University of Massachusetts Amherst

ScholarWorks@UMass Amherst

Citation Networks, Linguistics-Based Cues, and Logic-Based Approaches to Understanding What Persuades a Judge to Forsake Bias

James ben-Aaron
University of Massachusetts Amherst

CITATION NETWORKS, LINGUISTICS-BASED CUES, AND LOGIC-BASED APPROACHES TO UNDERSTANDING WHAT PERSUADES A JUDGE TO FORSAKE BIAS

A Dissertation Presented
by

JAMES BEN-AARON

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of:

Doctor of Philosophy

September 2016

Department of Political Science

CITATION NETWORKS, LINGUISTICS-BASED CUES, AND LOGIC-BASED
APPROACHES TO UNDERSTANDING WHAT PERSUADES
A JUDGE TO FORSAKE BIAS

A Dissertation Presented

by

JAMES BEN-AARON

Approved as to style and content by:

_____
John Brigham, Chairman

_____
Bruce A. Desmarais, Member

_____
James A. Kitts, Member

_____
Jane E. Fountain, Department Chair
Political Science

ACKNOWLEDGEMENTS

ABSTRACT

CITATION NETWORKS, LINGUISTICS-BASED CUES, AND LOGIC-BASED APPROACHES TO UNDERSTANDING WHAT PERSUADES A JUDGE TO FORSAKE BIAS

SEPTEMBER 2016

JAMES BEN-AARON, B.S., University of Massachusetts at Amherst
J.D. New England School of Law
Ph.D. University of Massachusetts at Amherst

Directed by: Professor John Brigham

Questions regarding what persuades jurists—and how legal decisionmakers actually do their work—are profound, motivating, and complex. The Public Law subfield has worked diligently to obtain empirically principled answers, but the gaps that remain provide an opportunity for this project to (hopefully) make a contribution. After discussing the nature of judicial decisionmaking, it is reasoned that rather than trying to understand jurists based upon the ways that their biases come into their work, a more effective approach is to isolate the occasions where they make unbiased decisions. In the interest of furthering the argument, a theoretical framework is offered that aims to isolate the major factors that will influence a jurist to "follow the law."

After a review of the state of the empirical study of judicial decisionmaking, three subprojects are presented, two of which tie directly to terms in the theoretical framework. The first is a novel effort to construct a network of case citations based upon specific language used in majority opinions. The second examines the propensity of Supreme Court Justices to cite to more "central" opinions when they are tending towards moderation in terms of ideology. The third subproject focuses on the often overlooked

difficulty that scholars have when attempting to state with definitive certainty what an "unbiased" legal opinion actually is.

These three subprojects are modest efforts to open new directions in research. Not all of the results that have been obtained fully square with the theoretical expectations that preceded them.

Note: Replication code and data for empirical analysis herein is available upon request.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

One of the primary focuses of the American Public Law subfield of Political Science has long been an effort to understand and account for the ways that preexisting individual ideology and biases affect the voting choices of appellate court judges and, to a lesser extent, the opinions that are subsequently authored in support of those votes. Much of that effort has focused on Supreme Court Justices and the occasions where bias has been observed to have had an impact on outcomes of cases (usually accounted for by examination of votes on the merits). Less consideration has been given to occasions where advocates successfully overcome ingrained ideologies and preexisting biases, and succeed in persuading appellate jurists to shift their positions based upon legal arguments. Woven throughout much of the previous work on the topic of judicial decision making is the presumption that members of the bench are typically not open to persuasion. This project starts from the position that it is the rarity of instances of judicial persuasion (with respect to significant matters) that makes them noteworthy, and therefore legitimate targets for in-depth study.

Real persuasion is a distinct and inexorable process whereby logic compels a realignment of one's worldview. Only if one has been moved either from one position on an issue to another, or from an agnostic position on an issue to some based position, has one been persuaded. That is not to say that one could not be re-persuaded at a later point in time back to a former position, or even to a third entirely new position, but some non-

1

trivial change must have occurred. That change must have been spurred by an argument (as opposed to caprice), and if one can find a way to refute the proffered line of reasoning without shifting to the next position, then the attempt at persuasion has failed (at least in the near term).

The first aim of this dissertation is to introduce a theoretical framework that advances our understanding of the elements that differentiate the legal arguments that are persuasive enough to compel judges to forsake their preexisting policy preferences and instead issue rulings that are firmly grounded in relevant statutes and established prior case law. Without this framework the ideas that are presented here are disembodied, and could be said to drift without purpose. Although the framework itself is theoretical, it is suggested that, with sufficient rigor, it would be possible to "plug in" values for each of the terms and to generate a probability for each specific new matter that comes before the Court. While scholarly work concerning judicial decisionmaking has considered a wide range of courts and administrative entities, much of the work done in Public Law has been focused on the U.S. Supreme Court. While an effort will be made throughout to specify which observations are of general applicability and which are specific to certain environments (federal courts, state courts, foreign courts, all courts), the reader will be informed by the specific label being used in a given instance—justice for the U.S. Supreme Court, judge for all U.S. Courts, jurists for the most general occasions that could go to any common law court). Toward that end it is argued that targeted research, especially research further exposing the role of citations in legal reasoning and regarding the ways that legal texts are interpreted, is necessary. The leverage obtained upon those areas would be the best first step towards solving the larger riddles that have confounded

us to this point (e.g., what species of legal arguments are most effective (Walton 2002)? how much variance is there in terms of what arguments persuade judges (Guthrie 2007)? are some individuals capable of suppressing their biases (Braman 2009)? can we develop methods of isolating those best suited to the work of judging (Knight 2009)?).

Answers in the behavioral area of scholarship can be notoriously difficult to nail down and the study of the judicial appellate reasoning process is no exception, as it is a black box with no obvious key. For matters that have been taken on by the Court, the votes are cast on the final outcomes are our first solid evidence of what happens inside the box, although it should be kept in mind that, for matters arriving under the discretionary jurisdiction of the Court, votes that determine which matters merit review (votes of certiorari) cannot be said to be dispositive of anything as we will generally be uncertain in the present time with respect to the motivation to accept or reject any given matter. While it is true that in the past some researchers such as Schwartz (1996) and Perry (1991) have been able to gain insights from documents that they have obtained from various justices, access to such internal writings is infrequently granted and does not appear to be a dependable resource upon which we can rely in any real sense.

Consider also that up or down votes on the outcome of a matter taken in isolation can serve only as a rough estimate of what has transpired. It could be argued that the queries asked and comments made during any oral arguments are the first evidence, but those interrogatives and accompanying statements are ephemeral. The justices who make them remain free to change their minds prior to the actual voting that follows. Nevertheless, oral arguments still have real relevance (Johnson, Wahlbeck and Spriggs 2006) and will be addressed in much detail in a later chapter herein. The actual opinion of

the Court (provided one is written) is the next potential source for insights. From the perspective of the academic researcher, it is probably best if an opinion putatively originates from the pen of an individual justice as the single unit of analysis is the simplest (although it is well established that even opinions that are attributed to a single justice are often the product of inputs from multiple other members of the Court (Schwartz 1996, but also scores of others)). In analyzing an opinion, the legal citations that are embedded within it are, collectively, the strongest signal for quantitative analysis (Cross, Spriggs, Johnson and Wahlbeck 2010), although the capacity to machine read and statistically analyze text is allowing statistical methods that are modeled on linguistic approaches to rapidly gain ground.

The convenience factor is not the justification for serious study of citation networks in this area. Rather it is the specificity of each call to a prior precedent—the reality that the work of the judge is to explain and expose the way that each earlier legal rationale dovetails with both the current fact pattern and the proper interpretation of the matter before the court—that direct our attention in this direction. Those signals collectively become our skein of thread marking the path through the labyrinth, as it were. Indeed, one of the distinctive features of our common law legal system is that it places a marked onus on appellate judges to present detailed justifications for their opinions. The degree to which these expositions are grounded in reason and are well buttressed with references to prior precedent naturally varies from jurist to jurist, but it is clearly a widespread professional norm to try to at least mount a colorable argument in support of a holding. Because the data that citations create can be gathered and analyzed

in a rigorous way to determine if their use was credible in each instance, they have real

value to those who hope to ascertain a deeper understanding of the legal system.

The analysis of citations, while central to the core of this project (Chapter 5), can

also be complemented by inquiry into other areas such as network centrality (Chapter 3),

and textual analysis (Chapter 6).

1.2 A General Theoretical Framework Concerning Persuasion

> "[With respect to understanding judicial decisionmaking] to the extent
> that the social science framework is found persuasive by the
> intellectual community at large, it will serve as both a guide and a
> constraint for how other scholars make . . . assumptions and employ .
> . . causal mechanisms."

—Professor Jack Knight[1]

A general theoretical framework of judicial persuasion is provided. Certain

caveats are necessary; to wit that this framework is presented strictly in an attempt to

provide theoretical clarity for the reader. Although effort has been made to cover all of

the main bases that combine to generate the observable outcomes, it is not possible to be

absolutely exhaustive—that is to cover all outcomes at all times for all justices and

judges. While it should be possible to provide empirical data and to "run" the equation,

the goal here—as with most theoretical frameworks—is not pure empirical certainty, but

rather to sharpen the focus, and to flesh out some of the nuances that are present; there is

value in this as a guidepost, but the project is not entirely staked to the absolute authority

of this representation.

---

[1] Knight 2009, at page 1556.

We open with the contention that the probability of a Justice who has been selected to author a front-of-book, majority opinion that does not reflect that individual's ideological bias can be estimated utilizing the following formula:

$$Pr(D_{Unbiased}|J_{IMRT} = (1 - M_S)(Risk)\left(\sum_{i=0}^{1}(C/P_{St.} + Cent._{CL} + C_{2nd.A} + C_{EJ})\right) \pm (Adv)$$

This theoretical framework ignores the effects of other judges if there is a panel. Such effects are real and are important to consider, but this model is the starting point; further work can take the peer effect into account.

The focus here is narrow, it is essentially to "turn the telescope around." Most work to date has essentially been based upon the assumption that individuals make rulings based upon bias and that opinions are subsequently efforts to mask that infidelity to the law. The circumstance under the microscope here is the exact opposite: This attack considers the (arguably less frequent) circumstance where a jurist becomes "cornered" by irresistible legal logic, and resigns themselves to voting and holding against their own inherent bias. The difference is significant and worthy of further attention: it is an instance where the exception to the rule sheds much light on the entire enterprise of judging. The terms within the theoretical framework are the following:

The Probability of an Unbiased Decision being issued

$$Pr(D_{Unbiased})$$

Given an Ideological Individual, and keeping in mind that not all individuals are equally ideological, variances must be considered. We may attempt to measure for Supreme Court Justices through each justice's Martin-Quinn, Segal-Cover (the lone ex ante measure of the three), and Epstein et al. scores (Martin and Quinn, 2002; Segal and Cover 1989; and Epstein et al., 2007) each of which attempts to account for their general

6

degree of ideological behavior (work on federal Circuit Court and state appellate court judges would need alternate scores to be developed).

$$J_I$$

Who is reasoning in a motivated fashion. The legitimacy of the common law courts is, to a large extent, grounded in the assumption that judges will act as neutral third-parties, and will dispassionately arbitrate in an unbiased fashion. Yet, if that is the case, and if judges are using only sound legal reasoning to reach their conclusions, how then do we explain the attitudinal forces that present so consistently when we "count the votes" and figure in the role of ideology? Borrowed from cognitive psychology by Segal and Spaeth (1996a), motivated reasoning is a "biased decision process where decision makers are predisposed to find authority consistent with their attitudes more convincing than cited authority that goes against their desired outcomes" (Braman 2009). It has been established that this gravitational "pull" is not necessarily something that an individual will be conscious is taking place (Kunda 1990). Braman (2009)" established experimentally that this tendency is exhibited by individuals with legal training placed in the role of the judge.

$$J_{MR}$$

Possessing an individual Judicial Temperament. Judicial temperament was defined by Jeffrey Rosen (2007) as the capacity to coexist peacefully with fellow judges, the ability to compromise, and the desire to keep institutional legitimacy paramount by setting aside individual ideological agendas. The term is used here as a catchall that captures the individual's proclivity towards following professional norms; no such measure is currently in place but a sound framework must take notice that not all

individuals will be equally invested in consciously attempting to make unbiased choices when on the bench.

$$J_T$$

The Matter Salience (the more salient a matter is to an individual, the more difficult it is for them to remain impartial).

$$M_S$$

Perceived Risks to Reputation posed by a biased opinion (the greater the risk that a biased opinion would compromise the reputation of the individual, and by extension the court, the greater the incentive to cleave to the law; it should be noted that only a minority of cases are likely to have any lasting, significant impact on an individual, or an institutional, reputation).

$$Risk$$

The Net Aggregate Strength of Cannons of Interpretations/Presumptions Aligned with the Individual's Ideologically Favored Outcome for Pertinent Constitutional Elements, Statutes, Regulations, and/or Ordinances + Net Aggregate Centrality of Relevant Case Law Aligned with Individual's Ideologically Favored Outcome (Positive Case Law) + Net Aggregate Strength of Cited Secondary Authority Supporting Individual's Ideologically Favored Outcome (Positive Legal Commentary) + Net Strength of Cited Extrajudicial Sources Supporting Individual's Ideologically Favored Outcome (Expert Testimony, Demographic Data, National Academies Research, etc.).

$$C/P_{St.} + Cent._{CL} + C_{2nd.A} + C_{EJ}$$

Net Effect of Advocates on Outcome (likely to be small on most occasions and also has the potential to move the outcome to be either more or less biased).

Each of these terms is considered in more depth in the next section.

1.3 Some Uncontroversial Assumptions Concerning the Justices

For the purposes of this theoretical framework, several assumptions are made regarding justices. First, it is assumed that each justice will have at least some ideological positions toward which they will tend to gravitate. Human beings, of course, exist on a continuum, and having political biases would appear to be about as normal as having been born with two lungs (recall Aristotle's incantation of man as "zoon politikon."). Stating that a justice will have some viewpoints regarding optimal legal outcome that are strongly influenced by their political beliefs is not the same thing as saying that those beliefs will inexorably dictate the holdings upon which that justice will finally settle. Rather, stating that those political biases exist is acknowledging that they are within the range of phenomena for which this model is designed to account. Second, while some rare individuals likely have the capacity to engage in the act of judging in a wholly detached and dispassionate fashion, the greater majority will engage in motivated reasoning, a biased cognitive process whereby the decision maker is predisposed to favor and to find more convincing sources of authority that are aligned with their own attitudes. In turn, motivated reasoning will also cause the individual to discount sources of authority that are at odds with their beliefs even when they are not fully aware that they are doing so (Braman 2009). Third, each justice will possess a "Judicial Temperament" somewhat akin to Lincoln's "better angels of our nature." This catch-all encompasses that individual's predisposition towards respect for the doctrine of stare decisis; their manifest interest in behaving in a way that they calculate will best maintain the larger public's

faith in both the integrity of the legal system at large, and the inherent fairness of the Court, and their deference to the doctrine of Separation of Powers that should to some degree override their inclination to legislate from the bench.

Note that these assumptions do not exclude the possibility of a justice who actively (albeit likely tacitly) attempts to exercise their political will through their votes and their authored opinions, who deliberately shuns valid precedent that disagrees with them, or who an impartial observer would conclude possesses a decidedly non-judicial temperament. Instead, these assumptions merely start from the proposition that most Supreme Court Justices will sincerely try to do their work well, and that they will aim to be a credit to the Court itself. Not all will succeed to the same degree in meeting those lofty goals, but the initial position is to give the benefit of the doubt at the outset and to not presume any inclination towards misbehavior (i.e., the willful expression of political bias) ab initio.

1.4 Matter Salience

The probability of a biased opinion emanating from a justice is contingent in large part upon how salient the issues raised by that particular matter are to that individual specifically (Unah and Hancock 2006; for salience to the Court at large see Baird 2004). That variable is captured by the Matter Salience term and resides on the right-hand side of the model from where it serves as something of a gatekeeper function. From a theoretical perspective, it is expected that the higher the salience of a given matter to a given judge, the higher the probability that the opinion that issues regarding that matter will exhibit a pronounced political bias. Such a response would generally be expected to be aroused by "hot button" social issues such as reproductive autonomy and the death

penalty, where race and gender issues are on the proverbial table, and could also be observed in the area of substantive economic due process. Patent law, probate law, and the notably obscure region of property law referred to as "future interests" all likely fall into this "too anodyne" category for many. It is entirely understandable that such disputes push even a highly dedicated, top-flight judge's interest down to a low level. Simultaneously, making such a blanket assumption could well be an error given, for example, the reflexive aversion to any form of taxation that grips a segment of extremely ideologically motivated individuals, an area that many would naively suppose was "too dry" to inspire much fervor.

At the same time, for each individual there are likely some legal matters that will be perceived as being are exceedingly technical in detail and dry in nature as they orbit around obscure and even esoteric concerns. Overall, as a noticeable portion of the matters that make it onto the docket of the United States Supreme Court will tend to elicit strong bias responses from the public in general, it would be sensible to presume that similar emotional responses will be elicited from the justices who must rule on them (Perry, 1991).

The model presented assumes that most judges will aim (and to some extent succeed) to remain detached towards their cases. It is further assumed that peculiar circumstances are required to cause a judge to become more invested in a pending case. Although those occasions where judges do find high levels of salience in the matter before them are the same ones that are most intriguing from the perspective of judicial scholars, Danelski (1965) used content analysis of statements made by justices prior to appointment to the Court. An alternate approach would be to topic model articles and

speeches delivered by individual justices to try to ascertain what areas appear to be salient enough to them to write and comment upon. Where a justice places themselves on the spectrum in terms of expressing ideology from the bench is clearly where an individual has the greatest opportunity to build up (or to diminish) their own reputation. By extension, each justice also contributes to (or potentially cause harm to) the public reputation of the Court itself.

In the present study, it is necessary that the focus will often be upon opinions authored regarding matters where the Salience term is high and those sections will be the default throughout. Other sections and sub-sections will be concerned either with random samples of opinions, or with the entire corpus of opinions from a specific interval.

1.5 Perceived Risk to Reputation

The second right-hand-side term, Risk to Reputation, is similar to the aforementioned Salience term in some important respects. At the outset it should be noted that because of the salient nature of most matters that make it to the Supreme Court's docket, likely every decision, whichever way it is decided, likely damages the Court's reputation in someone's eyes, and simultaneously enhances it in someone else's. While that situation is a given, clearly there are certain topic areas that inflame the passions of many, and such matters have the potential to cause significant swings in the Court's approval rating (which can be interpreted as a post hoc proxy for the level of risk that was latent in the topic area). The idea that Justices actively seek the approval of their audiences does not square well with the leading models of judicial behavior (Baum, 2006). Although the effects that can be caused by judges and justices strategically

considering the limits of what the "traffic will bear" is not the central focus of this work, though those effects are thought to be significant and this model does acknowledge them.

Most matters that come before a court, if handled with reasonable professionalism, will not significantly alter that jurist's reputation. Even at the Supreme Court level it is a really only a handful of opinions that, were one of them looked at in isolation, it would appear to have had a lasting negative effect on the reputation of the main author (consider Plessy v. Ferguson and Dred Scott v. Sandford as two prime examples). Further, given the spectrum of political opinions in the populace, decisions that lessen the opinion of the Court with some will, often boost it with others (this crosscutting effect was explored by Kritzer (2001) with regard to the overall impact of Bush v. Gore that sharply skewed by political party, but did not have a significant net effect).

It should be uncontroversial to posit that a Justice's reputation would be unlikely to suffer any profound or lasting damage due to fallout from an opinion that assiduously hewed to the black letter law (at least we should hope that that would be the case). It also seems reasonable to assert that serious, lasting damage to a Justice's reputation (or to a group of Justices voting in a block) would be much more likely to accrue as the result of an opinion that was transparently driven by political bias (although there other types of conventional political scandals could conceivably damage a justice's reputation if one were bought into the public's collective consciousness—graft, general criminal misconduct, marital infidelity and the like—one would need to go back to Justice's Abe Fortas' resignation under a cloud of ethics troubles to isolate a truly scandal-tainted end to a career on the Court).

Once that step has been taken, the other terms become critical. Because judges are concerned with both their own personal reputations and the reputation that the court system itself maintains (Baum 2006), the risks associated with issuing an overtly biased decision (in the extreme one that is indefensible—a clear error—one that runs afoul of H.L.A. Hart's "rules of recognition" as will be covered in the next chapter) will often preclude such a gross misstep. In any event, the process of legal education, the professional work that invariably precedes an individual making it to the bench, and the culture in which judges reside, all contribute to the socialization of judges to be inclined to rule in step with the law (Fleischer, 2008).

Even matters that do not garner national (or even local media attention) could be highly relevant within a judge's immediate professional cohort (obviously including other members of an appellate court who are casting votes on the same matter (Sunstein, 2003)) and larger professional cohort (other judges not familiar to that particular judge, but who are aware of the case in question, courthouse personnel, and members of the bar). Also the standing of that judge's courtroom in the opinions of members of other branches of government as well as with the individuals in that judge's immediate and extended social network are also implicated as having influence on that judge's reputation (Baum, 2006).

The great difficulty with the Risk to Reputation term is that is has yet to be quantified. Nevertheless, leaving it out entirely would corrupt the external validity of the model so it must be included, even if just as a placeholder for potential future work.

1.6 Advocate Effects

The fourth and final term on the right-hand side accounts for the impact that advocates have on the outcomes—as Justice Scalia has remarked, "a judge must remain open to persuasion by counsel" (Scalia and Garner 2012 at page xxx). Although that direct impact by advocates is not the main focus of the present study, extensive work has been done in this area with robust results indicating that advocates do have the potential to influence the outcomes of appellate matters (Collins, Corley, and Hamner 2015; Johnson, Wahlbeck and Spriggs 2006). Because of those findings the Advocate Effect term is added to the end of the framework to cover those effects on the eventual dispositions of judges as persuaded, dissuaded or influenced (mainly by briefs and oral arguments).

1.7 Citations Are the Nexus

This model provides a point of departure for this project. Refining it sufficiently to establish its veracity would likely be a life's work, but the more immediate goals are to make empirical progress towards establishing the critical nature of citations in support of persuasive legal arguments, the propensity of justices' opinions to become less central as they drift towards the fringes of the ideological spectrum, and the tendency of more complex language used to betray less adherence to established legal reasoning. Admittedly, this theoretical framework for judicial decisionmaking is greatly simplified when compared to the way that the real world is seen to work. Experience with complex models in political science, and other social science fields (e.g. economics and sociology), strongly suggests that efforts to design models that truly approach reality can create daunting requirements for data collection, and can still produce unreliable results

(how many predicted the last economic crisis? who can accurately predict when and where civil unrest will next strike?).

The aim is not to establish that this one single theoretical framework of judicial behavior is infallible. Rather, it is hoped that the explanation of judicial behavior proffered here helps to make clear the interplay among the various elements that have been explored by other Public Law scholars and, in turn, present a direction for future work that can generate the data that will then, in turn, allow the further adaptation and tuning of the theoretical framework that has been introduced.

1.8 Legal Citations

Charles Darwin commented (roughly) that efforts to understand the nature of first principles without first understanding human evolution were akin to "puzzling at astronomy without mechanics" (Boyd 1985). Roughly the same could be said regarding efforts to understand the judicial decision making process without first examining the role that legal citations play in the process.

For the purpose of establishing a metaphor, we can imagine a wall with thousands of light bulbs arrayed across its face, each one representing a prior legal opinion. A justice sits beyond our access on the opposite side of the wall in a sealed room with thousands of buttons arrayed on a wall, each of which is marked with a citation, and each of which lights a corresponding light bulb on our side.

Within that sealed room the justice considers the arguments presented in a case and forms impressions of what the outcome should be; impressions that can be further developed by interactions with fellow justices who are also beyond our access. If selected as the author of the majority opinion, our particular fictional justice will ostensibly start

out with the entire universe of prior opinions available to them as authority (each prior opinion, again, having a corresponding light bulb on our side of the aforementioned wall). It is also reasonable to suspect that the justice also has prior knowledge of a significant number of previously decided matters that they have "tagged" for citation at the first appropriate opportunity so as to expand their influence as deemed appropriate.

After some study, those prior opinions that potentially have specific application to the matter at hand will be isolated and, from that larger group, a smaller subset of prior opinions will be selected—those that are putatively sufficiently relevant to support the final position (or, in the alternative, those that are included for criticism or reversal). Once the Court's opinion is published, those buttons corresponding to the selected prior citations are "pressed" and the audience on the other side of the wall sees a corresponding bulb light up for each selected citation. It is from those constellations of lights that each opinion illuminates that many Public Law scholars have been attempting to discern what transpired within that justice's mind—the mind hidden within the black box on the other side of the wall.

Additional significant clues can assist in divining what transpired inside the black box; after all, the justices frequently write extended opinions explaining why and how they reached their conclusions. Depending upon how much stock we wish to place in the reliability of those writings, making further progress in understanding what transpired in a justice's mind is possible. Justices also occasionally give public talks, author law review articles, or write entire books that potentially shed some light on specific opinions, upon their reasoning in a given area of the law or just with respect to general legal philosophy. Nevertheless, for empirical analysis, it is the citations to prior opinions

that remain the obvious place to commence quantitative investigation. They remain the best available resource and, although it has been often pointed out that the area of legal citations in court opinions has been under studied (Cross, Spriggs, Johnson and Wahlbeck 2010), that observation has now been raised in enough articles that promptly went on to take the approach that that particular claim has finally overstayed its welcome.

Citations matter greatly in the context of legal persuasion because most (perhaps all?) justices utilize analogical reasoning "whereby they cite cases due to those case's legal relevance and authority . . . [c]ase citations thus represent a latent judgment by justices regarding the relationship of cited cases to the legal and factual circumstances in the cases they are deciding" (Cross, Spriggs, Johnson and Wahlbeck 2010). The literature that addresses the meaning of opinion citations can be divided into three categories: works that view them primarily as the basis for legal conclusions as dictated by stare decisis; works taking the position that opinion citations are merely utilized as cover for legal decisions that are in fact motivated by individual biases; and, a somewhat nuanced middle position that allows for the influence of both controlling precedents and ideological biases, each simultaneously contributing to the final outcome.

1.9 Citations and the Legal Model

The traditional legal model of judicial behavior—variously called mechanical jurisprudence or precedentialism—holds that what has been determined in prior precedents drives later decisions (this model obviously makes an allowance for those exceptional circumstances where the Court expressly abandons prior precedent and announces new law) (Spaeth and Segal 1999). It would also be within the ambit of this explanation to claim that the ways that statutes and regulations are to be properly read

and interpreted is through the reading of prior case law within which the recognized principles of interpretation and canons reside (Scalia and Garner 2012). Higher courts instruct lower courts with regard to how they (the lower courts) should proceed. The process, as observed in print, has each level of the judicial system hurling precedent back and forth and positing the question: "What did we do to get here in the first place?" This is more of a "monkey see/monkey do" effect than a "chicken-and-the-egg" quandary. If a lower court is found to have committed what is viewed as an obvious error, that lower court would merely be reversed and the decision itself remanded back with instruction to comply with direct instructions. Appellate courts are more likely to engage with matters that are not well settled, or easily resolved with minimal reliance on obvious principles of statutory or case law interpretation.

Although the Supreme Court is not subject to vertical stare decisis and is quite capable of reversing its own prior decisions, it has never summarily rejected citation to prior opinions, nor does it appear likely to do so at any point in the future. The Court's utilization of prior opinions has led to the development of theories regarding Supreme Court's ongoing adherence to stare decisis.

In an application of Occam's razor, some of those looking at judicial behavior through the lens of economic/rational choice theory have noted that reliance on prior legal authority effectively reduces the amount of effort that needs to be put into the process of authoring opinions (Epstein, Landes and Posner 2013). An equally mundane possibility is that the people who become justices are simply habituated to the norm of utilizing citations, and therefore do so in a rote fashion (Fleisher, 2008). Yet another possible explanation for the use of citations that fits into the rubric of the traditional legal

model is that there exists a quid pro quo of sorts and that justices respect the work of other justices with the expectation that other justices will, in turn, respect theirs (Landes and Posner 2013; 1976).

1.10 Citations and the Attitudinal Model

If citations are being artfully used to provide plausible deniability that a justice is merely making bias-based policy choices from the bench, it is possible that the process could be revealed. Segal and Spaeth (1996) attempted to demonstrate this possible explanation by tracing the votes of Supreme Court justices who had been in the minority and were later faced with the same legal issue. Their results appeared to show that few justices fell into line with the prior precedent that they had not favored (but theoretically should have accepted as precedent). Although it attained wide acceptance, this model has also been criticized by numerous others for shortcomings such as coding errors and failing to count summary dispositions (Cross, Spriggs, Johnson, and Wahlbeck 2010), and for its conception of the role of precedent and stare decisis in the decisionmaking process (Brenner and Stier 1996; Songer and Lindquist 1996).

1.11 The Middle Ground: Citations as Guidance

Less dogmatic observers have come to believe that citations do matter to justices, some for the legal positions that they establish, and others for the way that they slant towards an individual justice's political biases. Thus, under certain circumstances justices can tilt outcomes in favor of their own political beliefs, but those opportunities are constrained by the legal realities that are established and sufficiently well moored that they (the legal realities) cannot be moved. Of course appellate courts are not mandated to engage with legal matters that are obviously settled under well established law.

Taking the middle road gives some flexibility to the process, but it still fails to get us away from the "vote counting" rut. As was discussed in detail in the introduction, the observation has been made that researchers have not done enough work that treats the law itself as a dependent variable, as Hansford and Spriggs (2006 at page 4) comment:

> "Researchers working in [the attitudinal] tradition generally argue that the language in Court opinions constitutes the post hoc justifications for the outcome preferred by the justices. Thus, they recommend that scholars study "what justices do [i.e., their votes]" rather than "what they say [i.e., their opinions]". Although attitudinalists recognize that the "opinion of the Court . . . constitutes the core of the Court's policy making process", there continues to be an overwhelming tendency to study the justice's votes."[2]

That high level of attention to the ways judges vote, and the weight given to the premise that the votes themselves are the nearly exclusive result of each individual's ideological beliefs, have naturally produced reams of research on the ideological nature of individual votes in appellate cases and on the disposition of cases, as opposed to analysis of the announcement of legal policy that resides within majority opinions.

Outside of legal and bias based explanations of judicial behavior, competing explanations are somewhat scant. Economic models (as already introduced), and the strategic model (Maltzman, Spriggs, and Wahlbeck 2000), are probably the next most often cited (with the latter being arguably compatible with the attitudinal model). The more obscure pragmatic model (that a decision is made not based upon the legal language that is arguably relevant, but upon the anticipated effects that the decision will have), and the phenomenological model (wherein the focus is on the psychology that the individual

---

[2] Internal citations to Spaeth 1965, and Segal and Spaeth 2002, omitted.

doing the judging actually experiences is the focus (Rubin and Feeley 1996)), are more "off in the weeds."

