

THE SUPREME COURT AND THE JUDICIAL GENRE

Michael A. Livermore, Allen B. Riddell & Daniel N.
Rockmore*

Although the Supreme Court is a singular institution within the American judiciary, it remains recognizably a court. Among the many institutional characteristics that mark the Court as a court, perhaps the most defining is that it exercises its law-making function via justificatory opinions that accompany dispositions in individual cases. By conforming its lawmaking texts to the judicial genre, the Court distinguishes itself from other government actors, such as Congress and the president, that exercise government authority through other textual forms, such as statutes and executive orders. This Article presents the results of a quantitative analysis of Supreme Court opinions that measures the degree to which the content of the Court's opinions conforms to, or departs from, the judicial genre. With the opinions of the federal appellate courts as a baseline, we use topic modeling to estimate the degree of semantic distinctiveness of Supreme Court opinions and track changes in that distinctiveness over the second half of the twentieth century. We find that the Court has become measurably more distinctive over time. We further find that the divergence of the Court's opinions from the judicial genre is not due to the selection of an increasingly non-representative pool of cases for review. Rather, the Court is analyzing and writing about a similarly representative pool of cases, but in an increasingly idiosyncratic fashion. We accompany this quantitative analysis with a qualitative analysis of individual Supreme Court opinions that draws out the significance of this change.

* Michael A. Livermore is an Associate Professor at the University of Virginia School of Law. Allen B. Riddell completed this work as a postdoctoral fellow at the Neukom Institute for Computational Science and the Leslie Center for the Humanities at Dartmouth College. He is now an Assistant Professor in the School of Information and Library Sciences at the University of Indiana. Daniel N. Rockmore is the William H. Neukom 1964 Professor of Computational Science and Professor of Mathematics and Computer Science at Dartmouth College. He is also a Member of the External Faculty of The Santa Fe Institute. Our gratitude to Quinn Curtis, Michael Gilbert, Andrew Hayashi, John Setear, and participants of the 2016 Conference for Empirical Legal Studies in Europe for helpful comments.

TABLE OF CONTENTS

INTRODUCTION	838
I. GENREIFYING THE COURT	844
A. The Judicial Opinion as a Legitimizing Symbol	845
B. Defining Distinctiveness from the Ground Up	853
C. Sources of Difference	857
II. TOPIC MODELING OPINIONS.....	862
A. Topic Models.....	862
B. Data and Topics	865
C. Testing the Topic Model	869
III. THE INCREASINGLY IDIOSYNCRATIC COURT	871
A. Supreme Court Opinions as a Genre Apart	871
B. Subject-Matter Selection	874
C. Trending Toward Distinctiveness.....	876
IV. CASE STUDIES IN DIFFERENCE.....	881
A. 1950–1955	881
B. 2000–2005	884
C. Directions for Future Research.....	887
CONCLUSION	890
APPENDIX	893

INTRODUCTION

The U.S. Supreme Court is a singular institution within the American judiciary. It is uniquely powerful, sitting atop a judicial hierarchy in which lower courts are legally bound by its pronouncements. Its unique institutional features include a small docket, the ability to select the cases that it will decide, and a tradition of sitting as a whole, rather than in panels. The Court also plays a unique role in American political and social life: in the last several decades alone, its decisions have spurred social movements,¹ reshaped core cultural institutions,² and settled presidential elections.³

Although as an institution the Supreme Court is distinctive, it remains recognizably a court. The Supreme Court shares certain rituals with other U.S. judicial institutions, such as the black robe and gavel. It also shares many procedures with other courts, including adversarial hearings and restrictions on ex parte contacts. Perhaps most important, its mode of decision making is through case-by-case adjudication, typically in the course of hearing an appeal from a lower-court

1. See, e.g., *Roe v. Wade*, 410 U.S. 113 (1979).

2. See, e.g., *Obergefell v. Hodges*, 135 S. Ct. 2584 (2015); *Loving v. Virginia*, 388 U.S. 1 (1967).

3. See, e.g., *Bush v. Gore*, 531 U.S. 98 (2000).

decision.⁴ As such, when the Court creates, amends, or clarifies legal obligations, it does so not through directly stated rules (as in a statute or regulation), but through the justificatory documents that accompany a disposition in a particular case.

The court-like nature of the Supreme Court appears to be important for its institutional legitimacy. Scholars have long puzzled over the substantial reservoir of public prestige enjoyed by the Court, in light of its lack of democratic accountability.⁵ This “diffuse support” has been the subject of considerable theorizing and study,⁶ with leading accounts grounding the Court’s legitimacy in the popular association of judicial institutions with a principled style of decision making that is removed from the posturing and compromise that is routine in the political branches.⁷ Symbols that are associated with the judiciary, such as the courtroom and the gavel, appear to play a role in triggering positive attitudes toward the Court’s decisions, even among people who disagree with the outcomes of individual cases.⁸

Just as the Justices symbolically distinguish themselves from other government actors by donning their robes, the particular textual form in which the Court’s lawmaking power is embodied—the written opinion—is a hallmark of the exercise of judicial power. The judicial opinion can be understood as a genre of legal text that is “characterized by a particular form, style, or purpose.”⁹ Defining the concept of *genre* is notoriously difficult, and genre categorizations—whether in music, literature, or art—are often controversial.¹⁰ For our purposes, the concept of judicial genre is related to the role of judicial symbols in legitimating the Court as an institution. That is, the genre of the judicial opinion is among the defining characteristics of the Court. Even within the larger genre of “law writing,” judicial opinions are a distinct subcategory. Judicial opinions are easily distinguished in form, style, or purpose from the texts produced by other American legal institutions,

4. The only area in which the Supreme Court can issue general rules outside the context of deciding individual cases is for the procedures that govern the conduct of the courts. Rules Enabling Act of 1934, 28 U.S.C. §§ 2071–2077 (2012).

5. See generally ALEXANDER M. BICKEL, *THE LEAST DANGEROUS BRANCH: THE SUPREME COURT AT THE BAR OF POLITICS* (1962); Robert A. Dahl, *Decision-Making in a Democracy: The Supreme Court as National Policy-Maker*, 6 J. PUB. L. 279 (1957).

6. See *infra* Section I.A.

7. See, e.g., John M. Scheb II & William Lyons, *The Myth of Legality and Public Evaluation of the Supreme Court*, 81 SOC. SCI. Q. 928, 928 (2000) (reporting a survey of public attitudes toward the Court and correlating those with the “myth of legality,” which the authors define as “the notion that the Supreme Court’s decisions are based on legal principles rather than on political influences”); JOHN R. HIBBING & ELIZABETH THEISS-MORSE, *CONGRESS AS PUBLIC ENEMY: PUBLIC ATTITUDES TOWARD AMERICAN POLITICAL INSTITUTIONS* (1995). See *infra* for a more nuanced discussion.

8. See James L. Gibson, Milton Lodge & Benjamin Woodson, *Losing, But Accepting: Legitimacy, Positivity Theory, and the Symbols of Judicial Authority*, 48 LAW & SOC. REV. 838 (2014).

9. *Genre*, OXFORD ENGLISH DICTIONARY (2d ed. 1989) (second definition). See generally Robert A. Ferguson, *The Judicial Opinion as Literary Genre*, 1 YALE J.L. & HUM. 201, 202 (1990) (examining “the appellate judicial opinion as a distinct literary genre within the larger civic literature of the American republic of laws”).

10. See generally RICK ALTMAN, *FILM/GENRE* (1999); DAVID FISHELOV, *METAPHORS OF GENRE: THE ROLE OF ANALOGIES IN GENRE THEORY* (1993).

such as the regulations promulgated by administrative agencies or the statutes adopted by legislatures. When the Court exercises its power through texts that are recognizable as judicial opinions, it marks itself out as a judicial institution, triggering a set of positive associations that tend to enhance support for—or at least willingness to accept—its decisions.

The legitimacy-enhancing value of certain symbols, such as robes, is contingent on their association with judicial institutions. Were the Court to adopt overly distinctive garb that is not associated with the judiciary more generally—or even worse the attire of politicians—it would run the risk of undermining the symbolic value of its costume. Similarly, if its writings are easily distinguishable from the opinions of courts more generally, their power as a legitimating symbol could be diminished. At the margins, it may be difficult to know at what point a garment is no longer a judicial robe or a building is no longer a courthouse; similarly, a document’s status as a recognizable judicial opinion will often not be a straightforward binary distinction. Furthermore, the genre of the American judicial opinion has changed over time, and a document that falls well within the conventions and norms of one time may be highly distinctive from another period. Nevertheless, although uncontroversial genre classifications may be difficult, there is likely to be a more broadly recognizable spectrum of similarity and distinctiveness.

A growing body of research and commentary articulates and interprets changes and trends in the Court’s writings.¹¹ Two authors of this Article (Michael Livermore and Daniel Rockmore) along with Keith Carlson use quantitative methods to explore changes in writing style over time on the U.S. Supreme Court. That analysis finds that contemporaneous Justices tend to write in identifiably similar ways to each other.¹² Other scholars have examined change in certain stylistic features of the Court’s opinions, such as opinion length, as well as trends in language usage.¹³ Some stylistic changes, such as the rise of complex, multi-part opinions with several concurrences and dissents, have been broadly noted by scholars and commentators.¹⁴

11. See, e.g., WILLIAM D. POPKIN, *EVOLUTION OF THE JUDICIAL OPINION: INSTITUTIONAL AND INDIVIDUAL STYLES* (2007) (surveying changes in writing styles over the Court’s history).

12. Keith Carlson, Michael A. Livermore & Daniel Rockmore, *A Quantitative Analysis of Writing Style on the U.S. Supreme Court*, 93 WASH. U. L. REV. 1461 (2016).

13. See RYAN C. BLACK ET AL., *U.S. SUPREME COURT OPINIONS AND THEIR AUDIENCES* (2016) (examining how the Supreme Court alters language usage by context and audience); Frank B. Cross & James W. Pennebaker, *The Language of the Roberts Court*, 2014 MICH. ST. L. REV. 853, 865 (2014) (using “Linguistic Inquiry and Word Count” program to examine texts of Roberts Court opinions); Ryan C. Black & James F. Spriggs II, *An Empirical Analysis of the Length of U.S. Supreme Court Opinions*, 45 HOUS. L. REV. 621 (2008) (finding cyclic time trends in opinion length).

14. See, e.g., Morton J. Horowitz, *The Supreme Court, 1992 Term-Foreword: The Constitution of Change: Legal Fundamentality Without Fundamentalism*, 107 HARV. L. REV. 30, 98–99 (1993) (noting proliferation of “multiple opinions, designated in Parts, sub-parts, and sub-sub-parts”); Adam Liptak, *Justices are Long on Words but Short on Guidance*, N.Y. TIMES, Nov. 18, 2010, at A1.

In this Article, we extend this body of work by asking if the Court's writings as a whole have grown more semantically distinctive over the course of the twentieth century, as compared to the judicial opinions issued by other American courts. Our analysis is based on a large collection of opinions written by the Supreme Court and federal courts of appeals. We use the appellate court opinions as a baseline and examine whether Supreme Court opinions are becoming measurably different from that anchor point over time. The courts of appeals are a conservative point of departure for our analysis because if there were a set of opinions that could be expected to co-evolve with the Court, it would be those produced by the courts that the Court most directly supervises. If the Supreme Court is becoming more distinctive vis-à-vis the federal courts, then we can be fairly certain that there would be a similar change with respect to other U.S. judicial institutions.

Methodologically, our approach is intentionally interdisciplinary and draws from the field of computational text analysis, which among other things has developed approaches that address the problem of articulating the varying semantic content in a large set of documents—in this case, hundreds of thousands of court cases.¹⁵ Although Court observers have examined related issues in a qualitative way,¹⁶ traditional empirical methods that rely on reading and hand-coding opinions are ill-suited to studying nuanced differences in language usage that are difficult for human readers to systematically capture. Existing coding of issue area and case outcome provides valuable metadata on decisions, but there is an obvious loss of information as thousands of words of text are reduced to a handful of coded variables. The large number of cases to be processed also poses substantial resource challenges even if a reliable human coding methodology could be developed.

For this study, we exploit the technique of *topic modeling*, a computational and machine-learning approach to text analysis that is now a decade old and has spread to a number of academic disciplines, including history and political science.¹⁷ Scholars have only recently begun to apply topic models to legal texts, and a contribution of this project is to introduce the tool of topic modeling to the community of legal scholars in a familiar and important context—the analysis of

15. See, e.g., Sara Klingenstein, Tim Hitchcock & Simon DeDeo, *The Civilizing Process in London's Old Bailey*, 111 PROC. NAT. ACAD. SCI. 9419 (2014) (comparing transcripts of violent and non-violent criminal proceedings in eighteenth- and nineteenth-century English proceedings); James M. Hughes, Nicholas J. Foti, David C. Krakauer & Daniel N. Rockmore, *Quantitative Patterns of Stylistic Influence in the Evolution of Literature*, 109 PROC. NAT. ACAD. SCI. 7682 (2012) (examining changes in writing style in the major Western literary texts).

16. Ashutosh Bhagwat, *Separate but Equal?: The Supreme Court, The Lower Federal Courts, and the Nature of the "Judicial Power,"* 80 B.U. L. REV. 967 (2000).

17. See David M. Blei, Andrew Y. Ng & Michael I. Jordan, *Latent Dirichlet Allocation*, 3 J. MACHINE LEARNING RES. 993 (2003); Allen B. Riddell, *How to Read 22,198 Journal Articles: Studying the History of German Studies with Topic Models*, in DISTANT READINGS: TOPOLOGIES OF GERMAN CULTURE IN THE LONG NINETEENTH CENTURY 91 (Matt Erlin & Lynne Tatlock eds., 2014); Kevin M. Quinn et al., *How to Analyze Political Attention with Minimal Assumptions and Costs*, 54 AM. J. POL. SCI. 209 (2010).

judicial opinions.¹⁸ Our analysis is based on an application of a topic model to all of the text in opinions issued by the Supreme Court and appellate courts from 1951 to 2007. This is a corpus of almost 300,000 texts.

Broadly, topic models generate two outputs. First are the *topics*. These are probability distributions over the vocabulary in the corpus, meaning that they are lists of numerical weightings of the words in the corpus. The topics can be thought of as representing subject-matter categories, with the connection to the colloquial meaning of topic deriving from a gestalt based on a given distribution's most highly weighted words. A word cloud can provide a sense of the content of a document in a similar fashion, with the most highly weighted words as the largest words in the cloud; taken together, these frequent words give a sense of the subject matter encoded by the distribution.¹⁹ Individual documents are then characterized (i.e., encoded) as a distribution over the topics so that they are represented as comprising the various topics. The number of topics needed to "fit" a corpus will vary with the corpus composition.²⁰ Topic models enable the identification of latent groupings of texts (e.g., according to similarity of topic distribution) when applied to an unstructured corpus. They can also aid analysis of the semantic content of documents by reducing the number of dimensions from an unmanageable number of distinct words in the vocabulary to a manageable number of topics. This reduces the dimensionality of a dataset from many thousands of words to between 10 and roughly 100 topics. Topic modeling is now a relatively mature algorithmic technology and executes efficiently, even on a set of a few hundred thousand documents.²¹

On the basis of the outputs of the topic model, we conduct three analyses. First, we compare the distribution of topics in Supreme Court opinions with the distribution in the appellate courts to estimate a degree of difference of the semantic content in the two corpora (i.e., collection of documents). We find that there are significant and meaningful semantic differences between the two corpora. This finding provides quantitative evidence that the Court's opinions as a collection are recognizably distinct from the opinions of the subordinate federal appellate courts. We also find that the appellate court cases that are selected for review are an intermediary corpus that is distinct both from the general pool of appellate cases and from the Supreme Court corpus.

Second, we examine appellate opinions that are associated with cases selected by the Court for review to provide insight into the certiorari process. There

18. See David S. Law, *Constitutional Archetypes*, 95 TEX. L. REV. 153 (2016); Jonathan Macey & Joshua Mitts, *Finding Order in the Morass: The Three Real Justifications for Piercing the Corporate Veil*, 100 CORNELL L. REV. 99 (2014); Daniel Taylor Young, Note, *How Do You Measure a Constitutional Moment? Using Algorithmic Topic Modeling to Evaluate Bruce Ackerman's Theory of Constitutional Change*, 122 YALE L.J. 1990 (2013).

19. See Florian Heimerl et al., *Word Cloud Explorer: Text Analysis based on Word Clouds*, Presented at 2014 47th Hawaii International Conference on System Science.

20. See Law, *supra* note 18 (discussing measures of fit).

21. One widely used publically available topic modeling software package is the Machine Learning for Language Toolkit ("MALLET"), available at MALLET, UMASS Amherst, <http://mallet.cs.umass.edu/>.

is a considerable body of empirical legal scholarship that seeks to understand the factors that influence the Court's decisions about what cases to select for review.²² This research has identified several variables, such as the presence of amici, that are associated with the likelihood of a grant of certiorari.²³ We contribute to this literature by identifying the topics that are correlated, either positively or negatively, with certiorari. We also examine how the topics that are associated with certiorari have changed over the course of the study period. This analysis has significant potential to contribute to future empirical work through the use of *topic controls* that help researchers draw more solid causal inferences by accounting for unobserved variables that affect both the factor of interest and the likelihood of a grant of certiorari.²⁴

In our final quantitative analysis, we examine whether the differences in semantic content between the Supreme Court and the lower courts are changing over time. We find that, indeed, the Court's opinions are becoming measurably more distinctive. For both of these analyses, we estimate the contribution of the certiorari process to this effect by comparing cases selected for review to the general pool of lower court cases. We find that cases selected for review are topically distinctive from the broader pool, but that they have remained consistently distinctive over time. The Court does not appear to be selecting more unusual cases for review than it has in the past. The Court's opinions themselves, on the other hand, have been growing increasingly idiosyncratic compared to the opinions of the appellate courts.

We augment these quantitative findings with an interpretive exercise. For this analysis, we examine Supreme Court opinions from the beginning and end of the study period. To avoid the problem of cherry-picking, we use a quantitative technique to identify the most- and least-characteristic Supreme Court opinions based on their likelihood of being misidentified as an appellate court opinion. We then engage in a close reading of these opinions to help tease out the meaning and importance of the growing semantic distinctiveness of the Court's opinions.

This is a particularly important time to be examining practices that bear on the legitimacy of the Supreme Court. Over the past several decades, the work of the Court has become increasingly political and polarized.²⁵ Confirmations to the Court have become extremely high-stakes partisan contests, and are now routinely discussed during presidential campaigns. The refusal of the Republican Senate to hold hearings on President Barack Obama's nominee to fill Justice Scalia's vacant seat both highlighted and heightened the degree to which the Court has become submerged within partisan politics.²⁶ If the Court is perceived as just another political institution and less as a court of law (by either elites or the broader public), then its unique position as the arbiter of last resort with the ability to conclusively

22. See *infra* Section I.C.

23. See *infra* notes 82–93 and accompanying text.

24. See *infra* Section III.C.

25. Cf. James L. Gibson, *The Legitimacy of the U.S. Supreme Court in a Polarized Polity*, 4 J. EMPIRICAL LEGAL STUD. 507 (2007).

26. New York Times Editorial Board, Editorial, *The Stolen Supreme Court Seat*, N.Y. TIMES, Dec. 25, 2016, at SR18.

settle legal disputes may be compromised.²⁷ Our work suggests that in its writings—among its most important defining characteristics—the Court may be edging away from its roots.

The structure of the remainder of the Article is as follows. Part I provides motivation for our quantitative analysis and discusses the sources of the Court’s legitimacy and the potential role played by judicial opinions. It then explains our methodology for defining semantic distinctiveness and examines the two major pathways for the Court’s opinions to systematically differ from the opinions in the appellate courts. Part II explains topic models, discusses our data and the topics that are generated by our model, and reports the results of several tests of how well the model captures semantic content of the data. Part III reports our main results: that the Supreme Court’s opinions are becoming more distinctive, and this trend does not result from the Court selecting more unusual cases for review. Rather, the Court is writing about a similarly distinctive pool of cases in an increasingly distinctive fashion. Part IV engages in a qualitative analysis of the important sources of difference between the Court’s opinions and those of the appellate courts, based on samples of more- and less-distinctive opinions as determined by the qualitative topic-model-based measure. It also suggests potential avenues for future research. This Article’s conclusion discusses normative and methodological consequences of our study.

I. GENREFYING THE COURT

The concept of legitimacy, and specifically the elusive source of the Supreme Court’s legitimacy, is a perennial topic of legal and social-science scholarship.²⁸ At a time when public support for the Court has fallen,²⁹ and party polarization over Supreme Court nominees has reached new heights,³⁰ questions about the Court’s legitimacy retain their salience. If conforming to the judicial genre benefits the Court’s legitimacy, it is worth considering whether there are secular

27. See, e.g., Charles Gardner Geyh, *The Supreme Court Is Losing Its Luster*, NEW REPUBLIC (Mar. 11, 2016), <https://newrepublic.com/article/131451/supreme-court-losing-luster>.

