated that this effect would be most apparent in the decisions of judges working in courts of first instance.⁸⁴

b. Causal Mechanism

Wikipedia's impact on judicial behavior could be direct, whereby the judge or her clerks consult the website of their own volition or indirect, whereby the parties consult the website in constructing the arguments that form the basis of the judge's deliberations.⁸⁵ To establish the mechanism of influence, we must consider what direct and indirect consultation might mean for the content of the resultant legal writing.

One obvious way in which Wikipedia might be used as a legal research tool is to confirm the similarity of cases that, in view of some prior, secondary source, are anticipated to suggest a particular outcome to the case at bar. The secondary source might be the relevant section of an academic textbook, or, in the case of the judge/clerk, the parties' respective legal arguments.

The starting point of a confirmatory use of Wikipedia is a theory of the instant case that indicates a particular outcome, and which allows the lawyer to assess whether another, prior case is similar and might therefore provide a legal basis for that outcome. For the author of a legal argument exercising discretion on what to cite, those cases that Wikipedia confirms to be analogous would appear in support of the stated reasoning, whereas those which it indicates are distinguishable, or are points of contrast, would not be cited. This feature of the evidence of the manner of Wikipedia's influence on legal writing permits us to distinguish the operation of direct and indirect causal mechanisms in respect of the legal reasoning of judges.

It might be the case that, in affecting the content of their arguments, the use of Wikipedia by litigants' lawyers influences the reasoning of the judge who must respond to these arguments. Used by both parties to litigation in a confirmatory fashion, Wikipedia will help each to choose supportive precedents for their competing legal submissions. Just in virtue of its responsiveness to the parties' respective submissions, we would then expect the court's

⁸³ M Tushnet, *The Possibilities of Comparative Constitutional Law* 108 YALE L. J. 1225, 1304 (1999).

⁸⁴ See our pre-registered hypothesis and analysis plan, which is available on As Predicted: <https://aspredicted.org/download_pdf.php?b=cyT1zTihHlsPqLplqDIHzBP5G12zKslKGGUZWxD5N6S7k9nEn8&a=dk13V1RwaWN1VIJyY0gwaU56ZFOU0T09>.

⁸⁵ An alternative possibility is that the judge/clerks are influenced by media coverage of a case that itself relies on the Wikipedia case summary. In related research, we discover a positive correlation between a US Supreme Court case's inclusion on Wikipedia and its subsequent citation in the *New York Times*.

judgment to feature more citations to Wiki-summarized cases. Notice, however, that in responding to litigants' Wiki-induced case citations, the judge is responding to cases both supportive and unsupportive of her ultimate legal conclusions. Thus, if Wikipedia were employed in this fashion by the parties alone, then any increase in the citation of Wiki-summarized cases in the court's judgment should extend to positive and negative citations alike. Conversely, the judge/clerk might consult Wikipedia directly to confirm whether a previous case is suitably analogous to the immediate case to make it amenable to be cited in support of her stated reasoning. If it is the judge/clerk herself who is consulting Wikipedia in this fashion, then the increased references in her judgment to Wiki-summarized cases should instead consist in positive citations and not in negative citations.

Evidence that judges also tend to echo the text of Wikipedia case summaries in their judgments' own text would seem to exclude the prospect that they, or their clerks, conduct a subsequent review of the website's information by reference to the primary sources, i.e., the cases themselves. Thus, if Wikipedia were just a steppingstone to primary sources, then, given the assistance of the latter, we would not expect the text of Wikipedia to continue to matter. This would be true even if the relevant Wikipedia text were just a passage ostensibly quoted from the pertinent judgment. Should the paraphrased Wikipedia text extend even to the Wiki-author's analysis of the previous case, Wikipedia would influence how the judge writes and which concepts she is connecting. In that event, the judge would effectively have outsourced the structure of her argument to the internet.

If the legal model is accurate, then i) a case's inclusion on Wikipedia will not correlate with its judicial citation in a way that suggests a direct influence, i.e., with positive but not negative citations, and ii) the text of a case's Wikipedia entry will not correlate with the text of judicial opinions. Conversely, establishing either correlation would confirm the applicability of the moot court model.

V. EXPERIMENTAL DESIGN

Ireland offered an ideal legal system in which to analyze whether Wikipedia articles influence judges' legal reasoning. As lawyers working in the common law tradition, findings about the behavior of Irish judges might extend to those working in many other national systems. Equally, an absence of Wikipedia articles on Irish court decisions allowed us to measure the impact of the inclusion of a Wikipedia entry by means of a field experiment.⁸⁶

154 new Wikipedia articles on Irish Supreme Court cases were created, mostly by law students. Half of these (77) were randomly chosen to be uploaded to Wikipedia where they could be used by judges, clerks, barristers, or whomever else sought them out (the *treatment* group). The other half were held back and not put on Wikipedia. This second group provided the counterfactual basis of what would happen to a case absent a Wikipedia article about it (the *control* group). Both sets of articles were written so that linguistic analysis could be done on them.