Churning through legal opinions and converting the often abstract and dense expressions of judges into usable data can be technically daunting work. That complexity, no doubt, contributed (prior to the recent advent of machine reading of text coupled with computer algorithms that can cipher through large corpuses and extract statistical meaning) to the paucity of research in which judges actually explaining which arguments did, or did not, successfully persuade them of the truth of various points of law is featured as the dependent variable. The tide has turned somewhat in recent times with Hansford and Spriggs' aforementioned work on the interpretation of precedent, and Maltzman, Spriggs and Wahlbeck's aforementioned work on strategic interaction and the opinion-writing process coming to the fore.

1.12 Research Design

1.12.1 Chapter 3

Corresponding to the Cent.$_{CL}$ term in the theoretical framework (which attached the Net Aggregate Centrality of Relevant Case Law Aligned with Individual's Ideologically Favored Outcome (Positive Case Law)) the first research chapter seeks to approach and decipher the proper interpretation of opinion centrality, especially when that quantity shifts over time for an individual justice. Although the differences in opinion citation rates among U.S. Supreme Court Justices are marked enough to imply that substantive differences are in play, several possible explanations could account for the observed variation (Cross, Spriggs, Johnson and Wahlbeck 2010). Epstein, Landes and Posner (2013) included within those explanations differences in the work ethics of

the justices, individual variation with respect to commitment to stare decisis, and the belief of an opinion's author that deeply cited opinions will exhibit greater relevance (i.e., be cited more often over time) and greater legal vitality (i.e., exhibit greater authority over time by being cited more prominently and in more important opinions).

In this study the network of Supreme Court citations (Fowler and Jeon 2008) is ascertained for the selected time frame of 1946 to 2002. For each justice, the opinions that they authored are isolated and the centrality scores of those opinions are compiled. Next the ideology score of the authoring justice of each opinion is attached to enable analysis of correlation between the two values.

One of the strengths of network analysis is that it can expose information contained in indirect connections, and thereby enables investigators to make inferences regarding the latent space that is embodied by the whole of the network. Starting from the proposition that the most central opinions must (by the operation of the algorithm that creates the network) be the highest in overall legal relevance and vitality, a sufficient sample of opinions by each justice should contribute to understanding some of the motivations that underlie observed variations in opinion citation frequency and quality. While we do have overall centrality values for individual justices, these raw numbers are not overly informative in and of themselves.

Because many justices drift over their careers in terms of their ideology, a single mean centrality score could be hiding a story of a significant drift over time in terms of that justice's preference for selecting past opinions to which to cite. Thus, the centrality scores of a justice must be calculated for shorter intervals to expose any evolving changes in propensity for citing (or for not citing) to deeply embedded opinions. Consider the

following career ranges (by term) of Martin-Quinn Scores of individual justices over their

careers:

- Thomas: 2.69 to 3.87

- Stevens: 0.03 to -3.21

- Brennan: -0.62 to -3.74

- Rehnquist: 4.43 to 1.22

- Douglas: -1.43 to -6.46

- Marshall: -0.9 to -4.49

- Blackmun: 1.9 to -1.86

Theories regarding the utility of the centrality measure can be tested against these

migrating values.

It is argued that grounded, well formed legal arguments will be more heavily

reliant upon opinions that exhibit a combination of high relevance and high vitality.

Opinions that score highly in those areas are exactly the ones that will themselves be

expected to become more popular, and therefore more central, within the network as it

evolves over time (in this context note that only citations to other United States Supreme

Court opinions are considered). This is explainable as a form of homophily, with

"stronger" opinions sticking together. At the same time, less well formed legal arguments

will tend to rely more heavily upon legal authority that is more "fringe" and less widely

respected (i.e., those that become less central). These tendencies to rely more or less on

highly central opinions will be observable for individual justices.

It is then hypothesized that analyzing the opinion citation networks of individual

justices over time with regard to standard measures of centrality and prestige will reveal

that those who are paying closer attention to stare decisis will be those who are inherently

less ideological within that given time segment that is being considered, and who will tend to cite more often to the most central opinions.

If we focus on the "less ideological" as the causal variable, we would expect justices moving towards the 0.00 Martin-Quinn center point (from either side) to cite to more central opinions (and for their own opinions to become central in the network over time). If we focus on "citations to more central opinions" as the dependent variable we would expect central opinions to have emanated from justices who were at or near the 0.00 Martin-Quinn center point.

As many justices have migrated across the ideological spectrum over their careers (Martin and Quinn 2002), an internally valid test measuring drift towards or away from citation to more central opinions is anticipated to be a potentially significant result—provided that the observed directions of the drift in the majority of the cases matched the theoretical expectation that the more ideological a stance a given justice takes, the less inclined they will be to cite to central opinions.

The potential payoff from the assembly and analysis of citation networks of opinions authored by individual justices would be the illumination of a significant relationship between the level of ideology of a justice and the propensity for following prior central case law.

1.12.2 Chapters 4 and 5

[Chapter 4 is a brief exposition upon the capacity that Public Law scholars have to reasonably rely upon written statements made by Supreme Court Justices to at least somewhat accurately correspond to their actual, truthful thoughts. It is stressed that the work that follows in Chapter 5 is dependent upon specific written statements having

veracity, but is would nonetheless be objectively advantageous for the project if there could be some stock placed in the notion that justices are being preponderantly faithful when they relate that a legal argument has persuaded them on a point of law.]

The focus of the project now shifts to the theoretical framework's C/P_{St} term (the Net Aggregate Strength of Cannons of Interpretations/Presumptions Aligned with the Individual's Ideologically Favored Outcome for Pertinent Constitutional Elements) and an apparently novel approach to the construction of a citation network is deployed.

Fowler and Jeon (2008) stated that "Each judicial citation in an opinion is essentially a latent judgment about the case cited. When justices write opinions, they spend time researching law and selecting precedents to support their arguments. Thus, the citation behavior of the Court's provides information about which precedents serve important roles in the development of American law." They further asserted that the quantity and quality of judicial citations in Supreme Court majority opinions can be analyzed to help us understand how legal policies are formulated by the judiciary.

In this chapter a network of Supreme Court opinions that acknowledge persuasive argument is introduced and analyzed with all citing opinions added. Two other stratified sets of opinions are also generated with citing opinions added. The first matches the reference set by the United States Reporter Volume Number (thus if an opinion with the "we are persuaded" language was authored by Justice O'Connor in Volume 458 of the United States Reporter the matching node in the "by volume" set would have an opinion that was randomly selected from among the other opinions included in the same volume). The second matches the reference set both by the United States Reporter Volume Number and the justice authoring the opinion (thus if an opinion with the "we are

persuaded" language was authored by Justice Stevens in Volume 443 of the United States Reporter, the matching node in the "by volume & justice" set would have an opinion from Justice Stevens that also was randomly selected from among the other opinions that he authored that were included in the same volume).

The theoretical expectation is that in comparison to the entire population of all opinions decided by a given court, the density of ties among the opinions (the incidence of opinion-to-opinion citations as divided by the maximum number that is theoretically possible) and centrality (how important or well connected the vertices of the network are as determined by one or more related measurement approaches) should be higher among the network of opinions that explicitly mention judicial persuasion. Such results would be explicable if the authors of opinions that explicitly disclose persuasion were motivated to rely upon prior opinions that also disclosed persuasion (the primary disclosure set), or upon opinions that are tied to the primary disclosure set through direct citation. That reliance would stem from the presence of common structural expositions of the legal arguments that were presented, supported, considered, and ultimately deemed to be winning, in the prior opinions and their progeny.

Thus the hypothesis that is tested is the expectation is that edges (citations that are made among opinions) that are observed between two given vertexes (the opinions themselves), are more likely to form among a set composed of opinions that explicitly announce persuasion of the court on some legal point and the subsequent opinions that cite those persuasion opinions, than among the entire population of opinions. The formation of these edges would cause earlier opinions to accumulate a higher "indegree" which is nothing more than more accumulated "hits" from later opinions that cite to them

as precedent. Non-parametric testing of covariate effects are next utilized in order to expose the vertex level attributes that account for the structures observed in various legal opinion citation networks. Although not the focus of this project, it is further hypothesized that altering the selection process to harvest opinions that use less precise language (e.g., "the Court finds this line of reasoning convincing") could be added to develop a more complete understanding of the nature of the arguments that do convince the Court of legal arguments. In the alternative, situations where the Court declares the exact opposite, "We are not persuaded" would also be worth examination.

To advance this portion of the project a primary collection of United States Supreme Court opinions (1946-2008) that explicitly announces that the Court was persuaded of an argument's validity (hereinafter: The WAP data set) is aggregated utilizing a free-text search on the Westlaw legal research database. Further opinion-by-opinion review was then done to establish that the language located was the majority of the Court itself speaking (not, say, a direct quote from some other source), and that the term was not being used within a counterfactual argument. Each of these WAP opinions was then forward cite checked in the Westlaw database, and each subsequent citing opinion that returned was then entered into the edge list to form the full network.

The primary opinion issue—"primary" in this instance means that the main legal issue with which the given opinion is concerned—is the vertex level attribute with which statistical analysis (utilizing Quadratic Assignment Procedure (QAP), a nonstandard-error-based test of coefficient significance (Dreiling and Darves 2011)) was focused on as it is the most likely to determine the likelihood of an edge forming in the network. Several other control variables are also collected to further develop the model. The term

in which an opinion is decided has an observable effect on the likelihood of edge

formation as there is a significant positive relationship between the age of an opinion and

the number of citations that it has received. This is exactly what we would expect as

generally opinions will continue to accumulate citations over time (with non-negative

citations being more frequent overall). It should be noted that this rate of accumulation

tends to be somewhat steady for a period of terms and then eventually the rate will

decrease (barring some odd circumstance with an opinion becoming suddenly important

after a long dormancy).

The page length of the decision was added in because it is obvious that longer

opinions strongly tend to attract more subsequent citations (this particular observation is a

simple extension of the inevitability of any discrete legal pronouncement by the Supreme

Court all but inevitably drawing a citation from some future court later in time, and each

additional page in an opinion will inevitably draw out more discrete legal

pronouncements). In addition, the Majority Opinion Author, simply the justice who is

credited with authoring the opinion, is included. In some instances an opinion is

presented as being Per Curiam (indicating that the reasoning of all of the justices who

comprised the majority is presented as a unit). The author of an opinion is considered

relevant as some opinion writers could have a propensity to cite to their own prior

opinions more often. Per Curiam and unsigned opinions were coded as such. Whether an

opinion was "good" law (able to serve as precedent without qualification) or "bad" law

(having been overruled in at least some part) at the time that the sample was collected

was also added into the node level data set. Opinions that have been subsequently

overturned will likely fade into obscurity whereas opinions with positive histories that remain "good" law are more likely to continue being cited.

After the data was extensively cleaned, reshaped, and converted into a network object (a data frame composed of the network's nodes and the edges that are present that has been formed into a matrix that the **R** statistical software environment can recognize), the visualization of the network through extensive plotting was undertaken to assist in understanding the time dynamics of the citation network over time. Next, network density was calculated along with network indegee and outdegree measurement. WAP network indegree was then compared to the stratified sample network that matches only by term, and to the stratified sample network that is matched both by term and by opinion author, and comparisons are made via a Welch's t-test. Finally, the network was modeled to obtain estimates of the distribution of the network coefficients using the aforementioned QAP.

A few words should be said about the additional groups of opinions that are generated and analyzed. This second stage of data collection by necessity mirrored the first; however the selection of the opinions for the stratified comparison set aligned opinion-by-opinion with respect to the volume number of the United States Reporter, and also with respect and the opinion author, with the opinion selected within the Reporter based on a random number generated using R. So, for example, once an opinion in volume number 400 was observed to have the "we are persuaded" language, the by-volume stratified set was allowed to randomly match to any other non-WAP opinion in that volume, and the by-author stratified set was randomly matched another opinion not just by within the same volume, by also by the same author (in the rare instances that a

WAP opinion did not have an author match, the original WAP opinion was removed from the by-author analysis.

For each set, WAP, by-volume, and by-author, an edgelist and a node level data set must be complied integrating the Westlaw data (the status of an opinion as "good" law as of April 2013) and several variables taken from the Supreme Court Data Base (majority opinion author, term, and primary issue).

This laborious process was expected to contribute to the project in several ways. First, while there have been prior efforts to explore the citation corpus of the Supreme Court (Bommarito, Katz, and Zelner 2009), I am aware of no other publicly presented data set of Court opinions that is the product of an effort to explore the Court's process with respect to legal reasoning through the isolation of specific terms. This "move" of generating Court citation networks contingent upon specific language within the text of opinions has some further potential, as it remains plausible that one single study is not a sufficient basis upon which to judge the utility of this approach.

Beyond that premise, it would be ideal if it could be established that there is a higher density of ties among the WAP opinions than among the random opinion set and, more specifically, if the indegree of WAP opinions was significantly higher than of stratified opinion sets that do not contain the specific language. The documentation for the **sna R Package** acknowledges that "interpretation of quantiles for single coefficients can be complex in the presence of multicollinearity or third variable effects." That warning, especially if combined with the low Adjusted R-squared result from that computation, makes the interpretation of these data a less than certain undertaking with respect to the importance of opinion author, and opinion page length. At some point, the

random element of the selection process would be pushed aside by the deliberate limiting of the pool of available opinions with which to match.

Significant measurements for variables such as Opinion Writer Homophily, for example, would allow for the guarded claim that authorship of an opinion is not a necessary element for a model seeking to explain the observed effects. Of the other positive results for measurements of homophily effects such as Issue would accord well with all established understanding of court opinion citation networks. The number of pages in an opinion should also stand out as clearly significant as should the status of an opinion as "good law."

Of some concern is the reality that different justices could have favored alternate phrasing, and the search in this initial phase of the project was limited to a specific three-gram, that is, only the three particular words in sequence "we are persuaded." The Court has expressed approval of lines of reasoning through words other than just the three-gram that this pilot study was built upon. The project could grow to cover more terrain by adopting greater flexibility with respect to the language it accepts as signifying instances of persuasion. Additional samples of language that are roughly synonymous with the "we are persuaded" phrase such as "we find convincing"; "is compelling"; "takes a more credible position"; "the correct interpretation is"; "logic requires that we"; "the most precise"; and, "is more cogent"; could be identified through testing of possible wording in the Westlaw database. Once exposed, the analysis of those formations could be used to augment the further iterations of this study.

1.12.3 Chapter 6

The relevance of any theoretical framework will be undermined entirely if it cannot account for, or even fails to make observation of, a relativistic effect. With respect to the legal system, observers create a potentially disruptive reality because there is no clear consensus with respect to what constitutes actually "following the law." Conservatives will argue that conservative justices are "following the law" and progressives will argue with equal force that progressive justices reaching the exact opposite conclusion on the same matter are also "following the law." A theoretical framework that aims to determine the probability of a justice returning an "unbiased" opinion must account for the wide open space around the interpretation of what does, and what does not, constitute an unbiased legal opinion.

When considering the possibility that many of the decisions rendered by courts are motivated at least in part by political biases, Public Law scholars have long assigned considerable weight to the individual votes of judges in general, and most often those of the justices on United States Supreme Court (Pritchett 1948). More recent work has aimed to address the topic through the analysis of citations (Cross, Spriggs, Johnson and Wahlbeck 2010; Fowler and Jeon 2008). As noted supra, justices have been able to consistently dodge the bullet because there is "no neutral arbiter for the evaluation of adherence to stare decisis" (Cross Spriggs, Johnson and Wahlbeck 2010 at page 513), so each side has been free to claim that it is the one that has been consistently faithful to controlling precedents.

The various approaches that have been previously discussed herein (the formal legal, the attitudinal, and the mixed models) are uniform in that they have all attacked the

prominent questions with a focus on what takes place entirely (voting), or largely (the selection of citations), prior to the commencement of the opinion writing process. It is posited that a potentially gainful contribution towards solving the puzzle of how to accurately identify instances of political bias manifesting itself in court opinions would be afforded by the application of the tools used by linguists to detect deception in written text (Zhou 2004.; Newman, Pennebaker, Berry and Richards 2003). Likewise, if the similar linguistic tools employed to analyze spoken words to detect deception are used to interrogate sections of oral arguments wherein the justices utilize slippery logic in support of politically biased reasoning, those instances will also be quantifiably different from instances where the speaker is able to buttress their arguments with legally sound and unbiased reasoning. As studies have identified the capacity of individuals to somewhat accurately discern differences in veridicality (Feldman, Forrest and Happ 2010), another approach would be to utilize artificial intelligence to analyze recordings of the facial expressions of judges making statements during oral arguments, to software trained to detect deceptive expressions with the results then being matched to the levels of bias that are measured in the subsequent opinions (Tsiamyrtzis2007). For a study based on the capture facial expressions it would be necessary to use lower court judges because the Supreme Court does not allow cameras.

It is well established that engaging in deceptive activity often requires the maintenance of a greater cognitive burden (Zhou 2004). While authoring an opinion that is politically biased will most likely not entail telling outright lies per se, if one is making efforts to present the biased opinion as legally correct and unbiased, that process will necessarily demand commission of one or more acts of subterfuge with regards to the

legal reasoning that is presented. Psychology scholars have found that telling stories that are made up from falsehoods are quantitatively different from stories based on actually events (e.g., Johnson and Raye 1981; Vrij, Fisher, Mann and Leal 2006; Cf. Undeutsch 1989), thus it should follow that the writing in opinions and oral statements that are driven by political bias, will also be quantitatively different from the writing in opinions that are driven more purely by the unbiased reading of the controlling authorities. It is also expected that even facial expressions of judges making statements in support of politically biased positions could be susceptible to detection by subjects in controlled experiments.

It is hypothesized that both in court opinions and in oral arguments, the intellectual gymnastics required to present politically biased arguments that plausibly pass as unbiased legal positions will create a higher cognitive burden for the justice and the required increase in cognitive effort will be telegraphed in ways that can be detected through the application of extant linguistic tools that have been developed to detect untruthfulness as well as through observational studies designed to detect stress and lack of fidelity in speakers.

The detection of variations in text that was produced by a person under an increased cognitive burden by automated testing for Linguistics-Based Cues (LBCs) has been undertaken. In a 2004 study Zhou, et al., utilized twenty-seven cues that were clustered in nine linguistic constructs: quantity, diversity, complexity, specificity, expressivity, informality, affect, uncertainty, and non-immediacy that were then measured in a text source. That research group found that a systematic analysis of textual information could be of use in the discovery of deception. More recently Pennybaker has

offered the **LIWC 2007** software package that has a set of discrete dictionaries that are keyed to various linguistic cues.

Oral arguments are the ideal data set to work from for this study. There is no question about who is "speaking" (a justice as opposed to a clerk, or a different justice who has requested that specific language be injected into an opinion), the utterances are less "filtered" than court opinions, and the language is likely to be less dense than that found in written opinions which could be advantageous given that the highly technical nature of Supreme Court decision writing is more likely to be confounding to the software.

As machine reading allows for a broad sample of opinions to be assembled and tested—the results from the analysis can then be checked by author first against the composite political ideology scores of the justice. Various LBCs could then be developed and evaluated in order to determine the highest value cues.

CHAPTER 2

THE EMPRIRICAL STUDY OF JUDICIAL DECISIONMAKING

A summary is presented of the literature and research that had focused on judicial decisionmaking. Note is to be take on the contributions that behaviorists (particularly the attitudinalists) have made to our understanding of how judges work, and to the trend in the Public Law field to work with a blended model that mixes features of attitudinal, strategic, and formal legal mechanical explanations. At the same time, many have bemoaned the comparative lack of attention that has been given to the actual words that judges author, and to the sometimes broad and deep effects that those writings can have. Although the overall reality is that some progress has been made, and that "new" methodologies (network analysis, machine reading of text, topic modeling, linguistic approaches) have shown promise, it is the author's considered opinion that there are persistent gaps in our understanding that only neuroscience will be able to fill.

2.1 Introduction

Professor Jack Knight has authored and coauthored a number of thoughtful studies that zero in on both the positive and normative work that has been done on the courts. Those concerns have bled into two distinct areas of inquiry. The first area of concern can be framed as how social scientists "measure judges"—the conceptualization, operationalization, and explanation—of judicial decisionmaking. The second area of concern widens the focus to try to evaluate how the work done in the first area can be utilized to evaluate which judges meet the normative benchmarks by which they can be evaluated. Because Professor Knight (2009) has sharply framed the challenges that

judicial decisionmaking now confronts, his trenchant analysis has helped to frame this chapter. That analysis sharpens down to the point that "our positive explanations of judicial decisionmaking ought to significantly inform the normative assessments we make about the quality of this decisionmaking" (Knight 2009 at page 1531), a point that is frequently considered throughout this project.

Skepticism towards the suggestion that the empirical study of the courts would be a productive endeavor, and that such work can potentially account for and contribute to our understanding of how judges actually do their work, would be understandable. Professor Knight and his frequent coauthor Professor Lee Epstein have championed the argument that the justices of the Supreme Court in particular are invested in influencing the substantive nature of the law. It follows from that position that if social scientists dwell too closely on the ways that the votes on the merits turn out (that is on the ruling as it concerns the parties), and too little on the words and logic that are woven into the opinions that the Court authors (that is on the holding as it concerns American law generally), far too much substance is missed. The remedy is both straightforward on one the hand, and subtly elusive on the other.

Professors Knight and Epstein lament that their wishes would not have been fulfilled if all that their work led to was "studies designed to explain the decision to accommodate or bargain or to persuade or to vote in a particular way [. . .]" (Epstein and Knight 1998; page 185). They go on to pronounce that the best goal for future researchers would be to contribute to the understanding of how those choices "come together to explain the substantive content of law." Perhaps it is the product of haste, or of lackadaisical editing, but the lumping of persuasion in with accommodation or bargaining

is a curious choice. Although they are likely considering persuasion in the context of collegial courts, they are overlooking an important distinction between bargaining and being persuaded. One does not choose to be persuaded, as one chooses to make an accommodation, or to strike some sort of a bargain. If anything, being persuaded only occurs when there is some initial mental effort put into not being swayed, but the argument succeeds none the less.

Real persuasion, as previously discussed, is an inexorable process whereby logic compels a realignment of one's worldview. If one has been moved either from one position on an issue to another, or from an agnostic position on an issue to some biased position, one has been persuaded. Where Professors Epstein and Knight have made an error vis à vis the most forensic definition of persuasion, their sin still does not appear to be a fatal one to their larger point—that Public Law scholars must focus on the larger picture, must move well beyond dispositional votes in order to engage fully with the mechanics of decisions, and with how the structural elements that are formed into opinions are where the real next frontier is for the field (This claim is not unique by any means, and echoes of prior scholarship by John Brigham are more than apparent (Brigham 1978)).

A mere two paragraphs later Professor Knight invites persuasion back in with the observation that "The task of crafting persuasive opinions plays a central role in two aspects of the decisionmaking process; the justification of the legitimacy of the decision and the establishment of new law." (Knight 2009; at page 1533) Although this is persuasion that emanates from the bench outward to the world beyond chambers, the

nexus appears to be clear: No matter which end of the telescope we decide to look through, persuasion matters greatly in the context of judging.

Central to the arguments that are being covered here is the notion that it is crucial to keep track of the positive implications that using various models and measurement approaches will have on the effectiveness of the explanations that are derived. Simultaneously, it is equally crucial to not lose sight of the normative implications that extend out over subsequent assessments of the objective quality of the opinions that courts render so that meaningful assessments of the process can be made, and substantive suggestions for practical improvements can follow. If we do not ask the correct research questions, we could have been under the misimpression that all is perfectly, truly and globally well with judging (i.e., nobody has any issue with the process). Were that our benchmark then social scientist of all stripes would have no valid reason to pay any attention to it whatsoever. Pursuing the topic and finding legitimate friction points as we will from time-to-time, it is certainly beholden upon us to follow-up our criticisms with both effective research and constructive, sound suggestions aimed at correcting the problems upon which we have isolated.

We transition now to an examination of the merits of the various approaches that have been utilized to date. One viable razor to use in order to help in fathoming out the prior work is to consider separately quantitative efforts that have been built upon empirical measurement and statistical designs (mainly of the classical sort (e.g., Johnson, Black, Goldman and Treul 2009) but sometimes straying towards the Bayesian as when a justice's "priors" are weighted differently depending on the novelty of the matter (Posner 2004 at 345)). Alternatively, formal theoretical approaches to the study of judicial

behavior have been utilized, some of which have been steeped in mathematical models (Lax and Cameron 2007).

Beginning from the already often mentioned efforts to understand and explain how votes on the merits come to pass as they do, the empiricists have been on the frontlines of Public Law as it concerns judicial decisionmaking for a long time. Going back to Pritchett (1948), a great deal of the action in the field has taken place in this arena. The axis upon which this local galaxy turned from the start was the disagreement between the formal legal model of judicial behavior and the attitudinal model, a conflict so central that it will be mentioned in some form in each chapter herein. While the differences between the two will be explored in great depth in later chapters, here it is sufficient to state that within the formal legal model judges merely "followed the law" and did whatever precedent instructed them to do, whereas within the attitudinal model judges made decisions that were based upon their own internal agendas (a patois of individual biases, values, and ideology blended together and then reassembled so as to resemble sound legal reasoning).

As the attitudinalists went about their work their priorities were to demonstrate a persistent disconnect between the opinions and precedent, and to present evidence of significant correlation between ideology and outcomes. Votes on the merits became the obvious dependent variable for study. This attention to voting was fruitful, and it furnished an approach for comparing the decisions that courts (and mainly the Supreme Court) made to the results that would have been anticipated if a purely formal legal model were correct. The argument tilted back and forth and various other variables were considered as the field attracted new scholars (e.g.: inter-judge dynamics on collegial

courts (Murphy 1964), amicus curiae briefs (Collins 2004), and how agendas are set (Perry 1991)). The study of the Court as an institutional entity, embedded within the other institutions of the U.S. Government, and the United States as a whole, have also provided grist for the empirical mill (e.g., the use of language to evade Congressional review (Owens, Wedeking, and Wohlfarth 2013), the interactions between the Court and the Executive Branch (McGuire 1998), the effects of public opinion (Baum 2006)).

In considering the ongoing relevance of judicial decisionmaking, Professor Knight reemphasizes that the dispositional vote remains the primary metric for divining judicial motives but also highlights a 2006 study of tax law (Epstein, Staudt, and Wiedenbeck) as a beacon of clarity in the field. Whereas most work that reviews the mechanics of the legal system and the courts have looked at the attitudinal model through the lens of U.S. Supreme Court whose data sets are necessarily mainly concerned with constitutional law, Professor Knight contends that the relatively anodyne area of tax law that Professor Epstein and company utilized for their analysis opens a different door. The shift from contentious, rights-centered litigation towards less fraught, economic activity-centered matters was hoped to allow a clearer view of the reasoning done by the justices. Nevertheless, the results led the authors to conclude that ideology does indeed play a significant role in accounting for the decisions that are made in these sorts of matters (either for the citizen/corporate entity or the Internal Revenue Service).

Professor Knight expresses (in a limited and professional fashion) at least mild surprise at that outcome, calling the outcome "contrary to common intuition" (at page 1536). This response is perhaps slightly naïve given the extreme levels of ideological intensity that have defined the conservative relationship to both the federal government's

42

exercise of its taxing power generally, and the ongoing activities of its main revenue collecting agency specifically. While the article itself is indeed an outstanding example of innovation in empirical analysis, and of the possibilities that vote-based analysis is indeed capable of reaching, there is no reason to be surprised by the results. We also see a pivot from the empirical approach to note that formal theoretic research (again, held to be work done using mathematical models) has spun off in a different direction.

From their initial origins where methods were taken from the sorts of standard spatial models that had been utilized to study ideological distances in the votes of legislators (Martin and Quinn, 2002 at 138; in which the justice will "vote to affirm the decision of the lower court if the utility the justice attaches to the status quo is greater than the utility the justice attaches to the alternative regardless of the expected actions of the other actors"), Public Law scholars moved on to adapt a "case-space" approach that took notice of the difference between a ruling (as it applied to which party "won" or "lost," and a holding which would go on to apply to non-involved parties in subsequent legal conflicts (Lax and Cameron 2007). A meaningful distinction was noted between the way that legislators worked and the way that appellate courts worked where the bargaining went to the content of the legal rules that majority opinions promulgated and which, in turn, bound lower (and to an extent coequal) courts moving forward. Thus, scholars working in the formal mode began to conceptualize the choices that justices made in terms of the alterations that were made to the substantive law as a result of the substance of the opinion that was composed to clarify and support the outcome of the case itself.

The distinction then can be summarized as the empiricist reducing matters to the votes that decide the disposition of the case, and the formal theorists trying to drill down to the effects of the rule that is pushed out by the case. There will, of course, be exceptions to such a categorization, but Professor Knight is essentially on point with his analysis. He conceptualizes the difference as potentially being a product of divergent opinion about what judges really do, about what is actually important in the work of judges, and what social scientists are most able to engage with and account for in this area. Although we will move on to address some of the issues that are raised by the first two differences, with regard to the third possibility Professor Knight observed that to his knowledge "there have not been any serious efforts to translate the results of the case-space analyses into an empirically meaningful research agenda" (at page 1538).

Arguably the section of this project that concern itself with the network of opinions that cluster around specific language in U.S. Supreme Court opinions qualifies as an effort to pull substantive content of opinions into a rigorous, quantitative research plan. At the same time, while Professor Knight identifies the lack of empiricist efforts to bring elements of "judicial reasoning and substantive argumentation into their analysis in a systematic way" (at page 1538), this project aims to develop further so as to address those gaps in the field.

2.2 A Brief Consideration of How Judges Make Decisions

Several sources provide useful insights on the "business" of judging, with the insights provided directly by jurists themselves. Justice Antonin Scalia's Reading Law: The Interpretation of Texts (with Bryan A. Garner, 2012), David M. O'Brien's edited volume Judges on Judging: Views from the Bench (2004), and Judge Richard A. Posner's

How Judges Think (2008) are all excellent sources from which to begin an inquiry into the topic. In summary, each book tends to stick closely to the narrative that judges all but exclusively use the law—and not personal ideological biases—when arriving at their decisions.

As will be discussed in more detail at a later point, there is ample—again for emphasis: ample—reason to be not simply cautious, but outright skeptical, in the evaluation of any such sources. At the conclusion of the Preface to his almost six hundred page treatise on precisely how a judge should go about using Textualism to unravel cases before the courts, Justice Scalia (in a wise exercise of ex post facto legislation by fiat) retroactively excused himself from having had to follow his own stated methodology when he wrote, "Your judicial author knows that there are some, and fears that there may be many, opinions that he has joined or written over the past 30 years that contradict what is written here—whether because of the demands of stare decisis or because wisdom has come late." He further granted himself an infinite degree of freedom to completely ignore his own counsel prospectively by further stating "Worse still, your judicial author does not swear that the opinions that he joins or writes in the future will comply with what was written here—whether because of stare decisis or because wisdom continues to come late, or because a judge must remain open to persuasion by counsel" (Scalia and Garner 2012, at page xxx).