28. A Westlaw search of U.S. law-review articles published in the period 1960–2015 with the words *legitimacy*, *legitimate*, or *illegitimate* in the same sentence as *Supreme Court*, and excluding all articles with the word *child* or *children*, returned over 8,000 results. For an overview of recent trends in the relevant social science literature, see James L. Gibson & Michael J. Nelson, *The Legitimacy of the US Supreme Court: Conventional Wisdoms and Recent Challenges Thereto*, 10 ANN. REV. L. SOC. SCI. 201 (2014).

29. For example, in an annual Gallup poll of public confidence in the Court, fewer than 40% of respondents answered that they have a “Great deal” or “Quite a lot” of confidence in each of the past ten years. Prior to 2006, in only one other year since 1975 was confidence below 40%. Supreme Court, GALLUP, <http://www.gallup.com/poll/4732/supreme-court.aspx> (last visited Jan. 15, 2017).

30. Vikram David Amar, *The Grave Risks of the Senate Republicans’ Stated Refusal to Process any Supreme Court Nominee President Obama Sends Them*, JUSTIA, (Feb. 26, 2016), <https://verdict.justia.com/2016/02/26/the-grave-risks-of-the-senate-republicans-stated-refusal-to-process-any-supreme-court-nominee-president-obama-sends-them>.

trends in the degree to which the Court's opinions fit the genre's conventions.³¹ To test whether that is the case, we introduce a quantitative method of describing distinctiveness in collections of texts—what we will call *genrefication*—that begins with an agreed-upon collection of texts and relies as little as possible on strong assumptions about the defining features of a genre.³²

A. *The Judicial Opinion as a Legitimizing Symbol*

One of the classic paradoxes of the American political system is that in a country that purports to hold democratic values dear, final decision-making authority on some of the most hotly contested political issues is vested in a body that is almost entirely free from formal democratic accountability.³³ This paradox has spurred decades' worth of discussion and debate from a variety of perspectives.³⁴ Some philosophers have attacked the Court's peculiar status as inappropriate and undemocratic,³⁵ while others have argued that an independent arbiter can better protect substantive rights than majoritarian institutions.³⁶ Some legal scholars have claimed that the Court can protect democracy from its defects, while others have argued that the Court's premiere place in the political order is inconsistent with the

31. For a more detailed definition and discussion of the concept of genre, see *infra* Section I.B. The use of computational techniques to uncover genre formation (which generally require machine reading of tens or hundreds of thousands of texts) was first deployed in the field of literary criticism. Cf. MATTHEW L. JOCKERS, *MACROANALYSIS: DIGITAL METHODS AND LITERARY HISTORY* (2013); FRANCO MORETTI, *GRAPHS, MAPS, TREES: ABSTRACT MODELS FOR LITERARY HISTORY* (2005).

32. Cf. ALTMAN, *supra* note 10, at 24 (discussing genre in film and noting that “the genre itself is typically thought of as a corpus of films” rather than “a particular complex of topic and structure”). The term *genrefication* (and its alternative spelling *genrification*) does not appear to have been in use before 1980, after which it has enjoyed rapid growth. Ngram Viewer, Google Books, <http://bit.ly/2iEtlAG>. Altman refers to the *genrification process* as the “process of genre constitution and transformation” and emphasized the role of multiple actors, such as (for film) producers, distributors, and critics. ALTMAN, *supra* note 10, at 77. Early uses of the term include in a discussion of the crisis in the humanities and in a critical analysis of the comic Dick Tracy. See GARYN G. ROBERTS, *DICK TRACY AND AMERICAN CULTURE: MORALITY AND MYTHOLOGY, TEXT AND CONTEXT xvii–xviii* (1993); Jerry Herron, *The Genrification of Desire and Posthistorical Pastiche*, 16 *SUBSTANCE* 45 (1987). More recent uses of the term apply it to efforts to implement categorization systems in libraries other than the Dewey Decimal system to increase user friendliness. Kiera Parrott & Karyn M. Peterson, *Ditching Dewey: Hot Topic in Harford*, *SCH. LIBR. J.* (Nov. 20, 2013), http://www.slj.com/2013/11/reviews/ditching-dewey-hot-topic-in-hartford-aasl-2013/#_ (reporting interest in the subject at the 2013 American Association of School Libraries' annual conference).

33. See THE FEDERALIST NO. 78 (Alexander Hamilton) (defending life tenure for judges as an “excellent barrier to the encroachments and oppressions of the representative body”).

34. See, e.g., BICKEL, *supra* note 5. By 2015, over 3,500 law-review and journal articles cited to Bickel's 1962 book. Westlaw (search “Bickel, The Least Dangerous Branch” for dates 1962–2015) (conducted Jan. 16, 2017).

35. See Jeremy Waldron, *The Core of the Case Against Judicial Review*, 115 *YALE L.J.* 1346, 1348 (2006).

36. RONALD DWORKIN, *LAW'S EMPIRE* 356 (1986).

founders' vision of a popular role in constitutional interpretation.³⁷ From an empirical perspective, political scientists and sociologists have offered a variety of theories to explain the counterintuitive fact that the Court enjoys a high level of support by the public even though, from time to time, it reverses the policy choices of democratically elected branches.³⁸

In this discourse, the fraught word *legitimacy* is often raised in connection with the Court and its decisions.³⁹ Professor Richard Fallon refers to three ways in which this term is used: legal legitimacy, moral legitimacy, and sociological legitimacy.⁴⁰ Legal legitimacy turns on the legal permissibility of a government official's conduct.⁴¹ Moral legitimacy is "a function of moral justifiability or respect-worthiness."⁴² Sociological legitimacy, which as a concept has its roots in Max Weber's writings,⁴³ involves the relevant public's perceptions of the conduct of government officials. When the public "regards [acts or institutions] as justified, appropriate, or otherwise deserving of support for reasons beyond fear of sanctions or mere hope for personal reward," those acts or institutions are legitimate from a sociological perspective.⁴⁴ A weaker form of sociological legitimacy entails mere

37. Compare JOHN HART ELY, *DEMOCRACY AND DISTRUST, A THEORY OF JUDICIAL REVIEW* (1981), with LARRY D. KRAMER, *THE PEOPLE THEMSELVES: POPULAR CONSTITUTIONALISM AND JUDICIAL REVIEW* (2004).

38. JAMES L. GIBSON & GREGORY A. CALDEIRA, *CITIZENS, COURTS, AND CONFIRMATIONS* (2009); Or Bassok, *The Sociological-Legitimacy Difficulty*, 26 J. L. & POL. 239 (2011); Gibson & Nelson, *supra* note 28; Michael L. Wells, "Sociological Legitimacy" in *Supreme Court Opinions*, 64 WASH. & LEE L. REV. 1011 (2007) (discussing the Court's practice of "appearance management"); James L. Gibson, Gregory A. Caldeira & Lester Kenyatta Spence, *Measuring Attitudes Toward the United States Supreme Court*, 33 BRIT. J. POL. SCI. 535 (2003); John R. Hibbing & Elizabeth Theiss-Morse, *The Media's Role in Public Negativity Toward Congress: Distinguishing Emotional Reactions and Cognitive Evaluations* 42 AM. J. POL. SCI. 475 (1998); Jeffery J. Mondak, *Policy Legitimacy and the Supreme Court: The Sources and Contexts of Legitimation*, 47 POL. RES. Q. 675, 690 (1994); Gregory A. Caldeira & James L. Gibson, *The Etiology of Public Support for the Supreme Court*, 36 AM. J. POL. SCI. 635 (1992); James L. Gibson, *Understandings of Justice: Institutional Legitimacy, Procedural Justice, and Political Tolerance*, 23 LAW & SOC. REV. 469, 471 (1989).

39. See Jan G. Deutsch, *Neutrality, Legitimacy, and the Supreme Court: Some Intersections Between Law and Political Science*, 20 STAN. L. REV. 169 (1968).

40. Richard H. Fallon, *Legitimacy and the Constitution*, 118 HARV. L. REV. 1787, 1790 (2005).

41. "That which is lawful is also legitimate." *Id.* at 1794. Fallon notes that the inverse is not necessarily true, and distinguishes between "the substantive legal legitimacy of judicial rulings, which reflects their correctness or reasonableness as a matter of law, and their authoritative legitimacy or legally binding character." *Id.* at 1794-95.

42. *Id.* at 1796.

43. See, e.g., Max Weber, *Die drei reinen Typen der legitimen Herrschaft*, 187 PREUSSISCHE JAHRBÜCHER 1 (1922) (appearing later as Max Weber, *The Three Types of Legitimate Rule*, 4 BERKELEY PUBLICATIONS IN SOC'Y & INSTITUTIONS 1 (1958) (Han Gerth trans.)).

44. Fallon, *supra* note 40, at 1795.

acceptance and the failure of acts or institutions to generate active resistance or opposition.⁴⁵

In this Article, we are concerned with sociological legitimacy—that is, the actual perceptions of the public (justified or not) concerning institutions and official acts. This concept of legitimacy can be further refined by distinguishing between what political scientists refer to as *diffuse support*—which acts as a “reservoir of favorable attitudes or good will” toward an institution that facilitates acceptance or endorsement of even unfavorable decisions—and *specific support*, which is based on happiness with individual decisions.⁴⁶

There is some disagreement over whether the Court in fact enjoys diffuse support. From opinion surveys, it is relatively clear that the public in general has a reasonably favorable view toward the Court, and has no particular appetite for reforms that would fundamentally change its role.⁴⁷ Two basic explanations have been offered for this phenomenon. The first argues that the Court enjoys broad popular support because the policies that it promotes through its decisions are broadly popular—if the Court were to issue unpopular decisions, the level of support for the institution would decline.⁴⁸ Under this account, the Court’s current high level of support amounts to a stock of goodwill that it has built up over the years by doing what the public wants, but which is amenable to depletion in the face of decisions

45. *Id.* at 1796.

46. DAVID EASTON, *A SYSTEMS ANALYSIS OF POLITICAL LIFE* 273 (1965). This distinction was emphasized in Caldeira & Gibson, *supra* note 38, at 637 and Dean Jaros & Rober Roper, *The U.S. Supreme Court: Myth, Diffuse Support, Specific Support, and Legitimacy*, 8 AM. POL. Q. 85, 101 (1980). The distinction between diffuse and specific support is related to a different distinction introduced by political scientists between input and output legitimacy. Input legitimacy is based on the procedures used to arrive at a decision, while output legitimacy is based on the effectiveness of those decisions at promoting public well-being. FRITZ W. SCHARPF, *GOVERNING IN EUROPE: EFFECTIVE AND DEMOCRATIC?* 6, 27 (1999); Robert O. Keohane & Joseph S. Nye Jr., *Between Centralization and Fragmentation: The Club Model of Multilateral Cooperation and Problems of Democratic Legitimacy* 3 (John F. Kennedy School of Government, Working Paper No. 04/01, 2001). The input/output distinction was introduced in the context of transnational governmental institutions—most notably the European Union. The primary input feature of interest for the relevant discussion is democratic accountability, which has been viewed as problematic for these institutions.

47. Gibson & Nelson, *supra* note 28, at 205–06.

48. Brandon L. Bartels & Christopher D. Johnston, *On the Ideological Foundations of Supreme Court Legitimacy in the American Public*, 57 AM. J. POL. SCI. 184, 187 (2013).

that flout public preferences.⁴⁹ Under the second account, support for the Court is institutional rather than substantive, and is robust to unpopular decisions.⁵⁰

Various theories have been put forward for why the Court might have such institutional prestige. One is the “myth of legality” view, in which popular support for the Court is grounded in a widespread belief that the Justices reach their decisions through the neutral application of binding and dispositive legal authorities.⁵¹ In its strongest form, the public is theorized to believe that legal decision making is a value-free and technical exercise that does not involve meaningful discretion.⁵² If this view of the legal process is in fact widespread, then the Court’s legitimacy is resistant to disagreement with the outcomes of individual cases.⁵³

A more nuanced version of this hypothesis is that the public is aware and accepts that there is a degree of discretion involved in judging, but believes that the Justices exercise their discretion in a principled, public-regarding fashion, rather than strategically to benefit themselves.⁵⁴ This belief may be supported by a view that the Justices make a good-faith effort to respect the rule of law and conform their decisions to governing authority. Such a view would recognize an important role for the law in judicial decision making, while still allowing room for the values and

49. See Vanessa A. Baird, *Building Institutional Legitimacy: The Role of Procedural Justice*, 54 POL. RES. Q. 333, 334 (2001) (discussing theory that people keep a running tally of whether they view the Court’s decisions favorably); James L. Gibson, Gregory A. Caldeira & Vanessa Baird, *On The Legitimacy of National High Courts*, 92 AM. POL. SCI. REV. 343 (1998). This empirical account on the relationship between support for the Court and the popularity of its decisions also lines up with normative scholarship that defends the Court in the face of its purported “countermajoritarian difficulty” by arguing that the Court actually does follow majority preferences most of the time. BARRY FRIEDMAN, *THE WILL OF THE PEOPLE: HOW PUBLIC OPINION HAS INFLUENCED THE SUPREME COURT AND SHAPED THE MEANING OF THE CONSTITUTION* (2009).

50. Caldeira & Gibson, *supra* note 38, at 657; see also Robert H. Durr, Andrew D. Martin & Christina Wolbrecht, *Ideological Divergence and Public Support for the Supreme Court*, 44 AM. J. POL. SCI. 768, 774 (2000) (finding that public support the Court does respond to evaluation of outcomes, but that “the impact of temporary shocks is relatively short-lived”).

51. Scheb & Lyons, *supra* note 7, at 929.

52. Chief Justice John Roberts appeared to promote such a mechanistic view of the Court’s work during his confirmation hearing when he famously analogized the role of a judge to a baseball umpire. Confirmation Hearing on the Nomination of John G. Roberts, Jr. to be Chief Justice of the United States, Before the S. Comm. on the Judiciary, 109th Cong. 55–56 (2005).

53. Baird, *supra* note 49.

54. Gibson & Nelson, *supra* note 28, at 211; JOHN R. HIBBING & ELIZABETH THEISS-MORSE, *STEALTH DEMOCRACY: AMERICANS’ BELIEFS ABOUT HOW GOVERNMENT SHOULD WORK* 158–59 (2002). *But see* Vanessa A. Baird & Amy Gangl, *Shattering the Myth of Legality: The Impact of the Media’s Framing of Supreme Court Procedures on Perceptions of Fairness*, 27 POL. PSYCH. 597, 606 (2006) (finding that perceptions of the Court are negatively affected by reports of consideration of even public regarding political—rather than legal—factors).

policy preferences of the Justices to play a role in difficult cases.⁵⁵ This more nuanced view would be robust to the widely available and well-publicized evidence that—at least some of the time—Justices’ ideological views predict their decisions.⁵⁶ Holders of this view might accept decisions that they disagree with because they respect the Justices’ decision-making process.

Both the myth-of-legality hypothesis and the related, more realistic view are grounded in the judicial nature of the Court. It is the association of the Court with judicial forms of decision making (either legalistic or principled) that provide it with enhanced prestige. Of course, purported mode of decision making is not the only characteristic that separates courts from legislatures or bureaucracies. Other differences include case-by-case adjudication and the various procedural protections that are unique to courts—these too could play a role in generating acceptance with judicial decisions.⁵⁷ Although it may be difficult to tease out which of the many distinguishing characteristics of courts contribute most to institutional legitimacy, the more general point is that whatever institutional prestige the Court enjoys is derived from the judicial characteristics that it manifests.

In addition to what we might call these functional characteristics of courts are the symbols that differentiate courts from other decision-making bodies.⁵⁸ In particular, the “pageantry of judicial symbols”—which includes “the robes of judges, the honorific forms of address, and the temple-like buildings”—may play a role in activating a positive frame that reminds people of their latent favorable view toward courts.⁵⁹ Professors James Gibson, Gregory Caldeira, and Lester Spence refer to this effect as *positivity bias* and argue that judicial symbols function in this

55. Cf. Harry T. Edwards & Michael A. Livermore, *Pitfalls of Empirical Studies that Attempt to Understand the Factors Affecting Appellate Decisionmaking*, 58 DUKE L.J. 1895 (2009).

56. In recent research, Professors James Gibson and Michael Nelson distinguish between “ideological versus ‘politicized’” judicial decision making and “find that it is the latter that constitutes the gravest threat to the Court’s legitimacy.” James L. Gibson & Michael J. Nelson, *Reconsidering Positivity Theory: What Roles do Politicization, Ideological Disagreement, and Legal Realism Play in Shaping U.S. Supreme Court Legitimacy?*, 14 J. EMPIRICAL LEGAL STUD. 592, 595 (2017). See generally JEFFREY A. SEGAL & HAROLD J. SPAETH, *THE SUPREME COURT AND THE ATTITUDINAL MODEL REVISITED* (2002) (testing predictive power of ideological disposition for the behavior of Supreme Court Justices).

57. See ALLAN E. LIND & TOM R. TYLER, *THE SOCIAL PSYCHOLOGY OF PROCEDURAL JUSTICE* (1988); cf. Jason Sunshine & Tom R. Tyler, *The Role of Procedural Justice and Legitimacy in Shaping Public Support for Policing*, 37 LAW & SOC. REV. 513 (2003).

58. A focus on the interaction of symbols and court legitimacy has long interested scholars of the Court. See Gregory Casey, *The Supreme Court and Myth: An Empirical Investigation*, 8 LAW & SOC. REV. 385, 387 (1974) (citing MURRAY EDELMAN, *THE SYMBOLIC USES OF POLITICS* (1967)) (“Symbols can elevate the institution, setting it up as special, remote from ordinary skills and practices, difficult to check against daily experience, and unapproachable by the common man.”).

59. Gibson et al., *supra* note 8, at 838, 842.

way by signaling the difference between courts and other, less-favorably perceived official decision makers, such as Congress or agency bureaucrats.⁶⁰

Relatively little empirical work has been done to assess the importance of judicial symbols in affecting perceptions of the Court. One recent study by Professors James Gibson, Milton Lodge, and Benjamin Woodson examines how exposure to judicial symbols affects the level of support given to the Court and willingness to challenge the Court's rulings.⁶¹ In the study, one participant group was exposed to judicial symbols—a gavel, the Supreme Court courthouse, and the Justices in their robes—while the other was not, and a survey elicited information about their responses to various judicial rulings.⁶² In general, they found that exposure to these symbols enhanced levels of support, especially for those with relatively less prior awareness of the Court.⁶³

At the margins, it may be difficult to distinguish between functional and symbolic characteristics of courts. In their initial study, Gibson, Lodge, and Woodson focus on judicial symbols that have a fairly high degree of “pageantry.”⁶⁴ To the extent there can be a purely symbolic feature of the judicial role, robe-wearing and gavel-wielding seem like strong candidates. But it is possible that other characteristics of the Court that are more functional in nature could similarly trigger a positivity bias. For example, perhaps the ritual of oral argument serves a similar, positivity bias-triggering function, while at the same time (at least potentially) affecting substantive outcomes.⁶⁵

One of the most obvious distinguishing characteristics of courts is the form of the textual outputs through which their power is exercised and expressed. Judicial opinions are quite different from other textual manifestations of lawmaking, such as the statutes adopted by legislatures or the regulations promulgated by administrative agencies. Statutes and regulations take the form of more or less-clearly stated rules, whereas opinions consist of narrative explanations for a decision in a particular case. Indeed, one of the great challenges of the first year of law school is learning to decipher judicial opinions to extract the legal principles and rules that are coded within them.⁶⁶ The practice of issuing judicial opinions is among the most recognizable defining features of courts—especially appellate courts.

Opinions may play both a functional and a symbolic role. Functionally, the need to issue a justificatory opinion that is released to the public and subject to praise or criticism may influence decisions in individual cases. This effect may be

60. Gibson, Caldeira & Spence, *supra* note 38; JAMES L. GIBSON & GREGORY A. CALDEIRA, *CITIZENS, COURTS, AND CONFORMATIONS: POSITIVITY THEORY AND THE JUDGMENTS OF THE AMERICAN PEOPLE* 3 (2009).

61. Gibson et al., *supra* note 8, at 838.

62. *Id.* at 849–50.

63. *Id.* at 853–57.

64. *Id.* at 838.

65. *Cf.* David A. Karp, *Why Justice Thomas Should Speak at Oral Argument*, 61 *FLA. L. REV.* 611 (2009).