⁸⁶ Within the Irish State, the Supreme Court is the court of final appeal. Sitting below the Supreme Court, the Court of Appeal, established in 2014, hears appeals from the High Court. As the third superior court in Ireland, the High Court is a court of first instance that has full jurisdiction to hear questions of law or fact in civil and criminal cases, as well as the power to hear questions on the constitutionality of legislation. Below the superior courts, the Circuit Court and District Court are courts of first instance with limited jurisdiction. Like other common law judges, Irish judges are appointed by the head of state on the advice of the government.

The goal of the experiment was to see whether these Wikipedia articles would affect the legal decisions of Irish judges. We looked at two measures: (1) whether the cases that we wrote up were more likely to be cited in judicial decisions, and (2) whether the argumentation in court judgments echoed the semantic content of these Wikipedia pages.

a. Choosing the Sample

We first created a candidate pool of Irish Supreme Court cases on which a new Wikipedia article could be written, and then narrowed that pool to a set on which any effect that existed could be most accurately measured.

Open access to written judgments in Ireland is available through the Irish courts service website, which contains virtually all written Supreme Court judgments issued since 2001, and Court of Appeal and High Court judgments issued since 2014 and 2005, respectively.⁸⁷ However, when published in this way, the judgments are not assigned to any particular legal category or divided by legal topic. The judgments also do not contain an index of cases cited, a reference list setting out which cases the court followed, referenced, distinguished, or overruled, or a list of cases that have subsequently cited the judgment. As such, the Irish courts services website provides open access to the written judgment, but no editorial guidance on the content, relevance, or potential impact of that judgment.

To determine which cases to include in our analysis, we relied on case and citation information from the legal database JustisOne.⁸⁸ JustisOne is a subscription-based, legal information platform that publishes written judgments from the Irish courts in the form that they appear on the Courts services website, as well as law report versions where available. This platform categorizes cases by the areas of law that are relevant to the judgment. It also includes an index of cases cited within the judgment (as well as the treatment of those cases) and an index of cases that have subsequently cited the judgment together with information on how subsequent cases have treated the judgment. Treatment of cases (both within a judgment and in subsequent cases) can be positive (such as where the judgment was followed, affirmed, applied or approved), neutral (where the judgment was considered, referred to, explained, cited or discussed) or negative (where the judgment was not applied, not followed, distinguished or overruled).

To select the cases that would be included in this project, we chose seven categories of law used by JustisOne: administrative and constitutional law; asylum, immigration and nationality; crime and sentencing; family law; tort; practice and procedure; and banking and finance. These areas of law were selected based on the comparatively high prevalence of cases within each category and the potential for cases in these areas of law to establish a precedent that would be referred to in subsequent cases. To maximize the application of the doctrine of precedent, we restricted our analysis to Supreme Court cases as these would be

⁸⁷ THE COURTS SERVICE OF IRELAND, <https://www.courts.ie/judgments>, accessed September 14, 2021. The Irish Reports, which, together with the other main series of law reports in Ireland – the Irish Law Reports Monthly – report a (very small) selection of the more instructive, novel or legally significant judgments issued each year. These reports include headnote summaries of the case facts, judgment and relevant legal principles discussed, as well as a list of cases and legislative provisions cited. As commercially published reports, the Irish Reports and the Irish Law Reports Monthly are only accessible on a payment or subscription basis and are often subject to a delay in their publication; see RAYMOND BYRNE ET. AL., *BYRNE & MCCUTCHEON ON THE IRISH LEGAL SYSTEM* 545 (Bloomsbury Professional 2021). Other subscription-based services, such as LexisLibrary and Westlaw, also provide access to written judgments and editorial summaries.

⁸⁸ JUSTISONE, <https://app.justis.com/>, accessed September 14, 2021.

binding on the maximum number of subordinate Irish courts. We also excluded from consideration the nine Irish Supreme Court cases that already had Wikipedia articles written about them.