Leaving aside the late Justice's touch for light humor,[3] the perils of taking what elites put forth at face value remains a persistent concern (one that will be more fully addressed in the next chapter). Those on the benches certainly have good cause to maintain the veneer of respectability that accretes from the general public widely believing that they (those on the benches) are fair dealers who rely exclusively upon the black letter law when determining how their opinions should be written, and how their votes on the merits should be allotted. It is so much the case that Professor Keith Bybee made the novel but compelling argument that "[p]ublic skepticism about whether judges actually mean what they say is potentially corrosive, but it also points to an enabling dynamic that makes possible the exercise of legal power" (Bybee 2010, at page 7). Thus Bybee is making the argument that although we all know that judges lie about how they decide matters, we go along with such nonsense because common courtesy (i.e., not calling them on their bullshit) allows the world to keep functioning smoothly. The simple truth remains: we cannot just take judges at their words in this instance. The matters to which their decisions pertain are too global in terms of importance, and the strong impetus upon judges to keep the entire enterprise moving along demands that we must allow the metaphorical scales to fall from our metaphorical eyes as Saul did in The Book of Acts.

At another point in this text the argument will be made that we should be open to the possibility that, when a judge makes explicit reference to being persuaded on some

---

[3] Not to mention the propensity that Justice Scalia's had for writing in the persona of a flustered, nineteenth century schoolmarm with references to such antiquated terms as "jiggery-pokery", "[p]ure applesauce" (*King v. Burwell*, 576 U.S. __, 2015), and "argle-bargle" (*United States v. Windsor*, 570 U.S. __, 2013).

point of law, it would be reasonable to at least entertain the possibility that there is a kernel of truth in the claim. Further arguments will be made, but for the present time is should be sufficient to note at the least the vast difference in scale between those broad claims that are made with regard to the entire enterprise of judging, and those specific claims that are made with respect to narrow legal arguments that are laser-focused on discrete, atomic points of law. The difference in scale is critical in the case of these two widely dissimilar claims about judicial behavior.

Coming back to the topic of what judges and justices themselves have had to say about the act of judging, we shall be mainly considering Judge Posner's writings in this section (although many others have also written on the topic (Cf. Cardozo 1921; Kozinski 1992). While it can be said that Judge Posner simultaneously resonates with some of what his colleagues on the bench have written on the topic, it is also the case that he has, as Professor Knight puts it, also resisted "aligning too closely with any number of theoretical models in both the social sciences and jurisprudential literature that purport to set out the answer to the question of what determines judicial choice" (at page 1539 (italics in original)). Rather Judge Posner presents his own framework to aid in explaining the processes that drive judicial decisionmaking that can be useful here regardless of whether one decides to adopt it lock, stock, and barrel.

While Judge Posner does start out by elaborating upon nine theories of judicial behavior, he shortly concludes that while they are "overlapping" and "insightful", they are also "incomplete" and (in a wonderful understatement) that they collectively "make for an unwieldy analytic apparatus" (Posner 2008; at page 57). For the present purposes it is more useful to conceive of Judge Posner as offering a simple list of components that

must be accounted for when seeking to explain the work judges do. That list unpacks as causal factors the reasoning process, and opinion relevance. Part of the utility of this "stripped down" way of dissecting the process is that it is "loose" enough that it can be adapted to each of the three prominent regimes that are currently in place.

If one is exploring the formal legal model, one can emphasize the role of precedent and the norm of building opinions based on previous holdings. In the alternative, one who is an attitudinalist can reinforce the significance of jurists who defer to the precedential significance to the advantage of policy-driven outcomes. If one is in the middle between the two camps, one can view the process as drawing from each side and use the proffered framework as a bridge of sorts. With respect to this project, this heuristic explanation serves best as a de facto benchmark against which the current state of empirical research could be roughly evaluated, and to which the current project could eventually be compared, again, in a rough sense.

With regard to the causal factors that play into the processes that jurists must weigh—the undergirding structures that animate the reasoning process—in Judge Posner's view the relevant components are what precedent provides to the decision maker; the norms that have been established within the law and which all professionals recognize and rely upon (at least to some degree); and, what internal agendas and biases come into the picture. To move much beyond this point though requires some further discussion of how the law and those who reside within it in professional judicial capacities confront and cope with the inevitable gaps that are to be found, as system of statutes, case law, and regulations can ever encompass all of the potential outcomes that could conceivably give rise to conflicts that invite litigation as a channel for resolution.

That territory is squarely the domain of the famed Oxonian legal scholar H.L.A. Hart, remembered for many things, but perhaps mainly for his "chestnut" used far and wide to introduce judicial reasoning and the nuances of the textual interpretation of statutes. He began with the proclamation that "No person may bring a vehicle into the park" in order to illustrate that there will always be "debatable cases in which words are neither obviously applicable nor obviously ruled out" (1958 at page 607). In the fact pattern he rattles off a list of modes of transport and asks whether the no-vehicle rule should be applied in each instance (e.g., automobiles, bicycles, rollerblades, and a child's pedal-powered toy car). Because the fact pattern provides no further guidance, the object of the exercise can be achieved through a series of Socratic interrogatories utilized to badger One-Ls about the nuanced differences between, say, motorized golf carts and unicycles. The process also introduces the concept of a penumbra, and the idea that some laws are going to be relatively easy to interpret whereas some others are going to be devilishly difficult to adequately thresh out.

It will be necessary to circle back around to H.L.A. Hart again shortly, and his ideas regarding the reconcilability of legal arguments, but for now our attention must shift to the argument made by the attitudinalists that these gaps provide the opportunity for decision makers to take latitude and to consult "outside" sources to aid in their arriving at a conclusion when adjudicating such matters. Privately held political views, policy positions, outright biases, idiosyncratic opinions, and naked prejudices can all seep into matters. It is the opinion of Professor Knight that, in a complementary fashion, it would be desirable for the empirical social sciences focused on this area to make careful study of these components, and to attempt to develop effective methods for ferreting out

the aspects of the process that justices (and lower court judges) are likely to be invested in shielding from the eyes of the Academy.

With respect to the reasoning process that drives how decisions are made, Judge Posner has a clever approach to unpacking how individual justices (and judges) can actually vary their approaches from situation to situation. He begins by calling into doubt the straight legalist model while noting that although some likely act as legislators (to behave attitudinally) only after attempts to follow legal texts and legal precedents provide unsatisfactory results, others could change the sequence. If the matter can be reasonably argued to be controlled by some orthodox legal markers, and if not taking note of those guides would constitute some form of error, then the path forward should be clear. Naturally, to this point the justice would have had a number of advocates standing before him or her, metaphorically peppering (if not carpet bombing) the bench with all manner of guidance as to how the matter should be determined. The jurist is assigned the task of determining the merits of each argument, that is, the persuasive strength of the advocate's positions. If there is a clear and unambiguous "winner" then the matter can be disposed of if the jurist is content with that outcome. If there is a significant gap that needs patching because the matter is in the hazy penumbra of the otherwise controlling legal principles and statutes, or if the jurist is disinterested in following the precedent, a different outcome (and how to support it) will become the focus. This approach would fit into the mode of citations being determinant of the outcomes of cases.

In the alternative, Judger Posner submits, many jurists will reverse that sequence, as they are driven mainly by policy concerns (although those such as Professor Braman who hew to the motivated reasoning model would observe the possibility that this drive is

not a conscious one). The starting point for these individuals is the legislative ruling

itself, basing the voter (or opinion) not just on which party gets the ruling, but on the

effects of the holding that is announced with the verdict. After the choice is made, legal

guidance is consulted, and a calculus is computed that determines if that desired outcome

is completely precluded by existing statutes or case law, and if not, if the benefit of the

desired outcome outweighs the potential costs incurred by the refusal to stay within the

established legal boundaries. This approach would be seen as conforming to the use of

citations as a mask for attitudinal decisionmaking.

In the ultimate analysis Judge Posner maintains that most judges "blend" the two

approaches, as opposed to sequentially considering them. A reaction to a given matter

that is presented for review at the appellate level forms from an amalgamation of legal

materials, existing constraints, policy leanings, interactions with other jurists (if the

context is a multi-judge panel), and the various equitable doctrines that are implied.

According to Judge Posner's account the consideration of the matter is at all times

mediated by "temperament, experience, ambition, and other personal factors" (at page

85). He goes on to add that:

> "A judge does not reach a point in a difficult case at which he says
> "The law has run out and now I must do some legislating." He knows
> that he has to decide and that whatever he does decide will (within the
> broadest of limits) be law; for the judge as occasional legislator is still
> a judge."[4]

In turn it is suggested by Professor Knight that these observations regarding the

processes that combine to generate decisions form the foreground of the next group of

elements that social scientists in this area must scrutinize. Gathering data that helps us to

---

[4] At page 85.

understand the balancing of causal factors in the decisionmaking process, in parallel with the grouping of conditions that establish which path is followed in a given typical circumstances, should be a priority for researchers moving forward. The theoretical framework presented in the introduction to this project is a modest attempt to add some more specificity to Professor Knight's suggested empirical agenda.

To complete the triad, the relevance of judicial opinion writing must be briefly considered in the context of what these scholars have put forward. Although one might imagine that reliance on citations in Supreme Court opinions has always been the norm, the Court only really began to fully adopt the principles of stare decisis around the start of the twentieth century (Fowler and Jeon 2006). From that time forward the standard was that for significant opinions (loosely referred to by attorneys as "front of book" opinions because legal publishers tend to aggregate the longer, more substantive opinions in the front of published volumes) demanded that justices provide to the public some explanation of the legal reasoning process that informed their decisionmaking process.

There are several relevant points that this habit of authoring opinions requires one to examine so as to fully grasp its implications. The most significant is that the actual relationship between the explanation that the opinion provides and the actual biases that yielded it can be fully divorced from each other without the author's direct knowledge. While Professor Braman's (2009) adoption of motivated reasoning (with the subtext being that individuals are possibly not even aware of the real thought process that yields decisions, or of their own propensity to discount precedent that causes cognitive dissonance), Judge Posner also relies upon the unconscious as an explanatory variable in this model.

In a clever set of observations he explains that while issuing written opinions can theoretically serve as a backstop to expose mistakes that can emerge from biased decisionmaking which is prone to displaying a "compressed, inarticulate character" (Posner 2008; at page 110), that process is a less than perfect failsafe because while opinion writing does follow voting on the merits, and while justices do reserve the right to shift their votes on occasion before the opinion is finally rendered, that is not the typical way final votes are made. That observation allows us to put an edge on Judge Posner's second observation vis à vis opinions, namely that they are composed post voting which is relevant because ". . . most . . . do not treat a vote, though nominally tentative, as a hypothesis to be tested . . . at the opinion writing stage" (at page 110).

Pushing the argument further Judge Posner also maintains that it is material that opinions all have the potential to one day be regarded as precedential authority no matter what their undergirding causal mechanism is, even if raw emotion was at the root of the process. Rounding out his list, Judge Posner also considers the weight that is placed upon opinions, weight that conversely assures that opinions that go unreversed can come to be perceived as legitimate. When it comes time to decide what traits confer that legitimacy upon written opinions, Judge Posner emphasizes the interplay between the public nature of each opinion, and the capacity of the audience to comb through it in order to establish (or discredit) its specific adherence to the established principles of judicial decision making, and its general tendency to respect and align with legal formalism.

Simultaneously, and perhaps surprisingly, Judge Posner has no difficulty looking beyond the chains of formalism when evaluating what established legitimacy. His model does not require holding the opinion and something called "the law" up side by side to

establish congruence as the only way to establish the validity of an opinion. More flexible in his approach, Judge Posner allows for a wide range of mechanisms to be viable tools in a jurist's kit, and contends that jurist will be able to pass muster so long as the they keep from straying so far from the metaphorical Temple of the Law that their process is outside of the scope of what is recognizable as sanctioned processing of legal claims. Although, given that he is a judge, it is unsurprising that he argues for a lenient standard when evaluating the legitimacy of opinions (in fact, if both Professor Braman and Judge Posner himself are correct, it is possible that he was unaware of his own conveniently self-interested thought process).

From this point forward our focus can be sharpened to a point: Judge Posner's argument that it is necessary to look to the legal community in order to establish if an opinion, or a specific element of an opinion, actually passes muster as a legitimate product of unambiguously acceptable judicial decisionmaking. While this "crowdsourcing" approach is rational, we should pause for at least a moment to ask how often judicial opinions in general, and Supreme Court opinions specifically, are found to have been decided without proper reference to the consensus legal standards. While a few opinions are regularly offered as examples of poor decisionmaking (Citizens United v. FEC, Bush v. Gore, and McCleskey v. Kemp could collectively be thought of as the low hanging fruit in that orchard), the Court rarely reverses itself (although it has done so on such topics as sodomy statutes, separate but equal, and the death penalty), and is even more rarely corrected by the other institutions of the government.

A pivot is now necessary in order to move on to addressing how the social scientist should ask specific questions about the reasoning, logic, and rationales that

justices employ when writing their opinions, as those elements are what determines the likelihood that those opinions will be received as legitimate, and what actual changes the law itself is most likely to undergo. Prior to making that move a moment to reflect on Judge Posner and his analysis is in order. Professor Knight is content to make use of Judge Posner's breakdown of the three relevant categories for analyzing judicial decisionmaking and there is no crime in that. It is a fairly lucid approach, and one that I too have shamelessly cribbed. Perhaps its major asset is its simplicity, in the habit of a well trained legal mind, Judge Posner has stripped away superfluous details, and presented a parsimonious grouping that is simultaneously complete enough to do its assigned task and does not burden the reader with unnecessary clutter. It is a good base to work from because it is parsimonious, and also because there is really no other logical way to parse the topic out. Added to the economy of the program is the reality that as a Judge on the Federal Court of Appeals (Seventh Circuit), Posner certainly stands on solid ground from which to make principled stands with respect to what judges likely do and do not think.

Nevertheless, Judge Posner only speaks for himself, and his passionate championing of economic modeling and rational choice based explanations skews his views significantly. He takes the social sciences seriously, and his extensive list of scholarly publications has given much grist to the mill of academia, but he is limited by his narrow approach, and not always so absolutely lucid on matters that his statements

with respect to causality should be given any special deference (Cf. Judge Posner's email "debate" with Philosophy Professor Peter Singer regarding animal rights[5]).

In the area of judicial decisionmaking and the social scientific study of the courts we can safely rely upon Judge Posner as a guide to the general terrain features and the way that major landmarks are oriented with respect to one another. Following him, and his distinctive rational choice/economics based approach, carries concomitant baggage as it is prone to disregard much well established work by solid scholars.

2.3 The Struggle to Empirically Measure the Law

Of the three sets of factors that social scientists have concentrated on analyzing, the several determinates of choice—and the fundamental reasoning behind legal decisionmaking—have received the lion's share of the attention to date (Knight 2009, at page 1545). The point at which reason and choice intersect is the focus of the debate between the attitudinalists and the legal formalists—the point at which we would learn the most about the basic drives behind justices as they decide cases.

Much of the study of judicial reasoning converges on questions about what specific mental processes influence justices motivated by bias, seeking to gain insight regarding the various inputs that are weighed on the way to reaching a verdict. Granted then that judicial decisionmaking seeks to illuminate how judges make decisions, as previously mentioned, a great deal of that work has relied on using votes as the dependent variable (although often that raw data is converted into ideal points).

---

[5] >> http://slate.me/28Rlvaz << (accessed 21 June 2016) Wherein Judge Posner blithely asserts that non-human animals require no further protections under the law beyond what they are now provided as, given their current status as chattels, we need not fear harm coming to them because people have an economic disincentive to cause harm to their own property. Apparently the Judge has led a sheltered life during which he has never come across the spectacle of a demolition derby.

Those who have worked at the questions presented in the arena of judicial decisionmaking from the theoretical side have more often focused on the reasoning side (often through a game theoretic approach), but often with at least the implicit assumption that the attitudinalists have made a fairly convincing case. This work, which is rife with figures showing graphic representations of spaces and boundaries that represent negotiating positions on collegial courts and the various negotiating positions that justices take up among themselves, is much more focused on the material content of judicial writings, but has yet to yield a raft of material that can dovetail well into the empiricist's corpus (Knight at page 1545).

The question remains: how best to address the gap between the work done by empiricists so far in this field, and what goes on inside the "black box" of a justice's mind where the causal mechanisms that promulgate the substantive rules that define the content of our laws.

Before descending the allegorical ladder into those murky depths, this is a logical point to touch upon a nagging problem that needs some attention. With respect to making declarations regarding which decisions are biased, and which are well grounded in the law, it would obviously be useful to have some definitive guage that could tell us with certainty which decision is what. The ideal point approach relies upon the way that each successive opinion entered into the model "fits" in with each previous one, and it (within limits) should give us the relative relationships among rulings and, by extension, the justices who authored them. Well and good, but consider what happens when the institutions of our government, and the composition of the Court, swing in various directions. An extended period of (predominantly) conservative control of the legislative

bodies of the House of Representatives and the Senate, in tandem with a prolonged period of time during which the Court itself swings to the conservative side is bound to eventually have an effect on what it actually means to "follow the law."

Thus, if the majority of the legislation that is put into place is the product of conservative philosophies, and was intended to further conservative policies, and if the majority of Court opinions over an extended time period have been authored to advance conservative theories regarding legal matters, a real sea change has occurred. An opinion authored today that closely follows the black letter of the law could have been considered a radical, or even a fringe opinion, thirty years ago. Exactly how academics are to account for the possibility of the significant tectonic drift of the law over relatively brief periods of time will be addressed in more detail in the Chapter 7 of this project, but for now suffice it to say that we have a lot of careful reckoning to do in order to properly ascertain what is a faithful reading of the law, and what is a departure from prevailing norms. Understanding this change process is relevant to the main questions asked herein about persuasion because persuasion, when effective, must be used to "pull" a biased justice into the position of following the law faithfully. If we are uncomprehending what legitimately constitutes following the law, then we will be rudderless with regard to evaluating if persuasion has been applied properly (we certainly would not want to instruct advocates how to most effectively persuade judges to not follow the law).

Returning to the question of how empirical social scientists have addressed the "black box" of legal reasoning, it will be helpful to consider some of the approaches that can contribute to our understanding of the topic. To do this it is sensible to simply parallel some of the categories from the prior section.

With regard to the task of describing the things that justices do when making decisions we need not be overly concerned. Although a focus on the process by which opinions are created is an important part of describing the mechanisms of judicial decisionmaking, empirical social science is not necessarily bound up in this activity. The absolute accuracy of descriptions is not the criteria by which social science is evaluated. So long as descriptions are adequate to meet the requirements of explaining the judicial reasoning process and the normative implications of opinion writing, then the descriptions will be sufficient for their purposes (Knight at page 1456).

Turning to explanation, the task becomes more difficult, and it is necessary to bifurcate the analysis as different sorts of questions demand different approaches. Asking direct questions about causation (the most basic being "does bias determine outcomes", but there are other factors such as advocate effectiveness and the identity of the parties that can also fit into this sort of investigative query) has been the stock and trade of the attitudinalists. For such inquiries the use of votes on the merits has been sufficient for them to make substantial progress. Likewise, if the task is merely to compare the relative impacts of two different variables then, once again, this examination of causal connections should yield to vote-based measures.

In the alternative, other sorts of inquiries will require other sorts of evaluations. If the mission of the central question being studied begins to drift towards explaining "how" a given variable affects judicial decisionmaking (such as having formerly been a prosecutor, collegial relationships, or the force of prior precedent), the process becomes more complex. Simple reliance on the possibility that justices are following the law is going to fall short of providing necessary clarity in many circumstances. A more nuanced

evaluation of such a question requires an approach that blends both the formal legal model of following precedent, and the bias-driven model that concerns itself with attitudes and desired policy outcomes that a justice wishes to see. This wider view of the space in which decisions are made is going to be one that accounts for constraints and for the balancing that will be done in the furtherance of agenda advancement. Because of the accretion of complex and even obscure variables into the equation, simple vote counting will obviously not be sufficient to adequately plumb these areas of inquiry.

Spatial models have been employed to evaluate the "how" questions as more elaborate ones offer independent refinements, often predicated upon the standard that each judge has an ideal locus in the set of points within the case-space that is the precise place where that individual's preference for legal outcomes is situated. There are then the constraints of precedent to be considered, as they will often cause the final outcome to be pulled away from the ideal locus and the distance between those two points is a representation of the causal effects of the external world on the internal desires of the individual to effectively legislate from the bench.

Some logic can be applied to the various measures of distance that different cases create. A significant distance in an opinion that states reliance on prior precedent would indicate that the precedent was constraining. Mining more data will enable sharper and sharper explanations to be proffered, and quickly make evident the inadequacy of simple reliance on votes in circumstances such as these. The challenge will be in isolating exactly what alternatives reside in a justice's choice set, and then in effectively gathering that data.

The assessment of the objective quality of judicial decisionmaking will typically require some consideration in order to determine which specific forms of data will be necessary to make normative assessments. Less complex analysis, such as an inquiry as to whether or not a single variable is or is not part of the calculus of legitimate decisionmaking should be possible to accomplish using vote counts. It is self-evident that circumstances where a justice has a wider range of choices available will be more of a challenge to model, and that the underlying work stems from accurately—or reasonably accurately—accounting for those alternatives, a task that will be both specific to each situation and will also likely demand a keen grasp of the law itself.

In keeping with parallel structure, the normative nature of judicial decisionmaking must now be considered. We begin by disposing of the obvious case where some explicit rule binds an outcome; although not appellate decisions, sentencing guidelines are the obvious instance here as they effectively limit the menu of choices from which the decisionmaker has to select. In the more complex world of appellate opinions congruent limitations are also present—for example we do not anticipate justices to issue opinions that boldly contradict the Equal Protection Clause, or some well established principle of constitutional doctrine such as the Clear and Present Danger standard. Such a thing could transpire, but it would encounter widespread condemnation as it would be held to be clearly contra-normative. Indeed, even a simple vote in that direction would be sufficient to get the pitchforks up in the air, a clear indicator that the vote on the merits is a fair measure in such a stark circumstance.

Most of the work of appellate courts will be within Hart's penumbra, and thus require far deeper analysis. Again, though, we run headlong into the conundrum

regarding "whose interpretation of the law is actually the accurate and unbiased one." Professor Knight's analysis, which I track here without fully accepting, borrows directly from Judge Posner and does not really squarely address this fundamental concern. Rather the Judge skates around it, not wishing to get too close to the inky void that threatens to blot out his arguments entirely.

To some extent Judge Posner's position can be interpreted as a variant on the "rules of recognition" that were articulated by the aforementioned Professor Hart in his celebrated book The Concept of Law (1961). In Professor Hart's view the legal system, that is the constructed (constituted) program that administers laws and establishes what the courts can and cannot do, determines in a fundamental way what will be deemed to be legitimate at a given point in time (and that standard will obviously carry forward until such time as the legal system undergoes some significant shift in its posture). Professor Hart's rules of recognition are a layer of standards that govern how the primary rules of the law are to be interpreted, put into effect, and even challenged. In not being primary, these rules of recognition are secondary, and the accord between how jurists behave in the context of their work, and how these rules are construed and understood, establishes what will be held to be legitimate versus illegitimate.

To Professor Hart the rules of recognition can remain in force and be respected provided that the concerned community gives them some level of willful acceptance. If the public respects the legal institutions that are in place, then those institutions will have some latitude in which to operate. The value of the courts and the more general legal apparatus are in synchronicity with the observers, with the more influential and directly

concerned observers having proportionately more influence regarding the final disposition of these institutions than the casual or distant observers.

Professor Hart goes into more specifics, as he links the major elements of legitimacy to the specific instances where the courts must apply the primary rules to previously unseen fact patterns. The primary rules to which Professor Hart makes reference are, for all reasonable purposes when dealing with common law jurisdictions, the canons and principles that Justice Scalia and Garner attempted to at least partially codify in the aforementioned Reading Law: The Interpretation of Legal Texts (2012). The rules of recognition go in part to how well the jurists do at construing and interpreting that general body of legal reasoning to novel fact patterns. Nevertheless, because not every jurist is a textualist, the rules of recognition must also absorb the reality that some relevant observers take a less textualist approach, and could instead advocate for purposive readings of the law. Indeed, in sophisticated milieus there will be a range of interpreters each advocating for what the real rules of recognition dictate in a given set of circumstances. This is only to be expected as the capital "L" Law breeds complexity, and with growing complexity commentators will have more and more space in which to spin out their various interpretations.

The relevance of this theory to the present study is in that Professor Hart's reasoning holds that the legitimacy of a court's rulings is tied to what the majority of its professionally qualified commenters say about it. That viewpoint is arguably congruent with his legal positivism—the Law is what the King says it is; the legitimacy of the courts is determined by what the body of professionals who practice before it conclude regarding its merits. Further elaboration follows, but the reader must ask how this

program will work if the professionals in question are not themselves objective. If the defects in societies can be traced back to the individuals who are its members, what happens when corrupt professionals support a corrupt regime? An objective standard is needed.

As presented, those rules of recognition clearly provide jurists with a concise way to obtain and maintain legitimacy: so long as they play within the "rules"—that is they impartially apply their legal acumen to the decisions while respecting the principles and provisions that have been established within their institution—then their verdicts will be regarded as worthy of respect and obedience. By the same token, those rules of recognition also provide guidance for advocates who appear in court with respect to what types of arguments are made and evidence is tendered. This central concept of boundaries invites further scrutiny from various angles as they are certainly recorded and understood within the language that expresses them and that language—like all languages—is open to interpretation (Cf. Brigham, 1978, for an extended discussion of what does, and does not, make sense with respect to the Law and language).

The major trouble with Professor Hart is his reliance on the inherent fidelity of jurists to following the rules and respecting the norm of impartiality (Hart 1961 at pages 136-140). His expectation is that the publicly presented writing will expose the decisions that jurists make to sufficiently stringent analysis that they will strive (and succeed) at making their rationales watertight. This allows for the possibility of individual bias driving the decision, but only so long as the outcome can be masked in legitimacy by citation to prior precedent that has been "properly" interpreted within the widely accepted

rules of recognition that the relevant community has learned, and which has evolved to be a worthy standard for measuring the legitimacy of a decision.

Opinions become grails of sorts in this model. If the language utilized by jurists to ground a verdict can be found to agree with some prior precedent or with the standard rules for the interpretation of legal texts, and if the opinion is rendered within the recognized rules that are widely believed to control in such circumstances, legitimacy is conferred. If a superior court reverses a lower court this would be viewed as a strong signal that there was some defect in the process of the lower court, that the rules were not properly followed. This presumption would be held so long as a more superior court did not re-reverse the middle court's opinion overturning the lower court. Reliance on citation to prior authority, and upon the ongoing standing of an opinion as "good" law, should not be viewed as adherence to the formal legalism. Instead those quantifiable qualities—conformance to the norms of citing prior precedent and opinions remaining intact—are among the yardsticks that the legal community selects to test the rationale of an opinion to what that community deems to be legitimate (Knight 2009 at 1552-1553). That category of "legitimate" could be restricted to only opinions that are the product of mechanical legal reasoning, but it is an open question as to what each individual society will be willing to accept as a reasonable process for arriving at legal decisions. While a strong philosophical argument can be made that objectivity should prevail, that argument will not always be the one that wins out.

It should also be kept in mind that the majority of citations are at least colorably "on point," and that the number of opinions that are reversed is far less than the number of opinions that stand over long stretches of time. The United States Supreme Court only

takes on roughly eighty cases a year for full argument, and the majority of those do not

result in reversal, so it is clear that at least at the level of the Federal Appellate Courts the

vast majority of opinions are going to stand untouched (at least in the near term). It would

be unsurprising to find that similar patterns also hold true at all levels for state courts.

The distilled essence of all of this explanatory matter is that, at least until recently,

much of what empirical social scientists who work on have concentrated on this topic is

not particularly close to what could, in fact, help us to draw reasonable, impartial, and

more importantly principled, conclusions with regard to the actual quality of the judicial

decisionmaking process that we can observe in the real world. The narrow focus on the

votes of the justices is not sufficient to give us genuine insights about the outputs of the

courts. Simultaneously, it is conceivable that the empiricists who study judicial

decisionmaking and those who seek to assess the quality of justice that the courts

dispense should be able to find ample common ground. This assertion is plausible

because the changes that jurists effect on the law and the social outcomes that follow

from those changes are tightly intertwined. Professor Knight argues that "the data is

primarily the same in both cases: the arguments and the reasons that they employ in their

decisions are factors that affect both substantive content and judgments of legitimacy" (at

page 1553; footnote omitted).

2.4 To Where Should We Go Next?

In summary, grasping the larger picture of what is happening when courts issue

opinions would demand that we capture both the structural changes that the law

undergoes, and a parallel accounting of the claimed underpinnings that are relied upon in

those opinions. Professor Knight bemoaned that such rich datasets were not in evidence

in 2009, and almost a decade later it would appear that there is still much to moan about in that regard, although the use of machine reading of the language of the courts has been on the upswing. Nevertheless, the ongoing dearth of "thick" legal analysis motivates the researcher to question if the capture of such data is actually feasible.

When pondering that question Professor Knight asks if generalization, a fundamental component of social science explanations, can be reconciled with the levels of nuance and granular detail that is inherently characteristic of legal reasoning. His attempt at an answer with respect to the substantive content of the law, and how justices effect changes in the opinion writing process, calls again on the same familiar approaches that have been hammered on for the past several decades (case-space framework, measures of judicial ideology), with perhaps a few tweaks. In that passage Professor Knight draws particular attention to efforts to examine and describe ideological drift over time (Epstein, Martin, Segal, and Westerland 2007), an effort that is at the root of one of the chapters of this project. Professor Knight further suggests gathering the opinions of an entire population of judges on a particular topic, over a period of time and collating that into a set of "feasible" outcomes to then (presumably) give us the ultimate range of possible legal outcomes in that area.

In that suggestion Professor Knight is close to getting it right, but his arrow is still not quite inside the bullseye. His reach exceeds his grasp I think because he is (apparently) simply unaware of a significant dataset with which attorneys and judges are well acquainted—the Westlaw Keycite system. Although not the direct focus of this project, the Keycite database has, over the course of close to 150 years, codified and indexed the entirety of appellate law in the United States. Over ten thousand individual

categories, divided and subdivided into topics, capture each and every atomic point of law that has been announced. It is a large number of categories, but well within the computing capabilities that are now at hand. Networks can be built, by court, justice, and judge that can provide statistical and even visual reference to the substantive structures that give form and shape to our entire legal system, or just to specific subparts of that system. There is no doubt that this is the single largest, almost completely unexplored data set that will ever be available to the Public Law subfield—a veritable goldmine waiting to be exploited.

Also presenting a devilish problem for social scientists is the second battlefront that has been identified which considers the sources that are used to justify outcomes (most often announced in written opinions), and how those sources are then manipulated as grounds for decisions. Professor Knight proposes that different sources could be categorized which could allow the generalizable claims about each to be tendered (at page 1555), thus the place of the argument in the mosaic of authorities. The theoretical framework proposed in the first section of this project is somewhat aligned with this thinking (although at the time that I developed it I had not yet read Professor Knight):

$$C/P_{St.} + Cent._{CL} + C_{2nd.A} + C_{EJ}$$

As the reader will recall, the terms that are summed are: Net Aggregate Strength of Cannons of Interpretations/Presumptions Aligned with the Individual's Ideologically Favored Outcome for Pertinent Constitutional Elements, Statutes, Regulations, and/or Ordinances + Net Aggregate Centrality of Relevant Case Law Aligned with Individual's Ideologically Favored Outcome (Positive Case Law) + Net Aggregate Strength of Cited Secondary Authority Supporting Individual's Ideologically Favored Outcome (Positive Legal Commentary) + Net Strength of Cited Extrajudicial Sources Supporting

Individual's Ideologically Favored Outcome (Expert Testimony, Demographic Data, National Academies Research, etc.). The reader will notice that the second two terms, Net Aggregate Strength of Cited Secondary Authority Supporting Individual's Ideologically Favored Outcome (Positive Legal Commentary) and Net Strength of Cited Extrajudicial Sources Supporting Individual's Ideologically Favored Outcome (Expert Testimony, Demographic Data, National Academies Research, etc.) have been given scant attention here,. It is my estimate that while they are relevant to the decisionmaking process, they are less influential than the first two. Time constraints kept them out of this cycle, but it will be possible to develop them further in the future.