66. *Cf.* SCOTT TUROW, *ONE L* 31 (2d ed. 1988) (noting sensation for law student that reading judicial opinions is akin to “stirring concrete with [one’s] eyelashes”).

particularly powerful if there are strong conventions about appropriate and inappropriate justifications that are backed by some sanction (even if only reputational) or internalized by judges themselves.⁶⁷ Symbolically, judicial opinions may serve a role similar to robes or gavels by signaling courts' separation from the political branches.⁶⁸ Professor Michael Wells argues that the Justices take this symbolic effect seriously during opinion drafting by structuring their rhetoric to satisfy public expectations rather than offering a candid explanation for the Court's decisions.⁶⁹

Whether the practice of issuing recognizably judicial opinions actually reinforces the legitimacy of the Court, and if so what characteristics of opinions are responsible for that effect, are, of course, empirical questions. In a recent study, Professor Dion Farganis makes a start at addressing the question in an experimental setting by examining how exposure to different simulated Supreme Court opinions affect support for the institution among participants.⁷⁰ For the study, Farganis created three simulated judicial opinions on the question of state bans on same-sex marriage. The simulated opinions had the same outcome (upholding the bans), but differed in the supporting justification offered: one opinion relied on legal arguments; one relied on public-opinion polls; and one relied on religious texts.⁷¹ When later asked a series of questions meant to elicit their institutional support for the Court, participants who were exposed to the opinions containing the legal arguments reported higher levels of support than the participants exposed to the opinions that relied on opinion polls or religious texts.⁷² Although reliance on non-legal arguments did not entirely undermine support for the Court, there was a statistically significant and meaningful relationship between the content of the opinions and participants' later views about the Court as an institution.⁷³

The Farganis study takes a coarse-grained approach by generating simulated opinions that fall fairly far outside the norm and by testing the effect on support for the Court of one-off exposure to such unusual opinions. While the findings are evocative, they tell us relatively little about how smaller departures from the standard form of judicial opinions might affect the Court's legitimacy over time.

67. The question of whether, and how, law constrains judges is the subject of significant scholarly debate in law and political science. See generally Barry Friedman & Andrew Martin, *Looking for Law in All the Wrong Places: Some Suggestions for Modeling Legal Decision-making*, in WHAT'S LAW GOT TO DO WITH IT? WHAT JUDGES DO, WHY THEY DO IT, AND WHAT'S AT STAKE 143 (Charles Gardner Geyh ed., 2011) (reviewing relevant literature and offering a model in which law "channels" judicial decisions rather than dictates them).

68. Cf. PAUL W. KAHN, MAKING THE CASE: THE ART OF THE JUDICIAL OPINION 2–7 (2016) ("Every public authority performs in two dimensions: vote and voice. The weights of voice and vote stand in an inverse relationship when we compare legislators and judges.").

69. Wells, *supra* note 38 (arguing that the Court's need to conform to conventional expectations of judicial opinions undermines its ability to be candid).

70. Dion Farganis, *Do Reasons Matter? The Impact of Opinion Content on Supreme Court Legitimacy*, 65 POL. RES. Q. 206 (2012).

71. *Id.* at 208.

72. *Id.* at 210–12.

73. *Id.* at 210.

Indeed, in a follow-up study, Professor Chris Bonneau and colleagues used a similar experimental design to test whether the inclusion of non-legal arguments in addition to legal arguments resulted in a decline in institutional support, and found that extra-legal arguments resulted in no appreciable difference.⁷⁴ At the very least, the Bonneau study indicates that these effects may be sensitive to small changes, and that where they do exist, they may only be perceptible over longer time horizons.⁷⁵ Nevertheless, coupled with more general research on judicial symbolism, the existing state of relevant research raises the interesting prospect that the practice of exercising power by issuing documents that are recognizable as judicial opinions plays a legitimating role for the Court.

If issuing judicial opinions helps trigger support in part by demarcating the Court as a judicial institution rather than a political one, the ability to do so would be bound up with how well the Court's opinions conform to public expectations of the form more generally. By analogy, one might imagine a similar situation for the Justices' attire, with the symbolism most powerful in the case of an archetypical judicial robe and with diminishing effect as the garment becomes less robe-like and takes on characteristics of other garments such as the standard business suit associated with corporate managers or politicians. Similarly, it is possible to imagine an archetypical judicial opinion on the one hand, and on the other the kind of text that might be read from the floor of the Senate or published in the Federal Register. To the extent that the Court issues documents that read like judicial opinions, those documents effectively serve the symbolic function of separating the Court from other institutions. If opinions read more like floor speeches or regulatory preambles, they are less likely to have that effect.

The judicial opinion, then, can be thought of as a legitimating genre. By conforming to the norms and conventions of that genre, the Court marks itself as a non-political institution and in doing so triggers positive associations in the relevant public that reinforce feelings of support, even for that portion of the public that might disagree with a specific decision. Judicial opinions work alongside other functional and symbolic characteristics to generate a cumulative signal that separates the Court from the political branches in the public mind. Collectively, these defining features of the Court allow it to be deeply involved in contested political and social questions, while maintaining a level of support far above the other organs of government that are engaged in policymaking on similarly controversial issues.

One objection to including judicial opinions alongside other symbolic markers is that the general public is not exposed to opinions because few people actually read them. There are several possible responses to this objection. First, a substantial number of people may read at least portions of some judicial opinions, especially in very high-profile cases. Although it is certainly the case that the group of opinion readers does not make up a majority of Americans, the minority who do have some exposure to these opinions is not *de minimis*. Second, the people who

74. Chris W. Bonneau et al., *Evaluating the Effects of Multiple Opinion Rationales on Supreme Court Legitimacy*, 45 AM. POL. RES. 335, 349 (2017).

75. *See id.* at 357 (noting that "it is important to note that our study is unable to observe longitudinal effects").

read opinions tend to be a highly educated and relatively influential group that includes leaders in government, academics, practicing lawyers, and media-content producers (e.g., journalists, bloggers, television producers, and opinion-piece writers). Finally, Americans can interact with judicial opinions secondhand, and are exposed to characterizations of the Court's work or short snippets of quoted language.⁷⁶ For these reasons, the opinions of the Court cannot be thought of as esoteric texts with a highly limited audience and few effects on broader perceptions of the Court. Rather, they have sufficiently far reach such that they can, at least in theory, have an important legitimizing (or delegitimizing) effect.

B. Defining Distinctiveness from the Ground Up

The preceding section offers motivation for the work described in the remainder of this Article. Based on existing social-science research, there is reason to suspect that judicial symbols associated with the Court have some influence on its institutional legitimacy. The exact set of symbols that function in this way and how they contribute to or detract from public support remain open areas of inquiry. However, based on research into other judicial symbols, as well as the limited amount of research carried out on the effect of opinion characteristics on support for the Court, there is some evidence that opinions contribute to diffuse support by marking the Court as a judicial institution that is separate and different from the political branches. If this is the case, then Court opinions that conform to the norms and conventions of the broader judicial genre are likely to more effectively trigger those positive associations.

Determining how well the Court's opinions conform to genre conventions poses some thorny challenges. To understand whether, as a matter of ground truth, the texts issued by the U.S. Supreme Court *are* judicial opinions—either empirically, normatively, or aesthetically—would require some more general account of how correct genre categorizations can be made. These questions have interested (and vexed) scholars in a variety of fields for some time. The concept of genre and its applications has implications for cultural studies and humanities disciplines across a variety of domains, including literature, music, and art, and has accordingly been much discussed within those fields.⁷⁷ The more general study of categories and categorizations extends at least as far back as Aristotle, and has given rise to a sophisticated branch of analytic philosophy that puzzles over, for example, the ontological status of categories.⁷⁸

For our purposes, we can take a functional approach to genre that sidesteps many of these deeper questions. As described above, the genre of judicial opinions *functions* as a legitimating symbol by differentiating the issuing authority as a court, as opposed to a political body, such as a legislature or executive branch department.

76. See, e.g., Erica L. Green, *Justices' Ruling Could Shape Future of School Choice, Voucher Advocates Say*, N.Y. TIMES, June 28, 2017, at A16 (discussing *Trinity Lutheran Church of Columbia, Inc. v. Comer*, 137 S. Ct. 2012 (2017)).

77. See generally FISHELOV, *supra* note 10; ALTMAN, *supra* note 10.

78. See, e.g., RODERICK CHISHOLM, *A REALISTIC THEORY OF CATEGORIES* 3, 19 (1996); GEORGE LAKOFF, *WOMEN, FIRE, AND DANGEROUS THINGS: WHAT CATEGORIES REVEAL ABOUT THE MIND* 6 (1987).

The relevant features of the genre are those that trigger associations with the judicial (as opposed to political) exercise of power. The concept of genre stands in for a sociological effect in which exposure to judicial opinions increases diffuse support for the decision-making body.

At a high level of generality, this functional account allows theories about the sources of judicial legitimacy to be used to derive the genre's characteristics. This is the approach taken by Farganis. He begins with a theory about what separates judicial from non-judicial bodies in the public mind. From there, he uses that theory to generate the hypothesis that legal reasoning, as opposed to opinion polling or religious sentiment, is a relevant genre characteristic (although he does not use the term *genre*).⁷⁹ The study finds that use of legal reasoning does appear to enhance institutional support; meaning that it is a strong candidate for inclusion as a genre convention.⁸⁰ Building on this methodology, it might be possible to develop a more comprehensive account of the genre's defining characteristics.

Although there is value in this approach, there are some important limitations. Naturally, theories about the underlying sources of the Court's legitimacy are controversial, and drawing inferences from the theory to genre characteristics will be difficult at any but a very coarse level. In addition, some genre characteristics may have different symbolic meanings for different audiences; for example, legal experts may respond very differently to how arguments are presented than lay readers. The meaning of opinion characteristics may also change over time; for example, what read as appropriately neutral in the past may now strike readers as formalistic and baroque.

We take an alternative tack, using as our starting place a less controversial judgment about the members of the class of judicial opinions.⁸¹ Starting with a body of documents, we then examine whether the Court's opinions are distinguishable based on their semantic content, without imposing a theory about what is a relevant characteristic. The more general theory in support of this approach is that it is the *judicial* characteristic of opinions that generates their symbolic value. Conversely, if the Court's opinions are categorically distinct from other judicial opinions, then

79. Farganis's hypothesis is grounded in something like the myth-of-legality theory of the Court's legitimacy. Farganis, *supra* note 70, at 207.

80. *Id.* at 210–12.

81. Altman is critical of an approach to genre theory that “reduc[es] the notion of genre to a *corpus* of texts” and urges instead “treat[ing] genre as a complex *situation*, a concatenated series of events regularly repeated according to a recognizable pattern.” ALTMAN, *supra* note 10, at 84. For Altman's purposes, which are focused in part on the construction of (film) genre by different social actors, his criticism may have some force. We are less interested in whether anyone construes the opinions of the Supreme Court as occupying a distinct genre. Instead, we are using the concept of genre to capture the idea of distinctiveness of categories of texts within a corpus. The thinner concept of genre that we apply here is sufficient for our purposes and is not meant to contradict or undermine a thicker conception in which commentary outside of the corpus itself is part of a social process of constructing genre. Indeed, this thicker concept of genre may apply to the Supreme Court's opinions, if journalists or academics systematically read them under a different set of conventions and expectations than other judicial opinions.

their symbolic value is attenuated and they will have, at least potentially, less of a legitimacy-enhancing effect.

Our approach can be illustrated through a simple thought experiment. Imagine a hypothetical law student, walking the corridors of a law library. This law student notices on the floor a few pages torn out of the previous year's Federal Reporter. The document lacks information identifying the authoring court. The student tries to guess whether the opinion was written by the Supreme Court or an appellate court. How likely would the student have been to correctly guess the authoring court?

In engaging in this classification task, the (rational) law student should take account of the background probability of randomly drawing a Supreme Court case (perhaps the pages are found between bookshelves filled roughly equally with Supreme Court and appellate court reporters). The student should then update that probability estimate based on evidence drawn from the opinion. If the case dealt primarily with the First Amendment, the student might revise the probability upward based on the higher relative frequency of First Amendment cases in the Supreme Court. If the case dealt with state-law issues, the student might revise the probability downward to reflect the fact that Supreme Court cases that primarily interpret state law are fairly rare. If the opinion reached for a quotable phrase, or overtly invoked moral or political values, perhaps the student would lean toward attributing the case to the Supreme Court.

Whatever the reasoning capacity of our student, his or her ability to guess correctly will be related to the distinctiveness of Supreme Court opinions. At one extreme, if Supreme Court opinions were written in Latin while appellate court opinions were written in Greek, the classification task would be trivial. At the other, if the Supreme Court's docket were selected at random from all appellate court cases, and the Justices employed similar reasoning and writing styles to appellate court judges, it would be extremely difficult to improve on the prior probability estimate based purely on background frequency. If it is relatively easy to distinguish Supreme Court opinions, then on our measure, they depart from the more general genre of judicial opinions, which is defined according to the baseline corpus of appellate opinions.

Taken statically, it would be very difficult to interpret findings of distinctiveness or ease of classification, other than to note departure from pure chance. But a dynamic understanding allows for comparison between time periods. If our hypothetical student is better able to classify cases from 2004 than cases from 1954, it is fair to infer that the Supreme Court has grown more distinctive over time. This conclusion does not necessarily imply that the appellate courts have remained steady while the Supreme Court has veered off in uncharted territory. But it does mean that the Supreme Court has become more distinctive relative to the appellate courts.

Carrying out this experiment in real life by presenting students with randomly generated snippets of text and asking them to classify the documents would pose substantial technical and logistical challenges. First, it would require recruiting a team of students to read a very large number of cases, which would substantially increase the cost of the project. Students might also perform this task

differently from each other, focusing on various features of cases and coming to the task with more or less exposure to the law and insight into differences between the institutions. Intuition and gut instinct would likely play a role in ways that would make it difficult to articulate reasons why some cases appeared more Supreme Court like than others.⁸²

To avoid these problems, we employ a principled application of statistical, computational text analysis called *topic modeling*. Topic modeling is used to estimate the relative weight in an opinion text that is paid to a probability distribution of reliably co-occurring and typically semantically related words (referred to as *topics*).⁸³ These topics are determined from the corpus in the way described below.⁸⁴ This analysis is applied to three sets of documents: all opinions published in the Federal Reporters during the study period, the subset of those opinions associated with cases that were selected by the Supreme Court for review, and the subset of those opinions that were published in the U.S. Reports (i.e., Supreme Court opinions). The data generated by the topic model substantially reduces the number of dimensions needed to characterize the semantic content of cases, allowing us to engage in useful statistical analysis.

There are several advantages to the topic-model approach used here compared to conventional methods. Topic modeling allows us to analyze all of the text in a very large number of published federal court cases, a task that would be infeasible for human readers.⁸⁵ There are also two advantages of topic modeling compared to existing data sources, such as the Supreme Court Database (“SCDB”).⁸⁶ First, compared to human-coded categories, we deploy a relatively large number of topics, allowing a fine-grained estimate of the subject matter found in the text of opinions. Perhaps more important, topic models can decompose documents into weighted distributions, rather than lists, of the topics that they contain. This enables a more detailed, quantitative case comparison that is more closely linked to the actual text, because the relative weight that cases afford to a subject can be systematically and objectively estimated. Finally, compared to simple word-frequency measures, topic models are better able to capture semantically significant features of words, and they reduce the risk of accentuating trivial differences in the classification task.

This approach of defining the genre from the ground up (from the observed documents), rather than from the top down (based on a theory of judicial legitimacy), has some useful advantages. Perhaps most important is that it allows for a comprehensive approach that captures a broad array of semantic characteristics. The

82. Cf. Carl P. Simon & Eric S. Rabkin, *Culture, Science Fiction, and Complex Adaptive Systems: The Work of the Genre Evolution Project*, in *BIOCOMPLEXITY AT THE CUTTING EDGE OF PHYSICS, SYSTEMS BIOLOGY AND THE HUMANITIES* 279 (Gastone Castellani et al. eds., 2008).

83. Henceforth, we reserve the word *topic* for referring to its technical meaning within the statistical learning methodology of topic modeling.

84. See *infra* Section II.A.

85. For purposes of convenience, we use a large random sample of appellate court cases to represent the entire population. Even a large sample would be difficult for a team of human researchers to tackle.

86. For information on the Supreme Court Database, see *infra* Section II.B.

alternative is highly targeted, but creates the risk of missing important characteristics that fall outside the theory. By defining the genre with respect to a class of documents, rather than a set of characteristics that are generated *ex ante*, our approach remains open to unanticipated features of the genre.

A downside of our approach is that it creates the possibility of identifying distinctions without differences. So long as Supreme Court opinions fall within the genre of judicial opinions in relevant ways, small variations that allow an observer to correctly categorize those opinions may not be important for the Court's legitimacy. The Court's opinions could, in principle, carry out their symbolic function even if they are distinguishable. Because our approach allows a large degree of freedom to identify semantic differences, it is possible that some sham distinctions will emerge that do not connect, in important ways, with genre characteristics with genuine functional value.

We address this potential problem in two ways. First, by using a topic-model approach, we reduce the dimensionality of the corpus—effectively zooming out from individual words and focusing on broader subject-matter categories. This reduces the risk that unimportant individual words or semantic oddities associated with a single writer would skew the analysis. Second, we engage in a close reading of a selected group of opinions to determine if the quantitative differences that we identify seem to capture relevant characteristics that plausibly bear on the legitimacy-enhancing function of the judicial genre. We believe that use of the topic model, together with the qualitative analysis, substantially reduces the risk of focusing on sham distinctions. It is also worth noting that our analysis can be used as a rich source of hypotheses concerning genre characteristics that could be tested in an experimental setting akin to the Farganis study.

C. Sources of Difference

There are two basic mechanisms through which the opinions of the Court may come to be systematically different from those of the appellate courts that it supervises: the certiorari process and the process of opinion drafting. This Section discusses prior scholarship on these two sources of difference and examines their potential to contribute to the distinctiveness of the Court's opinions.

The hierarchical structure of the judiciary generates a vast winnowing of cases and issues before they reach the pages of the U.S. Reports. Each year, roughly one million cases are filed in federal courts.⁸⁷ In reporting year 2012, there were 35,302 federal appeals terminated on the merits, disposing of a number of cases roughly equivalent to 10% of the non-bankruptcy filings in the federal court.⁸⁸ The vast majority of these appellate dispositions were not accompanied by a published

87. Federal Judicial Center, *The Federal Courts and What They Do* 6, [http://www.fjc.gov/public/pdf.nsf/lookup/FCtsWhat.pdf/\\$file/FCtsWhat.pdf](http://www.fjc.gov/public/pdf.nsf/lookup/FCtsWhat.pdf/$file/FCtsWhat.pdf) (last visited Jan. 15, 2017). A substantial portion (80%) of federal cases are bankruptcy filings. *Id.*

88. ADMIN. OFF. OF THE U.S. CTS., ANN. REP. OF THE DIRECTOR: JUD. BUS. OF THE U.S. CTS., tbl. B-5 (2013). This figure does not include the U.S. Court of Appeals for the Federal Circuit.

opinion.⁸⁹ From this pool, several thousands of petitions for certiorari were submitted, with the Court granting just over 100.⁹⁰ The Court's control over its docket allows it substantial ability to influence its own agenda. In one study, Professor Richard Pacelle examined how the Court apportioned its "agenda space" among 14 policy areas and identified rough eras between 1933 and 1988 during which, in turn, economics, federalism, regulation, and civil liberties dominated the Court's agenda.⁹¹ The Court's choice of cases also appears to influence the composition of cases in the lower courts by affecting litigant behavior about what cases to bring and appeal.⁹²

Given the consequences of the Court's certiorari jurisdiction, it is not surprising that it has long been a subject of study by social scientists and academic lawyers.⁹³ Each step of the process has been investigated to some degree. At the petitioning stage, scholars have examined the effects of litigants' choices of whether to petition on the pool of cases before the Court⁹⁴ and have investigated the Court's ability to signal to the pool of potential petitioners the cases that it would like to hear.⁹⁵ Professors Ryan Black and Christina Boyd examine the influence of judicial clerks when making initial recommendations on cases to be granted certiorari.⁹⁶ A number of theories have been advanced concerning how the Justices themselves make decisions about which petitions to grant, including a theory that posits that they seek to correct errors in the lower courts;⁹⁷ that they act strategically, voting to grant certiorari when they believe that they would be successful at the merits stage;⁹⁸

89. *Id.* at tbl. S-3.

90. *Id.* at tbl. B-2.

91. RICHARD L. PACELLE, JR., THE TRANSFORMATION OF THE SUPREME COURT'S AGENDA: FROM THE NEW DEAL TO THE REAGAN ADMINISTRATION 15, 20 (1991).

92. Douglas Rice, *The Impact of Supreme Court Activity on the Judicial Agenda*, 48 LAW & SOC. REV. 63, 67 (2014); Mark Hurwitz, *Institutional Arrangements and the Dynamics of Agenda Formation in the U.S. Supreme Court and Courts of Appeals*, 28 LAW & POL'Y 321 (2006).

