

# Judicial Ghostwriting: Authorship on the Supreme Court

Jeffrey S. Rosenthal

Albert H. Yoon

Follow this and additional works at: <http://scholarship.law.cornell.edu/clr>

 Part of the [Law Commons](#)

---

### Recommended Citation

Jeffrey S. Rosenthal and Albert H. Yoon, *Judicial Ghostwriting: Authorship on the Supreme Court*, 96 Cornell L. Rev. 1307 (2011)  
Available at: <http://scholarship.law.cornell.edu/clr/vol96/iss6/11>

This Article is brought to you for free and open access by the Journals at Scholarship@Cornell Law: A Digital Repository. It has been accepted for inclusion in Cornell Law Review by an authorized administrator of Scholarship@Cornell Law: A Digital Repository. For more information, please contact [jmp8@cornell.edu](mailto:jmp8@cornell.edu).

# JUDICIAL GHOSTWRITING: AUTHORSHIP ON THE SUPREME COURT

Jeffrey S. Rosenthal & Albert H. Yoon†

*Supreme Court Justices, unlike the President or members of Congress, perform their work with relatively little staffing. Each Justice processes the docket, hears cases, and writes opinions with the assistance of only their law clerks. The relationship between Justices and their clerks is of intense interest to legal scholars and the public, but it remains largely unknown. This Article analyzes the text of the Justices' opinions to better understand judicial authorship. Based on the use of common function words, we find that Justices vary in writing style, from which it is possible to accurately distinguish one from another. Their writing styles also inform how clerks influence the opinion-writing process. Current Justices, with few exceptions, exhibit significantly higher variability in their writing than their predecessors, both within and across years. These results strongly suggest that Justices are increasingly relying on their clerks to write opinions.*

INTRODUCTION . . . . .	1307
I. METHODOLOGY . . . . .	1313
II. DATA . . . . .	1317
III. RESULTS . . . . .	1318
A. Variability Scores . . . . .	1318
B. Tests of Significance . . . . .	1327
C. Authorship Identification . . . . .	1332
IV. DISCUSSION . . . . .	1337
CONCLUSION . . . . .	1341
APPENDIX . . . . .	1342

## INTRODUCTION

*The reason the public thinks so much of the Justices of the Supreme Court is that they are almost the only people in Washington who do their own work.*

—Justice Louis D. Brandeis<sup>1</sup>

---

† Professor, University of Toronto, Department of Statistics, and Professor, University of Toronto Faculty of Law, respectively. We would like to thank Ed Cheng, John Goldberg, Pamela Karlan, Helen Levy, David Madigan, Todd Peppers, Richard Posner, Michael Trebilcock, Fred Tung, Artemus Ward, and Robert Weisberg. Professor Yoon would also like to thank the Russell Sage Foundation for its general financial support of this research. All remaining errors are our own.

<sup>1</sup> CHARLES E. WYZANSKI, JR., *WHEREAS—A JUDGE'S PREMISES: ESSAYS IN JUDGMENT, ETHICS, AND THE LAW* 61 (Greenwood Press 1976) (1965) (internal quotation marks omitted).

Imagine a job where each year one is required to evaluate over seven thousand files, closely evaluate approximately sixty to eighty cases, and write seven to ten lengthy published documents, all of which will become established law and be scrutinized by countless judges, lawyers, academics, and law students. Add one additional requirement: do this job well past retirement age into your seventies and eighties. Incredible as it sounds, the above describes the job of a Supreme Court Justice.

In contrast to the other branches of the federal government,<sup>2</sup> the Court is a small and closed environment. It consists of only nine Justices, who are collectively responsible for deciding which cases to hear and ultimately deciding those cases in the form of written opinions. Unlike the President or Congress, the Court operates with lean staffing. The Justices are each responsible for their own writing; their only assistance comes from their law clerks, for whom each Justice is currently allowed four.<sup>3</sup> Within each Justice's chambers, law clerks assist in evaluating certiorari (cert) petitions, preparing the Justices for oral argument, and ultimately producing written opinions. The law clerks are recent law graduates, typically among the top of their classes from a handful of elite law schools.<sup>4</sup>

Given the demands of the job, it is understandable that Justices seek to ease their work burdens. They have already taken some steps. For example, in 1972 the Court created a cert pool.<sup>5</sup> Whereas previously each Justice evaluated the entire docket of cases on their own, participating Justices in the cert pool now review petitions collectively, meaning that they share work across chambers. This division of labor dramatically reduces the number of petitions each chamber has to process. Moreover, in the cert process, it is widely accepted that the

---

<sup>2</sup> For example, each member of the House of Representatives is allowed eighteen permanent employees and up to four additional shared or part-time employees. *See* IDA A. BRUDNICK, CONG. RESEARCH SERV., RL 30064, CONGRESSIONAL SALARIES AND ALLOWANCES 3 (2009).