The more objective weight each brings, the greater the probability that a given justice will be compelled to follow the law. This rough sketch of the landscape corresponds to Professor Knight's plan, his further notion being that the social scientist had less to offer on the specifics of the law as it progresses, and more to offer regarding comparative study of which categories are generally taking the lead as substantive changes are put into effect.

While Professor Knight finally meditated upon the potential for social science to provide persuasive frameworks that help to explain judicial decisionmaking, and for those frameworks to essentially quarterback the general task of inquiry into this area, one follow-up proposal is in order. As it is currently practiced, fMRI and similar brain "imaging" procedures are far, far away from providing the sort of refined data that can truly revolutionize social science. Recent research (Eklund, 2016) found erroneous statistical assumptions built into the algorithms of several software packages used by fMRI researchers; those findings have, in turn, created doubts with respect to over 40,000

studies. As regrettable as that revelation is, the matter is not one to dwell upon because the development of technology will continue apace, kinks will be worked out, and the capacity to answer critical questions will come along in the next few decades. Until the point in time when we can place judicial decisionmakers inside of imaging technology that will provide us with legitimate scientific evidence regarding how those decisions are actually made, we will mainly be marking time.

In the meanwhile, one experiment does suggest itself. It is probable that the act of judging—that is the conscious process of fathoming out facts and precedent while consciously under the mantle of authority and with the knowledge that one's choices will be scrutinized according to rules of recognition—likely affects how decisions are made by individuals (as opposed to casually and anonymously allowing ones biases to steer the proverbial boat). Testing that supposition in an experiment is likely a worthwhile endeavor.

A convenience sample could be broken into two groups. The control group would be given difficult "cases" to decide—cases that deal with hot-button social issues or racial topics—but are told that their anonymity would be preserved, and that their decisionmaking process would remain unquestioned. The control group would be informed that their decisions would be scrutinized, and placed in circumstances that would create the impression that medical equipment would measure various brain activities as they made their choices. The reasons for the suggestion of a machine-based scrutiny is both because of the impression that will be presented is of complete objectivity in the process, and due to the expectation that most people will be less

inclined to believe they can manipulate the process and deceptively outsmart a piece of high tech hardware as opposed to a human expert.

A significant variation between the two groups would provide meaningful insight into the actual question of whether, and perhaps to what degree, being in the role of the judge changes the mental process of the person making the decisions. Just as the observer changes the experiment in the quantum theatre, it is likely that merely donning the robe changes the decider in the legal forum.

CHAPTER 3

AN EXPLORATION OF THE RELATIONSHIP BETWEEN IDEOLOGY AND

OPINION CENTRALITY ON THE U.S. SUPREME COURT

In their 2010 article "Citations in the U.S. Supreme Court: An Empirical Study of

their Use and Significance" Cross, Spriggs, Johnson and Wahlbeck state that:

> "The number of citations in a Supreme Court opinion is not randomly
> distributed but demonstrably varies according to a number of factors,
> including the individual justice authoring the majority opinion and the
> type of case. Operating from this beginning, future researchers may
> examine more closely the role of ideology in citation choices and how
> that differs among the justices."

They go on to offer that:

> "The quantitative results also may prove valuable in other studies of
> judicial characteristics. The Martin Quinn scores discussed above
> permit comparison of justice ideologies with our measures for
> citations."

This chapter describes an effort that engages directly with the possible

connections between an opinion's centrality and the ideology of the author. In a simple

design, the "drift" of Supreme Court justices over time with respect to ideology is

compared to the centrality of the majority opinions that they author. The theoretical

expectation is that the further the drift by a justice (in either direction) away from the

center point, the less central the opinions that they author will tend to be.

3.1 Introduction

This segment of the subproject examines the interplay between the level of

political bias of justices (measured using the Martin-Quinn score of ideology (Martin and

Quinn 2002)), and the authority, or relative importance, of the opinions that they authored

(measured using the Fowler authority score (Fowler and Jeon 2008)). If justices all tended to "stay put" with respect to their ideology over their careers on the Court such an analysis would be very simple, but many justices exhibit significant drift in ideology over time (Epstein, Martin, Quinn, and Segal 2007). Because of that propensity, it was necessary to manage the analysis at a granular, by-justice/by-term level, and to consider the authority score for each majority opinion that each justice in the sample authored. It was also necessary to consider the time delay effects that tend to suppress the authority of more recent opinions (as it naturally takes a period of many years for opinions to accumulate their subsequent citations).

We should care about the relationship between ideology and opinion authority for several reasons. First, we expect that justices themselves care about how their opinions will be regarded over time; an opinion that does not gain traction with later citations is, in effect, a dead letter. An opinion that is widely cited will tend to exert some influence on the future course of the law. Second, it would be useful if it could be established that, by striking a more ideologically extreme posture, justices sacrifice (at least to some degree) the long range significance of the opinions that they author. Third, it would be an intriguing finding—and one worthy of follow-up—if the relationship between opinion authority and author ideology was uncorrelated. With regard to the last point, a perfectly reasonable explanation for opinion authority to not correlate well with ideology would be if the determinant factor was the talent or ability of each justice to write compelling opinions. While such an explanation is a reasonable one on its face, everything is mild supposition up until the point that the relationship between authority and ideology has

been formally quantified and measured. Once that relationship has been evaluated, speculation can be replaced with concrete analysis.

Within the theoretical framework offered in Chapter 1, the motive of this subproject takes aim at the Cent.$_{CL}$ term on the right-hand side. The framework, although only intended to act as a rough guide to the terrain in this area of Public Law, is designed to estimate the probability that a justice (who is possessed by individual biases and is reasoning in a motivated fashion, but at the same time is of a given judicial temperament) will return an unbiased opinion. The Cent.$_{CL}$ term accounts for the Net Aggregate Centrality of Relevant Case Law Aligned with Individual's Ideologically Favored Outcome (Positive Case Law), with more central (authoritative) opinions having greater persuasive weight to most justices in most circumstances. Granted, a justice could disagree with an opinion that has a high level of authority, but to argue against such a "landmark" opinion is a very steep proposition. After all, it seems unlikely that a true firebrand who disagrees with a significant number of landmark opinions would be able to find a seat on the Court. As our understanding of the relationships between opinion authority and judicial response increases, this term will become more useful as a contributor to the prediction of judicial decisionmaking.

Absent an unambiguous, controlling statute (i.e., case law that is controlling in the present situation), a justice will rely upon relevant case law that has controlling authority (i.e., a previous U.S. Supreme Court opinion that is aligned with the issue in question and that announces the legal standard to be followed) to reach a decision. Even with respect to a circumstance where there a statute is controlling, somewhere back in the mists of legal time there is a piece of case law that directs how the statute is to be interpreted that

will guide the justice's analysis of the text (at least in theory). Justices take note of case law, but at the same time they also, no doubt, take note of the relative authority of each subsequent opinion that has cited to it. Also keep in mind that while some opinions live on to be heavily cited, others languish and may never, or only very rarely, be cited. Moreover, a few are cited mainly to be criticized, and a few others are even overturned in whole or in part. In short, not all opinions go on to be treated equally, and the relative status of an "opinion" must factor into how justices will treat it as time continues apace.

We could conceive of the continuum in the following fashion, from the most celebrated opinion to the most maligned (although not critical to the discussion, it should be noted that opinions can be struck only in part as well as in full):

- A heavily supported opinion with very little, or no, criticism;

- A well supported opinion with perhaps a smattering of criticism;

- A somewhat controversial opinion with some significant criticism;

- An opinion that has little, or no, subsequent history;

- An opinion still good in part, but that has been struck in part; and,

- An opinion that has been struck entirely, or almost entirely.

How a justice, or any legal professional, examines a prior opinion will almost certainly start with placing it—even unconsciously—in its approximate position along this spectrum. This sorting is not going to produce the exact same results as an analysis of network centrality (the next section), but in many cases it will provide some of the same information.

3.1.1 Network Analysis

Network analysis organizes network objects (nodes—in this instance opinions) in space with their relative positions determined by the connections that have formed among them (edges—in this instance citations). It would defy statistical possibility for all opinions in the U.S. Supreme Court to occupy the same exact position in the opinion citation network. One way of quantifying observed variations in position within a citation network is through the measurement of centrality (Newman 2007, Chapter 7). There are several different ways to evaluate the centrality of a given "node" (again, a member of a network), but the general gist of the concept is that being better connected to other nodes via edges (again, in this instance a citation) increases the centrality score of an opinion. If one has a network where the average number of connections among nodes is four apiece, but one exceptional node has connections to (shares edges with) fifty other nodes, that hyper-connected node will have the highest centrality score.

That variation in centrality carries over by extension to the authority scores developed by Fowler and Jeon (2008) and utilized here. Fowler and Jeon's approach takes notice of both outward citations (those citations that the primary opinion under consideration itself made), and of inward citations (citations from later opinions back to the primary opinion being considered). Thus, in the Fowler/Jeon scheme, two types of important opinions are given weight in the network: hubs and authorities.

A hub opinion cites to many other prior opinions, and in doing so helps to illuminate which opinions are the members of the constellation of legally relevant precedents for a given legal issue. These sorts of opinions become highly relevant in establishing where the history of a given legal issue can be found, and are helpful in tracing the development of a given area of the law. An authority is an opinion that is later

cited to by many subsequent opinions. An opinion that accretes subsequent citations grows in prestige because of its emerging popularity in the network. These sorts of opinions become relevant in establishing what the controlling law is in a given area and may also point out emerging trends in the relevant area of the law.

Most opinions will be observed to some exhibit both some hub characteristicsand some authority characteristics. In that respect, an individually authored, "front of book" opinion (i.e., a significant opinion that was given serious attention by the full Court) that does not cite to any other opinion would be an extreme outlier. Likewise, most opinions will eventually accumulate some authority characteristics (as the majority of Supreme Court opinions do go on to garner at least some citations over time). The most important opinions under the Fowler authority scoring protocol will be those that are both noteworthy for their propensity to cite to other previous popular opinions, and to also be cited widely by later popular opinions. Note that once an opinion has been published it is virtually impossible for it to increase outdegree, that is the number of prior opinions to which it cites.[6] At the same time, opinions when first published will exhibit zero indegree up until the time when a subsequent opinion cites to it. Over time some opinions will pick up significant inward citations, and will grow in authority.

Fowler and Jeon considered the proportional values of hub and authority scores to yield two equations that could be represented in matrix format. Those matrixes could be solved as convergences, and the resulting hub and authority scores were then converted

---

[6] In the rarest of circumstances the court may retract or rewrite portions of opinions, usually to correct some error that was made (consider the Courts retraction of *EPA v. EME Homer City Generation, L.P.*, 134 S. Ct. 1584 (2014), triggered by Justice Scalias' embarrassing mischaracterization of his own prior opinion in the uncorrected version). It is possible that such a "do over" could cause an opinion's outdegree to shift.

into a single Fowler Authority score that was used as the basic centrality value in this study.

3.2 Theory

With respect to causality, herein the expectation is that ideology contributes to the reasoning approach that an opinion's author takes, and that a more extreme ideology will exert a more pronounced effect. That expectation dictates that opinions by more moderate justices will be more intellectually palatable to more other justices over time (unless the Court as a whole moves towards an ideological extreme in which case it would be possible that ideologically formed opinions would likely gain in authority). That palatability (assuming that the Court does not stray into a pronouncedly less moderate mode) will tend to draw those moderate opinions to more central positions in the network as they garner citations. In turn the formation of those edges will boost those opinions indegree which in turn will contribute to their Fowler Authority score rising. In the alternative, as more ideologically extreme justices will tend to author less widely palatable opinions, the expectation is that those opinions will be less cited, and will subsequently reside further from the central part of the network where they will achieve lower Fowler Authority scores over time due to the paucity of edges forming amongst them and subsequent opinions.

This study was motivated by the reality that overall legal relevance and vitality give rise to an opinion's authority score (as defined by Hansford and Spriggs, 2006). Because many justices exhibit significant ideological drift over time (Epstein, Martin, Quinn, and Segal 2007), it would be problematic to rely upon a single mean Martin-Quinn score (2002)—a widely used measure of justice bias through a dynamic ideal point

estimation whereby the ideological extremity of justices is determined based upon "the company they keep." Martin-Quinn is a fourteen point scale with zero as its center point. Keeping in mind that the positive/negative scheme is intended to be arbitrary, a positive score (up to seven) indicates tendency towards conservative ideology (i.e., the justice tends to vote in the fashion of more conservative colleagues), and a negative score (down to negative seven) indicates a tendency towards liberal ideology (i.e., the justice tends to vote in the fashion of more liberal colleagues). Although some have exhibited relative stability during their terms on the Court (e.g., justices Thomas, Alito, and Murphy), for many justices just taking their mean career Martin-Quinn score fails to tell the story of significant transition over time in terms of that justice's preferences for settling matters. Although the reasons have not been fully proven, (Baum (2006) has postulated that pressure from liberal media may play a role), more often than not the shifts have been from the right/conservative side to the left/liberal side (Epstein, Martin, Quinn, and Segal 2007).

Consider the following career ranges of Martin-Quinn Scores of individual Justices:

- Thomas: 2.73 to 4.83

- Stevens: 0.03 to -3.21

- Brennan: -0.62 to 3.74

- Rehnquist: 4.43 to 1.22

Examining the entire Court as a single entity, and not tracing the variations in the authority scores of opinions authored at different points in a justice's career, runs the risk of missing potentially useful data. Cross, Smith, and Tomarchio (2006) have reported finding that ideological decisionmaking correlates with lower network cohesion.

Subsequently, and as described supra, Fowler, and Jeon (2008) described a method for the determination of authority and hub scores using network data derived using all Supreme Court majority opinions (30,288 in all) that were issued between 1754 and 2002.

In order to rank those Supreme Court opinions with respect to the precedential significance that each one exhibited, the authors aggregated the latent judgments in the network using the number of times that an opinion is cited, and a separate measure for the quality of the opinions that cite to it. The authors determined that the opinions that their approach identified as having higher authority scores were significantly more likely to be found on "landmark" decision lists that had been promulgated by political scientists and legal scholars and designated as being "important" and as having "salience."

Authority scores also appeared to be reasonably good at predicting which opinions would gain importance in the future. Because these Fowler-Jeon authority scores were generated without any dependence on the content of the decisions, they are without any ideological biases (The same study also tested the rise and fall of opinion precedent over time). More recently, Robinson (2010) utilized Epstein's Judicial Common Space ideology scores and network relevance data to test his hypothesis regarding the impact of opinions by the Rehnquist Court.

3.3 Hypothesis

The Null Hypothesis would be that the Fowler Authority scores of opinions distribute randomly, and that zero correlation would be observed between the authority of opinions and the Martin-Quinn scores of the justices who respectively wrote each. The hypothesis that this theory dictates is that justices whose Martin-Quinn scores for a given

term fall further from the center will author opinions that eventually become less central in the network (and obtain lower Fowler Authority scores) than opinions authored by justices whose Martin-Quinn scores place them closer to the center of that index. Martin-Quinn scores are taken as absolute values throughout because it is not considered to be relevant for this study whether a justice tends to be conservative or liberal. The focus here is on the effects of greater ideology, not on the effects of one variety of ideology as opposed to its opposite variety.

3.4 Data and Methods

The analysis of the justices was bounded at the early point by the start of the Martin-Quinn scoring system in 1946 and at the late point end by the Fowler Authority scoring system in 2002 terms. All of the attributed opinions for each justice were compiled from the 1946 through 2002 using Spaeth's Supreme Court Database (this fifty-six term segment ran from the seventy-seventh justice, Stanley Forman Reed, to the one-hundred-and-eighth justice, Stephen Breyer, and thereby encompassed just under thirty percent of the justices who had comprised the Court prior to 2003), and the corresponding Fowler Authority scores for each were used to create a by-term mean opinion authority score (simply extracting a full list of each justice's opinions each term and affixing the opinion's Fowler score and then calculating the mean). By-term Martin-Quinn scores were then sifted into the composite data, and also pulled out for by-justice/by-term analysis.

That aggregation of data provided a by-justice, by-term data set that allowed for a simple comparison of each individual's authority scores over time. Justices could be individually studied, compared to each other in pairs or smaller groups, and considered in

a global fashion. Because Martin-Quinn scores could not be assigned, per curiam opinions were eliminated from consideration.

An OLS regression was run over the entire span of data: 6,015 opinion authority scores both against Martin-Quinn scores and against opinion age. Fixed effects were next added in order to try to account for the tendency of a straight OLS model to simply draw a regression line through a cloud of data points but to reveal little about the individuals within the system. With a fixed effects estimator in use it is possible to get a better handle on the individuals while still running a single test. Often one of the major drawbacks of using fixed effects can be the loss of explanatory variables that do not vary by individual. In this instance at least, that is not a problem because there are no additional variables being utilized. Further work on this topic should start by including more variables to determine with greater accuracy what is accounting for the variance in authority scores.

3.5 Results

For the simple model each incremental unit of increase in the absolute value of a justice's Martin-Quinn score shaves 0.00027 from an opinion's authority score with significance at the $p < 0.001$ level (once again, the absolute value of the Martin-Quinn scores were used so this number could be interpreted for justices from both sides of the political spectrum). Given the mean authority score for the set of 0.00382 this represents a seven percent reduction effect (Table 3.1). Regression results are in accord with the hypothesis that a greater degree of political bias for a justice is apt to have a negative impact on the authority scores of the opinions authored by that justice.

Table 3.1: Full Court Combined OLS Regression

Coefficient	Estimate	p
M-Q score	-0.000267786	.00168**
Age	0.000103229	< .001***

Note: N = 6,011.

For individual justices the results are more difficult to determine. Twelve justices return results that are significant at the p < 0.001 level (see Table 3.2). Among those justices the impact of a one-increment move in the Martin-Quinn score ranges from Justice Marshall at 0.0029 to Justice Goldberg at 0.0059 (Justice Byrnes was omitted by the fixed effects model). Unfortunately, the direction of those impacts is not signed in the direction that the theory would expect.

Table 3.2: OLS Regression with Fixed Effects

Coefficient	Estimate	p
M-Q score	-1.004e-04	.495548
Age	0.00019312	< .001***
Reed	-.00006219	.885245
Frankfurter	-0.00012638	.885245
Douglas	0.00023387	.771989
Murphy	-0.00465647	.003357**
Jackson	-0.00240453	.024056*
Rutledge	-0.00303175	.083673
Burton	-0.00257796	.015502
Vinson	0.00034764	.759936
Clark	-0.00107021	.178525
Minton	-0.00335443	.005765
Warren	0.00412464	< .001***
Harlan	0.00193761	.021872*
Brennan	0.00555850	< .001***
Whittaker	-0.00082825	.568083
Stewart	0.00447514	< .001***
White	0.00361166	< .001***
Goldberg	0.00594075	.000143***
Fortas	0.00301441	0.043584*
Marshall	0.00292427	< .001***
Burger	0.00432199	< .001***
Blackmun	0.00313702	.000244***
Powell	0.00496310	< .001***
Rehnquist	0.00358487	< .001***
Stevens	0.00301512	< .001***
O'Connor	0.00337538	< .001***
Scalia	0.00313017	.003195**
Kennedy	0.00339301	.002789**
Souter	0.00300764	.017732*
Thomas	0.00337747	.007706**
Ginsburg	0.00331062	.015896*
Breyer	0.00337397	.020966*

Note: N = 5,981.

3.6 Conclusions

While the observed effect of the OLS regression is statistically significant and has some strength (a single step on the ideology scale moving the authority score seven percent), and in the direction that the theory predicted is should be, it could certainly be more pronounced.

A partial explanation as to why a stronger effect was not observed is that the selection process for writing opinions that the Supreme Court follows is subject to selection bias. Typically, for matters that are to have opinions written regarding their outcome, the Chief Justice will pick the author (provided that the Chief Justice is a member of the majority). If the Chief Justice is in the minority, then the right to select the author passes to the most senior justice who is a member of the majority. In the alternative, if the Court is heavily Balkanized (with three or more distinctly separate voting blocs), then the various factions will generally hash out amongst themselves who will write their opinion.

This system dictates a Chief Justice who is a member of a Court that has a majority that is aligned with him will be making a significant number of selections. The authority to assign opinions will next most frequently fall to the senior justice in the opposition bloc. Junior justices will rarely, if ever, have an opportunity to assign majority opinion writing duties. Each justice always retains the right to author a dissent, but dissents are not considered to be opinions and do not figure into this analysis.

Given those parameters, a pattern could easily form that would skew the results in a study of the relationship between centrality and ideology. A justice who is going to assign a significant number of opinions in a term is going to be mindful of the fact that, in order to function efficiently, the Court must balance the workloads of each justice. A

chief or senior justice who is going to be assigning fifteen to fifty opinions in a term is also going to be mindful of which justice is writing on what topics. As justices will be aware of the political ideologies of their cohorts, and of the degrees to which each will be willing to go to influence the law, strategic reasoning will doubtless come into play.

Presented with a group of justices with a range of ideologies, it would be unlikely that a selecting justice will place a justice with the most extreme viewpoint on a given legal topic in the position of writing the majority opinion concerned with that topic (unless perhaps that assigning justice were equally extreme on that issue). The reason is simply that it would lead to at a minimum four other justices wrangling to pull the more extreme justice back to where the "center of mass" was on that majority. If that strategy was in effect, the justice assigning an opinion would often aim to give the more extreme members of the Court opinions that dealt with less inflammatory issues, those where the extremism would be more easily managed.

This hypothesis could be tested if a large enough data set were gathered that went term-by-term and matched justices' Martin-Quinn scores with the legal topics with which they were assigned to author majority opinions. Some simple testing could establish if the topics were assigned randomly, or it there was a higher likelihood of "hot button" topics going to the more 'middle of the road" justices, and the more anodyne topics going to the more extreme justices.

That possible explanation is, of course, overly elaborate, although, it appears to have a kernel of truth to it. With respect to this proposed explanation, the null hypothesis that would need to be tested in order to dispense of it would be that each justice has an exactly equal chance of being selected for any opinion-writing task, regardless of their

political ideology, regardless how extreme (not unlike the null hypothesis of this subproject that ideology would have zero impact on eventual opinion centrality). That is an unlikely outcome, but the challenge is in seeing if the strength of the predicted significant effect is sufficient to account for the slightly muted results that have been observed.

A further avenue to explore is the cumulative effect of extended periods of governmental control by only one political pole. Under such conditions as legislation, executive action, and court doctrine accrete, they will collectively shift the reality of where the "center" is (McCarty, Poole, and Rosenthal 2016). In that scenario, one that has arguably been going on in the U.S. for much of the past thirty-six years, justices on the prevailing side (in this case the conservative side) will likely have their opinions become more central, whereas those on the side that is in retreat (the progressives in the modern era) will be likely to see their opinions be less central/exhibit lower authority scores. Moreover, using absolute values of Martin-Quinn scores, as was done here, will cause the mean of all scores to be located in a trough between the two means for each "side." Further work will be necessary to determine if this potentially confounding factor is actually having an impact on the results in this study.

The wrinkle in this particular corner of the study is that the opinions that were generated at the start of the so-called "Regan Revolution" have, after thirty-five years, only just begun to settle into what will likely be their (more-or-less) static values. A further wrinkle could also come about if Hillary Clinton wins the 2016 Presidential Election and is able to establish a putatively liberal majority on the Court that enjoys a long period of stability. The next president might have the opportunity to nominate as

many as four justices, replacing perhaps Ginsburg, Kennedy, Breyer and the still vacant Scalia seat. A return to a more Warren-like posture by the Court would, over time, be expected to boost the average authority of Warren Court opinions, and lower the average authority of Berger, Rehnquist, and Roberts Court opinions (although perhaps not for the 2015 term in which the Roberts Court appeared to take a rather progressive turn). Although centrality scores can and do tend towards becoming static over time, major upheavals in the alignment of the Court can alter matters to some extent.

Overall, given the results that are presented, this subproject is far from being a dead letter. There is more analysis that can be done to further develop the study, but what has so far been uncovered does advance understanding of judicial activities, and the role that ideology plays in the lives of opinions that the justices author.

CHAPTER 4

CAN WE TRUST WHAT THE ELITES TELL US?

Chapter 5 presents a study that built a citation network of Supreme Court opinions in which the majority made the statement "We are persuaded" with respect to some argument that had been offered. What follows in Chapter 4 is a brief discussion of the effects of taking that statement as being a truthful expression of the majority's actual, collective state of mind. Although one may question whether or not the justices in the majority are faithfully reporting their intellectual engagement with the presented legal rationales encountered, there are good reasons to take at face value what is proffered in this specific context.

4.1 Introduction

A significant section of this project concerns the citation network that emerges among Supreme Court opinions in which the majority makes the announcement "we are persuaded." The data sets were gathered and the statistical analysis that was performed generated empirical output. The process was mechanical as the inputs were subjected to manipulation and the outputs were dutifully logged and reported herein. The question of what particular normative significance those inputs hold, if any, is a different matter.

Anecdotally presenting a small sample of attorneys with the proposition that instances of the Court making the affirmative statement "we are persuaded" should be assigned at least provisional relevance—a preponderant probability that, indeed, some "thing" did literally persuade the justices in the majority to realign their beliefs in some non-trivial fashion—did not faze any of them. Several political scientists, on the other

hand, questioned this. The notion that an elite, on some rare occasion, perhaps actually says exactly what they mean—or that an investigator could rely to any degree upon the verity of such a statement by an elite—led to some genuinely spirited debates. I may also have failed to fully explain my position which, stated succinctly, was not that we should take all judges at their words in all circumstances, but rather that the Supreme Court's use of the "we are persuaded" language was a specific instance where doing so makes some sense.

Nevertheless, that gut-level revulsion is somewhat curious, as there is a fair amount of support in mainstream literature within the field of political science (as well as in economics) for exactly the proposition that when political actors speak, they are conveying at least some useful information about their views (Austen-Smith 1990— delving into a game theoretical model of committee decision making; Diermeier and Fedderson 2000—arguing that congressional hearings may not be informative to committees but may provide crucial information to the floor.; Black, Treul, Johnson and Goldman 2011—testing comments from the bench during oral arguments and finding them predictive of voting). The standard in psychological research is that explicit statements by subjects that they have been persuaded with respect to a subject—such as the ones collected for this study—are a reasonable basis for measurement of the phenomenon (Gerber, Gimpel, Green, and Shaw 2011—using a simple survey mechanism as evidence of persuasion, Bader 2005—simple statement of having been persuaded deemed evidence of persuasion taking place). But perhaps personal demons can be temporarily pushed aside—along with the literature just cited—and the question of

what weight should be given to "we are persuaded" statements that have been made by justices can be approached naively, so that some light can be shed.

Some of the heat that the question of relevance appears to generate is diffused by the reality that the "we are persuaded" statement need not be given any credibility whatsoever for the ensuing study to have baseline validity. The focus on the "we are persuaded" language (designated as the WAP data set) takes the set of majority opinions that invoke the term and next examines the network of citations that emanate out from that primary set. That examination involves comparing the indegree for citations from WAP opinions against other stratified sets of opinions (one set that is matched by U.S. Reporter volume number, the other matched by volume number and by opinion author; in each instance where more than one opinion could have been a "match" randomization was used). While the study itself was motivated by a theoretical model based upon the inferred meaning of the term, any other 3-gram could have the same operations performed upon it—the actual meaning—or lack thereof—that the majority ascribed to the words does not really impinge in an absolute sense upon the utility of the network of citations or the measurements that were recovered from the edgelists that were generated post hoc.

If we start from the position that, arguendo, the issue of actual meaning that the majority of the Court ascribes to "we are persuaded" should be addressed at the outset, then the political scientists would likely have, far and away, the best showing in the ensuing donnybrook. Starting the analysis of the question at the broadest level, beyond the bounds of our field, unmitigated skepticism—of the actual, philosophical brand of applied skepticism proffered by David Hume and his intellectual kin—suspends virtually all belief and brooks no dissent. The hard-core epistemological argument that none of us

can ever really "know" anything proves frighteningly difficult to upend. Just as it is such a challenge to rigorously establish that this is my hand (Moore 1925), that cats do not grow on trees (Judge Richard Posner in United States v. Andrea Hall and Richard Magnant, 854 F.2d 1036 (1988)), that a duck is not simultaneously a rabbit (Wittgenstein 1953), that ghosts do not exist (Douglas Walton 1989), or that the entire Universe was not fully formed and set into motion just this past Thursday (Russell 1921), we cannot ever truly know that a given use of the statement "we are persuaded" is not a subterfuge employed by clever justices to deceive readers when, in fact, the argument in question failed to sway them (the justice individually or the majority collectively) in any way.

4.2 Political Science Really, for Real, in the Real World

Not an extension of Pyrrhonism, the empirical end of political science instead aligns more with Popper's Fallibilism (1934). Once we collectively assent to the possibility that we can aim to "know" some things in a loose sense—so long as we are willing to revise our beliefs once new evidence is obtained—our possibilities open up considerably. In the present case we now strive to locate some clues that reveal that a justice is faithfully reporting the truth when they make the claim "we are persuaded." Those clues arrive in three main baskets: the logic of rational choice; the pattern of general structural support that is presented within the opinions that undergirds the claim; and the relative infrequence with which the claim "we are persuaded" itself has been made.

That set of arguments, which will be reached shortly, disregards a series of other approaches that, although arguably somewhat effective at isolating the verity of statements, are not practical for the considered application. Perhaps some day a number

sitting judges (or even retired ones) might willingly submit to fMRIs during the decision making process, but that time is still in the offing. Denied access to that goldmine of data, we must move forward with the tools that we do have available at this moment in time.

Green and Shapiro (1994) did not slay rational choice theory, but they certainly succeeded in neatly encapsulating many of the weaknesses that pervade it to the point that their description of the area as having "pathologies" is somewhat apt. For all of the shortcomings that rational choice has, conventional wisdom holds that (as with its close cousin, game theory) its best chances for offering sound, well grounded explanations of the real world tend to emerge when it is employed to examine small groups or rule-bound elites. Appellate courts in general and the Supreme Court in particular, fit that description well.

In the current context, if the Court's majority states that a specific argument has persuaded them with respect to a specific legal issue, they have made an affirmative decision to highlight that instance—to deliberately draw the reader's attention towards that particular thread. The justices of the Supreme Court live in an environment that is full of both formal rules and normative expectations with respect to the work product that they produce. Although there is no actual requirement that they must elaborate on their decisions—that is upon the simple by-justice votes that establish which side "wins" each matter that the Court hears—it is all but a mortal lock that each term opinions will issue that will be both detailed and at times voluminous. While it would be something of a surprise to discover that an individual who was appointed and confirmed to the Court happened to be uncomfortable with written expression, some have reputedly taken more

relish in the job of authoring opinions (Scalia) than others (Blackmun, Marshall) (O'Brien 2008)).