93. See, e.g., Fowler V. Harper & Alan S. Rosenthal, *What the Supreme Court Did Not Do in the 1949 Term—An Appraisal of Certiorari*, 99 U. PA. L. REV. 293 (1950); Glendon A. Schubert, *The Study of Judicial Decision-Making as an Aspect of Political Behavior*, 52 AM. POL. SCI. REV. 1007 (1958).

94. See, e.g., Maxwell Mak, Andrew H. Sidman & Udi Sommer, *Is Certiorari Contingent on Litigant Behavior? Petitioners' Role in Strategic Auditing*, 10 J. EMPIRICAL LEGAL STUD. 54 (2013); Jeff Yates, Damon M. Cann & Brent D. Boyea, *Judicial Ideology and the Selection of Disputes for U.S. Supreme Court Adjudication*, 10 J. EMPIRICAL LEGAL STUD. 847 (2013).

95. Vanessa A. Baird, *The Effect of Politically Salient Decisions on the U.S. Supreme Court's Agenda*, 66 J. POL. 755, 755 (2004).

96. Ryan C. Black & Christina L. Boyd, *The Role of Law Clerks in the U.S. Supreme Court's Agenda-Setting Process*, 40 AM. POL. RES. 147 (2012).

97. John F. Krol & Saul Brenner, *Strategies in Certiorari Voting on the United States Supreme Court: A Reevaluation*, 43 W. POL. Q. 335, 335 (1990) (citing earlier work on error-correction hypothesis).

98. Jan Palmer, *An Econometric Analysis of the U.S. Supreme Court's Certiorari Decisions*, 39 PUB. CHOICE 387, 392–93 (1982); Gregory A. Caldeira, John R. Wright &

and that the Justices focus on “cues” from external actors that indicate public policy salience or lower court errors.⁹⁹ The influence of actors outside of the Court has also been subject to scrutiny. Certiorari is more likely to be granted when the United States is a petitioner.¹⁰⁰ Even when the government is not a litigant, the Solicitor General is frequently consulted by the Court over certiorari decisions and appears to have substantial influence.¹⁰¹ The presence of amicus supporters and opponents increases the likelihood of a grant of certiorari.¹⁰² Litigant characteristics also appear to have an effect, with the participation of the “obscenity bar” increasing the rate of certiorari grants in relevant cases,¹⁰³ while self-represented litigants are unlikely to be heard.¹⁰⁴ Congress also appears to influence the decisions that are taken up by the Court.¹⁰⁵

Based on this prior work, Professors Jeff Yates, Damon Cann, and Brent Boyea conclude that “[a] wealth of judicial politics literature suggests that [J]ustices have an interest in taking on cases that are salient, resolve important legal conflicts, and, in fact, do map well onto [J]ustices’ distinct ideological preferences.”¹⁰⁶ Because the Court’s docket differs in systematic ways from the general pool of appellate cases, we should expect that the opinions the Court issues will be distinguishable based on the unusual characteristics of the underlying cases. If the Court uses its certiorari jurisdiction to focus its attention on certain legal questions (e.g., constitutional claims) while avoiding others (e.g., family law) then its opinions will naturally reflect that emphasis in its docket. Purely through the operation of the certiorari process, the body of Supreme Court opinions will reflect the issues that most capture the Court’s attention.

There are certainly normative arguments to be made that the Court should use its certiorari jurisdiction not only to correct errors in the lower courts or remedy

Christopher J.W. Zorn, *Sophisticated Voting and Gate-Keeping in the Supreme Court*, 15 J. L. ECON. & ORG. 549, 550 (1999).

99. Joseph Tanenhaus et al., *The Supreme Court’s Certiorari Jurisdiction: Cue Theory*, in JUDICIAL DECISION-MAKING 118–19 (Glendon Schubert ed., 1963).

100. Caldeira, Wright & Zorn, *supra* note 98, at 563.

101. Ryan C. Black & Ryan J. Owens, *Consider the Source (and the Message): Supreme Court Justices and Strategic Audits of Lower Court Decisions*, 65 POL. RES. Q. 385, 387 (2012); see David C. Thompson & Melanie F. Wachtell, *An Empirical Analysis of Supreme Court Certiorari Petition Procedures: The Call for Response and the Call for the Views of the Solicitor General*, 16 GEO. MASON L. REV. 237 (2009).

102. Gregory A. Caldeira & John R. Wright, *Organized Interests and Agenda Setting in the U.S. Supreme Court*, 82 AM. POL. SCI. REV. 1109, 1109 (1988).

103. Kevin T. McGuire & Gregory A. Caldeira, *Lawyers, Organized Interests, and the Law of Obscenity: Agenda Setting in the Supreme Court*, 87 AM. POL. SCI. REV. 715, 717 (1993).

104. See DORIS MARIE PROVINE, CASE SELECTION IN THE UNITED STATES SUPREME COURT 45 (1980).

105. Kenneth W. Moffett et al., *Strategic Behavior and Variation in the Supreme Court’s Caseload Over Time*, 37 JUST. SYS. J. 20, 26–27 (2016); Lee Epstein, Jeffrey A. Segal & Jennifer Nicoll Victor, *Dynamic Agenda-Setting on the United States Supreme Court: An Empirical Assessment*, 39 HARV. J. ON LEGIS. 395 (2002).

106. Yates, Cann & Boyea, *supra* note 94, at 852.

disunity between the circuits, both of which might arise randomly, but also to allocate its limited time to relatively more pressing legal questions.¹⁰⁷ If the Court's opinions are distinctive merely because they tend to focus on more important matters, perhaps that distinctiveness is not legitimacy-threatening, in which case the relevant genre conventions would bear primarily on how cases are discussed rather than the presence or absence of specific legal issues.

Alternatively, it is possible that certiorari selects cases that are more controversial or less amenable to being discussed within the confines of the genre. If this is the case, then the operation of the case selection process might ultimately affect the judicial character of the Court's opinions. Because of these two countervailing possibilities, it is difficult to say a priori the degree to which differences that arise from the certiorari process make the Court's opinions less recognizably judicial.

The second mechanism that could lead to differences between the Court's opinions and those of the appellate courts is the opinion-drafting process. Once certiorari has been granted, a case typically proceeds through merits briefing and oral argument, followed by drafting and editing.¹⁰⁸ During the drafting phase, versions of the majority opinion, and any concurrences or dissents, are circulated within the Court, spurring additional deliberations, occasional vote-shifting, and redrafting and editing.¹⁰⁹ All of these internal operations are governed by both formal rules and entrenched conventions.¹¹⁰ Most distinctly from the lower appellate courts, the Court always sits as a whole rather than in panels. This means that opinions serve as part of a running conversation between the Justices that has the potential to create a unique culture, especially during a period when the Court's membership is relatively stable.

There are many ways that the drafting process could lead to the Court producing opinions that are distinct from the appellate courts, even holding the underlying cases constant. The Court has considerable leeway to decide which of

107. For work discussing normative theories of how the Court should exercise its discretion, see Tom S. Clark & Jonathan P. Kstellec, *The Supreme Court and Percolation in the Lower Courts: An Optimal Stopping Model*, 75 J. POL. 150 (2013); Margaret Cordray & Richard Cordray, *The Philosophy of Certiorari: Jurisprudential Considerations in Supreme Court Case Selection*, 82 WASH. U. L.Q. 389 (2004); Samuel Estreicher & John E. Sexton, *A Managerial Theory of the Supreme Court's Responsibilities: An Empirical Study*, 59 N.Y.U. L. REV. 681 (1984); William J. Brennan, *The National Court of Appeals: Another Dissent*, 40 U. CHI. L. REV. 473 (1973).

108. Kevin T. McGuire & Barbara Palmer, *Issue Fluidity on the U.S. Supreme Court*, 89 AM. POL. SCI. REV. 691 (1995).

109. See, e.g., WILLIAM H. REHNQUIST, *THE SUPREME COURT* (2d ed. 2002); see also JEFFREY TOOBIN, *THE NINE: INSIDE THE SECRET WORLD OF THE SUPREME COURT* 264–66 (2007); BOB WOODWARD & SCOTT ARMSTRONG, *THE BRETHERN: INSIDE THE SUPREME COURT* (2005); EDWARD LAZARUS, *CLOSED CHAMBERS: THE RISE, FALL, AND FUTURE OF THE MODERN SUPREME COURT* (1999).

110. See REHNQUIST, *supra* note 109.

the legal questions presented in those cases to explore or emphasize.¹¹¹ When the Court grants certiorari, it frequently limits its review to specific questions. The process of bargaining may give special influence to swing Justices, who exercise considerable pressure on case outcomes and who may have quirky preferences.¹¹² The authors of majority opinions and the majority-coalition members may also have some ability to shape opinions for the Court, inserting their own perspectives or pet language.¹¹³ And, of course, Justices have wide leeway when authoring dissenting or concurring opinions to say what they want.¹¹⁴

During this process, the Justices face different incentives than lower-court judges because their decisions cannot be appealed and will serve as the final word on the legal questions that they decide. Justices may, accordingly, be freer in their language or view themselves as addressing a broader public or posterity. For example, when discussing a case concerning prayer in school, the Court may focus on first principles concerning the underlying values protected by the Establishment Clause while appellate courts may tend to focus instead on prior precedent or procedural elements of the case. The Justices may also make different choices in the

111. Ulmer refers to the phenomenon of “issue fluidity” in which the Court either “discover[s] and decide[s] an issue not raised by the petitioner . . . [or] suppresses and does not decide an issue posed by the petition[er].” Sidney S. Ulmer, *Issue Fluidity in the United States Supreme Court*, in *SUPREME COURT ACTIVISM AND RESTRAINT* 322 (Charles M. Lamb & Stephen C. Halpern eds., 1982). Although suppression is common and explicit, the degree of issue expansion on the Court is the subject of controversy. Compare McGuire & Palmer, *supra* note 108, with Lee Epstein, Jeffrey A. Segal & Timothy Johnson, *The Claim of Issue Creation on the U.S. Supreme Court*, 90 *AM. POL. SCI. REV.* 845 (1996), and Kevin T. McGuire & Barbara Palmer, *Issues, Agendas, and Decision-making on the Supreme Court*, 90 *AM. POL. SCI. REV.* 853 (1996).

112. See LEE EPSTEIN & JACK KNIGHT, *THE CHOICES JUSTICES MAKE* 76 (1998). Researchers have found that Justices do indeed switch their votes for strategic purposes. Forrest Maltzman & Paul J. Wahlbeck, *Strategic Policy Considerations and Voting Fluidity on the Burger Court*, 90 *AM. POL. SCI. REV.* 581, 591 (1996).

113. See Cliff Carrubba et al., *Who Controls the Content of Supreme Court Opinions?*, 56 *AM. J. POL. SCI.* 400, 409 (2012) (majority coalition affects content); Chris W. Bonneau et al., *Agenda Control, the Median Justice, and the Majority Opinion on the U.S. Supreme Court*, 51 *AM. J. POL. SCI.* 890, 903 (2007); Jeffrey R. Lax & Charles M. Cameron, *Bargaining and Opinion Assignment on the U.S. Supreme Court*, 23 *J.L. ECON. & ORG.* 276, 296–97 (2007). Different Justices may also have influence over different portions of the majority opinions. See James Khun, Matthew E.K. Hall & Kristen Macher, *Holding versus Dicta: Divided Control of Opinion Content on the U.S. Supreme Court*, 70 *POL. RES. Q.* 257, 265 (2017) (finding that median Justices tend to influence holdings, while majority-opinion authors influence dicta).

114. Several scholars have examined the causes of the growing levels of disagreement in Supreme Court decisions in the mid-twentieth century, one of the most striking differences between Supreme Court and appellate court decisions. Marcus Hendershot et al., *Dissensual Decision Making: Revisiting the Demise of Consensual Norms within the U.S. Supreme Court*, 66 *POL. RES. Q.* 467, 478–79 (2013); Gregory A. Caldeira & Christopher J.W. Zorn, *Of Time and Consensual Norms in the Supreme Court*, 42 *AM. J. POL. SCI.* 874, 900 (1998); Thomas G. Walker, Lee Epstein & William J. Dixon, *On the Mysterious Demise of Consensual Norms in the United States Supreme Court*, 50 *J. POL.* 361, 384–85 (1988).

language that they use, perhaps deploying certain rhetorical moves, such as personal anecdotes, metaphor, humor, or colloquialisms, that are less common in the lower courts. The nomination and confirmation process may also select for certain characteristics, such as partisan identification or flair with a pithy turn of phrase, that leave residual traces.

Given the unique processes employed by the Court, the distinctive nature of the Court's role, the audience that it addresses, and the peculiar nature of the Justices' bargaining, drafting, and editing processes, it would not be surprising if the types of reasoning or the language used in the Court's opinions differ from those that are used in appellate court opinions, even when the set of legal issues is the same. Some of these differences may enhance the prestige of the Court—for example, the Justices might systematically employ clearer or more persuasive writing than judges on the lower courts. However, under the theory of a legitimacy-enhancing judicial genre, the presence of identifiable differences would generally be problematic. Although individual Justices may even be correct that, in specific instances, their goals may be best promoted by deploying distinctive modes of analysis or non-judicial writing styles, over time consistent violations of genre conventions run a risk of negative legitimacy consequences.

II. TOPIC MODELING OPINIONS

Although there has been considerable quantitative analysis of the Court's behavior, focusing especially on how Justices "vote" in individual cases,¹¹⁵ quantitative analysis of the Court's opinions at any but the highest level of generality has, to date, been limited. Such analyses have historically depended on human researchers who code content, an expensive methodology to deploy.¹¹⁶ In this Part, we introduce topic modeling, a machine substitute for human coding that allows for a tractable means of translating the semantic content of texts into a form that is amenable to quantitative analysis. We then discuss the data that form the basis for our analysis and report the topics generated by our model and the results of several tests of how well the topic model performs in capturing relevant semantic features of the corpus under review.

A. *Topic Models*

There has been increasing interest in applying computational text-analysis tools to law.¹¹⁷ Recent advances in computational text analysis extend the use of simple word frequencies as representations of textual documents to sophisticated statistical models capable of drawing out relevant features of large collections of documents. These new approaches include a class of techniques referred to as *topic*

115. See, e.g., Timothy R. Johnson, Paul J. Wahlbeck & James F. Spriggs, *The Influence of Oral Arguments on the U.S. Supreme Court*, 100 AM. POL. SCI. REV. 99, 111–12 (2006).

116. See generally Mark A. Hall & Ronald F. Wright, *Systematic Content Analysis of Judicial Opinions*, 96 CAL. L. REV. 63 (2008) (discussing use of human researchers to generate quantitative variables).

117. See Carlson et al., *supra* note 12, at 1465–73 (discussing use of computational tools for legal scholarship).

models, which have proven especially useful to scholars in both the social sciences and humanities. The ability of topic models to quantitatively capture semantic features of very large corpora of legal documents has substantial potential to aid the work of empirical legal scholars in many domains. This Section introduces topic modeling and provides an overview of the outputs generated by this tool.

Raw word frequencies can provide a foundation for an efficient, if crude, measurement of similarity between two texts that can be used to suggest related documents and rank search results. A raw word frequency is simply a list of all words in a document along with a count of their occurrence.¹¹⁸ This information can be subjected to additional manipulation to draw out relevant characteristics within texts. For example, a simple weighted variation of word frequency calculates the *term frequency-inverse document frequency* (“tf-idf”). The tf-idf measure places relatively less weight on words that are very common within the corpus.¹¹⁹ This measure, and variants of it, was among the earliest effective sources of textual similarity metrics used in search engines to estimate the relevance of a document for a user query and formed the foundation of early document retrieval applications.¹²⁰

However, simple word frequencies, even if intelligently weighted, can still suffer from several drawbacks. For example, they are notoriously poor at disambiguating multiple meanings (e.g., “flag” as an emblem and “flag” as meaning “to distinguish”)¹²¹ so that two texts that are similar in terms of word frequencies could still be quite different in terms of content. They also do not provide useful information about a text’s membership in broader, more familiar categories (e.g., “cases concerning the First Amendment”).

A topic model offers one means of recovering accurate approximations of these higher-level categorizations without the expense (in time and resources) of hand coding. Given a textual corpus, a topic model produces topics, which in the technical topic-modeling sense are probability distributions over a vocabulary, where each word in the vocabulary is assigned a non-negative weight such that all weights sum to one. Each document is in turn summarized as a probability distribution over the topics. The highest-weighted words within a topic provide a sense of the subject matter that the distribution represents. For example, in Topic 14 generated by our model, the words “estate, trust, death, property, descendent, wife,

118. These frequencies can be normalized to one by dividing by the total number of words in the document.

119. The tf-idf weight is made up of the normalized term frequency (“tf”), which is the number of times a word appears in a document divided by the total number of words in that document multiplied by the inverse document frequency (“idf”), which is the logarithm of the total number of the documents in the corpus divided by the number of documents where the specific term appears.

120. Scott Deerwester et al., *Improving Information Retrieval with Latent Semantic Indexing*, in PROCEEDINGS OF THE 51ST ANNUAL MEETING OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE 25, 36–40 (1988).

121. Similarly, word frequencies alone would fail to distinguish between the use of the word “minor” in “minor parties” and “minor offense” from the use of “minor” in, for example, *Bethel School District v. Fraser*, 478 U.S. 675, 677 (1986), a decision concerning free speech in public schools.

interest” are weighted highly, which led us to hand label that topic as *trusts and estates*.¹²² Thus, the representation of a given document as a distribution over topics summarizes the document as weighted mixtures of intuitively understood themes.

These distributions—both of the topics and the words they contain—are produced as the best fit to an underlying generative probabilistic model for the observed simple-word frequencies. The canonical topic model is a latent Dirichlet allocation (“LDA”) mixed-membership model.¹²³ The LDA model posits some number of topics, which are distributions over the vocabulary, that account for all words observed in a corpus according to the following generative story: for each document in the corpus, a set of topic proportions (or “shares”) is drawn from a global probability distribution; then, each word in the document is drawn from a topic distribution in which the topic distribution in question is selected according to the document-specific set of proportions.¹²⁴

Topic models are often fit using an iterative algorithm known as variational approximation or by using a Markov Chain Monte Carlo (“MCMC”) approach.¹²⁵ In the case of the topic model, the parameters of interest are typically restricted to the topic-word distributions describing the association between topics and words, and the document-topic distributions that describe for each document the probability of finding words associated with each topic. The particular variant of MCMC used in our analysis is called a Gibbs sampler, which makes drawing samples from a high-dimensional posterior distribution tractable by considering a sequence of conditional distributions of smaller subsets of the parameters.¹²⁶

122. Topics are generally hand labeled.

123. The Dirichlet distribution is a family of continuous multivariate probability distributions commonly used as prior distributions in Bayesian statistics included in mixture models. Mixed-membership models are a statistical framework in which groups are modeled as mixture over a shared overarching structure. Edoardo M. Airoidi et al., *Mixed Membership Stochastic Blockmodels*, 9 J. MACHINE LEARNING RES. 1981 (2008); Elena A. Erosheva, Stephen E. Fienberg & John Lafferty, *Mixed-Membership Models of Scientific Publications*, 101 PROC. NAT. ACAD. SCI. 5220, 5220 (2004). For our purposes, *documents* are the groups (of words), with the *topics* as the shared overarching structure.

124. Each topic distribution is typically given a weakly informative uniform Dirichlet prior. Blei provides a general overview and formal description of an LDA topic model. David M. Blei, *Probabilistic Topic Models*, 55 COMM. ASS’N COMPUTING MACHINERY 77, 78–79 (2012).

125. A useful reference point for understanding the inferential machinery of topic models is the ideal point model of Supreme Court justices’ votes developed by Martin and Quinn. Cf. Andrew D. Martin & Kevin M. Quinn, *Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the U.S. Supreme Court, 1953–1999*, 10 POL. ANALYSIS 134, 152 (2002). Both topic models and ideal point models are Bayesian models, which use the strategy of latent variable augmentation to facilitate inference and make use of MCMC to estimate model parameters. The ideal point model confronts the challenge of translating from a continuous space (the embedding of the ideal points) to a discrete space of observed votes. In the case of the topic model, the underlying continuous space is the probability structure, and the discrete space is the observed space of word frequencies.