<sup>3</sup> *See* TODD C. PEPPERS, COURTIERIERS OF THE MARBLE PALACE: THE RISE AND INFLUENCE OF THE SUPREME COURT LAW CLERK 195 (2006).

<sup>4</sup> *See id.* at 30 ("Supreme Court [J]ustices pick from not only the best law schools but also the best students from these schools.").

<sup>5</sup> *See* ARTEMUS WARD & DAVID L. WEIDEN, SORCERERS' APPRENTICES: 100 YEARS OF LAW CLERKS AT THE UNITED STATES SUPREME COURT 117 (2006) (describing the creation of the cert pool during the Burger Court in response to the growing caseload). Most Justices joined the cert pool with a few notable exceptions, such as Justice Brennan and Justice Stevens. *See id.* at 121, 125. Most recently, Justice Alito has opted out of the cert pool. *See* Debra Cassens Weiss, *Alito Opts Out of the Supreme Court Cert Pool*, ABAJOURNAL.COM (Sept. 29, 2008, 8:42 AM), [http://www.abajournal.com/news/article/alito\\_opts\\_out\\_of\\_the\\_supreme\\_court\\_cert\\_pool/](http://www.abajournal.com/news/article/alito_opts_out_of_the_supreme_court_cert_pool/).

law clerks now are largely responsible for providing the initial assessment of the cases.<sup>6</sup>

But anecdotes abound that Justices have also increasingly delegated the responsibility of writing opinions to their clerks. One former clerk of Justice John Paul Stevens estimated that clerks generate “well over half” of the text in published opinions.<sup>7</sup> A 2006 historical account of Supreme Court clerks stated, “one can safely conclude that no other set of sitting Supreme Court Justices have delegated as much responsibility to their law clerks as those on the Rehnquist Court.”<sup>8</sup>

The purported degree to which Justices rely on clerks varies considerably. Justice Oliver Wendell Holmes wrote his opinions in long-hand, relegating his clerks to primarily nonlegal tasks akin to those of an administrative assistant.<sup>9</sup> Justice William Douglas maintained he wrote his own opinions.<sup>10</sup> Justice Stevens is reputed to draft his own opinions, as is Justice Antonin Scalia.<sup>11</sup> Other Justices, such as Justice Thurgood Marshall, were known to rely more on their clerks.<sup>12</sup> He was not alone; Chief Justice William Rehnquist confirmed that his clerks did “the first draft of almost all cases,” and in some instances the published decision was “relatively unchanged” from the draft.<sup>13</sup> Justice Harry Blackmun, in the words of one historian, “ceded to his law clerks much greater control over his official work than did any of the other 15 Justices from the last half-century whose papers are publicly available.”<sup>14</sup>

Should we care if Justices delegate the opinion-writing process to their clerks? The answer may depend on the degree to which it occurs. The import of an opinion, particularly from the Supreme Court,

---

<sup>6</sup> See WARD & WEIDEN, *supra* note 5, at 136, 142 (describing how from 1970 to 1972, the average number of cert petitions each clerk reviewed dropped from 634 to 257).

<sup>7</sup> Sean Donahue, *Behind the Pillars of Justice: Remarks on Law Clerks*, 3 LONG TERM VIEW 77, 81 (1995).

<sup>8</sup> See PEPPERS, *supra* note 3 at 191.

<sup>9</sup> See *id.* at 58–59 (describing Justice Holmes’ relationship with his law clerks).

<sup>10</sup> See WARD & WEIDEN, *supra* note 5, at 205 (citing Transcriptions of Conversations Between Justice William O. Douglas and Professor Walter F. Murphy, Cassette No. 3, (Dec. 20, 27 1961), [http://www.princeton.edu/~mudd/finding\\_aids/douglas/douglas3.html](http://www.princeton.edu/~mudd/finding_aids/douglas/douglas3.html)).