Irish Supreme Court cases available on JustisOne issued in each year from 2000 to 2017 within these seven categories were then analyzed. Since our most direct outcome of interest was subsequent case citations, within this set of cases we focused our analysis on instances where the average number of citations per year was higher because this made the measurement of effects more accurate. For example, if Wikipedia were to double the number of citations that cases receive, this would be relatively easy to detect for a case that typically gets five citations per year and thus jumps to ten. In contrast, it would be very hard to detect for a case that averages one citation per twenty years because during an experiment of only a few years the most probable outcome, either at the original citation rate or a doubled one, would be that it receives zero citations—and thus, it would be impossible to ascertain whether the rate had doubled. For this reason, we narrowed our sample to the most frequently cited cases within each category in each year.⁸⁹ By choosing our cases from amongst these more highly cited ones, we improve the experiment's signal-to-noise ratio. The treatment of each of these cases in subsequent citing cases (whether it was positive, neutral or negative) was also recorded.

b. Building the Treatment and Control Groups

The key procedural objective of a randomized control trial is to generate two groups—a treatment group that will get the intervention and a control group that will not—that are as similar as possible. When fully achieved, this eliminates the possibility that a factor other than the treatment is responsible for any change in outcomes.⁹⁰ In practice, this goal of perfect similarity (or “*balance*”) can be approached using a variety of techniques, including *randomization* and *stratification*.

Randomization provides a way to divide the sample (Irish Supreme Court cases about which Wikipedia articles could be written) into the treatment and control group. This procedure is valuable because each observation may have different *unobservables*, characteristics whose existence could affect the outcomes that we care about, but that are not visible to those conducting the experiment. For example, some cases might have been studied by Irish judges during their legal training, which could lead to those cases being more likely to be cited. To achieve balance, we want the incidence of this characteristic to be allocated evenly across the treatment and control groups. Since this characteristic is unobserved there is no way to intentionally divide the sample in this way. Thus, instead we rely on randomization to achieve that balance probabilistically. Fortunately, after randomizing a sample as large as ours the probability of a substantial mismatch is vanishingly small.⁹¹ While it is reassuring that randomization makes us highly unlikely to be led astray, it does not achieve our full goal. We not only want to not be wrong, we want to be sufficiently confident that we are right to conclude that our results are capturing real world phenomena.

⁸⁹ If there were not enough similar cases in a particular year x law type, we did not include those in the pool of potential articles to be written.

⁹⁰ John A. List, *Why economists should conduct field experiments and 14 tips for pulling one off*, 25(3) JOURNAL OF ECONOMIC PERSPECTIVES 3-16 (2011).

⁹¹ Eric W. Weisstein, *Central Limit Theorem* (MathWorld--A Wolfram Web Resource) <https://mathworld.wolfram.com/CentralLimitTheorem.html>, accessed September 14, 2021.

Stratification, or as it is sometimes called, 'blocking', helps guarantee that the balance between the treatment and control groups is even closer than randomization on its own would yield.⁹² This is done by splitting a heterogeneous sample into similar subsets and randomizing within each subset rather than the undivided whole. For example, a set of eight items of two types might be denoted {A,A,A,A,B,B,B,B}. This could be split (i.e. stratified) into 2-item subsets such as {A,A}, {A,A}, {B,B}, and {B,B}. Whereas randomization on the collective group will on average (but not always) lead both the treatment group and control group to have two As and two Bs, randomizing within each of the stratified samples guarantees this (one A from the first group, another from the second, one B from the third group and another from fourth). That is, stratification can help achieve balance on those characteristics that are observed by the experimenters.

To stratify effectively, we combined the citation information from JustisOne with other information we gathered on the presence of each case in various forms of media: the number of times a case was referred to on the website of RTÉ⁹³ (Ireland's national broadcaster), the website of the Irish Times⁹⁴ (a daily newspaper in Ireland), and in other public media sources.⁹⁵ As in our hypothetical example, we took the larger set of cases and sub-divided them into pairs. To be eligible to be a pair, cases needed to be decided in the same year and be part of the same type of law. For example, two "asylum, immigration and nationality" law cases from the year 2000. Within the potential pairs that met these first two conditions, we stratified cases based on finding their "nearest neighbor"⁹⁶ according to: # positive citations, # neutral citations, # negative citations, publication year, whether mentioned in RTÉ, whether referenced more or less than the median number of times in the Irish Times, and whether mentioned more than 10 times in other public media.⁹⁷ For 9 cases, no sufficiently comparable nearest neighbor case could be found and they were excluded from the experiment.

The resultant pairs of cases were then given to the article writers, mostly law students. For each pair, a single student would write both cases. This guaranteed that each author had an exactly equal number of articles in the treatment and control group. Put another way, by implementing this aspect of our experimental design we automatically stratified on the author characteristics, and thus our articles were also balanced in those ways too.