126. George Casella & Edward I. George, *Explaining the Gibbs Sampler*, 46 AM. STATISTICIAN 167, 173 (1992).

A primary advantage of the LDA topic model over earlier word-frequency approaches is that it accounts for probability of word co-occurrence, which helps to uncover characteristics of context. The LDA topic model also allows for the radical reduction in the dimensionality needed to describe documents within a corpus. A given group of documents may contain tens of thousands of unique words. To represent documents as frequency distributions over those words requires vectors with an equal number of components. The number of topics generated by a topic model is much smaller, typically fewer than 500. In the present case, we fit the corpus with 100 topics. The parameters in the topic model that are of interest are the document-specific distributions over topics (i.e., the topic proportions) and the topic distributions themselves. The document-specific distributions have an interpretation as the proportion of words in a document that were generated by each topic. The topic distributions are distributions over words and are typically described in terms of words with the highest likelihood of having been sampled from the relevant distribution. The LDA model tends to group together semantically connected words (such as “child” and relevant uses of “minor”) and has been shown in practice to succeed where other word frequency-based representations fall short.¹²⁷

While numerous incremental improvements to LDA topic modeling have emerged in the intervening years, the essence of the original model remains. The LDA topic model persists as a general industry standard for text analysis and serves, with minor variations, as a building block in more elaborate models of text data. More than a decade after the model’s introduction, researchers using topic models and closely related models may be found in almost every field where machine-readable text data is abundant. Topic models are now a familiar part of the methodological landscape in the human and social sciences, from political science to German studies.¹²⁸

B. Data and Topics

In the past, computational analysis of legal texts has been hampered by difficulties accessing the relevant data.¹²⁹ Although judicial opinions are not copyrightable, the commercial databases that provide digital access to these opinions are protected by terms-of-use agreements.¹³⁰ Public.Resource.Org, a

127. Blei, Ng & Jordan, *supra* note 17, at 1014–15. Other early methods for dimensionality reduction directed at recovering semantic representations of texts include (probabilistic) latent semantic indexing, principal components analysis, and mixture of unigrams. Kamal Nigam et al., *Text Classification from Labeled and Unlabeled Documents Using EM*, 39 MACHINE LEARNING 103, 103–05 (2000); Thomas Hofmann, *Probabilistic Latent Semantic Indexing*, in PROCEEDINGS OF THE TWENTY-SECOND ANNUAL INTERNATIONAL ACM SIGIR CONFERENCE ON RESEARCH AND DEVELOPMENT IN INFORMATION RETRIEVAL 50, 57 (1999).

128. Riddell, *supra* note 17; Quinn et al., *supra* note 17, at 224–26.

129. Olufunmilayo Arewa, *Open Access in a Closed Universe: Lexis, Westlaw, Law Schools and the Legal Information Market*, 10 LEWIS & CLARK L. REV. 797, 798–99 (2012).

130. See *Matthew Bender & Co. v. West Publishing Co.*, 158 F.3d 693, 698 (2d Cir. 1998) (describing judicial opinions in West case reporters as “[w]orks of the federal government . . . not subject to copyright protection”); Maureen A. O’Rourke, *Drawing the*

private, not-for-profit corporation, has created a publically accessible digital version of the Supreme Court and federal appellate court corpus based on the non-copyrightable information within the Westlaw database.¹³¹ CourtListener, an effort within the Free Law Project, has augmented the information released by Public.Resource.Org and created a user-friendly interface that is accessible to the public. We relied on CourtListener as the source for all the texts of the Supreme Court and appellate court decisions.

The set of Supreme Court documents used in this study includes the opinions associated with all formally decided full-opinion cases as defined in the SCDB.¹³² It covers the period between 1951 and 2007, inclusive. There are 7,528 documents in this set. The median document length for this set is 3,339 words. The set of appellate court documents used are all published opinions issued between 1951 and 2007, for a total of 289,550 documents. To reduce the computational burden of fitting the topic model, we randomly selected 25,000 documents from within the appellate court set. The number of randomly sampled opinions is large enough that we believe any parameter estimates discussed would be essentially unchanged were the number expanded.¹³³ In addition to the 25,000 randomly selected appellate court opinion documents, 4,180 appellate court documents that are associated with cases selected for review by the Supreme Court decisions are also included. In total there are 29,180 appellate court opinion documents.¹³⁴ The median decision length is 2,781 words.

To identify the set of cases that were selected for review by the Court, we gathered information from LexisNexis, which provides prior history and disposition

Boundary Between Copyright and Contract: Copyright Preemption of Software License Terms, 45 DUKE L.J. 479, 555 (1995) (“[D]atabase vendors routinely condition access to data on the licensee’s agreement to treat the data as if it were copyrighted when, in fact, it may not be.”); see also LexisNexis, Terms & Conditions of Use for the LexisNexis Services § 1(c) (2016), <https://www.lexisnexis.com/terms/general.aspx>.

131. This resource data has been used in prior studies of text usage in the federal courts and Supreme Court. See generally Daniel M. Katz et al., *Legal N-Grams? A Simple Approach to Track the ‘Evolution’ of Legal Language*, in PROCEEDINGS OF JURIX 2011: THE 24TH INTERNATIONAL CONFERENCE ON LEGAL KNOWLEDGE AND INFORMATION SYSTEMS (2011).

132. Specifically, using the Supreme Court Database coding, we include decision types 1 (opinion of the Court, orally argued), 2 (per curium, no oral argument), 5 (equally divided vote), 6 (per curium, orally argued), and 7 (judgment of the Court, orally argued). We exclude decrees and seriatim opinions.

133. Majority, dissenting, and concurring opinions are not treated separately. Rather, each document is made up of all of the opinions that were issued with respect to the disposition of a specific case.

134. While the resulting corpus of appellate court decisions is biased in that it has more appellate court decisions which were referenced by Supreme Court decisions than is actually the case, this imbalance is unlikely to influence the estimation of the topic proportions associated with each case because these cases are a small percentage of the total cases in the corpus.

fields for Supreme Court decisions.¹³⁵ Citations to federal appellate court decisions were identified by their Federal Reporter citation (either F.2d or F.3d). We verified the accuracy of the citations provided by LexisNexis by randomly sampling 50 decisions and manually checking the lower court citation. There are 724 decisions within our random sample of appellate court decisions that are associated with decisions in the Supreme Court corpus (i.e., appeals that were granted certiorari).

The vocabulary associated with the corpus comprises those words occurring at least 20 times in the entire Supreme Court corpus. There are 21,695 total words in the vocabulary. Some topic-model applications exclude function words (sometimes referred to as stop words) from the analyzed vocabulary.¹³⁶ Our analysis retained all words in the vocabulary, subject to the minimum-appearance constraint.¹³⁷

The LDA topic model was introduced over a decade ago, and a substantial amount of work has improved on the basic framework.¹³⁸ The more advanced variant that we use is a non-parametric topic model using the Pitman-Yor Process in place of the traditional Dirichlet distributions.¹³⁹ Topic models are partially supervised in the sense that they require an initial choice of the number of topics. More recently developed topic models attempt to be invariant to a sufficiently large number of topics, such that for a given corpus and a given probabilistic model there is, at least in theory, a “natural” number of topics for a given probabilistic topic model. Under such an assumption, specifying an excessively large number of topics will simply result in a large number of “empty” topics, which are associated with a negligible percentage of words in the corpus. Although it is also possible to use a form of “model selection” to help guide the choice of number of topics, it is not always necessary depending on the analytic task at hand. For purposes of the current

135. The accuracy of the citations provided by LexisNexis was assessed by collecting a random sample of 100 Supreme Court cases and comparing the citations with those found by a manual check. All citations were accounted for.

136. Function words include words such as *is*, *as*, *of*, *though*, and *around*. These words are frequently used in stylistic analysis of texts. See generally Carlson et al., *supra* note 12.

137. As would be expected from Zipf’s law, word frequencies within the vocabulary are strongly right-skewed. Zipf’s law, named after linguist George Kingsley Zipf, conjectures that in natural language use the frequency of a word is inversely proportional to its rank. If true for a given corpus, then the second most common word will be used half as frequently as the first and the third most common word will be used half as much as the second (and a quarter as much as the first) and so on.

138. Blei, Ng & Jordan, *supra* note 17, at 993.

139. Wray L. Buntine & Swapnil Mishra, *Experiments with Non-Parametric Topic Models*, in PROCEEDINGS OF THE 20TH ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 881, 882 (2014). This non-parametric version preserves the core heuristic of LDA while performing considerably better on held-out prediction tasks. The *hca* software we use is authored by Buntine and is open source. See Wray Buntine, *hca 0.63*, MACHINE LEARNING OPEN SOURCE SOFTWARE (Apr. 26, 2016), <http://mloss.org/software/view/527/>.

analysis, the number of topics is not central to our inquiry¹⁴⁰ and so we select a convenience number of 100 topics, which is large enough to capture a great deal of the semantic variability of the corpus and small enough to make fitting the topic model computationally straightforward.¹⁴¹

A complete list of the topics generated by the model is located in the appendix to this Article with the top words in each topic (i.e., the highest weighted words in the distribution) and the hand-coded topic labels.¹⁴² The top words for the first ten topics generated (the order in which topics appear is not meaningful) are presented in the following table:

Labels	Top Words
labor	union board labor employees employer nlr company bargaining relations national local act unfair
family	ms mrs did told husband time testified asked sexual home stated fact received mother daughter
elections	election political party candidates candidate campaign parties primary elections contributions ballot
narcotics	united drug cocaine government cir defendant conspiracy evidence drugs marijuana possession
immunity	immunity officers officer official police law county qualified officials city conduct rights liability
prisons	prison inmates inmate prisoner prisoners officials confinement conditions security jail amendment
procedure	motion district judgment appeal rule order filed summary party appeals judge final fed notice rules
medical	dr medical hospital mental treatment health care patient drug expert patients physician condition
criminal	trial defendant plea guilty indictment united jeopardy criminal double prosecution government
insurance	insurance policy company insured coverage insurer life ins policies liability loss judgment mutual

140. See Law, *supra* note 18, at 231–36 (drawing conclusions from model selection criteria).

141. Hanna M. Wallach, David M. Mimno & Andrew McCallum, *Rethinking LDA: Why Priors Matter*, in 22 ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 1973, 1981 (Yoshua Bengio et al. eds., 2009).

142. Most of the topics produced by the modeling admitted relatively straightforward labeling because the collection of most highly weighted words suggested a clear legal theme. A few topics were a bit more difficult to interpret and defied thematic labeling. To have a small number of unlabeled topics is common and even expected.

The labeling was done through the application of professional expertise and is fairly straightforward, in part because the top words map nicely to intuitive legal categories.

C. Testing the Topic Model

Demonstrating that topic models succeed in capturing semantic features of texts is itself a continuing area of research, particularly because there is no shared definition of semantic features. Traditional statistical measures (e.g., the ability to predict words in new or unseen texts from the same corpus) provide a means to evaluate a topic model's performance compared to previous approaches. But the capacity of a topic model to decompose a corpus into semantic themes or topics relies heavily on anecdotal testimony: many people observe that the inferred probability distributions over words (topics) tend to look like lists of semantically connected words. In addition to anecdotal evidence, a variety of indirect empirical tests involving human readers have been conducted. In these tests, readers' classifications of texts are compared with the representation provided by the topic model, or the topics themselves are directly evaluated.¹⁴³ The widespread use of topic models in research and industrial settings also offers some validation of their ability to produce parsimonious representations of large-text corpora.

The following discussion reports the results of two tests of how well the topic model we deploy captures relevant semantic characteristics of the Supreme Court and appellate court corpora. For this analysis, we first examine the general semantic fidelity of the topic model by examining how well topic distributions preserve information in the word frequencies for purposes of predicting hand-coded issue labels. Second, we examine whether the topic model is able to confirm several known features of the subject-matter distributions within Supreme Court and appellate court opinions.

One method of testing whether a topic model preserves information about the substantive issues found in opinions is by calculating what information about known classifications is lost when reducing a rich, high-dimensional representation of a corpus (i.e., word frequencies) to a low-dimensional summary (i.e., topic model).¹⁴⁴ One method for calculating this loss considers how well a supervised classifier performs at the task of placing documents in appropriate categories when provided with either of the following: (1) the low-dimensional summary provided by the topic model; or (2) the high-dimensional data of word frequency counts.

This method of validation is only available in situations where exhaustive expert labels are provided. Fortunately, the SCDB provides such an exhaustive set of labels for the Supreme Court.¹⁴⁵ The SCDB is the longest-running effort to use human researchers to collect metadata about Supreme Court decisions and opinions,

143. See, e.g., Jonathan Chang et al., *Reading Tea Leaves: How Humans Interpret Topic Models*, in 22 *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS* 288, 295 (Yoshua Bengio et al. eds., 2009).

144. This technique is used in Blei, Ng & Jordan, *supra* note 17.

145. Harold J. Spaeth et al., *Supreme Court Database*, Version 2014 Release 01 (2014), <http://Supremecourtdatabase.org>.

and the data collected in the SCDB has been used in dozens of studies of the Court.¹⁴⁶ We use the 14 issue-area labels used in the SCDB dataset (variable “issueArea”) as the reference categories and multi-class logistic regression as the classification algorithm. Classification accuracy is assessed using cross-validation; by turns 5% of the 7,503 Supreme Court decisions are “held out,” and the algorithm, given access to the remaining documents with the Spaeth labels, is asked to predict the labels. The classifier should, in theory, have an easier task at predicting the SCDB issue-area label using the original feature set (word frequencies of 21,695 unique words) than using a smaller set of topic shares. We find that the classifier is able to perform reasonably well using topic distributions alone, indicating that the topic model is capturing information that is highly relevant (in the sense of being predictively useful) to the expert issue-area classifications.¹⁴⁷

For a second test, we examine how well the model can capture the semantic similarity between pairs of Supreme Court opinions and appellate court opinions associated with the same case. For our measure of similarity, we rely on the symmetrized form of Kullback-Leibler (“KL”) divergence, a common measure of the difference between two probability distributions.¹⁴⁸ The procedure that we used for this test is as follows. For a given Supreme Court opinion, we identified the matching appellate court decision. We then calculated the KL divergence between the topic shares of the related opinion texts. For that same Supreme Court document, we then randomly drew 10,000 appellate-court-opinion documents with replacement and calculated the KL divergence between those opinions and the Supreme Court opinion. We then calculated the mean and standard deviation for the randomly generated distribution of divergences and estimated the likelihood that the divergence for the matched pair would have been randomly generated. Finally, we constructed a distribution of p-values derived by performing this test for each Supreme Court opinion with its matching appellate court opinion. The mean p-value is 0.03, indicating that we can safely reject the hypothesis that the topic shares associated with Supreme Court decisions are no more similar to the appellate court

146. See Hall & Wright, *supra* note 116, at 69 n.24; Lee Epstein, Jack Knight & Andrew D. Martin, *The Political (Science) Context of Judging*, 47 ST. LOUIS U. L.J. 783, 807 (2003).

147. Using the word frequencies for the entire vocabulary of 21,695 words, the mean out-of-sample accuracy for the issue classification was 0.73 (with a standard deviation of 0.05). Accuracy is calculated as the sum of true positives and true negatives, divided by total cases. What this means is that, with word count information for each Supreme Court opinion, the classification algorithm is able to successfully predict the issue category labels from the SCDB 73% of the time. This value is the mean of the distribution over many runs of training a logistic classifier and testing it on out of sample values. Using the topic proportions alone, which is only 100 rather than 21,695 variables, the mean accuracy of the classifier was 0.70—it turns out that very little information was lost, at least from the perspective of correctly predicting issue categories. Appendix Figure 1 illustrates this result.

148. See Hughes et al., *supra* note 15, at 7683; Carlson et al., *supra* note 12, at 1485. Following convention, we use the symmetrized version of KL divergence that is the average of the KL divergences of opinion A with respect to opinion B and opinion B with respect to opinion A. Hereinafter, when we refer to using the KL divergence we mean in this symmetrized form.

decisions being reviewed than to randomly selected decisions.¹⁴⁹ Given the background fact that the semantic content of Supreme Court opinions will be relatively more similar to the appellate opinion associated with the relevant case under review than a randomly drawn appellate court opinion, our result provides confidence that the topic model is successful at capturing relevant semantic characteristics of both cases selected for review and the accompanying Supreme Court decision.

III. THE INCREASINGLY IDIOSYNCRATIC COURT

This Part reports our main quantitative results. As discussed in Part I, our goals are to examine the degree to which the Court's opinions are distinctive when measured against the baseline of federal appellate cases, to identify the sources of any distinctiveness, and to estimate whether the Court's opinions have become more or less distinctive over time. To briefly summarize our results, we find that the semantic content of the Court's opinions—as estimated by our topic model—are measurably distinct. This difference appears to be due both to the selection of non-representative cases for review and to differences between how the Supreme Court and appellate courts discuss the same cases. In addition, there is a clearly identifiable time trend toward increased distinctiveness in the Court's opinions. This time trend is not apparent with respect to case selection, which indicates that although the Court is selecting cases with a similar level of representativeness as the past, it is writing about those cases in an increasingly idiosyncratic fashion.

A. *Supreme Court Opinions as a Genre Apart*

Our first analysis examines whether the topic model is able to identify aggregate-level differences in the semantic content of Supreme Court opinions and federal appellate opinions. To undertake this analysis, we rely on the topic share in each opinion, which is interpretable as the proportion of words in the document that is associated with a given topic. For each topic, we can examine the distribution of document-specific proportions across all Supreme Court opinions and all appellate court opinions and ask whether the distributions of these proportions are different.¹⁵⁰

For each topic, separate distributions of that topic are determined in both the Supreme Court opinions and the appellate court opinions. For example, for topic 1 produced by the model, we calculate the distribution of its use over the collection of Supreme Court opinions and then calculate the distribution of its appearance over the set of appellate court opinions. This procedure is repeated for each topic for a total of 200 topic distributions. For the statistical analysis, the observations are topic shares in each opinion. For each topic, the distributions are simply the frequency of opinions at any given topic-share level, which for each opinion can theoretically range from zero if a topic is not present at all to one when all words in the document are associated with that topic. Each of the topic

149. Appendix Figure 2 illustrates this result.

150. For this analysis, we use the Kolmogorov-Smirnov (“KS”) test, which is a non-parametric analog to Student's t-test (which only tests for a difference in the population mean).

distributions for the Supreme Court opinion can then be compared to the corresponding distribution for the appellate court cases.

We find that the distributions of shares in each of the two populations are unambiguously distinct for every topic.¹⁵¹ This finding indicates that, at a very basic level, Supreme Court opinions are recognizable and distinct from appellate court opinions. Given that there are significant differences between topic proportions for all topics, the distinctiveness of the Court's opinions is not due to a few stray words or the writerly oddities of an individual Justice: the coarse-grained topic-model analysis will ignore such small bore differences, and they are highly unlikely to be distributed across every topic. Instead, across the board we find visible differences using the relatively low-resolution topic-model approach.

There is also an intuitive match between the topics and the court that they are more associated with. For example, words associated with topic 91 tend to make up a greater percentage of words in appellate court opinions than Supreme Court opinions. Topic 91 includes a variety of words connected to contract law including the following: "contract," "agreement," "parties," "breach," and "terms." Words associated with topic 28, such as "amendment," "constitutional," "power," and "rights," are associated with constitutional law and appear more frequently in Supreme Court opinions. In the remaining ten topics with the largest differences in shares, the Supreme Court tends to have a higher share of words associated with statutory interpretation and constitutional rights, while the pool of appellate court opinions has a higher share of words associated with time and dates, testimony and evidence, proof, and appeals.

As discussed in Section I.C, the differences between the topics characterizing appellate court opinions and the Supreme Court opinions could arise through the following two separate processes: (1) the selection of cases through exercise of certiorari jurisdiction; and (2) opinion drafting. To distinguish between these two sources of difference, we can separate out the corpus of appellate court opinions that are associated with cases that were selected for review from the general pool of appellate opinions. In this way, we can estimate two different effects. First, we can examine how the certiorari process selects for particular topics. If some topics are overrepresented in the appellate court opinions associated with cases selected for review, this is an indicator that the certiorari process has targeted certain substantive areas. Second, we can disambiguate the certiorari process from the opinion-drafting process as a source of difference by in effect holding the underlying cases constant. We achieve this by comparing the Supreme Court's opinions with the appellate court opinions only for the cases selected for review.

This analysis begins by separating out the set of appellate court opinions associated with cases selected for review. Again, the 200 distributions are constructed, produced in pairs by generating one for each topic in each court-specific corpus, and comparisons are made between the distributions in each corpus. The results of this analysis confirm that both the certiorari process and the opinion-drafting process contribute to the distinctiveness of the Court's opinions. The corpus of cases selected for review is an intermediary corpus in the sense that they are more

151. The finding is significant with a $p < .0001$.

similar to the Supreme Court opinions than the pool of all appellate cases, but there remain identifiable differences. Recall that for *every* topic, there were statistically significant differences between their proportions in the Supreme Court corpus and all appellate courts. Reducing the comparator to only those appellate opinions associated with cases selected for review, the number of topics for which we can reject the null hypothesis of no dissimilarity with the same degree of confidence falls to only 29 out of 100. Both the processes discussed in Section I.C—case selection and opinion drafting—appear to be at work.

The topics that are overrepresented in one corpus or the other tend to follow intuitive judgments about the choices of emphasis made by the Court. Words associated with topic 21, which we label *Congress* include “act,” “congress,” “legislative,” “house,” and “senate.” Those words are more common in Supreme Court opinions. The same is true of topic 28, which we label *Fourteenth Amendment* and which includes words such as “process,” “amendment,” “protection,” “clause,” and “equal.” Words associated with topic 59, which we label *Appeals* include “appellant,” “district,” “circuit,” and “appellee.” Those words are more common in the appellate court opinions associated with cases selected for review than in Supreme Court opinions.