What can these simple observations tell us about the trustworthiness of mentions of persuasiveness? Although skill with the law and persistence with the pen are (we hope) requisite traits of individuals who are successful in securing seats on the Court, there is no reason to expect that justices take enjoyment in suffering through unnecessary writing tasks, nor that any of them would be in the habit of fashioning rods for their own backs. It is established that explaining their thinking in at times forensic detail is a norm of the Court. As claiming that some line of reasoning was notable for its persuasiveness naturally invites elaboration, it would then be logical to conclude that some sort of conscious thought process must weigh the cost in time and mental exertion to foregrounding those persuasive characteristics against the benefit of detailing the persuasive nature and power of a given argument.

This line of reasoning dovetails perfectly with the arguments made by Epstein, Landes, and Posner in their study The Behavior of Federal Judges (2013). Their approach was to model judges as participants in a labor market wherein they are both motivated (and constrained) by costs and benefits, some of which are pecuniary but many of which are not. The latter nonpecuniary costs (such as effort, criticism, and workplace tensions) and nonpecuniary benefits (such as esteem, influence, and leisure) become the focus of the three authors' positive analysis which seeks to make extensive use of data only to answer not how judges should decide their cases, but also how they do decide them.

Many of the later chapters in The Behavior of Federal Judges hone in on effort aversion which "includes both reluctance to work 'too' hard—that is leisure preference—

94

and reluctance to quarrel with colleagues (conflict aversion). Both are aspects of the 'quiet life' that is especially valued by persons in jobs that offer little upward mobility—and in the case of a federal judgeship involve virtually no downward mobility" (Epstein, Landes, and Posner; p. 7). Although there are some fundamental weaknesses in the model that the three authors put forward—mainly that it is rather generic and could be applied without modification to any profession from optometry to occupational therapy to stand-up comedy—the core of their conclusions dovetail nicely with the fundamental claim that it would be senseless to expend the energy necessary to highlight and expound at length upon the "persuasive" nature of an advocate's reasoning if that reasoning had, in fact, failed to be in any way persuasive to the Court.

With regard to the pattern of general structural support that is presented within the opinions that undergird the claims in every one of the instances sampled for the WAP data set, there followed a significant and diligent effort to illuminate how the persuasion was achieved. Far from being a mere token tossed off and then abandoned, there was a uniform effort to justify the decision to move to the position that was eventually asserted. Consider first the effect that would be achieved if the Court took up a new position without offering any justification. It would not be chaos (at first anyway), but the taking up of new positions—if presented as being random, arbitrary, or capricious—would before long certainly cause consternation on many levels and in many quarters. If the process continued, and the uncertainty accreted, eventually even societal chaos could be threatened.

Potential litigants would be less and less certain if their considered courses of action would be likely to sail them into the rocks, advocates would be unable to offer

guidance, or to properly plan arguments for trial. If it all comes down to the flip of a coin then the lower courts would be unable to properly interpret precedent, executive agencies would drift rudderless, and legislators would have no idea which way their legislation would play out. It is perhaps a trite observation, but the need not for legal certainty in every circumstance, but for some reasonably reliable bounding of potential legal outcomes from litigation, is essential for our society to function in both a reliable and a prolonged way.

The frequency, or perhaps the infrequency, with which the "we are persuaded" claim has been made is also somewhat telling. In the time period sampled, from the 1946 term to the 2008 term, the majority used the phrase 143 times. That is not a large number of instances yet, at the same time, it is not a vanishingly small number either. The distribution suggests its parsimonious application in circumstances when other language would have failed to capture the nuances of the thought processes that the Court underwent in analyzing the legal arguments that were presented (keeping in mind that, as an appellate court, the Supreme Court is limited to review of legal arguments, as opposed to reviewing evidentiary assessments such as the credibility of a witness or likelihood that a document is authentic).

The Court chooses its words carefully because the words that it uses have a genuine impact on the legal system. If the Court proclaims "we are persuaded," that assessment will be scrutinized because virtually all of the Court's writings invite scrutiny. Were the claim "we are persuaded" made with regard to an objectively unconvincing legal argument, commentators would be keen to investigate. Scholars and journalists would be forced to ask "For what reasons is such a flimsy rationale convincing to the

Court, especially when they went out of their way to try to highlight the force of the argument?" Clearly the statement "we are persuaded" has the potential to act as a double-edged sword. When that language is utilized, the Court is portraying itself as open to reason and intelligent discourse, but it is also inviting a critical review of its capacity to logically tease apart the reasoning that advocates who stand before it present.

4.3 Conclusion

Even if we decide to shunt off peer reviewed claims of established scholars that statements by elites can be substantive statements of truth (supra), there is no rationale for members of the Court to make deceptive claims regarding what is persuasive to them. Claims that persuasion has taken place have invariably been supported by the Court's own detailed explanation of the basis for having been persuaded in each instance. The relative infrequent use of the "we are persuaded" language point towards a selectivity borne of actual concern for veracity. The justices are invested in accurately flagging those legal arguments that have effectively altered their perspectives on relevant legal issues.

CHAPTER 5

THE "WE ARE PERSUADED" NETWORK

A network of Supreme Court opinions that acknowledge persuasive argument was introduced and analyzed with all citing opinions added. A second set of random opinions that closely match the persuasion set with respect to the times of publication was also generated with citing opinions also added, as was a stratified set that matched the reference set with respect to both the United States Reporter volume number and opinion author. A possible theoretical account is offered regarding why the set of opinions that discusses persuasion exhibits greater network density than was measured among a set of randomized opinions, with the opinions that mention persuasion being more popular than those that comprise the randomized set. A Welch's t-test lends support to the hypothesis that the difference in indegree between the two sets is not due to random happenstance. Lastly, Quadratic Assignment Procedure is performed in an effort to determine which covariates are most strongly implicated as contributing to the observed measures. While the initial results are favorable in that the expectation was that opinions that mention persuasion would be more popular, the tension between the positive and negative expectations that are generated by instances where that persuasion is discussed has created a knotty problem that remains open to further examination.

5.1 Introduction

All law schools seek to introduce the elements of persuasive legal argument to their students, especially in the first year writing class (Miller, 2004). Nevertheless, it appears that the majority of trial attorneys simply develop their own sense of which

approaches best fit their innate abilities and "wing it" from that point forward. The results

do not necessarily impress, or effectively persuade, judges (Scalia and Garner, 2008). The

opinions of appellate courts regularly devote effort to explaining the merits of the

arguments that advocates have made, and also to accounting for the reasoning that the

final decision was predicated upon (Kelly 1996). On the infrequent occasions that

appellate opinions go so far as to state that a given argument has succeeded in persuading

the Court, those arguments are most often given significant attention by the court.

Properly interrogated, those instances present opportunities for researchers to gain

insights regarding effective approaches to legal argument. This study employs network

analysis which has been used effectively to analyze the interdependence and the trans-

mission of data among subjects, individuals, groups, and institutions (Ward, Stovel, and

Sacks 2011).

Network analysis is useful because it can be used to effectively model the

influence of actors on each other, using a variety of data sources including social

relationships, records of transactions, archival data, and interrelated citations, among

others (Lazer 2011). The utility of network methodologies in the examination of legal

citation networks (where two opinions are tied if one cites the other) has already been

demonstrated. Notable efforts in this area include the aforementioned Fowler and Jeon

(2008), using network data to establish a novel measure of authority; Clark (2012),

demonstrating how opinion content results from strategic interaction between justices;

Desmarais (2010), theorizing that current decisions are used to correct perceived

ideological biases of past terms; and, Katz and Stafford (2008), arguing that social

structure—operationalized as the professional and social connections between judicial

actors—partially directs outcomes in the hierarchical federal judiciary. Scholars have only just scratched the surface when it comes to applications of these methods and this new toolkit has the potential to shed more light upon how judges can overcome their personal political biases, how members of the same court influence each other's decisions, and the role that advocates play in influencing the outcomes of trials and appeals.

The central hypothesis under study here is that opinions that announce instances of persuasion should gather more citations than those that do not, all else being equal. This effect is anticipated because the forms of argument that have been previously accepted by courts should attract the authors of subsequent opinions who are seeking support for their reasoning, regardless of the commonality of legal issues that are present. To test this theoretical expectation a matching research design is used, where opinions that mention persuasion are compared to a random sample of opinions that are otherwise similar.

5.2 Theory

Right at the outset, consider for a moment the ways that judges interact with their professional worlds. Judges sit on elevated platforms and they (the judges, not the platforms) are draped in robes that differentiate them from other court personnel, public servants in general, and pretty much everybody else for that matter. Other people must rise when a judge enters or leaves a court room; judges are properly addressed as "Your Honor"; and, judges wield power on a daily basis that outstrips that of most all other individual members of society. Clearly there is a large investment made in maintaining the superior power of judges over the rest of our population that is a natural extension of

the logic of our three branches of government, and the reality that the judiciary is the ultimate backstop on many questions of public policy and certainly with respect to the enforcement of criminal sanctions. When people who have attained that sort of power take the time to highlight and explain what they found to be persuasive, it would be quite natural that others, especially those invested in shaping the opinions of judges, would bother to take note.

What is considered here in support of that expectation builds upon: 1) the broad claim that legal persuasion itself is a worthwhile area of inquiry; 2) the assertion that many instances of legal persuasion of judges can be isolated; and, 3) the conjecture that network analysis can be helpful in the exploration of legal persuasion because of the inherent links between opinions are, in fact, part of the process by which legal persuasion is accomplished.

Before exploring these three premises a brief discussion of the possible reluctance of justices to announce having been persuaded is in order. Judges may not want to announce that they have been persuaded because: 1) they hold positions of high prestige and it is a risk for a person in a position of high prestige to concede that they have been "steered" or "led" to some conclusion by someone of lower prestige (as opposed to simply having inherently known the "correct" answer ab initio); and, 2) when we observe such instances of persuasion the judge often sends a clear signal to account for, and even pre-defend, the acceptance of an advocate's argument. If those two suppositions are generally correct, the next step is to offer an explanation for why opinions that announce instances of legal persuasion will, in turn, become popular and then attract notably more frequent citation than would be observed in a random sample of opinions from a similar

time frame. Before moving ahead with an account of what can cause an opinion citation network to vary in a predictable way based upon occasions where persuasive arguments have prevailed, the assumption that there are sound reasons that legal persuasion is a worthwhile area of inquiry must be examined.

One could naively expect that having justices who are open to persuasion is a desirable situation (a normative assumption) and that, under such a regime, opinions that allude to persuasion should accrue more citations. In the context of the courtroom there is tension between persuadability as an asset, and persuadability as a liability. Each is considered in turn, as each can have an impact on the popularity of opinions in the citation network; the nexus between the two is noteworthy from a philosophical standpoint. As we do not often explicitly consider the ways that various norms can be mediated and transmitted into data points that can be captured, examples of such should be noted when isolated and exploited when possible.

Because the advocate-driven persuasion of Supreme Court justices is one of the primary mechanisms by which established legal doctrine is transformed and new precedents are established, it should follow that instances where persuasion is observed should be of keen interest. As explained (supra), it has been the standard in psychological research that explicit statements by subjects that they have been persuaded with respect to a subject—such as the ones collected for this study—are a reasonable basis for measurement of the phenomenon (Edwards 1990, Millar and Millar 1990). This approach does not fully address the reality that accepting self-reporting of attitudinal changes could carry along inherent veracity issues. For now the extensive buttressing seen throughout the WAP opinions (opinion authors unpacking explanations regarding their acceptance of

an advocate's arguments) has been taken as adding some support to the cautious

supposition that the Court's claims of having been persuaded are reasonably sincere (or,

at the least, defensible).

In the alternative, consider a court system that is suddenly repopulated exclusively

by rigid, willfully unpersuadable judges, who uniformly rely exclusively upon their initial

impressions and preexisting biases when making legal decisions. Such a system would be

bound to deteriorate over time as prior precedent, derived from logic and reason, is

replaced by individual hunches and prejudiced holdings. Although it is well worth noting

that a range of views have evolved regarding the interplay between judicial decision

making and precedent, there is no clear consensus: Segal and Spaeth (1996), finding that

systematic support for stare decisis is exceedingly rare on the Supreme Court; Richards

and Kritzer (2002), finding that precedents are only drawn upon to support ideologically

charged Supreme Court rulings; McAtee and McGuire (2007), finding that in salient

cases Supreme Court justices are less amenable to legal argument; and Braman (2009),

presenting both qualitative analysis and experimental evidence supporting a model based

on motivated reasoning. The alternate mechanism that would enable change in such a

scenario is the confirmation (or election) of judges who are also biased, but whose biases

are counter to whatever established precedents happen to be standing.

That judges should be open to intelligently presented arguments and be swayed by

logic appears to be a normative assumption in the United States (78 percent of

respondents agreed that judges "should be free of political and public pressure." Justice at

Stake Survey, 2001). Braman (2009) observes that:

> "The legitimacy of judicial authority in our democratic system
> depends, in no small part, on judges' ability to be neutral third-party

arbitrators of disputes between parties. The suggestion that personal
biases may impact their decision making, even if unintentionally,
raises valid concerns about the fairness of distributive outcomes in
our legal system."[7]

At the same time, justices have an understandable interest in not being perceived

as gullible. Indeed, justices could be stuck in a proverbial "no win" position. When a

member of the public wants them to maintain the status quo, they expect judges that are

rigid and hidebound; when they disagree with the status quo, they want justices to be

persuadable on those matters. Since the public is split on many issues, the best defense

for the judiciary would be to maintain that "law is the law" and that they are just abiding

by the texts that they are interpreting. That facade of neutrality presents a natural

inclination to not admit to having been persuaded. It could be concluded that when judges

reveal that their thinking on a matter has been significantly reshaped by a legal argument,

they hope to be perceived as measured and thoughtful, and not as being easily deceived

(although the author speculates that a significant correlation would emerge if the

preference for rigidity in the justices were regressed against the tendency towards

authoritarianism in the general population; a different study for a different day).

Because judges must be concerned with impression management (Goffman,

1956), these concerns have real resonance for the institution of the court (indeed the

highly theatrical nature of many trials dovetails well with Goffman's dramaturgical

model of self-presentation). Goffman's work has been a touchstone within Sociology for

some time, although it has been fairly criticized for focusing too heavily on illicit

behaviors (Schlenker, 1985). Simultaneously, other elements of Goffman's work have

---

[7] At page 5.

appealing aspects for the proposed model of judicial caution with respect to acknowledgement of persuasion. In particular, his explanations of Idealized Performing, Maintenance of Expressive Control, Misrepresentation, and Mystification each dovetail nicely into the present explanation that so strongly relies upon the notions of jurists as being tightly wound-up in the minute details of reputation projection and management.

Examined in the context of the justice system, what is observed in general is what would be classified as a frontside, protective style of self-regulated impression management (Schlenker and Weigold, 1992), as written opinions can be correctly regarded as proxies for behavior in live human interactions. The efforts that constitute judicial impression management go well beyond simply the careful and conscious ways that they express themselves in the opinions that they author.

Because judges derive a significant part of their power from the projection of infallibility,[8] it is understandable that when they write of having been persuaded they will often make an effort to buttress the admission by explaining the careful consideration that allowed them to arrive at their ultimate conclusion. Indeed, this behavior is uniformly observed within the texts of the opinions that comprise the "We are persuaded" data set that has been collected for this project. An example is Harbison v. Bell, 556 U.S. 180, 183-189 (2009), which is discussed in detail, infra.

Scholars who study the common law are aware that later appellate opinions invariably "talk" to prior opinions that have confronted similar issues (Sunstein, 1999). That ongoing "conversation" is the shared cornerstone of the world's common law legal

---

[8] Recall Justice Robert H. Jackson's statement "We are not final because we are infallible, but we are infallible only because we are final" concurring in the result in *Brown v. Allen*, 344 U.S. 443, 540 (1953).

systems. (Kelly 1994). Those connections, made up of jurists seeking to build upon the collective wisdom of prior generations, then form the opinion citation network that creates the skeleton for our legal system.

As previously noted, the question that spurred this project was whether there is a judicial norm that favors the announcement of accepting legal arguments as persuasive. Granted, perhaps the claim that the "We are persuaded" language is a "norm" for the Court is something of a judgment call. From 1946 to 2008 the exact phrase was used 143 times—a little more than twice per term on average—which is often enough to qualify its use as a regular occurrence, while at the same time keeping it in a range that could fairly be described as infrequent. The phrase has been utilized consistently over time, and arguably with sufficient specificity, that scrutiny of it is appropriate. Keeping in mind that the Court is under no affirmative burden to disclose instances of persuasion, the reality that those exact words are called upon with regularity in order to flag circumstances where a particular event has transpired tell the reader that the invocation of the phrase is a conscious and measured act. Given that, it is reasonable to conclude that since the Court has seen fit to flag such pronouncements those who follow the Court ignore them at their own peril.

Granting, arguendo, that the use of "we are persuaded" is a norm (and may be treated as a network node attribute), it is to be determined if, once exhibited, it could be observed as contributing to the popularity (a marked tendency of other opinions to cite to the "persuaded" opinions more frequently than to opinions that do not disclose instances of persuasion) of opinions to a degree that the density of the network increased significantly. That would be the case if there are repeating patterns of legal analysis that

pervade similarly reasoned decisions. Those parallel lines of legal reasoning would then invite the citations to connect such correspondent opinions (citations here being treated as a pattern of behavioral social interaction within the four categories defined by Kitts (2014)). It would, however, appear that such a series of connections, especially if indirect, would elude the type of legal-topic centered, search-engine based inquiries that legal professionals typically perform (Mersky, Barkan, and Dunn, 2009). We would expect WAP cases to draw more cites than non-WAP cases, regardless of whether the later, citing cases explicitly mention persuasion.

In comparison to the entire population of all opinions decided by a given court, the density of ties among the opinions (the incidence of opinion-to-opinion citations as divided by the maximum number that is theoretically possible) and centrality (how many other opinions cite a given opinion) would be expected to be higher among the network of opinions that explicitly mention judicial persuasion. Such results would be explicable if the authors of opinions that explicitly disclose persuasion were motivated to rely upon prior opinions that also disclosed persuasion (the primary disclosure set), or upon opinions that are tied to the primary disclosure set through direct citation. That reliance would stem from the presence of common structural expositions of the legal arguments that were presented, supported, considered, and ultimately deemed to be winning in the prior opinions and their progeny. Naturally, further work on the question would seek to establish if these opinions are used by judges as blueprints that function as a second level of precedent and provides jurists with necessary support for a given conclusion.

Throughout this paper it is the convention that a majority United States Supreme Court[9] opinion that explicitly announces "we are persuaded" is eligible for consideration (excepting the occasions where the text is embedded in a quote from a different source). This acknowledges that any majority opinion is likely the end product of close collaboration among those in the majority, and that the representation of persuasion concerns a group of justices who are like-minded on the issue of what convinced them to reach the given conclusion (Schwartz 1996; also, for an actual example consider the extensive discussion of the persuasive elements of argument in Harbison v. Bell, 556 U.S. 180, 183-189 (2009), presented in Appendix A.

Supreme Court justices possess deep knowledge of relevant case law and it is assumed that they are attuned to the emergence of symmetrical lines of reasoning, even if those parallels evolved in topically disjointed matters. Given that the disclosure of persuasion looms as a doubled-edged sword, perceivable as either a sign of a reasonable mind, or of an overly credulous one, we should not be surprised to find judges making such a declaration to seek necessary support where it resides in previous opinions. Thus, judges and justices have several parallel incentives to bring in previous occasions where persuasion was implicated into the opinions that they author that go to the same circumstance. There is a logical basis for the supposition that judges seek to buttress their opinions in that it is rational to try to shield oneself from criticism by aligning with others who have previously taken the same course; indeed, this is one of the core elements of

---

[9] It is of note that almost just over twenty-five percent of the opinions in the WAP set were authored by Justice Stevens, a result that is possibly a reflection of the reality that Justice Stevens is undeniably a person who is quite comfortable in his own skin, and who therefore would reside within the group that Schlenker and Weigold (1990) identify as autonomous and independent.

the logic that established the common law tradition originally. There are psychological

reasons that could explain the formation of such an alignment, notably the desire to

appear to be in step with the prevailing justifications that others have taken (Kitayama

and Burnstein 1996). Moreover, there is a linguistic norm at work here as the data

exposes that over the past six decades numerous Supreme Court Justices have made use

of identical language when speaking for the majority and explicitly announcing that a

persuasive argument had been presented to the Court. While the decision to concentrate

upon Supreme Court opinions for this study is in line with the bulk of the literature on the

topic of judicial attitudes and while research arguing that precedent has little influence on

the Supreme Court is freely acknowledged (Segal, Cover 1989; Segal, Spaeth 1996), it is

manifest that precedent can be separate from persuasion—that a justice could be

persuaded without reliance upon prior precedent.

It is anticipated that a higher density of ties would be observed among appellate

opinions related only by the explicit disclosure of the court having been persuaded by

legal argument. At the outer limits such a result would perhaps reveal new and potentially

gainful approaches for advocates seeking to establish new law. Thus, judges who author

opinions that explicitly disclose persuasion are motivated to cite to prior opinions that

also disclosed persuasion, or to opinions that are in citation alignment (i.e., later opinions

that themselves cited to the prior opinion that disclosed persuasion), because of the

common structural explanations regarding how the legal arguments were presented,

supported, considered, and ultimately deemed winning to establish a blueprint that

functions as a second degree precedent and provide a jurist with necessary support for a

given conclusion. If the underlying model is correct, these citations would be made

109

among opinions that concerned justifiable, legally motivated instances of persuasion, although not all of the opinions would necessarily make use of the precise "We are persuaded" language.

5.3 Hypothesis

The central hypothesis presented is that the network of citations among opinions that use the term "we are persuaded" will be denser than the citation network among the entire population of opinions.

5.4 Data and Methods

A free-text search on the Westlaw legal research database collected a data set of United States Supreme Court opinions (1946-2008) that explicitly announce that the Court was persuaded of an argument's validity (the WAP data set). Further opinion-by-opinion review established that in each instance the language was the Court itself speaking (not a direct quote from some other source), and that the term was not being used within a counterfactual argument. Each of these WAP opinions was then forward cite checked in the Westlaw database, and each subsequent citing opinion that was returned was then entered into the edge list to form the full network.

The primary opinion issue—with "primary" in this instance meaning the main legal issue with which the given opinion is concerned—was collected for each opinion and made a vertex level attribute for Quadratic Assignment Procedure (QAP), a nonstandard-error-based test of coefficient significance which determines the likelihood of an edge forming in the network (Dreiling and Darves, 2011). Several other control variables were also collected to further develop the model. The term in which an opinion was decided has a measurable effect on the likelihood of edge formation as there is a

significant positive relationship between the age of an opinion and the number of citations it is has received (generally opinions will continue to rack up citations over time (with non- negative citations being more frequent overall), but that rate tends to decay (Black and Spriggs 2009)). The page length of the decision was added in as a control variable because—when the data set itself was being compiled—it immediately became obvious that longer opinions strongly tend to attract more subsequent citations (this particular observation is a simple extension of the inevitability of any discrete legal pronouncement by the Supreme Court all but inevitably drawing a citation from some future Court later in time and each additional page in an opinion will inevitably draw out more discrete legal pronouncements). The Majority Opinion Author, simply the justice who authored the opinion, was also included. In some instances an opinion is presented as being "Per Curiam," meaning that it is the opinion of all of the justices who comprised the majority presented as a unit. The author of an opinion is considered relevant as opinion writers have a propensity to cite to their own prior opinions more often. Per Curiam and unsigned opinions were coded as such. Whether an opinion is currently "good" law (able to serve as precedent without qualification) or "bad" law (having been overruled in at least some part) was added into the node level data set. Opinions that have been subsequently overturned—will likely fade into obscurity—whereas opinions with positive histories that remain "good" law are likely to keep being cited.

After the data had been extensively cleaned, plots were generated to assist in understanding the time dynamics of the citation network over time. Next, several descriptive statistics on the network were calculated. Finally, the network was modeled to

obtain estimates of the distribution of the network coefficients using the aforementioned

QAP.

While there have been prior efforts to explore the citation corpus of the Supreme

Court (Bommarito, Michael J., Daniel Katz, and John Zelner 2009), the author is not

aware of any data set of Court opinions that reference a particular key term.

5.5 Quadratic Assignment Procedure

Quadratic Assignment Procedure (QAP) is a hypothesis testing method for simple

and multiple regression models that are based upon dyadic data (Dreiling and Darves

2011); although not commonplace, the QAP methodology is well entrenched in political

science research (Grossmann, 201; Cha and Choi, 2012; Miles, Aflaki, and Petridou,

2015; Paik, Southworth, and Heinz,2007; and Peoples and Sutton, 2007). The complex

interdependencies exhibited by dyads (the micro-level units of analysis) create the basic

difficulty that frustrates efforts to make statistical inferences with respect to networks.

Because there is a probabilistic relationship among the edges of an observed network

such that network properties (e.g., transitivity, where an edge from a to b and also an

edge from a to c predicts an edge from b to c, thus, a friend of a friend is a friend tends to

be the rule) can be predictive of edge formation in a given instance based upon the values

of edges throughout the rest of the network, an opportunity to perform network inference

exists. Where misspecification threatens to undermine the validity of results, breaking the

correct specification of network effects is a possible approach. QAP takes that tack and is

able to function where there is only limited knowledge of network dependencies and

provides an empirical distribution of the coefficients simulated from a null hypothesis

model (Krackhardt 1988). In a QAP model, the dependent variable is an adjacency

matrix, a type of matrix which describes a graph by representing which vertices are adjacent to which other vertices. Thus, the i,j element of an adjacency matrix is 1 if case i cites to j and zero otherwise.

To run the QAP model, row and column values of the dependent variable matrix are randomly permuted while an independent variable is kept constant and an estimate is made after each successive iteration thereby providing estimates of the distribution of the coefficients (Dreiling and Darves 2011). The permuted data corresponds to the null hypothesis, so when the estimate is run the coefficients and statistics will be values from the empirical sampling distribution under the null model. If the original coefficient is not in an extreme percentile of the distribution under the null, or if the original coefficient is an extreme percentile of the distribution under the null model, the null hypothesis is rejected.

5.6 Two Alternate Sets of Opinions Are Built for Comparison

As a basis for comparison, an additional group of opinions were generated and analyzed. This second stage of data collection mirrored the first, however the selection of the opinions for this comparison set was aligned opinion-by-opinion with respect to the volume number of the U.S. Reporter, with the opinion selected within that Reporter based on a random number generated using R. Thus, if there was a single opinion from U.S. Reporter Volume 442 in the WAP set, and there were thirty-two opinions in Volume 442, a random number from one to thirty-two was selected and that opinion was added to the stratified/random (admittedly an oxymoronic sounding description, but the set is simultaneously each) set. As with the WAP set, each primary random set opinion was then checked for citing opinions and those were also added to the random opinion set. For

113

each set, WAP and stratified/ random, an edgelist and a node level data set were

compiled that integrated the Westlaw data (the status of an opinion as "good" law as of

April 2013) along with several variables taken from the Supreme Court Data Base

(majority opinion author, term, and primary issue). Next a secondary stratified/random

data set was created wherein each opinion was matched with the WAP data set both for

volume, and for opinion author. This data set was also forward cited with the later

opinions building the network.

5.7 Results

5.7.1 Visualization of the Networks and Measures of Degree

A notable feature of the directed WAP network is the prominence of "authorities"

within its structure. In the general case authorities are vertices that contain useful

information that then draw many citations from other vertices (the natural complement is

the "hub"-type of vertex, those that point to where the best authority is to be found within

the network (Newman, 2010)). In the present study, the opinions that exhibit the highest

levels of authority are those primary opinions (i.e., those opinions that were selected by

virtue of the thaumaturgic "we are" persuaded language) that were ultimately cited to by

a significant number of later opinions (although not every WAP opinion developed into

an authority).

$$2 \rightarrow A \leftarrow 1 \leftarrow 3 \rightarrow B$$

Figure 5.1: Thus, primary opinion A is cited to by secondary opinions 1 & 2 while
secondary opinion 3 cites to both primary opinion B as well as secondary opinion 1.

Figure 5.2 depicts a graph of the WAP network.
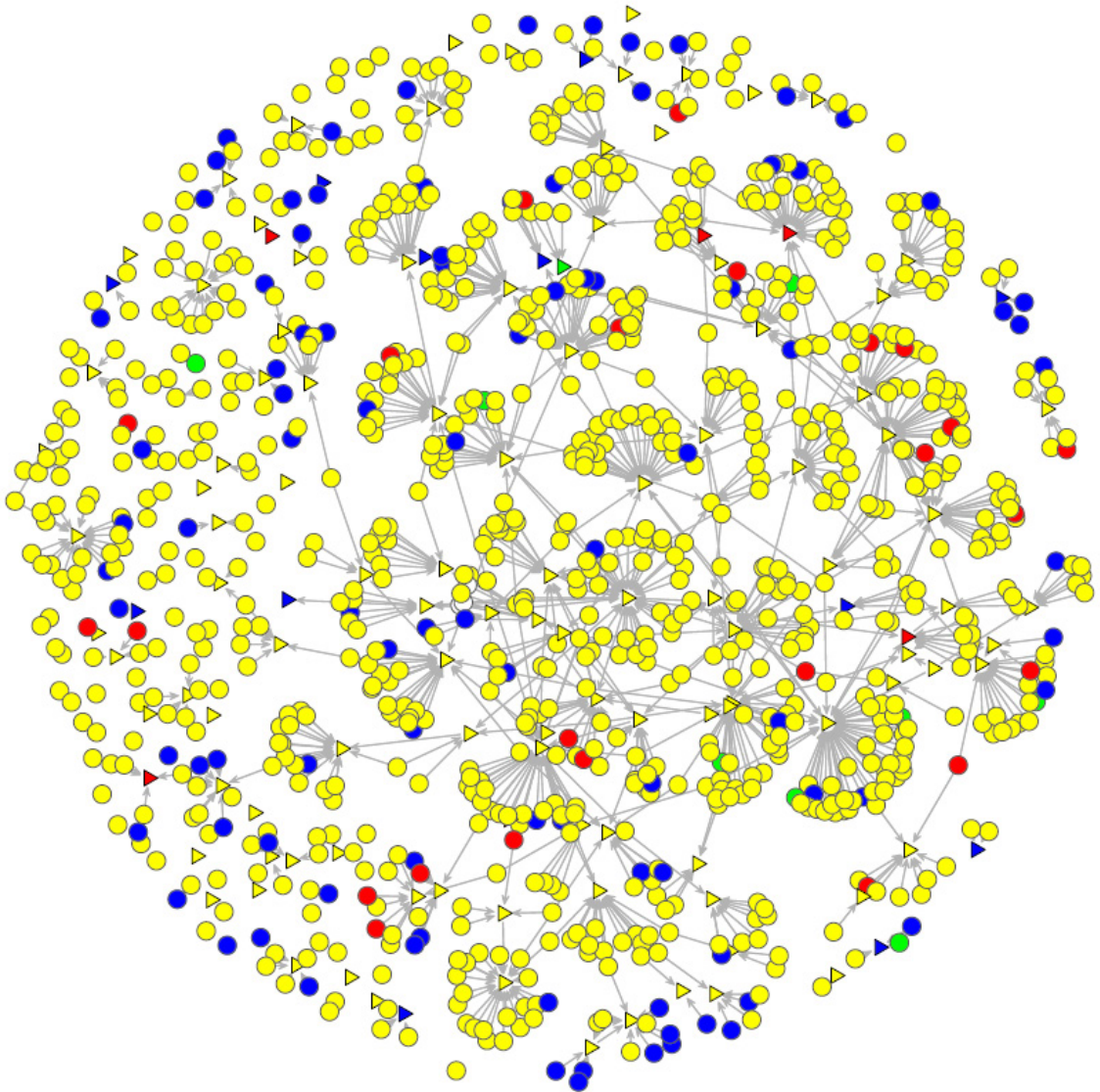
Figure 5.2: Opinions explicitly mentioning persuasion and all subsequent cases citing to them (red = "bad" law, yellow = criticized, but still "good" law, blue = some neutral history, green = "good" law, and white = no current history.). Triangles denote opinions from the primary "we are persuaded" set, circles denote subsequent citing opinions.