Given our large sample and the careful nature of our stratification, we would expect our randomization⁹⁸ to produce treatment and control groups that are highly similar. This was indeed true across many variables, including pre-treatment citations (shown below). Based

⁹² R.W. Sanson-Fisher, B. Bonevski, L.W. Green and C. D'Este, *Limitations of the randomized controlled trial in evaluating population-based health interventions*, 33(2) AMERICAN J. OF PREVENTIVE MEDICINE 155, 155-161 (2007).

⁹³ RTÉ IRELAND, <https://www.rte.ie/>, accessed September 14, 2021.

⁹⁴ THE IRISH TIMES, <https://www.irishtimes.com/>, accessed September 14, 2021.

⁹⁵ These Google searches were done using both case names and citations.

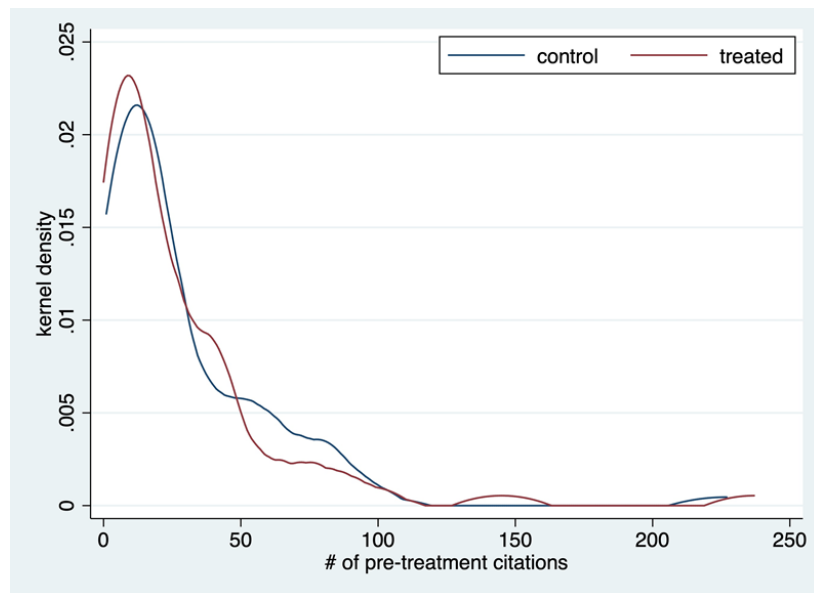
⁹⁶ Formally, nearest neighbor matching constructs a distance metric between each set of cases and selects pairs that are closest to each other. This mathematical formalism achieves the intuitive goal of finding cases that are similar across multiple dimensions.

⁹⁷ This was done using the quickblock package in the R programming language using studentized distances and variable weights on the covariates mentioned in the text of (0.1, 0.1, 0.1, 0.5, 0.1, 0.1, 0.1).

⁹⁸ Technically, we performed complete randomization (i.e. picking equal subsets for treatment and control) rather than individual randomization (i.e. each observation gets a 50-50 chance of being in treatment or control) because this guaranteed equal sample sizes which maximizes statistical power (Wei and Lachin, 1988); L.J. Wei and J.M. Lachin, *Properties of the urn randomization in clinical trials*, 9(4) CONTROLLED CLINICAL TRIALS 345-364 (1988).

on the observed balance between our treatment and control groups, we conclude that our stratification and randomization were successful.

Fig 1 Distribution of pre-treatment citations for Irish Supreme Court cases in our sample



c. Writing the Wikipedia Articles

For this project we created 154 draft Wikipedia articles on Irish Supreme Court cases. The process of creation was done in three waves. After each wave, a random half of the articles were added to Wikipedia and the other half held back. The first wave, in early 2019, was a pilot study to understand the article creation and addition process in which law faculty in Maynooth University wrote Wikipedia articles on 14 cases. In the second wave, in Spring 2019, undergraduate law students from Maynooth University created 8 articles as part of the civic engagement stream. These were published in late 2019. In the third wave, in Fall 2019, a cohort of graduate students wrote 132 articles as part of a professional development seminar. These were published in early 2020.

For all student-authored papers, logistical support was provided by faculty, as guided by previous related work.⁹⁹ This included induction sessions and fortnightly editing sessions in a computer lab. We also developed a suite of electronic resources to help students with Wikipedia editing.