The distinctiveness of the Court is not due, then, only to the fact that it selects a non-representative set of cases to review. Given the reality of certiorari jurisdiction, and the fact that the Court has given no indication that it uses that jurisdiction for the purposes of randomly sampling cases for purposes of supervision, it would be surprising if at least some topics were not over- or under-represented in the Court’s opinions. The substantial differences between the Court’s opinions and the appellate court opinions—when the underlying cases are held constant—provides insight into how the Court chooses to discuss the legal issues in those cases. These differences provide more convincing reason to believe that the Court’s opinions make up something of a genre unto themselves that is distinct from the more general judicial genre. The topic-model method provides a means of quantitatively articulating that distinctiveness.

As mentioned above, it is somewhat difficult to calibrate expectations concerning the degree of distinctiveness that we observe for the Court’s opinions, especially given the fact that the quantitative measures of semantic similarity that we use are not easily given an intuitive, qualitative interpretation. We know that there are measurable non-random differences, but the meaning of those differences is elusive without some point of departure to provide broader context. In Section III.C, we provide that context by examining time trends in the distinctiveness of the Court’s opinions, which allows comparison between the contemporary Court and earlier times. This temporal analysis can tell us, at the very least, whether today’s Court has moved closer to or farther away from the appellate courts. Before moving to that analysis, we examine temporal trends in opinions associated with cases selected for review to refine our understanding of how the Court has exercised its certiorari over time.

B. Subject-Matter Selection

To undertake this analysis, we used a simple logistic regression model to predict certiorari based on topic shares. Observations in the model are the set of appellate court opinions, coded according to whether their associated cases were selected for review. The predictor variables are the topic shares in the opinions. We also include indicator variables to control for year-fixed effects and also included an interaction term between year and topic proportions to examine whether there are time trends in the topics that tend to be more highly represented in the opinions associated with cases selected for review.¹⁵² With year interactions, there are six configurations of potential effects that we can discover. Topic prevalence can be associated with either (1) higher or (2) lower likelihood that a case has been selected for review; in addition, these tendencies can be (a) static, (b) increasing, or (c) decreasing.

higher likelihood (1)	1a	1b	1c
lower likelihood (2)	2a	2b	2c
	static (a)	increasing (b)	decreasing (c)

We find that roughly half (53 out of 100) of the topics are significant predictors of whether the associated case is selected for review.¹⁵³ This is a noteworthy result and has important consequences for study of the certiorari process. As discussed in Section I.C., there is a significant literature within empirical legal studies that examines the variables that determine whether the Supreme Court is likely to grant review for a specific case. Prior analyses have focused on variables that are external to the text of the appellate court opinion.¹⁵⁴ Our analysis allows for a quantitative estimate of the relationship between semantic features within lower court opinions and the likelihood of a grant of certiorari. At the very least, our finding indicates that future analyses could use topic proportions or a similar mechanism that encodes semantic content to control for possible unobserved variables.

For example, there is a robust literature supporting the finding that the presence of amicus briefs is positively correlated with a grant of certiorari.¹⁵⁵ From existing analyses, however, it is difficult to establish that amicus briefs play a *causal* role because it is possible for a third unobserved variable to influence both the probability of a grant and the decisions of amici. Controlling for the semantic content of the associated lower court opinions can help (although not fully remedy) this inferential problem. In an ideal experimental setting, the analyst would be able to secretly take control of the amicus process and treat a randomly selected group of certiorari petitions (and only those petitions) with amicus petitions. This ideal

152. The full results are presented in Appendix Figure 3.

153. The topics are significant either alone or in interaction with year.

154. See generally Caldeira & Wright, *supra* note 102; Matthew Hall, *Experimental Justice: Random Judicial Assignment and the Partisan Process of Supreme Court Review*, 37 AM. POL. RES. 195 (2009).

155. See Caldeira & Wright, *supra* note 102, at 1122.

experiment is—to put it mildly—unlikely, but it would allow the analyst to draw very solid conclusions about the causal role of amicus briefs. By including topic proportions as control variables, the analyst would be able to roughly hold the semantic content of the lower court opinion constant for purposes of isolating the influence of the presence of an amicus brief.¹⁵⁶ Such a topic-controls model would more closely approximate the ideal experimental context by absorbing unobserved variables that are correlated with the semantic content of the lower court opinion, likely including many of the characteristics of the underlying dispute.

In addition to finding correlations between topic proportions and likelihood of a grant of certiorari, we find that the mix of topics associated with Supreme Court review has changed over time. Of the 53 significant topics, 17 have year interactions of some sort. This ratio indicates that the topics in the Court's docket are, to some degree, a moving target. A substantial number of topics are subject to temporal trends, but at the same time, there is some degree of subject-matter stability in the selection process.

Topics that stably increase the likelihood of being in the pool of selected cases (configuration 1a in the above notation) during our study period include certain constitutional issues, such as *speech* (topic 76); topics that relate to inter-branch relations, such as *Congress* (topic 21) and *statutory interpretation* (topic 35); and some substantive areas that appear to be of durable interest, such as *antitrust* (topic 23), *education* (topic 95), and *Native American affairs* (topic 17). Two clusters of topics that stand out for having relatively stable reduced likelihood (configuration 2a) are state law-related topics and criminal-law-related topics. The former includes *family law* (topic 1), *trusts and estates* (topic 34), *negligence* (topic 51), and *contracts* (topic 91). Criminal law topics include *narcotics* (topic 3), *surveillance* (topic 47), *conspiracy* (topic 70), and perhaps to some degree *automobiles* (topic 16).

We identify two topics that tend to be selected for review more often than one would otherwise expect with an increase in this tendency over time (configuration 1b). These are *death penalty* (topic 27) and *state-federal relations* (topic 40). For the death-penalty topic, in 1979, an increase in the proportion of words in an opinion that are associated with the death-penalty topic from one standard deviation below the mean to one standard deviation above the mean is associated with a 1.9% increase in the probability that the case in question is selected for review. In 1995, by contrast, such an increase in the proportion of words associated with the topic is associated with a 2.6% increase in the probability that the decision is selected for review.

Two topics that have declining tendencies to increase likelihood (configuration 1c) are *due process* (topic 28) and *discrimination* (topic 31). For the discrimination topic, in 1979, an increase in the proportion of words in an opinion

156. Prior work on the certiorari process does account for semantic features of lower court cases in certain ways; for example, by coding for the presence of certain legal subjects, such as civil rights, that are correlated with selection for review. See, e.g., *id.* at 1117–18. Existing hand-coded information on issue area provided by the Supreme Court Database provides some information on semantic content, but the information density of topic proportions is higher, leading potentially to more useful controls.

that are associated with the discrimination topic from one standard deviation below the mean to one standard deviation above the mean is associated with a 1.4% increase in the probability that the case in question is selected for review. In 1995, such a shift in the proportion of words associated with discrimination does not increase the probability of review appreciably.

Three topics that decrease likelihood, with an increasing tendency in this direction (configuration 2b) are *insurance* (topic 9), *sentencing* (topic 19), and *energy* (topic 66). In the case of insurance, in 1979, an increase in the proportion of words in an opinion that are associated with the insurance topic from one standard deviation below the mean to one standard deviation above the mean is associated with a 1.6% decrease in the probability that the case in question is selected for review. In 1995, such a shift is associated with a 1.8% decrease in the probability that the decision in question is selected for review (a meaningful decrease given that the baseline odds against a case being selected have increased considerably between 1979 and 1995 due to a shrinking Supreme Court docket). Underrepresented topics that headed in the opposite direction (configuration 2c) include *patents and invention* (topic 61).

In the certiorari process, the topic distribution in cases is a significant predictor of the likelihood that a case will end up being heard and decided by the Supreme Court. In this process, there are certain legal areas that are consistently favored, others that are consistently disfavored, and many in which predictive power has shifted over time. Some of the time trends have reinforced overrepresentation (e.g., death-penalty cases) or underrepresentation (e.g., insurance cases) while the predictive power of other topics has gradually declined (e.g., discrimination cases).

C. Trending Toward Distinctiveness

We now turn to the question of whether the distinctiveness of the Court's opinions has remained constant, or has changed over time. In Section I.B, we defined the judicial genre with respect to a baseline corpus of documents. To the degree that the opinions of the Court are distinctive compared to the baseline corpus, they can be thought of as forming a separate genre. To motivate this definition, we imagined a prediction task in which a hypothetical law student was asked to guess the authoring court for a randomly selected opinion. If the law student finds that task easy, then the two corpora have distinctive genre characteristics. Furthermore, if the ease of prediction increases over time, we can be confident that the collections of topics are becoming more distinct with respect to each other.

For the analysis in this section, we build on this motivating thought experiment through a machine-learning algorithm that mirrors the prediction task given to the hypothetical law student. Using only the information contained in the topic distributions, the goal of the algorithm is to predict whether a randomly selected opinion has been drafted by the Court. From the results described in Section III.A, we know that with respect to the entire body of Supreme Court and appellate court opinions, prediction of some degree of accuracy will be possible—indeed, there are statistically significant differences between the two corpora in the proportions for every topic. The more interesting question that is the focus of this section is whether the corpora have been growing more distinctive over time.

To answer this question, one could start by comparing topic proportions on a yearly basis to examine whether there are changes over time. If a given topic begins by being comparatively overrepresented and its presence in the Supreme Court corpus increases over time, one might conclude that the corpora are becoming more distinctive. The problem with this approach is that there are many topics that are over- or underrepresented in Court opinions with levels of distinctiveness that may change considerably in a nonlinear fashion over time. Some measure to aggregate these many changes is required to arrive at a global assessment of distinctiveness.

One hypothetical way to attempt to do so would be to develop some measure of central tendency and then ask whether those measures are growing closer or further apart. For example, it would be possible for each year to estimate the topic proportions in each corpus and then estimate their level of difference using a measure such as KL divergence. However, this method is problematic because the topic proportions for the corpus need not be particularly common in individual opinions. Attempting to normalize the distributions or develop alternative measures of central tendency creates problems of interpretation and requires an unappealing degree of data manipulation.

A prediction task resolves this difficulty by generating a single metric of distinctiveness that is based on predictive accuracy of the classifier algorithm. Individual opinions are treated as they are—no attempt is made to tease out higher-level characteristics of the corpora that can be compared over time. In addition, predictive accuracy has an intuitive interpretation as a measure of distinctiveness.

To generate this metric, we begin by limiting the corpus of appellate court and Supreme Court opinions to a single year. Because the number of decisions in each year varies considerably—there were far more appellate court opinions in 2000 than in 1960—we randomly sample year-specific corpora of equal sizes. We then hold out 50% of the appellate court and Supreme Court opinions and train a basic logistic-regression-classification model using the remaining opinions. The only information that the classification model uses is the topic proportions in the opinions. Once the classification model has been fit, we evaluate it for accuracy on the 50% that was held out. This task is repeated many times, each time randomly sampling the 50% of cases that are held out. From this procedure, we construct a distribution of predictive accuracy for the classifier for that year. We repeat these same steps for each year in our sample, providing a means of evaluating whether the distributions change over time. For convenience, we aggregate this information by decade.

There is some risk that a naïve classifier will become quite good at prediction based on relatively insignificant differences, for example, differences in the usage of a few characteristic words such as the Justices' names or the courthouse address. A high degree of predictive accuracy if based on these small differences would not necessarily imply substantial and meaningful distinctiveness. Use of only the topic model proportions as the basis of prediction reduces this risk. There is a substantial level of aggregation involved in moving from all words to 100 topics, and this aggregation lowers the risk that trivial differences will substantially affect the success of the classifier. The *loss* of information associated with topic modeling

helps reduce the risk of accentuating minor differences. If we find that that the model has an easier time predicting whether opinions are authored by the Court based on topic model proportions alone, we can say with a reasonable degree of confidence that the two corpora are growing more distinct from each other in a meaningful way.

Figure 1 displays the results of this analysis. We confirm that the logistic-regression classifier performs reasonably well in predicting the difference between appellate court and Supreme Court opinions. Based on the results discussed in Section III.A, the finding is not surprising because we already know that there are significant differences in topic proportions in the two corpora. More interestingly, we find that prediction is improving considerably over time. The center of the distribution of the accuracy of held-out prediction starts at roughly 80% in the 1950s, but over time increases to over 95% in the 2000s. This is a highly significant result.¹⁵⁷ By the end of the study period, a quite simple classifier using only topic proportions achieved nearly perfect prediction. These results quite clearly indicate that Supreme Court opinions are growing more distinctive compared to those in the appellate courts. From this analysis, we know that the mix of topics present in each corpora in each year provides increasingly more information about the identity of the authoring court, in the sense that the classifier improves over time.

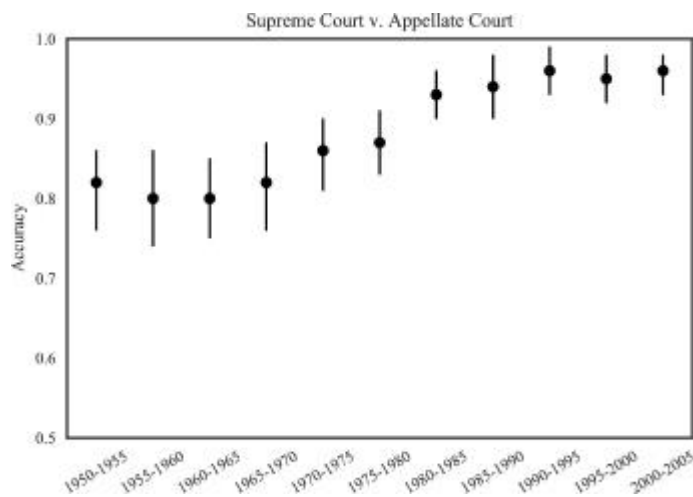


Figure 1: Prediction of Supreme Court Opinions

As discussed in Section I.C, both case selection and opinion drafting could result in Supreme Court opinions that are distinct from the general-pool appellate court opinions. From prior work, we know that there is an important winnowing effect during the certiorari process, and the cases that come before the Court are far from randomly drawn. The analysis discussed in Section III.B finds that the topic proportions in the opinions associated with cases selected for review are highly distinctive, with the certiorari process favoring some topics at the expense of others. At least part of the reason that the Court's opinions are distinctive is that they are based on a non-representative set of cases. It is worth considering whether the

¹⁵⁷ Pearson product-moment correlation between year and accuracy is 0.79. The error bars for Figures 1 and 2 indicate the range between the 10th and 90th percentiles.

contribution of the certiorari process to the distinctiveness of the Court's opinions is growing, declining, or remaining relatively flat. The converse question is whether the opinion-drafting process has changed over time such that, holding the underlying cases constant, the Court's discussion of those cases has become increasingly distinctive over time.

To investigate these two questions, we carry out the same logistic regression classifier analysis on three different corpora: the set of all appellate court opinions, the set of appellate court opinions associated with cases selected for review, and the Court's opinions. We then carry out two sets of analyses, using the cases selected for review set as an intermediary corpus. We first examine whether the Court is using its certiorari power more aggressively than in the past in the sense of selecting cases that are more distinct from the pool of all appellate court cases. If so, we should find that the performance of the classifier would increase over time. We then examine the opinion-drafting process by analyzing whether the Court's opinions are growing more distinctive *vis-à-vis* the intermediary corpus of appellate opinions associated with cases selected for review. In essence, this analysis holds the underlying legal issues constant to determine whether the Court is discussing those issues in a more distinctive fashion.

The results of these two analyses are reported in Figure 2. We do not find any evidence that there is any change over time in the representativeness of the group of cases being selected for review. As examined in Section III.B, the mix of cases selected by the Court for review does change and topics come in and out of favor, but the aggregate level of difference (as estimated through the prediction task) has not increased. Although this analysis cannot rule out the possibility that a more sensitive textual analysis would identify some temporal change, we fail to find any such effect using the same model that identifies an overall growth in the distinctiveness of the Court's opinions. We can therefore say with confidence that the increasing distinctiveness that we identify is not caused by a change in the level of representativeness in the cases selected for review.

Since the Court's opinions are growing more distinctive, yet the underlying cases selected for review are not, the natural inference is that the Court's opinions must be becoming more distinct from the appellate court cases selected for review. We confirm this conclusion, finding that, when comparing Supreme Court opinions and the intermediary corpus of appellate opinions associated with cases selected for review, the performance of the classifier improves over time. Starting with accuracy centered at roughly 70% in the 1950s, performance increased to well over 90% by the 2000s. The lesson from this analysis is that, although the cases selected for review in recent years are no more distinct from the pool of all appellate court cases than in the past, the way that the Supreme Court analyzes and discusses the legal issues presented in those cases has grown increasingly idiosyncratic over time.

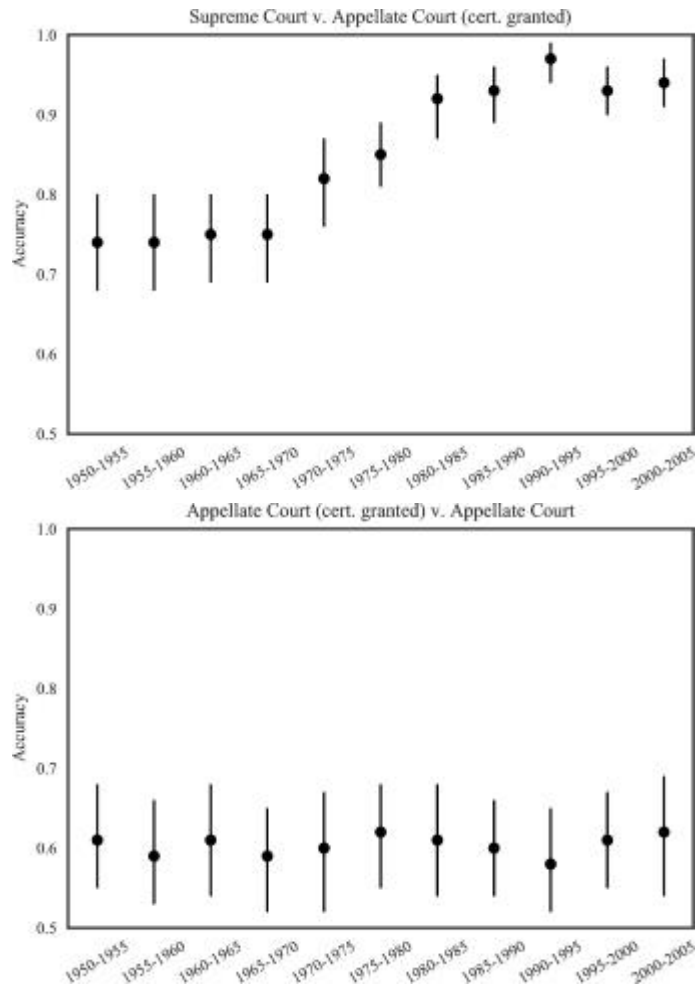


Figure 2: Prediction of Supreme Court and Appellate Court Opinions

This finding is quite striking and indicates that, at least according to the measure developed and discussed above, the Court's opinions conform less well to the genre of judicial opinions than in the past, and this change is due to the opinion-drafting process in the Court. Opinions written by the Supreme Court are more characteristic and easily identifiable than in the past; they are, on their face, less obviously associated with the opinions drafted by the appellate courts.

Although use of topic proportions helps mitigate the risk that this quantitative trend is the result of trivial differences, it is difficult in principle to fully reject this possibility. Since it is unlikely that the legitimacy-enhancing effect of conforming to the genre of the judicial opinion would be genuinely undermined by such trivial effects, it is worth investigating further into the source of the Court's increasing semantic idiosyncrasy. For this analysis, we leverage the results from our naïve classifier to engage in a qualitative analysis of a selected pool of cases to

provide an interpretation of the meaning of this growing distinctiveness and its potential importance for the Court's legitimacy.

IV. CASE STUDIES IN DIFFERENCE

As mentioned above, scholars and commentators have noted the growing stylistic distinctiveness of the Court's opinions for some time.¹⁵⁸ Other scholars have identified fine-grained temporal changes in the content of the Court's opinions; for example, Farganis notes that "since the Court's 1954 decision in *Brown* . . . the [J]ustices have made seventy-one . . . references to the Court's institutional legitimacy, compared with just nine in the 164 years up to that point."¹⁵⁹ The value of the analysis in the preceding Part is to apply quantitative techniques to a very general measure of the semantic content of Supreme Court opinions that can be compared against a recognizable baseline. The topic model and classifier can identify distinctiveness and changes in levels of distinctiveness, in a way that is relatively free from biases and potential blind spots that might arise from a top-down, theory-driven approach that begins with some a priori account of the judicial genre. One downside of our naïve approach is that the results can be somewhat difficult to qualitatively interpret: Supreme Court opinions are growing more recognizable and distinctive, but how?