Proceeding with the construction of the graph in a term-by-term fashion revealed

the dynamics that transpired over time produced the final network that is described. Prior

to being converted into a plot, various attributes of the WAP data stood out quite clearly.

Variation in terms of the number of opinions that cited to each of the primary opinions is

presented with the actual count ranging from 0 to over 50 citing opinions in some

instances. Also, the number of non-negative cites is far greater than the number of

negative ones.

Figure 5.3: A side-by-side histogram comparison of the indegree totals of the primary
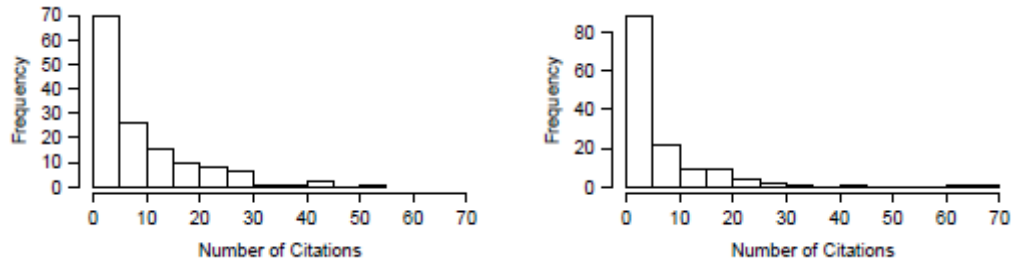WAP opinions and the indegree totals of the primary stratified/random opinions. The
primary WAP opinions drop more gradually than the random opinions. Also note that the
random opinions appear to have two extraordinarily popular, outlying opinions that
would have narrowed the differences in the density measurements between the two sets.

As one would expect, more recent opinions tend to have fewer cites, as is also the

case with opinions that have been overturned, although the latter condition appears to be

less determinative. There are occasional opinions in the secondary group that did connect

to multiple primary WAP opinions through indirect pathways.

A further exploration of the indegree (a count of the number of later opinions that

"reach back" to cite a given prior opinion) and outdegree (the complementary count of

the number of earlier opinions to which an opinion has cited) helps to reveal the structure

of the network.

In terms of distributions, the average opinion in the WAP Network exhibits a

small number of citing opinions (i.e., each opinions indegree) and each successive

incremental step up in the number of citing opinions generally follows a decaying

function. While the primary opinions in the WAP Network do cite to other opinions,

those were not collected in the data set so there are a number of opinions that do not cite

116

to any others in the Network. A significant number of opinions included cite to one other opinion, and a decaying function is observed to the right-hand side of the peak. A further exploration of the indegree (a count of the number of later opinions that "reach back" to cite a given opinion) and outdegree (the complementary count of the number of earlier opinions to which an opinion has cited) helps to reveal the structure of the network.

Turning next to the observed means, the average primary opinion in the WAP network (i.e., the opinions that actually contain the "we are persuaded" language) exhibits just over nine citing opinions (a mean indegee of 9.2). Each successive incremental step up in the number of citing opinions generally follows a decaying function. The primary random opinions have a smaller mean indegree of 6.7. This indicates that the primary opinions with the "we are persuaded" language exhibit greater popularity than the by-volume matched set does.

The secondary stratified/random set, containing opinions matched both by volume and opinion author is compared to the WAP set in Figure 5.4.
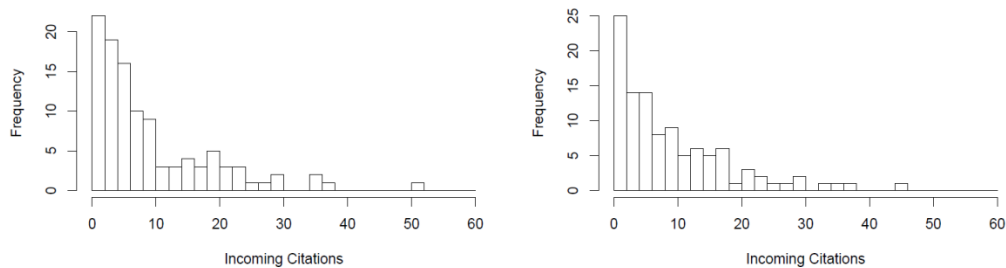

Figure 5.4: A side-by-side histogram comparison of the indegree totals of the primary WAP opinions and the indegree totals of the secondary stratified/random opinions. Here, the primary WAP opinions drop less gradually than the secondary stratified/random opinions.

117

5.7.2 Variation in Edge Density and Welch's t-Test Results

The WAP data set is approximately one-third denser than the random data set:

- Random Set Density: 0.00063

- WAP Set Density: 0.00082

- Difference: 0.00019

For raw comparison purposes, the highest issue-based density measured is for opinions clustered around reproductive autonomy at 0.0062, just under a tenfold increase from the random set (exhibiting a p-value of 0). The WAP cases were expected to draw more cites than non-WAP cases, regardless of whether the citing cases explicitly mention persuasion. As previously noted, although the difference in density is not staggering between the random set and the WAP set, it is present. The phrase "we are persuaded" would logically be featured in legal disputes that would naturally yield longer opinions. QAPY (a QAP permutation test using y-permutations) and QAPX (a permutation test using x-permutations) results over 1,000 iterations are affixed as a table in Appendix B. The phrase "we are persuaded" would logically be featured in legal disputes that would naturally yield longer opinions. Opinion Writer Homophily effects appear to not be significant, allowing the guarded claim that authorship of an opinion is not a necessary element for a model seeking to explain the observed effects. Of the other measured homophily effects, Issue is clearly the strongest which accords well with all established understandings of court opinion citation networks. The number of pages in an opinion also stands out as clearly significant as does the status of an opinion as "good law" (although to a lesser extent).

The Welch's unequal variances t-test comparing the indegree of the WAP data set to the secondary stratified/random by-author data showed that the variation in mean

indegree between them was within the statistically expected range (exhibiting a p-value

of 0.6772). Thus, there is no significant difference between the two.

5.8 Conclusions

For raw comparison purposes, the highest issue-based density measured is for

opinions clustered around reproductive autonomy at 0.00627, just under a tenfold

increase from the random set. As previously noted, although the difference in density

between the random set and the WAP set could be seen as minor, it is present. With

sparse networks such as Supreme Court citations, the differences in density are expected

to be limited in range. It was noted that the documentation for the **sna R** package

acknowledges that "interpretation of quantiles for single coefficients can be complex in

the presence of multicollinearity or third variable effects." That warning combined with

the low Adjusted R-squared result from that computation make the interpretation of these

data a less than certain venture.

CHAPTER 6

TESTING SUPREME COURT ORAL ARGUMENTS FOR COMPLEXITY

The written opinions of Supreme Court justices have been subjected to complexity testing (Owens and Wedeking 2011; Cross and Pennebaker 2014), statement made a t oral arguments have not. The later corpus is attractive because statements made by justices at oral arguments are certainly less guarded and more spontaneous than those made in written opinions. For this article the underlying hypothesis is that the utilization of slippery reasoning by the justices in support of politically biased outcomes, even if it is the product of a motivated reasoning process (Braman 2009), is often going to require the use of more complex language in order to properly cloak the underlying reasoning. To test the hypothesis, statements made by justices at oral arguments between 2009 and 2013 were analyzed using two established measures of linguistic complexity (the Linguistic Inquiry and Word Count program and the Flesch-Kincaid Score) and those results are then compared against the composite political ideology scores of each of the justices. The results found do not establish overwhelming support for the hypothesis. This finding may be due to the fact that the measures of complexity are relatively blunt, and this outcome suggests the need to develop a more refined measure of the complexity of legal language.

6.1 Introduction

In the struggle to ascribe political bias to the courts, the justices have been able to consistently dodge the proverbial bullet because there is "no neutral arbiter for the evaluation of adherence to stare decisis" (Cross, Spriggs, Johnson, and Wahlbeck

2010).[10] As a result one side (of the classic progressive/conservative divide) has been

able to implacably claim that it alone has been free of ideological biases and, ergo, has

been consistently faithful to controlling legal precedents. A possible explanation for this

is that scholars have long shied away from the law (or the language of the law) as a

dependent variable. There has been a high level of attention given to the way that judges

vote. Along that path much weight has been assigned to the premise that the votes

themselves are the nearly exclusive result of each individual's ideological beliefs, and

that line of work has produced ample research on the ideological nature of individual

votes in appellate cases and on the disposition of cases. While some of that work has

been quite impressive (Segal and Spaeth 1996), an alternative approach would be to

undertake an in-depth analysis of the announcement of legal policy both within majority

opinions (Cf. Owens and Wedeking 2011, discussed infra), and as stated at oral

arguments (the aim of this subproject).

Other than legal, strategic, and bias-based explanations of judicial behavior,

competing accounts of what is going on deep in the thickets that make up court opinions

are somewhat scant. Churning through legal opinions and converting the often abstract

and dense expressions of judges into workable data has long been technically daunting,

and exceedingly laborious. That complexity has, no doubt, contributed to the paucity of

work in which extracted judicial writings and/or statements are featured as the dependent

variable. The tide has turned somewhat with Hansford and Spriggs' work on the

---

[10] *Stare decisis* being the Latin term for a court "navigating by the stars" (i.e., a court being obedient to prior precedent that is, or should be, controlling with regard to the facts of the matter in controversy. Courts that wish to go against *stare decisis* will generally do so by establishing that the present facts are sufficiently different from the prior facts to justify not following the prior precedent.

interpretation of precedent (2006), Walton's exploration of the structure of legal reasoning (2002), and Maltzman, Spriggs and Wahlbeck's work on strategic interaction and the opinion-writing process coming more to the fore (2000).

Machine reading of text is being utilized to further rectify the situation. Attention has turned not only to the opinions of the Supreme Court (Rice 2015), but also to amicus briefs (Corley, Collins and Calvin 2011; Sim, Routledge and Smith 2014) and to what the justices say during oral arguments. Black, Treul, Johnson and Goldman (2011) subjected a unique corpus of over eight million words spoken by justices that spanned thirty years of oral arguments to machine reading. They concluded that the use of more "unpleasant language towards one of the sides at oral arguments was a prior signal of that side being less likely to prevail." That study added to prior work establishing that the statements and behavior of justices at oral arguments could be predictive of the eventual disposition of the matter (Johnson, Wahlbeck and Spriggs 2006).

While the use of predictive algorithms for the machine reading analysis of words spoken by justices could have a payoff, the "move" that this chapter attempts to make is to pivot from the treatment of statements made at oral arguments as augurs of eventual outcomes to instead searching them for the often subtle spoken signals that linguists refer to as Linguistics Based Cues (LBCs) that have, among other things, prooven useful in the detection of higher cognitive loads for speakers (Khawaja, Chen, and Marcus 2012).

Taking this approach is decidedly not making the claim that deliberate deceit taking place on the Court, nor is the claim being made that the justices are being deliberately obscure or abstruse when speaking from the bench. Although the later claim has been put forward with the supposition that strategic obfuscation increases the costs of

review by supervisory institutions and thereby enables the Court to evade effective

congressional oversight (Owens, Wedeking, and Wohlfarth 2013). At the same time

Corley and Wedeking (2014) have made the claim that more certain language raises the

likelihood that a Supreme Court opinion will receive positive treatment by lower courts.

What is being explored here is the notion that when any speaker gets away from

the simple case (for a justice that would be merely applying the cannons of interpretation

to obtain a fair reading the language of a legal writing, be it a contract, a will, a statute, a

regulation, a controlling court opinion, the Constitution, or any other legal document that

requires construction) and begins the process of stepping to a more abstract form of

argument (such as purposivism[11]), then the forms of reasoning that they offer change.

The aim of this project is to test the language of the Court in order to detect if the

transition from the direct interpretation (of a legal concept or text) to an indirect

interpretation (of the same) requires of the speaker a generally higher level of abstraction

and an increase the frequency with which more complex thoughts must be expounded,

and if those requisite mental gymnastics will tend to force the speaker to utilize more

complex rhetoric that can then be detected with the use of both standard text complexity

tests and more specialized machine reading software.

Making this journey requires crossing a minimum of two bridges. First,

establishing whether the machine reading the corpus is sophisticated enough to determine

reliably when a speaker has engaged in a significantly higher level of cognitive burden

---

[11] Defined by Scalia and Garner (Scalia2012) as "The doctrine that a drafter's 'purposes', as perceived by the interpreter, are more important than the words that that the drafter has used; specif., the idea that a judge-interpreter should seek an answer not in the words of the text but in its social, economic, and political objectives" (at page 438).

than is the norm for the forum that they inhabit. Second, entertaining the notion that there could be a reasonable, theoretical basis to believe that moving away from a grounded, textualist approach to legal interpretation and moving towards a purposive approach that will require the adjudicator to engage in a more complex level of analysis to justify their legal conclusions (other types of approaches area also possible, e.g., consequentialism which predicates the decision upon the anticipated outcome that it would likely produce for the parties).

6.2 The Validity of LIWC 2007 and the Flesch-Kincaid Reading Ease Score

Part of the machine reading in this study was done by the Linguistic Inquiry and Word Count 2007 program (LIWC). A recent search (March 2015) for the program on Google Scholar returned over 3,600 citations that mention the program (although that is not to say that every last one of them is necessarily a positive one). The validity and reliability of LIWC on a variety of its indicators has been established by several studies (e.g., Alpers et al. 2005; Bandum and Owen 2009; Cohen 2012; Kahn, Tobin, Massey, and Anderson. 2007). Pennebaker—one of the creators of LIWC—published a peer reviewed meta-study of 121 articles that employed LIWC (Tausczik and Pennebaker 2010) with a favorable outcome for the program. With respect to the use of LIWC to evaluate cognitive complexity, the appendix of Owens and Wedeking (2011; at pages 1055-1057; see also Abe 2011), provides an extensive exposition regarding the measure.

The alternate scoring system used herein, the Flesch-Kincaid readability index (Kincaid, Fishburne, Rogers, and Chissom 1975), is a relatively well-established test. Flesch-Kincaid attempts to register the level of difficulty of a sample of text and convert that measurement to a "reading ease level." The mechanics of the Flesch-Kincaid score

are relatively simple and operate off of the average lengths of sentences and the average number of syllables in the words that comprise the text being analyzed. The exact formula for the Flesch-Kincaid reading–ease test is (Calderón, Morales, Liu, and Hays, 2006):

206.835 - 1.015(total words/total sentences) - 84.6(total syllables/total words).

It should be noted that (perhaps counter intuitively) the scale for Flesch-Kincaid assigns higher scores (to a maximum of 100) to simpler, more easily readable text and lower scores (to a minimum of 0) to more complex, difficult text.

6.3 An Example of Complexity in the Law

Regarding the notion of moving away from the "black letter" law to a more remote and abstracted construction, consider the following thought experiment: You are an appellate judge in a jurisdiction that has made homicide a crime (defined by the Model Penal Code § 210.1 as "the act of purposely, knowingly, recklessly, or negligently causing the death of another human being"). For the purpose of this example this jurisdiction does not have a physician assisted suicide exception. A case comes before you the stipulated facts of which state that a physician prescribed an overdose of barbiturates for a patient who was suffering from terminal, end-stage throat cancer. The facts further stipulate that the patient had "pleaded and begged for an end to their severe and undignified suffering." The patient followed a protocol provided by the physician and, in doing so, ended their own life. No issue of patient competence looms, and the lower court returned a guilty verdict.

If one is a textualist, the matter is easy to dispose of—you apply the law to the facts and conclude that the physical has committed a homicide because he physician knowingly and purposely took steps that caused the death of a human being. Q.E.D.

If one is purposive, and if one believes that those facts tell the story not of a murder, but rather of a merciful deed that was done out of compassion—one that falls outside the ambit of the statute—and if one is not particularly concerned about pitchfork-wielding crowds howling for the hides of activist judges, one's reasoning could be more nuanced and could well lead to the opposite conclusion. One could cite to external materials that go to legislative intent, arcane case law from exotic foreign lands such as Europe, and speak at length of a "living" body of law that "must draw its meaning from the evolving standards of decency that mark the progress of a maturing society."[12] With some hard work, one would be able to cobble together a coherent legal argument (with which some, if not many, would agree), supporting the conclusion that the statute was likely never envisioned to be applied to circumstances such as those presented, and that the physician should not face the same harsh judgment and punishment that a coldblooded, spree killer would.

The point of the example is that one would have a lot more intellectual heavy lifting to do to make the latter argument than to make the former. That greater level of effort would create a greater cognitive load, and that expenditure of additional intellectual horsepower should be detectable through the use of LBCs. Thus, where individual justices must stray further from the black letter law to reach the verdicts that they feel are required, where political biases get played out through the words of the courts, and where

---

[12] *Trop v. Dulles*, 356 U.S. 86, at 101 (1958).

the engagement of motivated reasoning is what drives the legal equation, those patterns should be detectable and isolatable using the correct tools and algorithms.

It should be noted that the search for complexity in legal reasoning could return false negative results in the case of a justice developing new or biased law. If we allow ourselves to slip in the role of Max Weber scholars for a moment we can use that lens to examine the making of new law in the realm of Substantive Irrationality: the outcome is based upon a pre-decision of what is best, but no effort is put into giving a rational explanation for why the path was chosen (think of a parent separating two children who are fighting and sending both to their rooms without determining if one party provoked the other, and no rational explanation offered for the punishment—"Because I'm the parent!"). A justice could simply announce new law with minimal explanation, but this approach would be unlikely to muster a majority. Indeed, it appears that justices will go to great lengths to avoid being construed as operating under the banner of Substantive Irrationality. So, while the possibility should be acknowledged that judicial activism could elude the effort to ferret it out based on complexity, in practical terms that risk is vanishingly small.

6.4 Theory

The search for clarity and complexity in the written opinions of the Supreme Court has been ongoing. A top level concern to political scientists is whether the utterances of elites can be taken at face value. In the field of political science (as well as in economics) it is generally uncontroversial to maintain that when political actors speak, they are conveying at least some useful information about their views (Austen-Smith 1990a; Diemeier and Fedderson 2000; Black, Treul, Johnson and Goldman 2011).

Although words communicate ideas, scholars are aware of the meanings that go beyond the literal interpretations, and in the present study the actual elements of reasoning, and the logical steps that are traced, are of subordinate importance when compared to the questions that are being asked about which arguments hew to the law as it is given to the courts, and which arguments stray away from fair readings, and perhaps even from stare decisis.

Owens and Wedeking (2011) presented one of the early machine read, systematic examinations of clarity in Supreme Court written opinions and determined that there were a range of styles among the justices (Breyer and Scalia being the clearest, Ginsburg the most complex); that there was no significant correlation between ideology and complexity; that dissents tended to be clearer than majority opinions; that criminal procedure cases tended to produce the clearest writing; and, that opinions announcing Court precedent tended to feature the most complex writing. The fact that those findings were based on written opinions bears further examination.

As mentioned previously, the written opinions of the Supreme Court are generally the product of multiple authors, even if only one justice is noted as the author. With only a few exceptions (notably Justice Brennan and Justice Powell (Cross and Pennebaker 2014)), most justices rely at least to some extent upon their clerks, especially in the production of early draft opinions. Justices also borrow from both the briefs offered by the litigants, from amicus curiae briefs (Collins 2004), and frequently from each other, often inserting particular language in order to mollify a particular concern that a colleague has expressed through one of the vast number of memoranda that circulate through the Court's chambers (O'Brien 2008).

It is not particularly surprising that a range of styles can be detected given that opinions are an admixture of contributions, but most have a primary author, and that author's propensity for simplification (or disregard for it) will no doubt show through. Nor should it be surprising that dissents tend to be less complex—they are almost invariably shorter, can be terse at times, and have the luxury of focusing on only a few issues where the author feels that he or she has the best of the argument. It is logical too that a justice will be reluctant to devote vast amounts of time to authoring complex dissents, as there is little immediate gratification realized (Cf. Epstein, Landes, and Posner (2013) for a rational choice based examination of why a jurist would not want to invest a lot of energy into these dissents (a simple enough situation for a rational choice explanation to potentially have merit)).

Naive reasoning provides some possible clues that might account for the finding that opinions concerning criminal procedure tend to be less complex. Criminal procedure is a rule-based area of the law, and the fact patterns that tend to emerge from that corner of the law are much more driven by real world events that can be easily described than by the sort of arcane minutiae that is locked up in treatises (as is often the case with tax law, for example). Likewise, we should expect new precedent to be more complex—any other finding would be highly counter-intuitive. This leaves the assertion that there is no link between complexity and ideology.

The primary point to focus upon with regard to this study is that Owens and Wedeking (2011) are not considering rhetorical clarity (the "readability" of a statement). Their concern is the "cognitive clarity" of a text, and by cognitive clarity they mean a measure that is centered upon the differentiation and integration of elements (i.e., does

the author perceive varying viewpoints? Does the author recognize relationships and connections among the various perspectives and dimensions that they do, or do not, decide to explore?). As Owens and Wedeking describe the matter, it appears that the primary way to boost the cognitive clarity "score" of a text is to build up layers of authority and counter-authority. Thus, the more one acknowledges counterfactual arguments, the more one's level of cognitive clarity is expected to rise.

Defining greater cognitive complexity as going to greater lengths to illuminate the various facets of a topic is a defensible approach. At the same time, when using the LIWC program such a metric will have difficulty detecting that dynamic when a rhetorical argument is made that black is white, so long as that argument refuses to take the tack that there is more than one side to the coin. Moreover, finding that this type of complexity is not correlated with ideology is explicable. Justice Scalia was rated as the least complex writer and Justice Ginsburg the most complex. That these two real-life buddies (prior to Justice Scalia slipping this mortal coil) had different approaches to jurisprudence is common knowledge and the results here merely suggest that Ginsburg spends more time going into detailed analysis of the various sides that emerge from her analysis, while Scalia was at times content to simply state what his own side was, and to dismiss arguments that opposed his own by giving them short shrift.

6.5 Written Opinions Differ from Oral Arguments

Whereas the Court's published opinions are made up of composite text with multiple authors (understood here to be language incorporated from briefs, arguments inserted to placate non-authoring justices, passages composed by clerks, etc.), utterances made during oral arguments are "straight from the horse's mouth," so to speak. Also,

even when compared to public statements made by members of the Supreme Court at speaking engagements, it is also reasonable to assert that a significant part of what is said at oral arguments is likely to be both directly relevant to the matters litigated before the Court, as well as being unrehearsed and, perhaps at times, uttered without the benefit of prolonged reflection.

While it is almost certain that the justices put in some level of preparation prior to the day the case is heard, and likely have prepared questions, they have no way to predict where the advocates will wander, and certainly some significant portion of what is said by the justices must be extemporaneous in nature. That spontaneous nature is especially attractive from the perspective of a researcher seeking to isolate verbal cues that are flags for an individual's state of mind under a given set of circumstances. Given these general contours, it is somewhat surprising that statements from oral arguments have not been subject to more extensive analysis (Cf. Ringsmith and Johnson 2013; Johnson Wahlbeck and Spriggs 2006; Black, Treul, Johnson, and Goldman 2011).

Another area where oral arguments should be susceptible to useful measurement and analysis is the difficulty that justices likely experience disguising, on the fly, occasions where they are engaged in motivated reasoning. The process of sidestepping the objectively controlling authority would be the end result of motivated reasoning, an organic and unconscious process that leads individuals to discount sources of authority that are at odds with their biases, while simultaneously leading them to give greater weight to sources of lesser authority that align with their biases (Braman 2009). Newman, Pennebaker, Berry, and Richards (2003) have suggested that the process of creating a false story should consume cognitive resources. Their arguments in support of their

research are all well and good, but there is no evidence that justices making ideologically biased pronouncements are consciously being deceptive. Rather, the suggestion is that straying further away from the black letter law demands that greater cognitive efforts be expended to reconcile the slippery logic used with the most pertinent, legal authorities that an objective jurist would consider controlling in the situation.

6.6 Data and Methods

In order to analyze the data, the corpus of transcribed Supreme Court oral arguments going back to 2004 was scraped from the Court's webpage where they are posted,[13] cleaned, and converted to plain text files. Digitized records of oral arguments going back to 1979 are available, however, in the years from 1979 to 2003 the individual justices are not identified and the transcripts merely record any questions emanating from the bench as having been asked by "The Court." Records also exist of oral arguments going back to 1968, but those are on microfiche and their quality has been found to be poor enough that their digitization would be unreliable (Black, Treul, Johnson, Goldman 2011). It is hoped that at some point efforts will be made both to add the names of the individual justices to the 1979 to 2003 oral arguments, and to digitize the prior oral arguments that are on microfiche as having that data would allow further exploration to move forward.

For this subproject it was most straightforward to simply aggregate the terms from the year 2009 to the year 2013. The statements made by attorneys arguing from the bar were stripped out of the text files, and the statements of each individual justice were

---

[13] https://www.supremecourt.gov/oral_arguments/oral_arguments.aspx

aggregated. Once compiled, these by-justice text files could then be further subdivided, say by term, by type of main case issue, by case name, or by statements containing specific words or phrases. As alluded to elsewhere, Justice Thomas was all but silent in that cycle, speaking only ten words which is insufficient for meaningful analysis in this context. Justice Stevens (who left the Court after 2009) was quieter than the other non-Thomas justices, speaking a mere 8,000 words in that term. Because this volume of text is fairly minimal for analysis using the tools employed here, and because a single term provides no basis for panel data comparison, Justice Stevens was also dropped from the set.

Newman, Pennebaker, Berry, and Richards (2003), advise that the use of correlation of individual word categories with:

- Words associated with causation;

- Words associated with insight;

- Words associated with discrepancy;

- Words associated with inhibition;

- Words associated with tentativeness;

- Words associated with certainty;

- Words associated with inclusivity;

- Words associated with exclusivity; and,

- Words associated with negation.

Whether these nine categories are representative of a single, underlying concept is a real question. The nine indicators were standardized into a single quantity for each justice in each term. This process mimics Owens and Wedeking's (2011) method of running the group through an explanatory factor analysis. In doing so they obtained a

one-factor solution that was judged fit to rely upon as sufficient support for the premise

that all of the indicators are tied together as part of an underlying dimension that

embodies cognitive complexity. The explanatory factor analysis has been taken as a

reasonable gauge of the reliability of the program's methodology.

With regards to the parallel use of the Flesch-Kincaid Score of Reading Ease

there is less to be said. As presented, supra, the Flesch-Kincaid approach is a

straightforward calculation:

206.835 - 1.015(total words/total sentences) - 84.6(total syllables/total words).

McCall and Crabbs (1961) validated Flesch-Kincaid Score of Reading Ease

against their McCall-Crabbs Standard Test Lessons in Reading. Flesch-Kincaid Score of

Reading Ease results are also highly correlated with other readability formulas (Fry and

SMOG at .96 and .95, respectively (Meade & Smith, 1991). A possible confounding

factor in the use of an automated Flesch-Kincade Reading Ease test is that, due to

machine reading recognizing each instance of a period as the end of a sentence,

"abbreviations, numbers with decimals, and bullets may lower the [Reading Grade Level]

RGL and underestimate text difficulty" (Friedman and Hoffman-Goetz 2006).

6.7 Results

A linear regression model for panel data is the obvious choice for the basic

statistical analysis of this data.[14] In the first analysis the Flesch-Kincaid scores for all

justices in each year (2009-2013) were employed as the dependent variable with the

absolute values of each justice's Martin-Quinn scores as the independent variable (as the

---

[14] The **plm R** package was utilized for this analysis.

theoretical position is that changes in the extremity of ideology cause variation in the complexity of the utterances that justices make from the bench).

A linear model for panel data was also employed to regress the composite LIWC complexity scores of each justice as the dependent variable against the absolute values of their by-term Martin-Quinn scores (a repeat of the previous model as the theoretical position is that changes in the extremity of ideology cause variation in the complexity of the utterances that justices make from the bench). As this is panel data, fixed effects were added in order to try to account for the tendency of a straight OLS model to simply draw a regression line through a cloud of data points but to reveal little about the individuals within the system. With a fixed effects estimator in use it is possible to get a better handle on the individuals while still running a single test. Often one of the major drawbacks of using fixed effects can be the loss of explanatory variables that do not vary by individual. In this instance at least, that problem does not arise because there are no additional variables being utilized. With further time and effort, it will be possible to break down the larger mass of text and to isolate and consider some additional variables (e.g., analysis by legal topic area, by whether the given justice ended up in a majority, or whether the opinion that emerged was unanimous), but this was a pilot effort and such improvements will have to wait for later attack.

As a baseline, the standard score for each justice's five terms was simply summed from the standardized scores for each LIWC dimension. It should be noted that, with respect to the justices in the aggregate, there was not a great deal of variation in their total "scores" over time. In fact, there was actually startling uniformity by justice from term to term. As people exhibit individual "ways" of talking, we would not expect huge variation

if circumstances were similar for each measurement, but even with regard to categories such as total words used, the variations were remarkably small. The grouping gets even tighter by dropping Justice Kagan's first term from 2010, with an average standard deviation of 2.51 and an average standard error of the mean of 1.16 (this conforms to theoretical expectations as most first-term justices are notably more reticent than veteran justices tend to be). Thus, if there is any value to be derived from this concatenated all categories "score", it appears to tell us that year-to-year the justices are all fairly consistent in the language that they use.

Shifting the focus from how each justice chooses their words from year-to- year, it is also possible to examine how the justices vary among themselves. The values of all LIWC categories were summed (netting out Justice Thomas due to his acute reticence). Over the five years considered, an average standard deviation of 7.56 is observed along with an average standard error of the mean of 2.52. Although those numbers are not indicative of a huge variation in the use of language by the justices, more is noted than the tendency observed for single justices over time, thus, justice-to-justice variation in use of language is observed. This tendency is greater in certain term years such as 2010 and 2013 where the standard deviations approach 10 and the average standard errors of the mean rest at around 3.3. At the outer limits we can, for example, compare Justice Ginsburg's 2013 term "score" of 263.63 to Justice Breyer's 2011 term "score" of 297.23. If this "score" has any value, it is telling us that the justices are relatively stable in terms of what they say individually from term-to-term, and that they are different from each other in their use of language, both within each term, and over time.