We provided students with detailed article design guidelines. The faculty articles from the pilot study served as exemplars, and we emphasized the importance of secondary sources and the creation of strong article leads to ensure that the resultant articles would pass Wikipedia's new article creation screening process. We also required students to include a Wikipedia infobox for each case (the summary boxes found at the right-hand side of many Wikipedia articles). Infoboxes improve the appearance of the article but crucially they also embed metadata that allows search engines such as Google to draw upon their content. Before

⁹⁹ Brian McKenzie et. al., *From Poetry to Palmerstown: Using Wikipedia to Teach Critical Skills and Information Literacy in a First-Year Seminar*, COLLEGE TEACHING 140, 140-147 (2018).

publishing any student work, we also vetted the articles for copyright violations using the freely available tool Earwig.¹⁰⁰

d. Integrating the New Articles

Of the 154 Wikipedia articles that we created for this experiment, 77 were uploaded to the Wikipedia website as the treatment group. We then managed this integration process to ensure that they were not incorrectly flagged for copyright violations because of the use of direct quotations from cases, as well as to ensure that they were viewed as “notable.” The second is not automatic because there are no blanket criteria — not even being a Supreme Court decision — that guarantees that court cases will be notable by Wikipedia guidelines (in contrast, for example, any individual who has appeared in the starting line-up of a game in a “fully professional soccer league” is automatically considered notable¹⁰¹).

Once published, we made two small but important additions to each article. First, we added a “short description” to each article: “Irish Supreme Court case.” This appears as a small line of text under an article’s title. This is important for the visibility of Wikipedia articles on mobile platforms.¹⁰² We also added several categories to each article: “Supreme Court of Ireland cases,” “[year of case] in case law,” “[year of case] in Irish law,” and the area of law, e.g., immigration, criminal, constitutional. Categories are important both for search engines and internal Wikipedia linking.

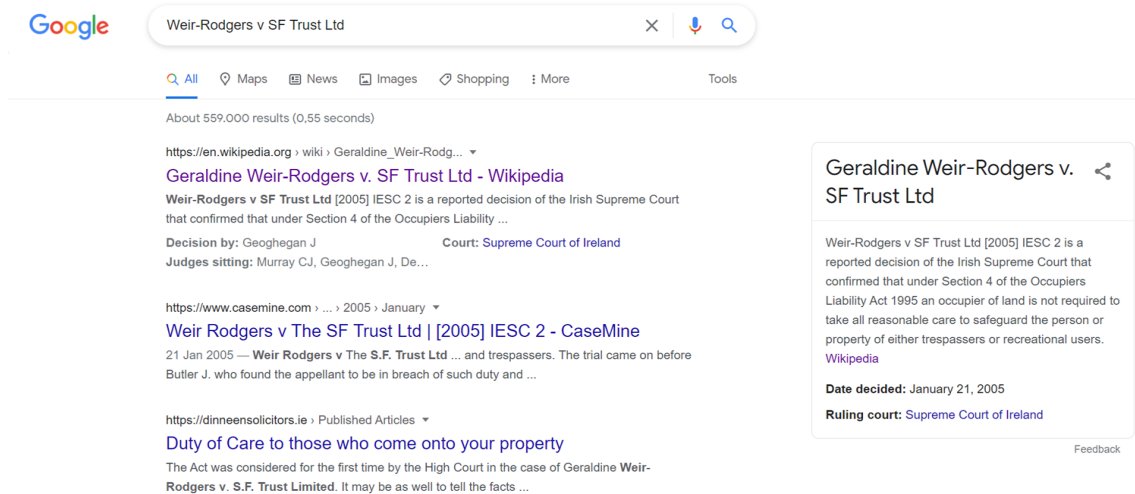
The combination of infoboxes, categories, and short descriptions resulted in a high level of visibility for our articles on various search engines. Our Wikipedia articles were the first search result on Google, Bing, or DuckDuckGo in almost every case when searched by decision title or just the citation. More impressively, internet search engines (Google, Bing, DuckDuckGo) now pull text and information from our article leads and infoboxes to create so-called “knowledge panels,” summary boxes to the right of the search results, through which one can click directly into the relevant Wikipedia entry.

Fig 2 Screenshot of Google search results (13 September 2021) for *Weir-Rodgers v. SF Trust Ltd.*

¹⁰⁰ CopyVio Detector, <https://copyvios.toolforge.org/>.

¹⁰¹ Wikipedia, *Wikipedia:Notability (sports)*, [https://en.wikipedia.org/wiki/Wikipedia:Notability_\(sports\)](https://en.wikipedia.org/wiki/Wikipedia:Notability_(sports)), accessed September 14, 2021.

¹⁰² Wikipedia, *Short description*, https://en.wikipedia.org/wiki/Wikipedia:Short_description, accessed September 14, 2021.



Our articles received a total of 56,733 views through January 16, 2022.