This Part serves as an initial remedy to this difficulty by engaging in an interpretive exercise that is informed by our quantitative results. For each five-year period in our study, we ranked every Supreme Court opinion according to the likelihood that it would be misidentified by the classifier as an appellate court opinion. In essence, we ranked the Court's opinions according to how distinctively Supreme Court like they are based on our measure. For the two five-year periods corresponding to the beginning and end of our study period, we identified the two most- and least-characteristic Supreme Court opinions. The following discussion is based on a reading of those cases. Unlike the prior Parts, here we depart from statistical analysis in favor of a qualitative approach that is exploratory and somewhat intuitive in nature. Our goal is to leverage the computational analysis above to provide grist for human interpretation that can both shed light on our quantitative results and point the direction for future research.

A. 1950–1955

The two least characteristic Supreme Court opinions—that is, the most easily mistaken for appellate court opinions by our classifier—from the early 1950s are *United States v. Borden Co.* and *Alison v. United States*.¹⁶⁰ The first involved an appeal from an antitrust case concerning a "conspiracy to restrain and monopolize sale of fluid milk."¹⁶¹ The second involved the appropriate tax year to deduct losses

158. See *supra* notes 13–14; *cf. supra* note 114 (examining growth of dissenting opinions).

159. See Farganis, *supra* note 70, at 207.

160. *United States v. Borden Co.*, 347 U.S. 514 (1954); *Alison v. United States*, 344 U.S. 167 (1952).

161. 347 U.S. at 515.

associated with embezzlement.¹⁶² Both are short—just a few pages in the U.S. Reporter—and neither has generated a substantial body of case law: *Borden* has been cited in a total of 128 cases, while *Alison* has been cited a mere 40 times.¹⁶³ Nor has either case been the subject of much academic commentary: *Borden* has been cited in 41 law-review articles; *Alison* has been cited in 11.

Justice Clark’s opinion in *Borden* was for a unanimous Court; Justices Black and Jackson did not take part in the decision. One of the questions in the case was whether the district court incorrectly ruled on the inadmissibility of certain evidence of antitrust violations.¹⁶⁴ The Court declined to rule on the evidentiary questions on prejudicial error grounds. After “considering as much of the evidence as is before us,” the Court found that, even considering the excluded evidence, there was inadequate support for the government’s case.¹⁶⁵ A second question in the case is one of procedure: whether the district court erred in refusing to issue an injunction sought by the government concerning certain pricing behavior because that court, in an earlier proceeding brought by private parties against the same companies, had issued a similar injunction to the same effect.¹⁶⁶ The Court held that the district court “abused [its] discretion in refusing the Government an injunction solely because of the existence of the private decree.”¹⁶⁷ More than 50% of the words in the opinion are associated with three topics: topic 13 (one of the procedure topics); topic 55 (which was difficult to label, but appears to be a topic of expository words); and topic 23 (antitrust). *Borden* has been cited relatively recently both on the requirement of prejudicial error and the issue of multiple injunctions in antitrust litigation.¹⁶⁸ Justice Black’s opinion in *Alison* was not for a unanimous Court; Justices Douglas and Burton dissented. However, the dissenters did not write separate opinions, and so the document’s text contains only Justice Black’s words. Half of the words in the document are associated with either topic 55 (expository words) or topic 14 (tax). The most recent citation to the case at the circuit court level occurred in 1984 for the proposition that “whether a loss has been sustained depends upon the taxpayer’s reasonable expectation of recovery from the wrongdoer.”¹⁶⁹

The two most characteristic cases (according to our measure) from the early 1950s are *United States v. Williams* and *Brannan v. Elder*.¹⁷⁰ *Williams* involved the interpretation of criminal provisions within Reconstruction Era civil-rights acts (codified at 18 U.S.C. §§ 241, 242), which prohibit conspiring or acting under color

162. 344 U.S. at 168.

163. Citation counts throughout this Part are based on Westlaw KeyCite searches conducted on January 27, 2017.

164. 347 U.S. at 516.

165. *Id.*

166. *Id.* at 517–19.

167. *Id.* at 520.

168. *Rosenfeld v. Oceania Cruises, Inc.*, 682 F.3d 1320, 1333 (11th Cir. 2012) (citing *Borden* on prejudicial error); *Howard Hess Dental Labs. Inc. v. Dentsply Int’l, Inc.*, 602 F.3d 237, 249 (3d Cir. 2010) (citing *Borden* on antitrust injunctions).

169. *Miller v. C.I.R.*, 733 F.2d 399, 402 (6th Cir. 1984).

170. *United States v. Williams*, 341 U.S. 70 (1951); *Brannan v. Elder*, 341 U.S. 277 (1951).

of state authority to deprive persons of their constitutional rights.¹⁷¹ *Brannan* involved the right to employment in the federal government for veterans under the Veterans' Preference Act of 1944. The more characteristic cases are somewhat longer, with *Williams* spanning roughly two dozen U.S. Reporter pages, but they are not very long by contemporary standards. *Williams* has been cited in just over 100 cases and has been discussed to some degree in the law-review literature, with 93 articles citing to the case. *Brannan* has been cited in only 27 cases and three law-review articles.

The *Williams* opinion reads, to some extent, as distinctive in ways that might match a popular understanding of a characteristic Supreme Court opinion. The decision was a split plurality, with four Justices dissenting and Justice Black concurring. The top three topics in the opinion were topic 28 (Fourteenth Amendment), topic 55 (exposition words), and topic 70 (conspiracy). Justice Frankfurter's plurality opinion included a two-page appendix which tracks the history of the statutory text of the relevant provisions from their adoption during the Reconstruction Era. The dissent includes some biting language characterizing Justice Black's approach of avoiding the case on *res judicata* grounds as "too facile for the facts."¹⁷² The substantive holding of the plurality opinion, which limited the application of the relevant criminal provision to "interference with rights which arise from the relation of the victim and the Federal Government, and not to interference by State officers with rights which the Federal Government merely guarantees from abridgement by the States,"¹⁷³ has not stood the test of time—it was "effectively overruled" 15 years later in *United States v. Price*.¹⁷⁴ Scholars have also criticized the plurality opinion's characterization of political conditions during Reconstruction as relying on an outdated (even for the time) historiography that was "abandoned by historians" as "incomplete, unbalanced, and often racist."¹⁷⁵ *Brannan* reads less obviously Supreme Court like than *Williams*. The words in Justice Clark's opinion (Justice Black dissented, but did not write separately) are primarily associated with topic 55 (exposition), and then topic 60 (mails) and topic 68 (employment). The *mails* topic may seem out of place, but the word "veterans" is among the most heavily weighted words in that topic, which may pick up on employment preferences within the postal service. The fourth most prevalent topic in the opinion is topic 21 (Congress), which includes words like "act," "legislative," and "history." Indeed, the opinion does include some discussion of legislative history, and it was cited in a law-review student note the following year in a prescient discussion of a then-

171. For background on 18 U.S.C. §§ 241, 242 and their origins in the Reconstruction Era, see Frederick M. Lawrence, *Civil Rights and Criminal Wrongs: The Mens Rea of Federal Civil Rights Crimes*, 67 TUL. L. REV. 2113 (1993).

172. 341 U.S. at 95.

173. *Id.* at 81–82.

174. *United States v. McDermott*, 918 F.2d 319, 325 (2d Cir. 1990) (quoting *United States v. Price*, 383 U.S. 787, 798 (1966)).

175. See Eric Foner, *The Supreme Court and the History of Reconstruction—and Vice Versa*, 112 COLUM. L. REV. 1585 (2012), 1585, 1592, 1594 n.43 (citing *Williams* as example of the Court's reliance on flawed history).

embryotic movement against the practice of relying on legislative history in the course of statutory interpretation.¹⁷⁶

The overall impression of reading the most and least recognizably Supreme Court like opinions from the early 1950s is that the topic-model-based measure serves as an at least rough proxy for meaningful opinion characteristics associated with the judicial genre. The slightly stinging language in dissent and the relatively more detailed discussion of legislative history in the more characteristic case conforms to intuition, as does the mix of relatively more technical subjects (tax, antitrust, and procedure) and the less contentious language in the opinions more likely to be mistaken as issuing from an appellate court. At the same time, an observer familiar with contemporary opinions of the Court finds the more judicial character of even the relatively more characteristic Supreme Court opinions somewhat striking. Even in *Williams*, which generated deep disagreement on the Court on questions of both statutory interpretation and how best to understand a highly divisive period in the nation's history, the language is primarily technical and legalistic, with only muted rhetorical flourishes.¹⁷⁷ This is quite unlike how one might imagine an opinion on a similarly divisive subject might be written today.

B. 2000–2005

At the end of our study period, the two Supreme Court documents that best conformed to the judicial genre—the ones that were least distinguishable from appellate court opinions—were *Central Green v. United States* and *Yarborough v. Gentry*.¹⁷⁸ The first concerns the ability of a pistachio farmer to recover damages from the federal government associated with damage arising from a federally owned water project.¹⁷⁹ The second concerns an ineffective assistance of counsel claim in a habeas case.¹⁸⁰ As with the less characteristic opinions from the 1950s, both *Central Green* and *Yarborough* are relatively short: Justice Stevens's opinion in *Central Green* spans 15 U.S. Reporter pages; the *per curiam* opinion in *Yarborough* spans 9. In neither case was there a dissent. *Central Green* has been cited in 58 cases; *Yarborough*, by contrast, has been cited in 3,380 cases. Despite these considerably different degrees of attention in the lower courts, academic commentators have relied in roughly equal proportions on the two opinions: *Green* has been cited in 49 law-review articles; *Yarborough* in 55.

The large number of citations that *Yarborough* has received is likely related to the fact that it is less recognizable as a Supreme Court opinion: it deals with matters such as habeas petitions and ineffective assistance of counsel claims that

176. See Note, *A Re-evaluation of the Use of Legislative History in the Federal Courts*, 52 COLUM. L. REV. 125 (1952).

177. For example, Justice Douglas notes in his dissent that the relevant statutory provisions “are companion sections designed for the protection of great rights won after the Nation’s most critical internal conflict.” 341 U.S. at 87.

178. *Cent. Green Co. v. United States*, 531 U.S. 425 (2001); *Yarborough v. Gentry*, 540 U.S. 1 (2003).

179. 531 U.S. at 426–27.

180. 540 U.S. at 5.

make up a substantial portion of the docket in the appellate courts. It is telling that among *Yarborough*'s citations, only 169 come from federal courts of appeals; the vast majority—nearly 3,000—of the case citations to the opinion are in federal district courts. There is even specific language within *Yarborough* that appears particularly prone to being quoted. There are 439 cases that directly quote the following language from *Yarborough*: “The Sixth Amendment guarantees reasonable competence, not perfect advocacy judged with the benefit of hindsight.”¹⁸¹ It is not difficult to understand why this pithy phrase has become popular in district courts that may feel overwhelmed by ineffective assistance of counsel claims in which there are at least arguable mistakes or unwise decisions on the part of defense lawyers. Another interesting feature of *Yarborough* is that two pages in this relatively short opinion are given over to quoting, in its entirety, the closing arguments in the petitioner's criminal trial which served as the basis for a finding by a Ninth Circuit panel that counsel had been ineffective. In that quoted language, there are a fair number of words and subject matters that are likely relatively uncommon in Supreme Court opinions but that may be more common in appellate court opinions (especially that include similar quotes).¹⁸² Accordingly, the predominant topic in *Yarborough* is topic 90, which includes language associated with testimony, such as “mr,” “did,” “said,” “know,” and “yes.” The second most prevalent topic was 71 (ineffective assistance).

The most prevalent topic in *Central Green* is topic 25, which is associated with Native American affairs, but which also includes several words associated with water law—including “water,” “river,” “rights,” and “waters.” This subject-matter commonality is unsurprising, given the overlapping history of litigation in the two areas. The second most prevalent topic we labeled as *precedent* because it includes words that courts often use when discussing prior precedent. Indeed, *Central Green* includes a considerable discussion of the difference between dicta and holding and the relative weight of each. The third most prevalent topic is 35, *statutory interpretation*, and Justice Stevens spends a considerable portion of the opinion construing the statutory liability limitation at issue in the case. Interestingly, each of the substantive topics in the case (25 and 35) is relatively more common in the pool of cases selected for review,¹⁸³ suggesting that they are of special interest to the Court, but it may be Justice Stevens's careful attention to prior precedent (as a portion of the total number of words in the opinion) that renders this opinion less-obviously distinguishable from appellate court opinions.

The two most easily distinguishable Supreme Court opinions in the final five years of our study period are *Branch v. Smith* and *Smith v. Doe*.¹⁸⁴ The first case concerns application of the Voting Rights Act to a congressional redistricting plan in the state of Mississippi.¹⁸⁵ The second concerns a claim that the Alaska Sex

181. *Id.* at 8.

182. For example, the opinion quotes the following statement from defense counsel: “I don't know if thievery and stabbing your girlfriend are all in the same pot.” *Id.* at 4.

183. *See supra* Section III.B.

184. *Branch v. Smith*, 538 U.S. 254 (2003); *Smith v. Doe*, 538 U.S. 84 (2003).

185. 538 U.S. at 258.

Offender Registration Act violated the ex post facto clause of the Constitution.¹⁸⁶ *Branch* involved multi-part majorities, concurrences, and dissents with the Justices joining each other in complicated partial alignments; the opinion ran a total of 54 pages. *Smith v. Doe*, although shorter (a bit under 30 pages) also generated two dissents and two concurrences. *Branch* has been relied on to some degree in the lower courts (it has been cited in 139 cases) but has been discussed somewhat more in the academic literature (150 law-review articles cite the case). *Smith v. Doe*, on the other hand, has generated a substantial amount of both citation and commentary, with over 1,200 cases and over 460 law-review articles citing it.

Particularly striking is the amount of negative treatment that *Smith v. Doe* has received from other courts. Some states have explicitly departed from the Supreme Court's interpretation of the Constitution when interpreting similar state constitutional provisions.¹⁸⁷ Other state courts, even when ostensibly following *Smith v. Doe*, have declined to extend its logic to new cases.¹⁸⁸ Although state courts are broadly free to interpret their own constitutions, the Court's opinion in *Smith v. Doe* appears to have carried relatively little in the way of persuasive authority.¹⁸⁹ Even the Alaska courts eventually struck down the same law challenged in *Smith v. Doe* based on the state constitution's ex post facto clause, noting that although the state court has not yet had any reason in prior cases to depart from the Supreme Court's interpretation of the ex post facto clause, it had "the authority and, when necessary, duty to construe the provisions of the Alaska Constitution to provide greater protections than those arising out of the identical federal clauses."¹⁹⁰

The prevailing topics in *Branch* are topic 86 (voting), topic 35 (statutory interpretation), and topic 21 (Congress). Each of these topics is closely associated with the Supreme Court so it is not very surprising that this opinion would be identified by the classifier as distinctive. In *Smith v. Doe*, the largest number of words are associated with topic 26 (briefs) followed by topic 52 (sentencing) and topic 41 (precedent). Although the sentencing topic is quite prevalent in the appellate courts, it may be that the discussion of sentencing-related issues is distinctive. For example, the prevalence of topic 26 may involve the number of cross references to various portions of the opinion as the Justices consider, endorse, or dispute the various arguments presented in either the majority, concurring, or dissenting

186. 538 U.S. at 89.

187. See, e.g., *Doe v. Dept. Pub. Safety & Corr. Servs.*, 62 A.3d 123 (Md. 2013) (construing Maryland state constitutional provision prohibiting ex post facto laws and striking down challenged sex-offender registry law); *Wallace v. State*, 905 N.E.2d 371 (Ind. 2009) (construing Indiana state Constitution ex post facto clause and striking down sex-offender registry law in as applied challenge).

188. See, e.g., *Riley v. N.J. State Parole Bd.*, 98 A.3d 544 (N.J. 2014) (examining features of New Jersey law and finding that they are sufficiently punitive to fail scrutiny under *Smith's* ex post facto analysis).

189. See *State v. Letalien*, 985 A.2d 4 (Me. 2009) (applying Maine Constitution); *State v. Williams*, 952 N.E.2d 1108 (Ohio 2011); *Starkey v. Okla. Dept. Corr.*, 305 P.3d 1004 (Okla. 2013); *Doe v. State*, 111 A.3d 1077 (N.H. 2015); *Jensen v. State*, 905 N.E.2d 384 (Ind. 2009).

190. *Doe v. State*, 189 P.3d 999, 1005 (Alaska 2008).

opinions. The Court's discussion of shaming punishments of the colonial period, and citation to various historical sources on the subject, are also fairly unusual compared to the run-of-the-mill appellate-court sentencing decision.¹⁹¹

Reading these four sample cases from the end of the study period, the relatively more characteristic cases appear distinctive in ways that roughly accord with expectations of how a typical Supreme Court opinion differs from a typical appellate court opinion. Certainly the issues are of greater public salience and likely to be subject to more political contestation and controversy. Comparing changes over time is more difficult. But trends toward greater reliance on values-oriented language and styles of discussion that particularly highlight disputes between the Justices and explicitly criticize the majority or dissenting opinions may be apparent. Given the limited sample examined here, it is difficult to draw any firm conclusions, but the following section proposes potential ways in which future research could translate these general impressions into testable propositions.

C. Directions for Future Research

Literary scholar and leading digital humanist Franco Moretti has coined the phrase *distant reading* to characterize the now-growing practice of applying quantitative tools to the study of literary texts.¹⁹² As might be expected, this development has not been uncontroversial, and a lively debate has broken out in the field about the relative merits and demerits of quantitative *distant* reading and qualitative *close* reading. This discourse is reminiscent of the debates that broke out in the legal academy in the late twentieth century as quantitative tools from the social sciences began to work their way into law scholarship.¹⁹³ Over the subsequent several decades, legal academics from both camps largely have been able to reconcile themselves to, and find the value in, alternative methodologies.¹⁹⁴

Quantitative approaches to empirical legal studies have typically relied on close human reading of texts. The goal has primarily been to standardize such readings, record them at scale, and then apply statistical tools to examine the results.¹⁹⁵ This Article takes the natural additional step of applying new statistical tools to recently available electronic resources to directly treat text as data. The result is a new vantage point, even more distant than prior empirical legal scholarship, that allows for the quantitative estimation of macro-level trends that are impossible to perceive for even the most dedicated group of human readers.

To take appropriate advantage of these new techniques, it is important to understand what they offer and what they miss. In a qualitative assessment of the distinctiveness of U.S. Supreme Court opinions, it is easy to focus on certain obvious characteristics, such as the length of opinions, the number of footnotes, the relative

191. 538 U.S. at 97–99.

192. See MORETTI, *supra* note 31, at 1.

193. See, e.g., Lawrence M. Friedman, *The Law and Society Movement*, 38 STAN. L. REV. 763 (1986).

194. For a recent assessment of the state of empirical legal studies, see Jeffrey J. Rachlinsky, *Does Empirical Legal Studies Shed more Heat than Light? The Case of Civil Damage Awards*, 29 RATIO JURIS 556 (2016).

195. See Hall & Wright, *supra* note 116.

lack of consensus, and the presence of politically controversial issues. These are, indeed, potentially important characteristics, some of which can (and have) been subjected to empirical study. The topic-model approach that we used will capture some of these characteristics—especially with respect to subject matter. But it does not attend to other characteristics such as opinion length and may catch others only obliquely via correlation with semantic content, such as the presence of dissents. At the same time, by focusing extremely narrowly on the words within an opinion (as represented by topics) the model is able to represent some features that are important in their own right, and that may also serve as a very general proxy for a wide range of opinion characteristics.

Subsequent work can build on our findings and the techniques used in this Article. The two most important advantages of the computational text analysis approaches that we apply are the ability to analyze a very large number of documents and the related ability to detect small or subtle differences that would elude a human reader. Here, we exploit these advantages to analyze a large corpus of appellate court decisions to construct a baseline against which the Supreme Court's opinions can be compared. By grounding our analysis in this baseline, we avoid issues that would arise in comparing the Court only to itself. For example, a baseline can help interpret the meaning of the fact that the Court, in the years after its *Brown* decision, referenced its own legitimacy much more frequently than in the past.¹⁹⁶ One possible explanation for this phenomenon is that the *Brown* decision (or other decisions at the same time) and the social reaction to it caused the Court to become concerned with its standing. But it is also possible that the Court was simply mirroring broader social trends: indeed, starting in the mid-1950s—roughly contemporaneously with the Court's *Brown* decision—the word legitimacy appears to have become considerably more popular within general usage.¹⁹⁷ Of course, the fact that the Court's concerns could be related to a broader phenomenon does not undermine its importance; rather, it provides some potential insights into the causal mechanisms at play.