Standard scores were calculated for each of the nine categories by justice and term (the individual score (x) minus the global average score (μ) for the full Court all divided by the standard deviation for the full Court (σ)) thus:

$$\text{Standard score} = (x - \mu) / \sigma$$

Using this standardizing into a single quantity approach, the mean complexity score is equal to 0, with the range for this data set extending from a maximum value of 6.7 and the minimum value being -5.9 (as Owens and Wedeking were working with written opinions, and were working with data that ran over a period of time fivefold the one considered here, they naturally observed a wider range that ran from roughly twenty to negative twenty).

The Flesch-Kincaid Score of Reading Ease does not require any permutation to be used. Scores for the justices ran from the mid-sixties to the mid-seventies. Again, it is critical to recall that for this system a higher score means that the language is easier for a reader to comprehend.

6.7.1 Regressions

This preliminary result suggests that further analysis of the data set could yield a profile of the type of language that is favored by more moderate justices, and a second set of terms that justices more prone to voting at the ideological extreme are likely to employ. Flesch-Kincaid based analysis determined that the panel linear regression model was significant at $p < .001$ level, exhibiting reasonable support for the hypothesis that expounding upon more ideological positions leads to slightly more complex use of language (each one unit increase in the absolute value of the justice's Martin-Quinn score would lower the Flesch-Kincaid Score of Reading Ease score by 3.6 points.

The results for the regression of the LIWC standardized scores are less helpful. Although significant at the p < .0 1level, the LIWC Cognitive Complexity score moved in a negative direction which is counter to the theoretical expectation. These results are not fatal to this project because, as discussed supra, the LIWC Cognitive Complexity score is not a measure of rhetorical complexity, but rather focuses on the clarity of the text that is analyzed.

6.8 Conclusions

For all of the excitement that machine reading of text has introduced to the various fields that comprise the behavioral social sciences, a fair dose of caution should be taken with any claims that are presented based upon such analysis. LIWC is able to achieve reasonable levels of reliability when detecting the four major dimensions that it is designed to search out (linguistic dimensions, psychological processes, relativity, and personal concerns) (Friedman 2008). Beyond those major areas, the assorted sub-areas that it searches are concatenated in various configurations designed to fathom out different aspects of speech, and what sentiments are lurking beneath the surface. LIWC was worth trying in this instance, but it was an imperfect solution to the task to be accomplished.

It could be argued that baseline questions regarding the ideological biases of the justices have receded in terms of the levels of interest that they generate, and the field of judicial politics is (for the most part) characterized by scholarship which assumes that any justice's writings and utterances from the bench will be tethered to his or her Martin-Quinn ideal point. For a given matter we begin by inquiring how the applicable law aligns with each justice's ideal point, and proceed from that step to further, more

nuanced, evaluations of the Court's behavior (say taking stock of strategic considerations).

While it is possible that the empirical evidence that LBCs are able to provide can give us insights regarding what has taken place in some circumstances, it is my suspicion that we will need a tool capable of more subtle analysis than LIWC to parse out levels of complexity in legal arguments. Human coders from legal backgrounds, as discussed in the final chapter, should be helpful here. Following that effort, transition to using the same analysis to try to predict the outcomes of future opinions should follow. Linguistic analysis could also be a resource for researchers to utilize in taking apart the text of laws, to next determine the level of bias latent within their structures.

CHAPTER 7

NEXT STEPS

While progress has been made with respect to the central questions that are posed by scholars who examine judicial decisionmaking—and while it is hoped that this project might constitute at least a tiny contribution to the field—there is much further to go. This chapter is focused on the next logical steps that can be taken to advance my own research agenda and to hopefully make further forward progress in the coming years.

7.1 Introduction

Two terms from the theoretical framework have been studied, and the question of what constitutes "following the law" has been raised and investigated. The following discussion is intended to set the table for a discussion regarding the next steps that this research project should take.

7.2 A Different Taxonomy

A slightly different approach to the analysis of judicial decisionmaking considers the three lines of attack developed to try to fathom out the judicial decision-making process. One area that has been studied is the effects of exogenous variables that contribute to the eventual votes and opinions that emerge; the next is the study of language and citation patterns, and the last is the study of the forms of argument that are used by judges in their legal reasoning.

7.3 Exogenous Variables

It is axiomatic that when a decision must be rendered, a justice's mental energies engage with a host of external influences and that the subsequent cognitive process produces various outputs that are available for analysis (votes, opinions, public statements). Notable external factors that appear to exert influence upon (or at other times constrain) the Court that have been the topic of recent research include the lower courts (Corley, Collins and Calvin 2011), the media (Baum 2006), amicus curiae (Collins 2004), the executive branch (Black and Owens 2012), and public opinion (Casillas, Enns and Wohlfarth 2011). In short, various scholars have been developing a complex and multi-layered set of models that, while disjointed, provide us with a series of potentially helpful, but largely independent, insights regarding what does, and what does not, account for the Supreme Court's term-by-term outputs.

7.4 Language and Judging

An entirely different branch of work is being developed by scholars examining the language of Supreme Court Justices. Several studies have already been mentioned in other chapters; a brief recap follows. Written opinions have been the focus of many such research projects. Cross, Spriggs, Johnson, and Wahlbeck (2010) have advanced our understanding with regard to the importance of citations to the Supreme Court's decision-making process. Owens, Wedeking and Wohlfarth (2013) utilized the Flesch-Kincaid Reading Ease Index to evaluate a random sample of Supreme Court opinions from between the 1953 term and the 2008 term and emerged with the claim that the justices attempted to obfuscate their language in instances where they face increasing ideological distance from pivotal legislative actors. Corley and Wedeking (2014) approached the

question of certainty in the language of the Supreme Court and presented the claim that there is a correlation between an increase in the certainty of the language that the Court uses and the likelihood that lower courts will treat the Court's decisions positively.

Oral arguments have also attracted attention from scholars. Black, Treul, Johnson and Goldman (2011) have examined the emotional content of the words used by Supreme Court justices to interrogate attorneys at oral arguments, determining that the use of more unpleasant language toward one side reduces the likelihood of that side prevailing both in terms of the individual justice's votes and in the ultimate opinion issued. Oral arguments were also the source material for two separate studies that have both claimed that the quality of the arguments tendered influences the justice's eventual decisions (Johnson, Wahlbeck and Spriggs 2006; McAtee and McGuire, 2007). Of particular interest is the ongoing work by Owens and Wedeking (2011) with regard to the language that is used in Supreme Court opinions. No link was found between the level of clarity of the writing and the justice's ideology, but dissents of all stripes tended towards less complexity than majority opinions (note also that some justices found—particularly Justice Douglass and Justice Scalia—great relish in the authorship of taut solo dissents where they could roam about the landscape free of the burdens of consensus building). Most significantly, opinions that establish new precedents were determined to render more convoluted law.

This last result makes intuitive sense. Upholding an existing statute or following stare decisis and simply proclaiming that we will continue to follow declared law should be a relatively simple matter to explain. Engaging in judicial review or fully changing the course of and charting a new direction for the law would appear to be an inherently more complex undertaking as arguments must be selected, justified, buttressed and explained.

Jurists are supposed to be reluctant to veer from the familiar, be it by overturning a statute or by rewriting the common law; if prior precedent is being abandoned reasons must be provided, and if new law is being put in place the reasoning must be put forward and supported. All language necessary to make arguments in support of declaring a law unconstitutional or in support of making the passage from prior law to new law invites various forms of complexity; all of the various legal arguments for and against the prior and the new law raise the level difficulty of the challenge. Moreover, a court comprised of many members will frequently have differing opinions regarding the reasoning, even where there is ultimate agreement on what the outcome should be. There are several wrinkles to this particular finding that will need to be addressed further, but for now the reader is asked to kindly note the finding that making new law necessarily invites expanded explanation and the theory that undergirds that principle. This assertion goes to the most frequently observed case. It is plausible to imagine a court that grows frustrated with an especially "overgrown" area of the law striking down a swath of a statute and substituting a cleaner, more limpid set of standards but no actual occasion of this behavior springs readily to mind.

With regard to their methodology, the primary point to focus upon is that Owens and Wedeking were not considering rhetorical clarity (the "readability" of a piece of writing). Rather, their concern was the cognitive clarity of the opinions, and by cognitive clarity they meant a measure that is centered upon the differentiation and integration of elements (i.e., does the author perceive varying viewpoints? does the author recognize relationships and connections among the various perspectives and dimensions that they have decided to explore?). More specifically, their concern was Pennebaker's measure of

cognitive complexity, as opposed to Integrative Complexity or Hermann's Conceptual Complexity to name just two of the many measures that are found in the literature (Gideon Conway, Gornick, and Houck 2014).

As Owens and Wedeking describe the matter, it appears that the primary way to boost the cognitive clarity "score" of a text is to build up layers of authority and counter authority, as the more one acknowledges counterfactual arguments, the more one's level of cognitive clarity will rise. Thus, a justice who writes an extensive and deliberative opinion that expands upon more than one line of legal reasoning is likely to obtain a higher complexity score for that writing than a justice who is dismissive of contrary viewpoints, or who glosses over fine details in his or her opponent's position. In a more recent study, Cross and Pennebaker (2014) utilized the LIWC program to analyze the corpus of opinions produced by the Roberts Court with similar outcomes.

Defining greater cognitive complexity as going to greater lengths to illuminate the various facets of a topic is a defensible approach. As the same time such a metric will not be able to capture the dynamic when a rhetorical argument is made that black is white, so long as that argument refuses to take the tack that there is more than one side to the coin. Moreover, a finding by Owens and Wedeking that this type of complexity is uncorrelated with ideology is unsurprising. Prior to his death, Justice Scalia was rated as the least complex writer and Justice Ginsburg the most complex. Although they were somewhat famously friends, that they had different approaches to jurisprudence is common knowledge and the results here merely suggest that Ginsburg spends more time going into detailed analysis of the various sides that emerge from her analysis while Scalia was oftentimes content to simply state what his own side was and to dismiss arguments that

opposed his own by giving them short shrift. While this work is no doubt trailblazing, there are further aspects of the clarity question that still need to be further explored.

7.5 Forms of Argument

Logicians have long worked on judicial decisionmaking. Professor Douglas Walton has been a leading exponent in this arena in his Legal Argumentation and Evidence (2008). Innovative in approach, but too far afield to delve into here, Walton essentially argued that legal controversies were settled by way of dialogical argument and that both the "form" of an argument and its specific context were crucial. Through this lens various truths emerge, such as the reality that certain forms of argument that are considered logical fallacies can survive, and even thrive, in the courtroom setting. It is a fascinating and radical study, but not one that is central to this discussion.

7.6 Moving Forward

The key next step, and what we should actively be focusing on in order to try to untangle this Gordian Knot of judicial behavior, is to uncover effective ways to better understand the verbal outputs of the courts so that we can "reverse engineer" the process (understanding that the language that the Court uses is the best evidence that we have of their thinking). Approaching the puzzle from the perspective of both an attorney and a Public Law scholar, there is one particular gap, the role of personal bias, which is in need of some urgent attention. A logical next step for this project would be to attempt to fill in that gap.

7.7 Objective Number One

As discussed in the previous chapter, while scholars have struggled to ascribe political bias to the Court, the justices have been able to consistently dodge the proverbial bullet because there is "no neutral arbiter for the evaluation of adherence to stare decisis" (Cross, Spriggs, Johnson, and Wahlbeck 2010). On occasion assertions of fidelity to fair minded and principled professionalism are doubtless accurate, but how do scholars go about objectively and empirically establishing the truth or falsehood of such claims? A potentially good indicator of such behavior would be locating votes by justices that follow previously established legal rules with which they are known to disagree. Segal and Spaeth's 1996 article "The Influence of Stare Decisis on the Votes of the United States Supreme Court Justices" did just that finding that in the time period covered (from the Warren Court until the article was prepared) for "Landmark" opinions (defined in the article at 976) only Justice Powell and Justice Stewart "show any systemic support for stare decisis at all."

We have several measures of political bias for individual Supreme Court justices such as the Martin-Quinn score (Martin and Quinn; 2002), the Segal-Cover score (Segal, Cover; 1989), and the Epstein et al. score (Segal, Epstein, Cameron, and Spaeth; 1995), but we have no measure that helps us to assess how closely justices hew to the capital "L" Law in various situations. A qualitative approach would be reasonable in a particular example—one could laboriously deconstruct the relevant statutes and assess the legal arguments mounted by each side in a stepwise fashion for each relevant point of law. With some applied effort, and with great attention to nuance, it appears probable that a bespoke approach to assessing judicial fidelity to strict interpretation could potentially bear fruit. Needless to say, such an approach would be intensely consumptive of time and

effort. The aim of this project is to develop an alternative linguistic analysis approach that would allow researchers to quickly reach adequately robust conclusions regarding the presence or absence of bias (defined as the abandonment of fidelity to the law) in a given section of a court opinion without having to invest vast resources in processing the relevant text.

Two approaches would be utilized and compared. In the first-case attorneys (construed as graduates of an American Bar Association accredited law school who have successfully passed the bar exam in at least one state) would be given blind samples of justice statements from U.S. Supreme Court oral arguments and assigned to code sections of text by assigning ordinal values that range from (say) a zero value for low fidelity to extant law to a (say) higher set of values for statements that hew more closely to extant law (because of the specialized nature of Supreme Court arguments and the broad nature of the legal ground covered by the Court over any given term, it is thought that individuals with the extensive training provided by a full legal education will have the best success at delineating the often dense and difficult reasoning of justices and of accurately and defensibly coding the signals that justices give with regard to stance on the various legal issues that come before the Court.) Text sections would be presented in blind format with the name of the justice redacted, and inter-coder reliability checks would be implemented to establish internal validity for scores.

The work would be challenging, and would require great concentration. At first some of the coding work that the human coders do would likely be driven by "gut" reactions regarding the level of bias that a justice was expressing in a given section of text. In time, a directory of terms and words that are suspected of betraying bias would be

accumulated. Ongoing refinements would develop the project and enable more accurate isolation of instances of ideological leakage into the judicial decisionmaking process.

The statements made by the justices at oral arguments are the optimal choice for this analysis. The written opinions of the Supreme Court are generally the product of multiple authors, even if only one justice is noted as the author. With only a few exceptions (notably Justice Brennan and Justice Powell (Cross and Pennebaker 2014)), most Justices rely at least to some extent upon their clerks, especially in the production of early draft opinions. Justices also borrow from both the briefs offered by the litigants, from amici briefs (Collins 2004), and frequently from each other, often inserting particular language in order to mollify a particular concern that a colleague has expressed through one of the vast number of memoranda that circulate through the Court's chambers (O'Brien 2008).

Oral arguments are taken "straight from the horse's mouth," certainly the justices put in some level of preparation prior to the day of the advocates making their points and doubtless have some prepared questions, but while a justice could have some success anticipating where an advocates will wish to steer the discussion, reviewing transcripts reveals that advocate sometimes wander, and the problem of prediction is compounded by the "hot" nature of the bench itself. Notwithstanding the possibility of advocates taking unexpected detours, or other justices interjecting new and novel questions, certainly some of what emanates from the bench must be extemporaneous in nature. That spontaneous nature is especially attractive from the perspective of a researcher seeking to isolate verbal cues that are flags for an individual's state of mind under a given set of circumstances.

At this point an example is in order. The following text is an excerpt from the statements made by a Supreme Court justice during the oral arguments made for a case that focused on Miranda warnings:[15]

"There's no suggestion that there's -- that the statements are not voluntary. The suggestion is that they may have violated Miranda. What if he said, "do you want to remain silent?" And he doesn't answer either one. You say you don't have to invoke your rights, but Butler also says that you can impliedly waive them. You don't have to expressly waive them. Right … And all he has to do is say: 'I don't want to talk to you. It's over.' I don't understand how they create the ambiguity. Well, I guess this gets back to a question I had earlier. I thought there was no dispute on this record that there was no involuntariness. We are talking about a violation of the technical, important but formal, Miranda requirements. But that's correct -- and that's where -- how I read Butler; you have to look at the circumstances. And you're saying no, you don't look at any circumstances; they have got to ask the question and he has to waive. The other circumstances are irrelevant. Well then, if yes, why are you talking to me about 2 hours 15 minutes, what they are doing? You say that circumstances don't matter. That issue is not in this case, though. As I understand it, you've lost at every stage on the voluntariness and have not renewed that, correct? This is a Miranda case; it's not a Fifth Amendment case … Before they can say anything more, they have to get a waiver. So it's 30 seconds if they go on, before they -- if they sit there for how long before -- how long do they have to ask, 'do you want to waive?' Could -- could you describe a situation where you think there would be an implied waiver? Well, I thought that -- that doesn't sound implied. That sounds express to me … So what -- what is an implied waiver case? Well, that's right. Now, getting back to Ms. Saharsky's point, she said if you prevail, you have to overrule Butler. And it seems to me that that's the point we're at. So, there's -- so, there's no implied waiver with respect to the right to remain silent?"[16]

Although it has some embodied ambiguity, this language would best be coded as

an example of a justice who wishes to have the Court render an opinion that takes steps

away from a prior precedent. This categorization is chosen because the justice initially

asserts that the holding in Butler establishes an implied waiver of the right to remain

silent, but at the end of the text that the implied waiver is in jeopardy. Because there

would be no way to resolve that conflict without somehow crafting a new statement

regarding the way implied waiver of Miranda rights cases are to be handled, the text is

then classified as breaking away from the established precedent. In the final holding that

followed these arguments, the majority did explicitly create a further limitation on the

rights that had been created in the Miranda decision. I am considering the opinion to have

been a limitation of prior law because the Court used specific language to explain the

nature of the new interpretation and it would have been unreasonable to expect that any

lower court would have reached the exact same conclusion without that guidance[17]. The

primary reason that this text would be coded as seeking to alter the existing law is the

general negative tone with regard to the controlling precedent. Now, consider a section of

the statements from the other side in the same matter:

> "How do we -- how can we imply waiver? Meaning if all he said was,
> yes, I want them in, that's much different than saying, if someone had
> asked him, do you want to leave, and he shakes his head no. The latter
> might imply to me that he waived, but the former certainly would be
> neutral. So what do we do with our case law that says that you can't
> infer waiver simply from the confession? I mean, we have said that.
> So that's pretty clearly established statement -- by the Court. Well, I
> think certainly in -- in Butler, if someone in their confession says, I
> know I don't have to talk to you, but I want to, that that would be
> using those words. So how can you say -- How can you say that an
> appeal to someone's religious position after 2 and a quarter hours is a
> voluntary waiver? You want to change the Miranda rule to say: Tell

---

[17] *Miranda v. Arizona*, 384 U.S. 436 (1966).

someone their rights, and unless they explicitly say 'I don't want to talk to you.' then they implicitly under virtually any circumstance haven't. That's what you believe the rule in Miranda and Butler and Davis sets forth? There wasn't -- there wasn't silence in Butler. There was an express 'I want to talk to you.'"[18]

This section of text would be coded as an example of a justice who wishes for the Court to affirm the prior holding that was under attack. As this is a single example there is far too little data here to actually inform us regarding the habits of justices who wish to adhere to, or significantly alter, the Law. The aggregation of many, many such examples isolated through a cross-checked process designed to assure internal validity would allow the machine analysis of the linguistic characteristics (specifically the sentiment that is being expressed) common to this type of legal reasoning. That machine analysis extracts a "profile" that then makes the classification of other text a fast, efficient and accurate process.

"And why isn't the -- the most sensible way to deal with the problem that you are raising, the one that the Chief Justice suggested, to permit as-applied proportionality challenges that take into account the particular circumstances of the juvenile in question, rather than this per se rule that you are advocating, which would deprive the State of Florida from reaching the judgment that there are some -- there are some juveniles, some individuals who are short of their 18th birthday, who cannot -- who deserve imprisonment in -- life imprisonment without parole? Some of the actual cases that -- in which this sentence has been imposed in Florida involve factual situations that are so horrible that I couldn't have imagined them if I hadn't actually seen them … But do you know anybody who is willing to say that, as a categorical matter, that -- you know, the 18th birthday is the magical date for every single person? Because the Court, up to this point, has said that death is different, and the rules -- the Eighth Amendment rules in capital cases are entirely different from the Eighth Amendment rules in -- in all other cases. If we -- you know, if we abandon that, then one of two things has to happen, either the rules for noncapital cases have to change dramatically, or the rules for capital

---

[18] Edited statements of Justice Sonya Sotomayor during the oral arguments for *Berghuis v. Thompkins*, 560 U.S. 370 (2010), that took place on March 1, 2010.

cases have to change dramatically, unless death is different, in fact. I know you are not asking for that, but that -- isn't that where this, logically, is going? If death is not different, then there should be uniform rules across the board. Why does it say that? Why doesn't it just say that, in this particular case, what this individual has done is so bad that, even if this person can be rehabilitated and would not present a danger to -- to society at age 60 or 70, that this person is -- should be sentenced to life without parole? That's -- that's what it means for an adult offender."[19]

In this text the justice clearly wants to leave the law as it was at the time, such that it allowed the assignment of life without parole to minors. It was a fight that this justice lost in this instance, but it does not take a great deal of effort to pick out some of the language that explains his position. The justice's first line of attack is to suggest that allowing states to sentence minors to life with an option to allow challenges is the best solution. The Justice then moves on to buttress his point by citing to the atrocity of some of the criminal acts minors have committed. For his next salvo the Justice points to the way that he Court has isolated the death penalty in relation to minors and argues that to ban life sentences for minors will either lead to major changes with regard to other noncapital cases or to major change for capitol cases (i.e., to make this change in the law will disrupt the balance in the system). Lastly, the justice asks rhetorically why the language of the statute cannot simply mean that if a crime is sufficiently heinous then the state has the option to impose a life sentence (as opposed to interpreting the statute to have some other meaning, such as meaning that the offender can never be rehabilitated).

It should be noted that the American public is split on how they feel the work of judges and justices should be done. In response to a poll by the Associated Press and the

---

[19] Edited statement of Justice Samuel Alito during the oral arguments for *Graham v. Florida*, 560 U.S. 48 (2010), that took place on November 9, 2009.

152

National Constitution Center (2012) that asked about judicial decision making thirty-eight percent were of the opinion that "Judges should interpret the laws as narrowly as possible, taking into account only what is clearly the intention of the lawmakers," whereas fifty-six percent responded that "Judges should interpret law broadly, taking into account the broader interests of the nation." There is no effort in this project to make any normative argument about whether a propensity to changing or modify the Law is an objectively "good" thing to do, or whether a propensity for leaving the Law intact is an objectively "bad" thing to do. It should be fairly uncontroversial to assert that even some of the most extreme actions taken by the Court, where major blocks of legislation have been struck down now are viewed as normatively "good" (with Brown v. Board of Education[20] likely being the leading example of same—although the Brown decision was legitimately, and thoughtfully, criticized by Professor Derrick Bell (2004) not so much in spirit as for having been ultimately and predictably harmful to the cause of racial equality), and that some instances of the Court taking a hands-off approach to existing laws have come to be criticized (with Lochner v. New York[21] being a leading example of same—although the Lochner decision is not without its champions, notably Bernstein (2011)). The basis of this project is rather to develop a system whereby we derive the measure of a given individual's propensity to make changes—or to not make changes—to the Law as that is the most conspicuous dimension of judging yet to be explored.

---

[20] 347 U.S. 483 (1954).

[21] 198 U.S. 45 (1905).

7.8 Objective Number Two

Over the past decade, machine reading of legal materials has gained some traction (Evans, McIntosh, Lin, and Cates 2007). Two machine reading approaches would be utilized in tandem for this project. The first, Latent Dirichlet Allocation (LDA) analysis would be performed on each individual justice's statements from oral arguments.

As with much of the ongoing research, the goal in LDA is to isolate the latent topic structures of text through the construction of an unsupervised topic model—one that is unable to impose topic and feature structures (Blei, Ng and Jordan 2003). These approaches are designed to excavate the latent topics of texts and to facilitate the subsequent organization of the subparts in accordance with those topics (so if "elements of a felony" is found to be a top level topic, the subparts of the text that share that topic can be aggregated). These topic models have been used to make recent inroads in political science and in the public law subfield (Hopkins and King, 2010; Grimmer, 2010, Rice, 2012; and Denny, 2015). Such algorithmic approaches vary somewhat, but the constant theme is the division of each document's topic classifications using probabilistic assessment with several desirable features (blindness to ideological categorization, topics selected only based upon the language within the text, and the capacity to reanalyze subsections to further divide topics into more granular levels).

The emergence of the mixed-membership LDA model has further advanced to field due to its capacity to asses documents as the products of multiple latent topics if necessary (Blei and Lafferty, 2009). Further, LDA allows the estimation of the proportionality of each topic within each with inferences calculated based on posterior probabilities.

Returning to the research design, justices would be grouped by ranges of Martin-Quinn scores for this analysis, and results would be utilized to determine the extent to which the languages used by the justices varies based upon their distance from the Martin-Quinn score middle-point. The statements of the justices from the human-coded end of the study would be broken into two groups by legal fidelity scores and the text with higher legal fidelity score would be tested against the text with lower legal fidelity score again using LDA analysis to once again compare the topic structures. Side-by-side comparison of the topic structures from the texts that were divided based upon the speakers' Martin-Quinn score for that term would be made against the texts that were divided based upon the human coding to assess the general congruence or incongruence of the two methods. A second pass would compare topic structures of statements from justices clustered near the zero-point on the Martin-Quinn scale with those clustered at the higher end of the range.

The LDA results are expected to be "rough," and should provide a "high altitude" survey of the topics that are significant and common to the text samples provided. Because it is unsupervised, LDA analysis will provide a range of topics, not all of which will be germane to this project. It will be necessary to cull out topics that are not of use to this project. As the crux of this project is the interplay between the role of ideology and fidelity to the law, a useful result would be, for example, finding that a topic such as "change" or "reverse" was frequent in both the human coding and LDA results and that the frequency of such a topic skews higher in justices who have Martin-Quinn scores that are further away from the neutral zero point.

Ideally each approach (human coding and LDA) would provide a similar grouping of topics from method-to-method, but with distinct variation between high-ideology and low-ideology categories. Terms associated with higher-and lower-ideology categories could then be utilized to develop dictionaries, and when fed into the LIWC program, would allow virtually immediate answers for scholars seeking objective reporting on the tendency of a given piece of legal text to hew closely to the law or to veer away from the law (provided the program was tuned to seek out rhetorical clarity). Such an unbiased reporter is missing from our current toolkit and such a method, if properly developed with sufficient checks for robustness in place, would be of use in the study of judicial attitudes, judicial psychology, legal interpretation, and of courts as an institution (naturally, the research design would incorporate a protocol for the addition of emerging opinions from the Court so as to continuously refine the data).

Although LDA itself has become less challenging to implement (especially through the use of the "MAchine Learning for LanguagE Toolkit," unsupervised LDA itself would have some limitations for this application that should be acknowledged. In the main, LDA produced its best results with vast numbers of documents. As of 2016 around 7,000 separate documents could be produced (11 available terms with approximately 80 cases each being interrogated by 8 justices in most instances (only 8 justices because Justice Thomas virtually never participates)). While that total is adequate, LDA does gain dependability as the number of documents with which it is presented increases, with over 10,000 being optimal. Also, as with many machine reading options LDA does not provide for sentiment, so linguistic expression of concepts such as sarcasm and irony cannot be given any special treatment by the underlying program.

156

The second approach would be to implement Joint Sentiment/Topic (JST hereinafter) modeling which is a further adaptation and extension of LDA that allows for sentiment analysis (Lin and He 2009; Lin, He, Everson and Rüger 2012). At this time sentiment analysis through machine reading (the automated detection of subjective data such as opinions, attitudes, and emotions expressed in text) would be best described as an aspirational goal of the machine learning community, but constant improvements are being made. JST, if it can be properly calibrated and optimized to approach the language of the Supreme Court, appears to have significant potential for the application envisioned herein, especially when the model can be utilized in a semi-supervised fashion, i.e., when a domain-independent sentiment lexicon is provided.[22]

7.9 Objective Number Three

It is understood that each individual placed in the position of power that a justice or a judge holds is assumed to have a certain capacity to properly execute the functions which that station demands. Each will possess, to one degree or another, a "judicial temperament" that in its best expression should be considered somewhat akin to Lincoln's "better angels of our nature"—an inclination to rise to the occasion and do the objectively correct thing. The judicial temperament could be taken to be the individual jurist's predisposition towards behaving in a professional with respect for the doctrine of stare decisis; their manifest interest in assuring the integrity of both the legal system at large and the inherent fairness of their own courtroom; and, their deference towards the doctrine of separation of powers that should to an extent override any inclination to

---

[22] A C++ implementation of JST authored by Lin and He has been made available on GitHub which will be employed to implement this approach.

"legislate from the bench." It should be uncontroversial to assert that, as with many other attributes that humanity collectively exhibits (vis., intelligence, empathy, athletic prowess), it is highly unlikely that each of us possesses the precisely same quantity of judicial temperament.

While it is well understood that a majority of citizens wish to have a federal judiciary that is committed to fairly and evenhandedly applying legal standards,[23] we have oddly made virtually no progress in the development of any sort of rigorous, impartial testing to determine if a given individual possesses a capacity to curb the tendency to reason in a motivated fashion, and rather to deliver unbiased judgment. Absent any organized, scientific approach it is likely that the haphazard selection process that is currently employed is yielding a suboptimal roster of judges and justices. Looking towards our future appointees, we can be confident that while some will be talented and inherently "good" at the job, others will be lacking. We have rigorous testing programs in place for many professions (commercial airline pilots, surgeons, even commercial divers), however the terminal test that most judges will have passed is the bar exam which is not designed to evaluate the individual's inherent capacity to ignore their own biases and to instead follow stare decisis.

Although an eventual solution is well in the offing, with the development of a well designed linguistic test that could be applied to legal writings to aid in determine the individual's level of commitment to following the law ex ante, we would be one step closer to effectively screening our judges.

---

[23] That judges should be independent intelligently approach arguments and be swayed by logic is arguably a normative assumption in the United States (78 percent responded that judges "should be free of political and public pressure." *Justice at Stake Survey*, 2001).

APPENDICES

APPENDIX A

DISCUSSION OF THE PERSUASIVE ELEMENTS OF ARGUMENT IN

HARBISON V. BELL, 556 U.S. 180, 183-189 (2009)

The majority opinion authored by Justice Stevens in Haribson v. Bell held, in

substantive part, that federal law gave indigent death row inmates (who had been

convicted of violating specific federal statutes) the right to federally appointed counsel to

represent them in post-conviction state clemency proceedings when the state has declined

to do so. The action in the matter centers around three sections of the United States Code:

18 § 3599 Counsel for Financially Unable Defendants which makes provision for legal

representation and allied professional assistance of indigent defendants facing the death

penalty in state court and in subsequent proceedings such as appeals under both 28 §

2254 State Custody-Remedies in Federal Courts; and, 28 § 2255 Federal Custody-

Remedies on Motion Attacking Sentence proceedings. Haribson also ruled on whether a

certificate of appealability (COA) is required to appeal an order denying a request for

federally appointed counsel pursuant to § 3599, concluding that said COA was not

necessary.