VII. RESULTS

As discussed, to assess Wikipedia's impact on judicial decisions, we test for two types of influence: (1) whether the creation of a Wikipedia article on a case leads to that case being cited more often in judicial decisions; and (2) whether the text of judicial decisions is influenced by the text of the corresponding Wikipedia article.

In both cases, we use an ordinary least squares regression analysis with a difference-in-differences framework that further insulates us against any remaining differences that might exist between the treatment and control groups.¹⁰³ This estimator combines the intuitive appeal of comparing treatment vs control groups, but compares only their change as a result of the treatment (rather than their level). In practice this means that, e.g., for the citation analysis, we are comparing how much citation behavior *changed* for the treatment group (first difference: before vs after) and how that compares with the change that happened for the control group (second difference: treatment vs. control).¹⁰⁴

a. Citation Behavior

Overall, we find that the addition of a Wikipedia article increases the number of citations received by that case by 0.064 per month.¹⁰⁵ Thus, the addition of a Wikipedia article is substantially increasing citations to these Supreme Court cases in subsequent Irish court cases. Moreover, this result is statistically significant, meaning that it passes the accepted standards for distinguishing a real effect from one that could arise by chance.¹⁰⁶ The following graph illustrates the clear change in citation behavior. The difference in average

¹⁰³ JOSHUA ANGRIST AND JÖRN-STEFFEN PISCHKE, *MOSTLY HARMLESS ECONOMETRICS: AN EMPIRICIST'S COMPANION* 227-242 (Princeton Univ. Press 2009).

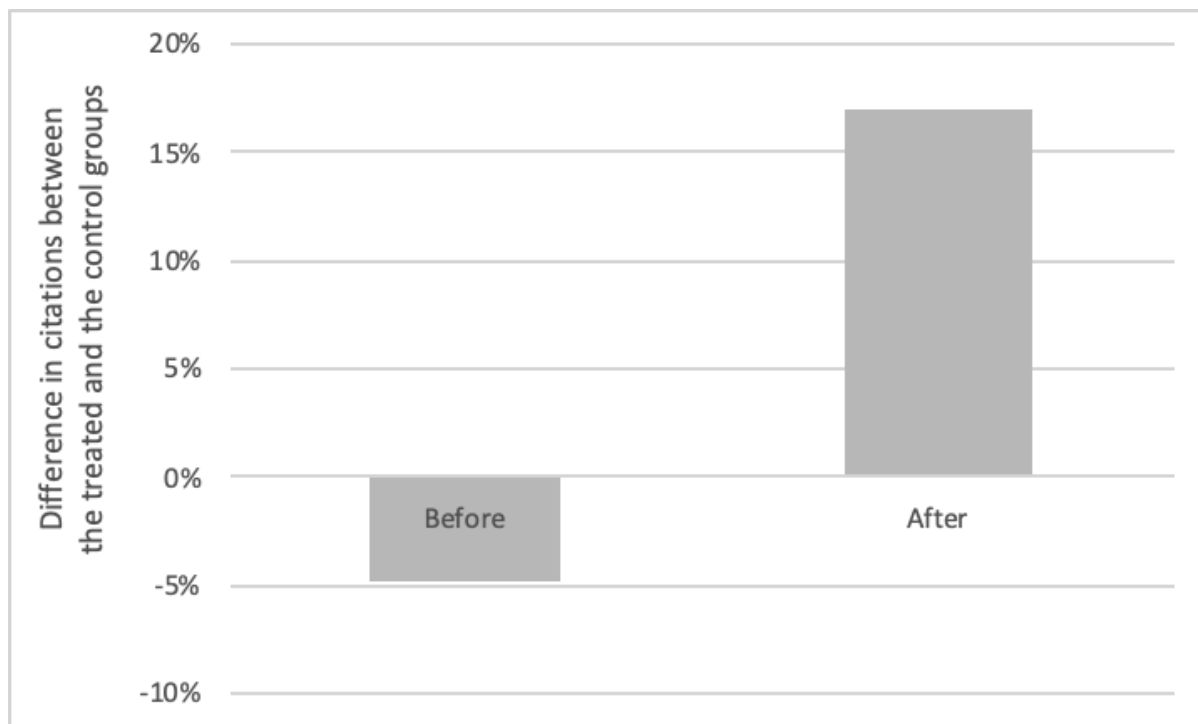
¹⁰⁴ For all the regression cases below we check robustness using a negative binomial formulation. We find qualitatively similar results with comparable statistical significance.

¹⁰⁵ Regression: citations = $\beta_0 + \beta_1 * \text{AFTER} + \beta_2 * \text{TREATMENT} + \beta_3 * \text{TREATMENT} * \text{AFTER}$, with block fixed effect using ordinary least squares with robust standard errors. Reported coefficient is the estimate for β_3 . P-value for the coefficient on β_3 is 0.026.