There are several directions for future empirically oriented scholarship on the judicial genre and the Supreme Court. Some that are suggested by our qualitative analysis are the relationship between distinctiveness—either on our measure or some other—and the downstream life of an opinion, such as the likelihood of being cited in the future, the degree of political controversy generated by an opinion, or its legal stability (i.e., the probability of being overturned). There may be upstream effects that are worth investigating, such as the role of amicus briefs or the involvement of the Solicitor General's Office in affecting the content of opinions. Naturally, there may be some partisan or ideological effects that are worth investigating as well.

Our work also suggests that our notion of distinctiveness and computational text-analysis approaches more generally can shed light on the interaction of the Supreme Court and appellate or state courts.¹⁹⁸ Two questions that stand out from the earlier qualitative analysis are, first, whether opinions that

196. See Farganis, *supra* note 70, at 207.

197. Ngram Viewer, Google Books, <http://bit.ly/2jkc3FU>.

198. Cf. Rice, *supra* note 92.

conform to the judicial genre tend to be cited more by the appellate courts; and, second, whether such opinions tend to be quoted at length more often. The causal story for these possibilities is relatively straightforward. Supreme Court opinions that are more judicial in character are more useful in terms of relating to issues often found in the appellate courts. Thus, they are more likely to be cited. In addition, the language used in those types of opinions may be a more comfortable fit for an appellate court decision, and therefore be more likely to be quoted at length. There are also important and interesting outstanding questions on how the Supreme Court affects opinion content in the lower courts—for example, whether interpretive methods (and associated word choices) that are adopted by the Court affect how lower courts reason through and write about legal questions. Tracing the diffusion of ideas, as embodied in text, through the judicial system is a particularly promising avenue of computational legal studies.

More broadly, the approaches that we deploy here can be used to compare interactions between the Court and other social institutions that generate text. These institutions include political bodies, academics, and even producers of popular culture such as blogs. Again, these different corpora could be compared to examine the diffusion of ideas as well as the extent to which the subject matters of controversy within political institutions or popular discourse are reflected within Supreme Court opinions. Previous work using topic models to examine the Congressional record can provide a starting place: building on the notions of divergence developed in this Article, it may be possible to span institutional settings to examine the interplay between Congress and the Supreme Court at a textual level.¹⁹⁹

Future studies could also refine or alter the approach to classification of the judicial genre that we introduce in this Article. In Section I.C we explain some of the advantages of the relatively naïve approach that we use here. But methods that combine elements of top-down theory, without the imposition of overly constrictive models, may be useful complements to both the theory-driven (such as that used by Farganis) and the naïve (deployed here) approaches.²⁰⁰ For example, it may be possible to separate language that is associated with substantive categories (e.g., contracts, torts, and sentencing) from language referencing to sources of authority (e.g., the Constitution, statutes, and precedent) and from language deploying different analytic techniques (e.g., legislative history, rules v. standards, values-laden). Computational techniques that may show promise for such a hybrid approach would include a combination of hand-coding and machine-learning classification, advances in topic modeling, such as the structured topic model that infers topics from metadata associated with documents, and topic models that incorporate parsing

199. See Quinn et al., *supra* note 17. For an example of an effort to compare a political measure across executive and legislative institutions, see Joshua D. Clinton et al., *Separated Powers in the United States: The Ideology of Agencies, Presidents and Congress*, 56 AM. J. POL. SCI. 341 (2012).

200. See Farganis, *supra* note 70.

information on the role that words play in a sentence (rather than relying simply on word frequencies).²⁰¹

No doubt in the coming years advances in natural language processing will create additional opportunities to study legal texts. New horizons include the application of advanced artificial-intelligence techniques (such as deep neural networks) to the tasks of summarization and text generation. The implications for law scholarship from these new techniques could be quite substantial, as interdisciplinary teams of legal and computer science experts apply advanced machine-reading techniques to process and interpret a potentially vast realm of legal and law-related texts. These experts have the ability to draw conclusions not only about the development of the law and legal institutions, but also the broader interaction of law and society.²⁰² But, as this Article has demonstrated, researchers need not wait for new advances before useful work can be done.

CONCLUSION

In this Article, we use topic models to formalize a notion of the judicial genre that is based on the summarized semantic content in a reference corpus of documents; in our study, opinions in the federal appellate courts. We then compare the semantic content of two sub-corpora—Supreme Court opinions and appellate court opinions associated with cases selected for review—to examine whether there are distinctive semantic characteristics that distinguish them from the broader judicial genre. We find that, indeed, the Court appears to use its certiorari power to select cases that are non-representative of the work of the appellate courts, and that the Court's opinions are even more distinctive than the opinions associated with cases selected for review. We then examine the topics that predict the likelihood of a case being granted certiorari, and find that some topics have remained predictive over the study period, while other topics have waxed and waned in their predictive power. Our findings suggest that the Court's attention is drawn to different subjects at different times, but that there are other legal questions that are of abiding interest. Future empirical legal scholarship on the certiorari process should consider whether including the semantic content of appellate court decisions (either expressed via a topic model or not) could be a useful mechanism to control for unobserved case characteristics. Overall, we interpret our finding as consistent with the claim that the Court's opinions, to some extent, make up a distinctive genre of legal texts, and that the cases selected for review are a kind of intermediary corpus standing between the Court and the other appellate court opinions.

We also examine the evolution of the distinctiveness of the Court over time, both quantitatively and qualitatively. We find that the Court's opinions become

201. See Margaret Roberts et al., *Structural Topic Models for Open-Ended Responses*, 58 AM. J. POL. SCI. 1064 (2014) (deploying structured topic model that incorporates information on political affiliation); Jordan Boyd-Graber & David Blei, *Syntactic Topic Models*, in ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 185 (D. Koller et al. eds., 2008).

202. See, e.g., Michael A. Livermore, Vladimir Eidelman & Brian Grom, *Computationally Assisted Regulatory Participation*, 93 NOTRE DAME L. REV. (forthcoming 2018) (examining relationship between agency ideology and the public comments they receive).

measurably more distinctive over the course of our study period; we also find that the opinions associated with cases selected for review are not becoming more distinctive. We interpret these findings to suggest that the Court is selecting a similarly non-representative pool of cases to review, but is analyzing and writing about them in an increasingly idiosyncratic fashion. We undertake a qualitative analysis of these findings by comparing the most- and least-characteristic cases, as measured by the likelihood that our classifier would misidentify them as cases selected for review rather than Supreme Court opinions. We argue that our quantitative approach appears to pick up on differences between opinions that track with intuitive notions of what distinguishes a Supreme Court like opinion from a run-of-the-mill appellate court opinion. Based on this qualitative review, we suggest directions for future research into the evolution of the judicial genre and the application of computational text-analysis tools to legal documents.

As discussed in Part I, the Supreme Court's exercise of the judicial *power* through the judicial *genre* may play an important legitimating role for an institution that lacks democratic accountability. The Court's writings appear, over the second half of the twentieth century, to have gradually drifted away from their roots. It is not obvious that this development is a normative problem, but it is not obvious that this is *not* a normative problem either. The place of the Supreme Court and its relationship to the courts it supervises and the political branches is an endless source of controversy. It would require further analysis—empirical, political, and philosophical—to fully assess the normative weight of the results reported above. Certainly, it would be helpful to have an explicit, normative conception of the Supreme Court's judicial role and how that role relates to the way in which the Justices analyze and discuss the cases before it. Additional empirical estimates that provide a finer-grained analysis of how the Supreme Court differs would also be useful: there is a wide range of sources of distinctiveness in texts, from relatively meaningless semantic quirks to an entirely specialized approach to analyzing the law. Although we are relatively confident that the use of the topic-model approach and our qualitative analysis avoid the problem of sham differences, further analysis is needed to fully capture what this trend of distinctiveness means for the role of the Court in the American judicial system.

In addition to these substantive findings, this Article helps demonstrate how topic models can be used to open up new potential avenues of empirical legal studies. In the past, quantitative analysis of law has traditionally been hampered by the lack of attractive mechanisms for estimating case characteristics or the legal features of opinions.²⁰³ Topic modeling provides a promising avenue through which to estimate legal variables because it is oriented toward the semantic content of cases, which is exactly where the law would be expected to show up. It also avoids some of the pitfalls of human readers, including error, bias, and, most important, time and attentional limits. By naïvely characterizing the semantic content of cases based on a very general set of assumptions about the structure of the corpus, topic

203. See generally Tom S. Clark & Benjamin Lauderdale, *Locating Supreme Court Opinions in Doctrine Space*, 54 AM. J. POL. SCI. 871 (2010) (using citation to develop more nuanced ideological variables for decisions); Barry Friedman, *Taking Law Seriously*, 4 PERSP. ON POL. 261 (2006); Emerson Tiller & Frank B. Cross, *What is Legal Doctrine?*, 100 NW. U. L. REV. 517 (2006).

models provide a quantitative and computationally tractable method to represent the text of the law. The corpus of the law—the published case law in the state and federal reporters, and other legal texts as well—is an enormous and rich dataset, and topic models provide an effective means of capturing important characteristics of that data that can be subjected to analysis. Topic models, and computational text analysis more generally, represent an exciting new frontier in empirical legal scholarship that has substantial potential to improve understanding of the law.

APPENDIX

Appendix Table: Topic List and Top Words

This table lists the ten most probable words for each topic in the 100-topic model of Supreme and Appellate Court opinions. Extremely common words (sometimes referred to as stop words) are not displayed. Categorizations are based on the authors' subjective judgments, for illustrative purposes. Not all topics were easily identifiable with standard legal categories.

Topic	Label	Characteristic Words
0	Labor	union board labor employees employer nlrp company bargaining relations national local
1	Family 1	ms mrs did told husband time testified asked sexual home
2	Elections	election political party candidates candidate campaign parties primary elections contributions
3	Narcotics	united drug cocaine government cir defendant conspiracy evidence drugs marijuana
4	Immunity	immunity officers officer official police law county qualified officials city
5	Prison conditions	prison inmates inmate prisoner prisoners officials confinement conditions security jail
6	Procedure 1	motion district judgment appeal rule order filed summary party appeals
7	Medical	dr medical hospital mental treatment health care patient drug expert
8	Crim pro 1	trial defendant plea guilty indictment united jeopardy criminal double prosecution
9	Insurance	insurance policy company insured coverage insurer life ins policies liability
10	Standard of proof	evidence record fact finding findings burden standard proof support facts
11	Crim law	offense crime criminal conviction united felony convicted defendant intent firearm
12	6th 7th cir.	ohio illinois ill michigan chicago state oklahoma indiana tennessee kentucky

Topic	Label	Characteristic Words
13	Procedure 2	district order injunction relief decree preliminary orders injunctive pending courts
14	Tax	tax income commissioner taxpayer revenue internal irs taxes year code
15	Precedent 1	cir district id did issue circuit does omitted review claim
16	Automobile	car vehicle truck driver automobile ford stop motor vehicles highway
17	Evidence	evidence testimony trial defendant witness rule witnesses cross defense united
18	Precedent 2	ct ed supra cases denied held united cert id cf
19	Sentencing	sentencing sentence district defendant guidelines offense united level guideline departure
20	Federal powers	power constitution congress law authority judicial president constitutional united powers
21	Congress	act congress cong legislative rep sess history committee section stat
22	Local government	city ordinance housing public property park county area building council
23	Antitrust	antitrust market price competition act sherman trade prices business united
24	Confessions	confession police miranda right defendant statements rights
25	Indian/Native American affairs	indian tribe lands united water reservation indians tribal land river
26	Briefs	respondent petitioner appeals id petitioners respondents ante opinion ibid supra
27	Death penalty	death penalty capital sentencing murder sentence state mitigating evidence circumstances
28	Due process	process state amendment right constitutional protection clause rights fourteenth equal

Topic	Label	Characteristic Words
29	Family 2	child children abortion parents state family juvenile parent minor statute
30	Disability	disability alj claimant benefits secretary work evidence disabled medical ada
31	Discrimination	discrimination racial race white black discriminatory equal women minority district
32	Environment	epa environmental water agency permit waste air project corps site
33	Aviation	appellants air convention aircraft airport flight airlines international united faa
34	Trusts and estates	estate trust death wife decedent spouse husband property divorce law
35	Statutory interpretation 1	statute language congress interpretation statutory meaning does provision term section
36	Military	military army war navy service united force security defense national
37	Banks	bank loan credit banks national account funds loans deposit savings
38	Statutory interpretation 2	sec section shall act provides person provisions cir pursuant subsection
39	Mining	land property lease coal use oil taking compensation mining owner
40	State federal	state federal law california cal statute laws supreme code courts
41	Precedent 1	cases does majority supra fact course particular make need example
42	Fees	fees attorney fee award attorneys costs district litigation party prevailing
43	Securities	securities exchange fraud rule act stock investment sec sale purchase

Topic	Label	Characteristic Words
44	Communications	commission fcc cable communications television service public access broadcast broadcasting
45	Valuation	cost value costs rate total market percent method year increase
46	Search warrant	search warrant fourth united officers amendment arrest probable police cause
47	Surveillance	king telephone gambling united conversations order electronic surveillance wiretap
48	Juries	jury trial verdict defendant instruction error judge instructions evidence jurors
49	Self-incrimination	grand privilege criminal united jury contempt witness amendment fifth prosecution
50	Jurisdiction	jurisdiction federal action district claims courts claim suit forum jurisdictional
51	Negligence	negligence plaintiff liability defendant duty accident injury cause negligent law
52	Sentencing	sentence parole release probation imprisonment term imposed years time sentences
53	Administrative law boards	board review hearing administrative decision judicial agency proceedings action process
54	Deadlines	time period notice date limitations filed days statute year day
55	[Unidentified]	said question shall opinion mr ed cases supra think page
56	Employment discrimination	discrimination employment title vii plaintiff eeoc employer employee age position
57	Administrative law regulation	regulations secretary agency regulation authority safety act standard standards requirements
58	Police testimony	police evidence testified car identification house robbery officer time man

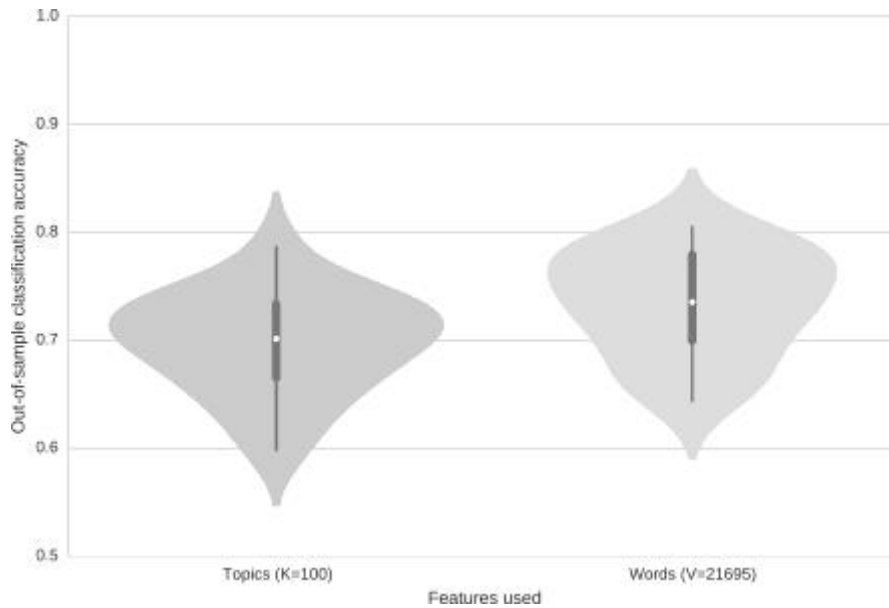
Topic	Label	Characteristic Words
59	Appeals	appellant united district cir circuit judge appellee appeals atty appeal
60	Mails	service board postal mail united local veterans letter classification selective
61	Patent/invention	patent claim claims invention infringement art patents prior use said fed
62	Corporations	corporation stock company business corporate corporations shares partnership control assets
63	Procedure filings	defendants plaintiffs plaintiff complaint claim district claims defendant appellees cir
64	Press	public press publication news newspaper times amendment statements article media
65	Tax 2	tax state taxes exemption taxation exempt property income business clause
66	Gas/power/oil	commission gas power rate order natural ferc public company rates
67	Habeas	petitioner state habeas petition federal corpus claim writ relief conviction
68	Employ	employees employee work employment employer compensation pay workers act time
69	ERISA	plan benefits erisa fund pension benefit retirement plans employee trust
70	Conspiracy	united conspiracy count government counts indictment defendant fraud cir evidence
71	Ineffective assist	counsel trial attorney defendant right defense assistance hearing judge lawyer representation
72	Rail	commission carrier railroad carriers transportation icc service rates interstate commerce
73	Maritime	vessel maritime ship admiralty act cargo vessels jones marine seaman

Topic	Label	Characteristic Words
74	Sales	sales sale products goods business sold product price trade food
75	8th cir.	district missouri johnson kansas iowa st arkansas mo miller minnesota
76	Speech	speech amendment public government free content expression regulation protected commercial
77	Federal powers 2	commerce state interstate pre clause congress regulation federal act power
78	5th cir.	texas florida louisiana tex georgia virginia state alabama brown la
79	Obscene material	sexual davis obscene material obscenity adult conduct young statute film
80	Testimony	did time smith stated fact told received judge testified record
81	Labor arbitration	arbitration agreement union bargaining collective contract labor dispute parties arbitrator
82	Damages	damages award liability law punitive injury claim plaintiff recovery tort
83	Political organizations	members committee association member organization party activities membership bar communist
84	Property	property lien sale transfer forfeiture title judgment real law right
85	Immigration	united foreign alien immigration deportation aliens ins country bia attorney
86	Political/voting	district voting county districts state vote plan population political election
87	1st 2d cir.	new york city jersey judge state massachusetts mass law second
88	Bankruptcy	bankruptcy debtor trustee creditors claim creditor chapter debt code claims

Topic	Label	Characteristic Words
89	Entitlements	program services benefits funds secretary assistance federal programs title social
90	Testimony 2	mr judge did said know yes don record asked going
91	Contracts	contract agreement parties contracts party terms breach agreements contractor agreed
92	FOIA/disclosure	information disclosure documents records report agency public investigation discovery
93	Copyright	copyright use mark trademark trade work infringement confusion product works
94	Precedent	rule law majority decision supreme opinion cases judge application new
95	Education	school religious schools education board students district public student religion
96	Federal claims	government united federal act claims claim private district public general
97	Class action	class standing action injury members plaintiffs settlement certification claims interests
98	Sovereign immunity	action federal private rights suit civil act immunity congress right
99	Supreme Court	opinion cite id dissenting tion ing united ante stevens supra

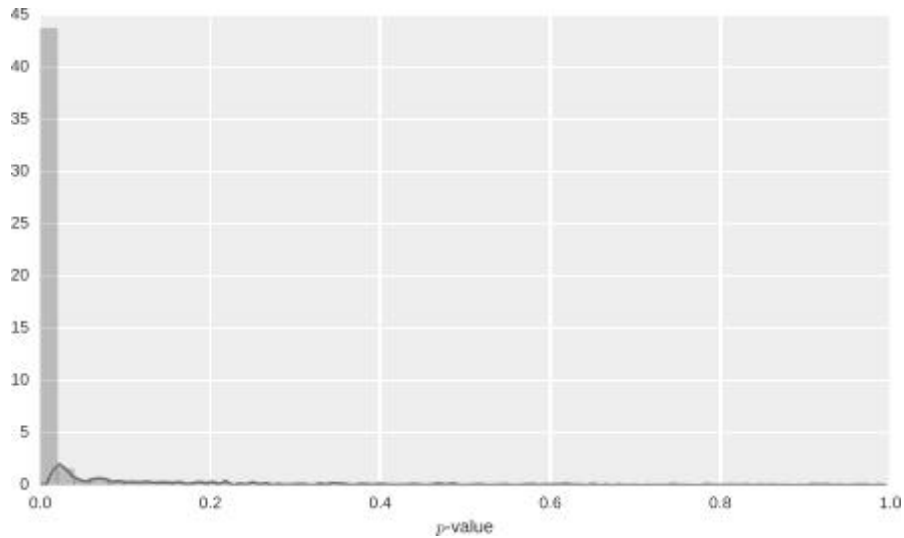
Appendix Figure 1: Prediction of Supreme Court Database Issue Tags

Distribution of out-of-sample accuracy in classifying Supreme Court opinions according to Spaeth issue labels, using, alternatively, topic proportions and word frequencies.



Appendix Figure 2: P-values for Matched Supreme Court–Appellate Court KL Divergences

The topic model can distinguish an appellate court opinion associated with a case under review from a randomly selected appellate court opinion, when compared to the matching Supreme Court opinion.



Appendix Figure 3: Topics with a Statistically Significant Relationship to Certiorari

For a substantial number of topics, the proportion of words in opinion documents that were associated with those topics are predictive of whether the associated case was granted certiorari. Several topics also interact with time, so that their influence either increases or decreases over the study period.