As is often the case, there are some gaps in the language of the statute and any

reasonably clever individual could quite easily argue both sides of the issue. While it is

agreed that § 3599 extends the guarantee of counsel to both § 2255 (federal) and § 2254

(state) court defendants, subsection (e) of § 3599 is somewhat vague about whether the

Federal Government is obligated to provide assistance of counsel to indigent state court

defendants with respect to their post-conviction state court clemency appeals where the

state has denied such assistance. In charitably describing the statute as vague, what is meant is that the section is not absolutely explicit about this particular point. Because the connection is rather implied, a colorable argument was made to the effect that the federally appointed counsel was available to represent the state court defendant (a term that was applied even after conviction) for appeals made through federal channels but not for appeals made directly for clemency at the state level.

What § 3599 does is unambiguously state that once appointed to represent a state defendant, federally funded counsel "shall also represent the defendant in such ... proceedings for executive or other clemency perhaps available to the defendant." The Court stated explicitly that it was persuaded by Haribson's argument that hinged upon the meaning of the word "available" in § 3599(e):

> "Because state clemency proceedings are "available" to state petitioners who obtain representation pursuant to §§ (a)(2), the statutory language indicates that appointed counsel's authorized representation includes such proceedings."

This was in no way a close decision. The majority opinion was authored by Justice Stevens who was joined by Justices Kennedy, Souter, Ginsburg, and Breyer. Chief Justice Roberts and Justice Thomas, J., filed opinions concurring in the judgment while Justice Scalia filed an opinion concurring in part and dissenting in part, in which Justice Alito joined. Nevertheless, the Court invested a significant effort in explaining in great detail the justifications for ruling in Haribson's favor. Scalia made some efforts to undermine the basic premise that the statutes, as drafted, provide a continuous safety net for indigent state defendant seeking writs of habeas corpus in response to capital punishment sentences handed out for federal violations (the notion that federal crimes can be charged in state courts is central to the action of Haribson and, although the notion

161

could surprise some readers, suffice it to say—without entering into a prolonged exposition regarding criminal procedure—that they can be).

The opinion itself did rouse Justice Scalia into taking a hardhearted swat at the majority in a brief dissent. The points scored by Scalia were few when contrasted with some of his other efforts during his long tenure on the Court. In truth, the avowed textualist Scalia appeared to understand that his side had the less robust argument from a linguistic standpoint and, perhaps for that reason, appeared unwilling to go to any great lengths in support of his own cause despite his longstanding commitment to an unencumbered, and oft utilized, capital punishment option.

QUADRATIC ASSIGNMENT PROCEDURE TESTING

QAPX (a QAP permutation test using x-permutations) and QAPY (a permutation test using y-permutations) results over 1,000 iterations for the WAP data.

Table B.1: Quadratic Assignment Procedure Testing

	QAPX	QAPY
Intercept	32.40 (1.000)	32.40 (1.000)
Term (sending)	6.51 (0.000)	6.51 (0.000)
Term (receiving)	-6.54 (0.000)	-6.54 (0.000)
Term (homophily)	-6.58 (0.000)	-6.58 (0.000)
Author (homophily)	0.00076 (0.428)	0.00076 (0.428)
Issue (homophily)	3.45 (0.000)	3.45 (0.000)
Status (homophily)	-0.29 (0.004)	-0.29 (0.004)
WAP (sending)	-7.64 (0.000)	-7.64 (0.000)
WAP (receiving)	13.48 (0.000)	13.48 (0.000)
WAP (homophily)	-6.40 (0.000)	-6.40 (0.000)
Pages (sending)	0.02 (0.000)	0.02 (0.000)
Pages (receiving)	0.04 (0.000)	0.04 (0.000)
Pages (homophily)	-0.02 (0.000)	-0.02 (0.000)

APPENDIX C

CASES CITED

Berghuis v. Thompkins, 560 U.S. 370 (2010)

Brown v. Allen, 344 U.S. 443 (1953)

Brown v. Board of Education, 347 U.S. 483 (1954)

Bush v. Gore, 531 U.S. 98 (2000)

Citizens United v. FEC, 558 U.S. 310 (2010)

Dred Scott v. Sandford, 60 U.S. 393 (1857)

EPA v. EME Homer City Generation, L.P., 134 S. Ct. 1584 (2014)

Graham v. Florida, 560 U.S. 48 (2010)

Harbison v. Bell, 556 U.S. 180 (2009)

King v. Burwell, 576 U.S. __ (2015)

Lochner v. New York, 198 U.S. 45 (1905)

McCleskey v. Kemp, 481 U.S. 279 (1987)

Miranda v. Arizona, 384 U.S. 436 (1966)

Plessy v. Ferguson, 163 U.S. 537 (1896)

Trop v. Dulles, 356 U.S. 86 (1958)

United States v. Andrea Hall and Richard Magnant, 854 F.2d 1036 (1988)

United States v. Windsor, 570 U.S. __ (2013)

BIBLIOGRAPHY

Abe, Jo Ann A. 2011. "Changes in Alan Greenspan's language use across the economic cycle: A text analysis of his testimonies and speeches." *Journal of Language and Social Psychology* 30(2): 212—223.

Alpers, Georg W., Andrew J. Winzelberg, Catherine Classen, Heidi Roberts, Parvati Dev, Cheryl Koopman, and C. Barr Taylor (2005). "Evaluation of Computerized Text Analysis in an Internet Breast Cancer Support Group." *Computers in Human Behavior* 21 (2): 361—376.

Associated Press, National Constitution Center. Associated Press/National Constitution Center/GfK Poll, Aug, 2012. USAP.091512G.R04. Gfk Roper Public Affairs & Corporate Communications. Storrs, CT: Roper Center for Public Opinion Research, iPOLL, accessed Aug 27, 2015.

Austen-Smith, David. 1990. "Credible debate equilibria." *Social Choice and Welfare*, 7 (1): 75—93.

_____, David. 1990a. "Information Transmission in Debate." *American Journal of Political Science* 34 (February): 124152.

Baird, Vanessa A. 2004. "The effect of politically salient decisions on the US Supreme Court's agenda." *Journal of Politics* 66 (3): 755—772.

Bandum, Erin O'Carroll, and Jason E. Owen (2009). "Evaluating the Validity of Computerized Content Analysis Programs for Identification of Emotional Expression in Cancer Narratives." *Psychological Assessment* 21 (1): 79—88.

Barrett, Paul H. 1974. Early writings of Charles Darwin. In Gruber, Howard E., *Darwin on man. A psychological study of scientific creativity; together with Darwin's early and unpublished notebooks. Transcribed and annotated by Paul H. Barrett, commentary by Howard E. Gruber. Foreword by Jean Piaget*. London: Wildwood House.

Baum, Lawrence. 2006. *Judges and their audiences*. Princeton: Princeton University Press.

Bell, Derrick. 2004. *Silent Covenants:* Brown v. Board of Education *and the Unfulfilled Hopes of Racial Reform*. New York: Oxford University Press.

Benson, Robert W. and Joan B Kessler. 1986. "Legalese v. Plain English: An Empirical Study of Persuasion and Credibility in Appellate Brief Writing." *Loyola of Los Angeles Law Review* 20:301.

Bernstein, David E. 2011. *Rehabilitating Lochner: defending individual rights against progressive reform*. Chicago: University of Chicago Press.

Black, Ryan C., and Ryan J. Owens. 2012. *The solicitor general and the United States Supreme Court: Executive branch influence and judicial decisions*. New York: Cambridge University Press.

_____, Ryan C., and James F. Spriggs II. 2009. "The Depreciation of Precedent on the U.S. Supreme Court." Conference on Empirical Legal Studies.

_____, Ryan C., Sarah A. Treul, Timothy R. Johnson, and Jerry Goldman. 2011. "Emotions, oral arguments, and Supreme Court decision making." *Journal of Politics* 73 (2): 572—581.

Blei, David, and John Laferty. 2009. Topic Models. In *Text Mining: Classification, Clustering, and Applications*, ed. A. Srivastava and M. Sahami. Boca Raton: Chapman and Hall/ CRC Press.

Blei, David, Andrew Ng and Michael Jordan. 2003. "Latent Dirichlet Allocation." Journal of Machine Learning Research 3: 993—1022.

Bommarito, Michael J., Daniel Katz, and John Zelner. 2009. "Law as a Seamless Web? Comparison of Network Representations of the United States Supreme Court Corpus (1791- 2005)." Proceedings of the 12th International Conference on Artificial Intelligence and Law (ICAIL 2009). Available at SSRN: http://ssrn.com/abstract=1419525 or http://dx.doi.org/10.2139/ssrn.1419525

Brader, Ted. 2005. "Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions." *American Journal of Political Science*, 49 (2): 388—405.

Braman, Eileen. 2009. *Law politics, and perception: How policy influences legal reasoning*. Charlottesville: University of Virginia Press.

Brenner, Saul, and Marc Stier. 1996. "Retesting Segal and Spaeth's Stare Decisis Model." *American Journal of Political Science* 40 (4): 1036—1048.

Brigham, John. 1978. *Constitutional Language: An Interpretation of Judicial Decision*. Westport, CT: Greenwood Press.

Butts, Carter T., Mark S. Handcock, and David R. Hunter. (March 1, 2012). network: Classes for Relational Data. R package version 1.7-1. Irvine, CA. http://statnet.org/

Bybee, Keith J. 2010. *All Judges Are Political—Except When They Are Not: Acceptable Hypocrisies and the Rule of Law*. Palo Alto: Stanford University Press.

Calderón, J.L., Morales, L.S., Liu, H. and Hays, R.D. 2006. "Variation in the readability of items within surveys." *American Journal of Medical Quality*, 21 (1): 49—56.

Cardozo, Benjamin, and Andrew Kaufman. 2010 [1921]. *The nature of the judicial process*. New Orleans: Quid Pro Books.

Casillas, Christopher J., Peter K. Enns, and Patrick C. Wohlfarth. 2011. "How public opinion constrains the US Supreme Court." *American Journal of Political Science* 55 (1): 74—88.

Cha, Eun-ho, and Junho Choi. 2012 "Social Network Analysis on Presidential Administration Formation." *정보와사회*/23: 1-34.

Clark, Tom S., and Benjamin Lauderdale. 2012. "The Genealogy of Law." *Political Analysis* 20: 329—350.

Cohen, Shuki J. 2012. "Construction and Preliminary Validation of a Dictionary for Cognitive Rigidity: Linguistic Markers of Overconfidence and Overgeneralization and their Concomitant Psychological Distress." *Journal of Psycholinguistic Research* 41 (5): 347—379.

Collins, Paul M. 2004. "Friends of the Court: Examining the Influence of Amicus Curiae Participation in U.S. Supreme Court Litigation." *Law & Society Review* 38 (4): 807—832.

_____, Paul M., Pamela C. Corley, and Jesse Hamner. 2015. "The Influence of Amicus Curiae Briefs on U.S. Supreme Court Opinion Content." *Law & Society Review* 49 (4): 917—944.

Corley, Pamela C., Paul M. Collins, and Bryan Calvin. 2011. "Lower Court Influence on US Supreme Court Opinion Content." *Journal of Politics* 73 (1): 31—44.

_____, Pamela C., and Justin Wedeking (2014). "The (Dis)Advantage of Certainty: The Importance of Certainty in Language." *Law & Society Review* 48 (1): 35—62.

Cramner, Skyler J. and Bruce A. Desmarais. 2011. "Inferential Network Analysis with Exponential Random Graph Models." *Political Analysis* 19(1): 66-86.

Cross, Frank B., and James W. Pennebaker. 2014. "The Language of the Roberts Court" *Michigan State Law Review*: 853.

Cross, Frank B., Thomas A. Smith, and Antonio Tomarchio. 2006. "Determinants of cohesion in the Supreme Court's Network of Precedents." *San Diego Legal Studies Paper* 07—67.

_____, Frank B., James F. Spriggs II, Timothy R. Johnson, and Paul J. Wahlbeck. 2010. "Citations in the U.S. Supreme Court: An Empirical Study of their Use and Significance." *University of Illinois Law Review*: 489.

Csardi G, Nepusz T: The igraph software package for complex network research, Inter-Journal, Complex Systems 1695. 2006. http://igraph.sf.net

Danelski, David J. 1965. "Values as variables in judicial decision-making: Notes toward a theory." *Vanderbilt Law Review* 19: 721.

Denny, Matthew J., James ben-Aaron, Hanna Wallach, and Bruce Desmarais. 2014. "Topic-Conditioned Hierarchical Latent Space Models for Text-Valued Networks" working paper.

Desmarais, Bruce A. 2010. "The Self-Correcting U.S. Supreme Court." Ph.D. diss. University of North Carolina at Chapel Hill.

Diermeier, Daniel, and Timothy J. Fedderson. 2000. "Information and Congressional Hearings" *American Journal of Political Science* 44 (January): 51—65.

Dreiling, Michael, and Derek Darves. 2011. 'Corporate Unity in American Trade Policy: A Network Analysis of Corporate-Dyad Political Action." *American Journal of Sociology* 116 (5): 1514—63.

Edwards, Kari. 1990. "The Interplay of Affect and Cognition in Attitude Formation and Change." *Journal of Personality and Social Psychology* 59 (2): 202—216.

Eklund, Anders Thomas E. Nichols, and Hans Knutsson. 2016. "Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates." *Proceedings of the National Academies of Science* 113(28): 7900—7905.

Epstein, Lee and Jack Knight. 1998. *The Choices Judges Make*. Washington, D.C.: CQ Press. Newman.

_____, Lee, William M. Landes, and Richard A. Poster. 2013. *The Behavior of Federal Judges*. Cambridge: Harvard University Press.

_____, Lee, Andrew D. Martin, Kevin M. Quinn, and Jeffrey A. Segal. 2007. "Ideological Drift Among Supreme Court Justices: Who, When, and How Important?" *Northwestern University Law Review* 101: 1483—1541.

_____, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, & Organization* 23: 303—325.

Evans, Michael, Wayne McIntosh, Jimmy Lin, and Cynthia Cates. 2007. "Recounting the Courts? Applying Automated Content Analysis to Enhance Empirical Legal Research." *Journal of Empirical Legal Studies* 4 (4): 1007—1039.

Farnsworth, Ward. 2005. "Signatures of Ideology: The Case of the Supreme Court's Criminal Docket." *Michigan Law Review* 104 (67): 5—20.

Feldman, Robert S., James A. Forrest, and Benjamin R. Happ. 2002. "Self-Presentation and Verbal Deception: Do Self-Presenters Lie More?" *Basic and Applied Social Psychology* 24 (2): 163—170.

Fleisher, Madeline. 2008. "Judicial Decision Making Under the Microscope: Moving Beyond Politics Versus Precedent.' *Rutgers Law Review* 60: 919

Fowler, James H., and Sangick Jeon. 2008. "The Authority of Supreme Court Precedent." *Social Networks* 30 (1): 16—30.

Friedman, Barry. 2006. 'Taking Law Seriously.' *Perspectives on Politics* 4 (2): 261—276.

Friedman, Daniela B., and Laurie Hoffman-Goetz. 2006. "A systematic review of readability and comprehension instruments used for print and web-based cancer information." *Health Education & Behavior* 33(3): 352—373.

Friedman, Mike. 2008. "Religious fundamentalism and responses to mortality salience: A quantitative text analysis." *The International Journal for the Psychology of Religion* 18 (3): 216—237.

Fruchterman, T.M.J., and E.M. Reingold. 1991. "Graph Drawing by Force-directed Placement." *Software-Practice and Experience* 21 (11): 1129—1164.

Gerber, Alan S, James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2011. "How large and long-lasting are the persuasive effects of televised campaign ads? Results from a randomized field experiment." *American Political Science Review* 105 (1): 135—150.

Gideon Conway III, Lucian, Kathrene R. Conway, Laura Janelle Gornick, and Shannon C. Houck. 2014. "Automated Integrative Complexity." *Political Psychology* 35(5): 603-624.

Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York: Garden City.

Green, Donald P., and Ian Shapiro. 1994. *Pathologies of Rational Choice Theory: A Critique of Applications in Political Science*. New Haven: Yale University Press.

Grimmer, Justin. 2010. "A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases." *Political Analysis* 18: 1—35.

Grossmann, Matt. 2011. The Long Great Society: Macro Politics, Governing Networks, and American Policy History." In *APSA 2011 Annual Meeting Paper*.

Guthrie, Chris, Jeffrey J. Rachlinski, and Andrew J. Wistrich. 2007. "Blinking on the bench: How judges decide cases." *Cornell Law Review* 93: 1—43.

Hansford, Thomas G., and James F. Spriggs II. 2006. *The Politics of Precedent on the U.S. Supreme Court*. Princeton: Princeton University Press.

Hart, H.L.A. 1958. "Positivism and the Separation of Law and Morals." *Harvard Law Review* 71: 593—629.

_____, H.L.A. 1961. *The Concept of Law*. Oxford: Oxford University Press.

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. 2002. "Latent Space Approaches to Social Network Analysis." *Journal of the American Statistical Association* 97(460): 1090- 1098.

Hopkins, Daniel, and Gary King. 2010. "A Method of Automated Non-Parametic Content Analysis for Social Science." *American Journal of Political Science* 54 (1): 229—247.

Johnson, Timothy R., Paul J. Wahlbeck, and James F. Spriggs II. 2006. "The Influence of Oral Arguments on the U.S. Supreme Court." *American Political Science Review* 100 (February): 99—113.

Johnson, Marcia K., and Carol L. Raye. 1981. "Reality monitoring." *Psychological review* 88 (1): 67—85.

Justice at Stake Survey, Oct, 2001. Retrieved June 28, 2013 from the iPOLL Databank, The Roper Center for Public Opinion Research, University of Connecticut. http://www.ropercenter.uconn.edu.silk.library.umass.edu/data access/ipoll/ipoll.html

Kahn, Jeffrey H., Renee M. Tobin, Audra E. Massey, and Jennifer A. Anderson. 2007. "Measuring Emotional Expression with the Linguistic Inquiry and Word Count." *American Journal of Psychology* 120 (2): 263—286.

Katz, Daniel Martin, and Derek Stafford. 2008. "Hustle and flow: A social network analysis of the American federal judiciary." *Physica A: Statistical Mechanics and its Applications* 389 (19): 4201—4208.

Kelly, J.M. 1996. *A Short History of Western Legal Theory*. New York: Oxford University Press.

Khawaja, M. Asif, Fang Chen, and Nadine Marcus. 2012. "Analysis of collaborative communication for linguistic cues of cognitive load" *Human Factors: The Journal of the Human Factors and Ergonomics Society* 54 (4): 518—529.

Kincaid, J. Peter, Robert P. Fishburne Jr, Richard L. Rogers, and Brad S. Chissom. 1975. *Derivation of new readability formulas (automated read- ability index, fog count and flesch reading ease formula) for navy enlisted personnel.* No. RBR-8-75. Naval Technical Training Command Millington TN Research Branch.

Kitayama, Shinobu, and Eugene Burnstein. 1996 "Social Evidence, Persuasion, and Group Decision Making." In *Persuasion: Psychological Insights and Perspectives*, Sharon Shavitt and Timothy C. Brock, editors. Needham Heights, MA: Paramount Publishing.

Kitts, James A. 2014. "Beyond Networks in structural theories of exchange: promises from computational social science." *Advances in group processes* 31: 263—298.

Knight, Jack. 2009. "Are Empiricists Asking the Right Questions About Judicial Decisionmaking?" *Duke Law Review* 58: 1531—1556.

Kozinski, Alex. 1992. "What I ate for breakfast and other mysteries of judicial decision making." *Loyola Los Angeles Law Review* 26: 993.

Krackhardt, David. 1988. "Predicting with networks: Nonparametric Multiple Regression Analysis of Dyadic Data." *Social Networks* 10: 359—381.

Kritzer, Herbert M. 2001 "The Impact of Bush v. Gore on Public Perceptions and Knowledge of Supreme Court." *Judicature* 85: 32—38.

Krivitsky, Pavel N., and Mark S. Handcock. 2008 "Fitting Position Latent Cluster Models for Social Networks with latentnet." *Journal of Statistical Software* 24 (5): 1—23.

Kunda, Ziva. 1990. "The Case for Motivated Reasoning." *Psychological Bulletin* 108: 1018—1035.

Landes, William N., and Richard A. Posner. 1976. Legal Precedent: A Theoretical and Empirical Analysis. *The Journal of Law and Economics* 19: 249.

Lazer, David. 2011. "Networks in Political Science: Back to the Future." *PS: Political Science & Politics* 44: 61—68.

Lax, Jeffrey R., and Charles M. Cameron. 2007. "Bargaining and Opinion Assignment on the U.S. Supreme Court." *Journal of Law, Economics, & Organization* 23: 276—292.

Levi, Edward H. 1949. *An Introduction to Legal Reasoning*. Chicago: University of Chicago Press.

Lin, Chenguhua, and Yulan He. 2009. "Joint Sentiment/Topic Model for Sentiment Analysis." In *Proceedings of the 18th ACM conference on Information and knowledge management*, ACM: 375—384.

_____, Chenghua, Yulan He, Richard Everson, and Stefan Rüger. 2012. "Weakly-supervised joint sentiment topic detection from text." *IEEE Transactions on Knowledge and Data* Engineering 24 (6): 1134—1145.

Maltzman, Forrest, James F. Spriggs II, and Paul J. Wahlbeck. 2000. *Crafting Law on the Supreme Court: The Collegial Game*. New York: Cambridge University Press.

Martin, Andrew D., and Kevin M. Quinn. 2002. ''Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999.'' *Political Analysis* 10 (2): 134–153.

McAtee, Andrea, and Kevin T. McGuire. 2007. "Lawyers, Justices, and Issue Salience: When and How Do Legal Arguments Affect the U.S. Supreme Court?" *Law & Society Review* 41 (2): 259-278.

McCall, W. A., and L.M. Crabbs. 1961. *Standard test lessons in reading*. NewYork: Teachers College Press.

McGuire, Kevin T. 1998. "Explaining Executive Success in the U.S. Supreme Court." *Journal of Politics* 57 (February): 187—196.

Meade, C. D., and C.F. Smith. 1991. "Readability formulas: Cautions and criterias." *Patient Education and Counseling*, 17: 153—158.

Mersky, Roy, Steven Barkan, and Donald "Duck" Dunn. 2009. *Legal Research Illustrated* 9th ed. Eagan, MN: Foundation Press.

Miles, Lee, Inga Narbutaité Aflaki, and Evangelia Petridou, eds. 2015. *Entrepreneurship in the Polis: Understanding Political Entrepreneurship*. Farnham, Surry: Ashgate Publishing, Ltd.

Millar, Murray G., and Karen U. Millar. 1990. "Attitude change as a Function of Attitude Type and Argument Type." *Journal of Personality and Social Psychology* 59 (2): 217—228.

Miller, Wentworth. 2005. *The Legal Essay Exam Writing Primer*, Seventh Edition. Coraopolis, PA: Wentworth Miller.

Moore, George Edward. 1925. *A defence of common sense*. London: George Allen & Unwin Ltd.

Murphy, Walter F. 1964. *Elements of Judicial Strategy*. Chicago: University of Chicago Press.

Newman, M.E.J. 2010. *Networks*. New York, NY: Oxford University Press.

Newman, Matthew L., James W. Pennebaker, Diane S. Berry, and Jane M. Richards. 2003. "Lying words: Predicting deception from linguistic styles" *Personality and Social Psychology Bulletin* 29 (5): 665-675.

O'Brien, David, Editor. 2004. *Judges on Judging: Views from the Bench*. Washington D.C.: CQ Press.

_____, David. 2008. *Storm Center: The Supreme Court in American Politics*. New York: WW Norton.

Owens, Ryan J., and Justin P. Wedeking. 2011. "Justices and legal clarity: Analyzing the complexity of US Supreme Court opinions" *Law & Society Review* 45(4): 1027—1061.

_____, Ryan J., Justin P. Wedeking, and Patrick C. Wohlfarth. 2013. "How the Supreme Court Alters Opinion Language to Evade Congressional Review." *Journal of Law and Courts* 1 (1): 35—59.

Paik, Anthony, Ann Southworth, and John P. Heinz. 2007. "Lawyers of the right: Networks and organization." *Law & Social Inquiry* 32 (4): 883—917.

Peoples, Clayton D., and James E. Sutton. 2007. "Congressional bribery as state-corporate crime: a social network analysis." *Crime, Law and Social Change* 64 (2-3): 103—125.

Perry, H.W. Jr. 1991. *Deciding to decide: adgenda setting in the United States Supreme Court*. Cambridge: Harvard University Press.

Porter, Mason, Jukka-Pekka Onnela. 2009. "Communities in Networks." *Notices of the AMS* 56(9): 1082-1097.

Popper, Karl. 1959 [1934]. *The Logic of Scientific Discovery*. New York: Routledge.

McCarty, Nolan, Keith T. Poole, and Howard Rosenthal. 2016. *Polarized America: The dance of ideology and unequal riches*. Cambridge: MIT Press.

Posner, Richard A. 2004. *Frontiers of Legal Theory*. Cambridge: Harvard University Press.

_____, Richard A. 2008. *How Judges Think*. Cambridge: Harvard University Press.

Pritchett, C. Herman. 1948. *The Roosevelt Court: A Study in Judicial Politics and Values, 1937—1947*. New York: Macmillian.

R Development Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org.

Rice, Douglass. 2015. "Measuring the Issue Content of Supreme Court Opinions through Probabilistic Topic Models" *Political Research Quarterly* tbd.

Richards, Mark J., and Herbert Kritzer. 2002. "Juisprudential Regimes in Supreme Court Decision Making." *American Political Science Review* 96 (2): 305—320.

Ringsmuth, Eve M., and Timothy R. Johnson (2013). "Supreme Court Oral Arguments and Institutional Maintenance" *American Politics Research*: 1532673X12471756.

Robinson, Robert. 2010. "The Relative (Un)Importance of Rehnquist Court Decisions." *Politics & Policy* 38 (5): 907—936.

Rosen, Jeffrey. 2007. "The Supreme Court: Judicial Temperament and the Democratic Ideal." *WAsHBuRn Law Journal* 47: 1.

Rubin, Edward, and Malcolm Feeley. 1996. "Creating Legal Doctrine" *Southern California Law Review* 69: 1989

Russell, Bertrand. 1921. *The Analysis of Mind*. London: George Allen & Unwin Ltd.

Scalia, Antonin and Bryan A. Garner. 2008. *Making your case: The art of persuading judges*. St. Paul, MN: Thomson/West.

_____, Antonin and Bryan A. Garner. 2012. *Reading Law: The Interpretation of Law Texts*. St. Paul: Thomson/West.

Schwartz, Bernard. 1996. *Decision: how the Supreme Court decides cases*. New York: Oxford University Press.

Schlenker, Barry R. 1980. *Impression Management*. Monterey, CA: Brooks/Cole.

_____, Barry R., and Michael F. Weigold. 1990. "Self-consciousness and self-presentation: being autonomous versus appearing autonomous." *Journal of Personality and Social Psychology* 59: 820—828.

Segal, Jeffrey A., and Albert D. Cover. 1989. "Ideological Values and the Votes of U.S. Supreme Court Justices." *American Political Science Review* 82 (2): 557—565.

_____, Jeffrey A., Lee Epstein, Charles M. Cameron, and Harold J. Spaeth. 1995. "Ideological values and the votes of US Supreme Court justices revisited." *Journal of Politics* 57 (3): 812—823.

_____, Jeffrey A., and Harold J. Spaeth. 1993. *The Supreme Court and the attitudinal model.* New York: Cambridge University Press. *American Political Science Review* 40(4): 971-1003.

_____, Jeffrey A., and Harold J. Spaeth. 1996. "The Influence of *Stare Decisis* on the Votes of United States Supreme Court Justices." *American Journal of Political Science* 40 (4): 971—1003.

_____, Jeffrey A., and Harold J. Spaeth. 1996a. "Norms, dragons, and Stare Decisis: A Response." *American Journal of Political Science* 40 (4): 1064—82.

_____, Jeffrey A., and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. New York, NY: Cambridge University Press.

Si, Luo, and Jamie Callan. 2001. "A statistical model for scientific readability." In Proceedings of the tenth international conference on Information and knowledge management ACM 574-576.

Sim, Yanchuan, Bryan Routledge, and Noah A. Smith. 2014. "The utility of text: The case of Amicus briefs and the Supreme Court" arXiv preprint arXiv :1409.7985.

Songer, Donald R., and Stefanie A. Lindquist. 1996. "Not the Whole Story: The Impact of Justices' Values on Supreme Court Decision Making." *American Journal of Political Science* 40 (4): 1049—1063.

Spaeth, Harold J., Lee Epstein, Andrew D. Martin, Jeffrey A. Segal, Theodore J. Ruger, and Sara C. Benesh. 2012 Supreme Court Database, Version 2012 Release 01. URL: http://Supremecourtdatabase.org

_____, Harold J., and Jeffrey A. Segal. 1999. *Majority Rule or Minority Will: Adherence to Precedent on the Supreme Court*. New York: Cambridge University Press.

Staudt, Nancy, Lee Epstein, and Pater Wiedenbeck. 2006. "The Ideological Component of Judging in the Taxation Context" *Washington University Law Review* 84: 1815-1820.

Sunstein, Cass. 1999. *One Case at a Time: Judicial Minimalism on the Supreme Court*. Cambridge: Harvard University Press.

Tausczik, Yla R., and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods*." Journal of Language and Social Psychology* 29 (1): 24—54.

Tsiamyrtzis, Panagiotis, Jonathan Dowdall, Dvijesh Shastri, I. T. Pavlidis, M. G. Frank, and P. Ekman. 2007. "Imaging facial physiology for the detection of deceit." *International Journal of Computer Vision* 71 (2): 197—214.

Unah, Isaac, and Ange-Marie Hancock. 2006. "US Supreme Court Decision Making, Case Salience, and the Attitudinal Model." *Law & Policy* 28 (3): 295—320.

Undeutsch, Udo. 1989. "The development of statement reality analysis." In *Credibility assessment* Dordrecht, Netherlands: Springer.

Vrij, Aldert, Ronald Fisher, Samantha Mann, and Sharon Leal. 2006. "Detecting deception by manipulating cognitive load." *Trends in cognitive sciences* 10 (4): 141—142.

Ward, Michael D., Katherine Stovel, and Audrey Sacks. 2011. "Network Analysis and Political Science." *Annual Review of Political Science 2011* 14: 245-264.

Wittgenstein, Ludwig. 2001 [1953]. *Philosophical Investigations*. Oxford: *Basil Blackwell Publishing*.

Walton, Douglas N. 1989. *Informal Logic: A Handbook for Critical Argumentation*. New York: Cambridge University Press.

_____, Douglas N. 2002. *Legal Argumentation and Evidence*. University Park: Pennsylvania State University Press.

Wasserman, Stanley and Katherine Faust. 1997. *Social Network Analysis*. New York, NY: Cambridge University Press.

Zhou, Lina. 2004. "Automating Linguistics-Based Cues for Detecting Deception in Text-based Asynchronous Computer Mediated Communication." *Group Decision and Negotiation* 13: 81—106.