¹⁰⁶ For example, if one were attempting to see if a coin was double-headed, it would not be conclusive if three successive heads were flipped - even with a fair coin that arises occasionally. But if ten successive heads were flipped, one could be highly confident that the coin was not fair.

monthly citations between the treated and control groups is -4.8% in the pre-treatment period, and 17.0% in the post-treatment period. Hence, we estimate that adding a Wikipedia article increases judicial citations to those cases by 21.8%.

Fig 3 Difference in the average number of citations for court decisions that had a Wikipedia page (treatment) vs. those that did not get one (control)¹⁰⁷



The effect comes nearly entirely from changes to the citing behavior of judges issuing decisions in the High Court, a court of first instance. We see only small and statistically insignificant changes to the citing behavior of Supreme Court or Court of Appeal judges, suggesting no measurable effect. But for High Court judges we see a statistically significant increase in the number of citations received by that case by 0.050 per month.¹⁰⁸ This reveals that not all judges are affected, only those in courts of first instance, where caseloads are heaviest and the demand for expedience is greatest.

We can also distinguish the types of citations that are being changed. The clearest distinction is between positive and negative citations. The number of positive citations increases by 0.023 per month and this result is statistically significant,¹⁰⁹ whereas for negative citations the effect is much smaller and statistically indistinguishable from zero. Moreover, we can conclusively reject that positive and negative citations are being affected equally,¹¹⁰ meaning that discretion is being applied in where Wikipedia is being used.

¹⁰⁷ As estimated using the regression already described and converting to percentage terms by dividing by the average number of pre-citations per month for the whole sample.

¹⁰⁸ Court-type regressions: citations by that court = $\beta_0 + \beta_1 \text{AFTER} + \beta_2 \text{TREATMENT} + \beta_3 \text{TREATMENT} \text{AFTER}$, with block fixed effect using ordinary least squares with robust standard errors. Reported coefficient is the estimate for β_3 . P-value for the coefficient on β_3 is 0.021.

¹⁰⁹ Citation-type regressions: citations of that type = $\beta_0 + \beta_1 \text{AFTER} + \beta_2 \text{TREATMENT} + \beta_3 \text{TREATMENT} \text{AFTER}$, with block fixed effect using ordinary least squares with robust standard errors. Reported coefficient is the estimate for β_3 . P-value for the coefficient on β_3 is 0.017.

¹¹⁰ Citation regression, but appending positive and negative citations rows. The coefficient on $\text{TREATMENT} \text{AFTER} \text{POSITIVE}$ is statistically significant with P-value: 0.07.

The effect on neutral citations is less clear. Despite neutral citations comprising 86% of all the citations that we observe (positive is 12% and negative 2%), the uncertainty around our estimate is large. In particular, we observe an increase of citations of 0.036 per month, but that result is statistically insignificant. Put another way, there is so much noise around the estimate of the effect on neutral citations that there could plausibly be no effect or there could be an effect as big as the one on positive citations.

b. Argument Structure

The citation results viewed so far would be consistent with a story of our Wikipedia articles being used by judges as merely a stepping-stone to other, more authoritative sources that might be used to verify the relevant information. If this were true, then, once this definite information had been located by the judge, there would seem to be little advantage to relying on the text of the Wikipedia article in the actual drafting of the judgment. Accordingly, to test the stepping-stone possibility, we performed a linguistic comparison similar to that used in prior work to establish that Wikipedia causes changes to the content of chemistry journal articles.¹¹¹ Thus, using natural language processing, we correlated the linguistic content in the judicial decision that cites the case with that of the relevant Wikipedia article.¹¹²

We find that textual similarity does increase, showing a statistically significant effect.¹¹³ Since our articles include direct quotations from cases, it is possible that this effect arises merely because these same case quotations are appearing in the judicial decisions. To test this, we removed the quotations from the Wikipedia articles and retested for similarity. The results remain steady and statistically significant, revealing that the contextualization of the case by law students on Wikipedia is itself influencing judicial reasoning.

c. Discussion

The standard legal model of adjudication implies that Wikipedia does not replace judges' expert application of the law. On this account, a case's coverage on Wikipedia will not lead a judge to cite it, let alone to alter her legal analysis. On the moot court model of adjudication, in contrast, Wikipedia does indeed displace judicial expertise. The data suggest that the moot court model accurately describes a systematic and important feature of contemporary Irish legal adjudication—at first instance, at least. Specifically, this model is indicated by our finding that a case's inclusion on Wikipedia influences positive but not negative citations.

¹¹¹ Neil Thompson and Douglas Hanley, *Science is Shaped by Wikipedia: Evidence from a randomized control trial*, MIT Sloan Research Paper No. 5238-17 (February 13, 2018), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3039505, accessed September 14, 2021.

¹¹² Specifically, we use a bag-of-words approach with the word vector defined by usage in pre-treatment cases. We then apply a term-frequency, inverse document frequency weighting, which upweights words used more often in a particular document and downweights by the log of the frequency of use across all documents. This makes common words like 'the' much less important in the analysis and content words more important. We then take the cosine distance between articles as our measure. A range of work, including Kenneth Younge, and Jeffrey M. Kuhn, *Patent-to-Patent Similarity: A Vector Space Model*, INNOVATION & MANAGEMENT SCIENCE EJOURNAL (2016) and Omid Shahmirzadi et. al., *Text Similarity in Vector Space Models: A Comparative Study*, SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3259971 (2018), accessed September 14, 2021 illustrate the success of this approach in predicting human interpretations of similarity and the superiority/comparability of this method versus other approaches.

¹¹³ Regression: $\text{CosineSimilarity} = \beta_0 + \beta_1 * \text{AFTER} + \beta_2 * \text{TREATMENT} + \beta_3 * \text{TREATMENT} * \text{AFTER}$, with block fixed effect using ordinary least squares with robust standard errors. Reported coefficient is the estimate for β_3 . P-value for the coefficient on β_3 is 0.023.

The circumstances in which a case's inclusion in Wikipedia causes it to be cited as a precedent for the judge's proposed disposition—but not to be cited as a point of contrast—are those in which a judge consults Wikipedia directly to confirm that a case is analogous to the instant case (see Section IV, above). Accordingly, a case's inclusion on Wikipedia correlates with its judicial citation in a way that suggests a direct influence. Similarly, we find that the text of judicial opinions correlates with that of Wikipedia case entries, and that this correlation extends beyond passages quoted from the pertinent case to encompass Wikipedia's legal analysis itself. Taken together, these results indicate that, sometimes, judges are deferring to Wikipedia rather than applying their own legal expertise as they craft decisions.

For the moment this effect in Ireland has a natural ceiling, because the number of Irish court judgments on Wikipedia remains small. But in other jurisdictions, such as the United States, many more cases have already been added to Wikipedia. This suggests that, across many common-law jurisdictions, Wikipedia (and, by extension, the legal analysis of unknown internet users) might already be playing an important role in shaping judicial decisions.

VIII. CONCLUSION

Wikipedia is an enormous store of easily accessed, often correct knowledge about the world. It is natural that many users would turn to it for convenience and expediency. The challenge faced by judges is that availing themselves of this efficiency comes at an important cost: devolving responsibility for interpreting the context and importance of legal decisions to an unknown cadre of Wikipedia article authors. At best, this risks inexperienced analysis. At worst, it exposes judgments to bias and the tampering of external actors. Faced with such a high cost, and the resultant threat to legitimacy, one might hope that judges would forgo using Wikipedia as a source of legal information. Yet judges (and their clerks) are human and are confronted with overfilled dockets that demand expediency, and so they might be tempted to use Wikipedia. Our findings show that they do.

Using the benchmark of empirical proof, a randomized field experiment, this research shows that judges do rely on Wikipedia – indeed, the presence of a Wikipedia article (in our case, those written mostly by law students) generates an increase in citations equivalent to 25% of pre-intervention levels. We further show that it is not just that Wikipedia is being used as a conduit to more authoritative legal sources (although that may also be happening); rather, the argumentation in judicial decisions is itself influenced by the argumentation presented in Wikipedia.

Together, our findings highlight a gap that needs to be filled: judges need an easily accessible source of knowledge that is also authoritative. Policy-wise, this could be addressed by buttressing the reliability and review of Wikipedia content by including legal professionals as supervising editors to certify page quality, or by augmenting the content of authoritative but less-broad sources, and using those for the provision of legal information about particular jurisdictions. This latter approach has been successfully adopted in fields such as philosophy – for example, with the Stanford Encyclopedia of Philosophy. A possible model in the legal context is the US-based Oyez Project, which, in collaboration with Justia, offers free synopses of recently published decisions of the U.S. Supreme Court and of the Court of Appeal that, unlike those of Wikipedia, are authored or editorially supervised by legal professionals. Our experiment reveals that initiatives along these lines would be valuable and that they might help protect the foundational expectation of how legal decisions ought to be

made: that they should be based on carefully-considered expertise about what laws and precedents say, and never be at the mercy of internet ghost-writers.