<sup>11</sup> See EDWARD LAZARUS, *CLOSED CHAMBERS: THE RISE, FALL, AND FUTURE OF THE MODERN SUPREME COURT* 271 (Penguin Books 2005) (1998) (“[O]nly Justices Stevens and Scalia made it a regular practice to participate in first drafts.”).

<sup>12</sup> See Paul J. Wahlbeck et al., *Ghostwriters on the Court?: A Stylistic Analysis of U.S. Supreme Court Opinion Drafts*, 30 AM. POL. RES. 166, 172 (2002), available at <http://apr.sagepub.com/content/30/2/166>; see also Mark Tushnet, *Thurgood Marshall and the Brethren*, 80 GEO. L.J. 2109, 2112 (1992) [hereinafter Tushnet, *Brethren*] (“Marshall relied more heavily on his law clerks for opinion writing than did the other Justices during the early years of his tenure, but his practices were not wildly out of line with those of the others on the Court.”).

<sup>13</sup> See BERNARD SCHWARTZ, *THE ASCENT OF PRAGMATISM: THE BURGER COURT IN ACTION* 38 (1990) (internal quotation marks omitted).

<sup>14</sup> David J. Garrow, *The Brains Behind Blackmun*, LEGAL AFF., May/June 2005 at 28.

stems less from the identity of the prevailing party than from the reasoning that accompanies the decision.<sup>15</sup> Even those who defended delegation of work to law clerks, such as Chief Justice Rehnquist, cautioned that each “Justice must retain for himself control not merely of the outcome of the case, but of the explanation of the outcome.”<sup>16</sup> Delegation, if taken too far, can threaten the integrity of the Court.

Judicial authorship raises important questions about the relative roles of Justices and their clerks. The extent to which this principal-agent relationship<sup>17</sup> advances the Justice’s interests depends both on the clerk’s competence and preferences. Clerks, while typically excellent law students from elite law schools, are also usually recent graduates. This bimodal age composition on the Court—well-seasoned Justices and inexperienced law clerks—lacks a middle cohort with work experience in the government or the private sector. Justice Ruth Bader Ginsburg noted that while clerks “save us hours upon hours of labor[,] . . . most of them are also young and in need of the seasoning that experiences in life and in law practice afford.”<sup>18</sup> This staffing structure stands in stark contrast to those of the executive and legislative branches, which are both replete with experienced staff.<sup>19</sup>

The ideological preferences of the Justice and her clerks may also diverge. While liberal and conservative Justices generally hire like-minded clerks, the clerkship process, in which applicants apply to all nine Justices and are expected to accept the first offer, may bring clerks of different ideologies within the same chambers.<sup>20</sup> Thus, even when the Justice dictates the broad direction of an opinion, such as the prevailing party and the general reasoning, the clerk may still exercise considerable influence.

Any meaningful discussion of judicial delegation, however, first requires a deeper understanding of judicial authorship. Questions of authorship itself are nothing new, addressing works as old as those of

---

<sup>15</sup> Justice Oliver Wendell Holmes is attributed as saying, “I am always suspicious of an advocate who comes before the Supreme Court saying this is a court of justice; it is a court of law.” EUGENE W. HICKOK & GARY L. MCDOWELL, *JUSTICE VS. LAW: COURTS AND POLITICS IN AMERICAN SOCIETY* (1993) (unnumbered page in front matter).

<sup>16</sup> BERNARD SCHWARTZ, *DECISION: HOW THE SUPREME COURT DECIDES CASES* 52 (1996).

<sup>17</sup> For a general discussion of the principal-agent model, see Joseph E. Stiglitz, *Principal and Agent*, in 3 *THE NEW PALGRAVE: A DICTIONARY OF ECONOMICS* 966, 966–71 (John Eatwell et al. eds., 1987).

<sup>18</sup> Adam Liptak, *A Second Justice Opts Out of a Longtime Custom: The ‘Cert. Pool,’* N.Y. TIMES, Sept. 26, 2008, at A21, available at <http://www.nytimes.com/2008/09/26/washington/26memo.html?scp=1&sq=A%20Second%20Justice%20Opts%20Out&st=cse>.

<sup>19</sup> For example, as of April 2009, 42% of President Obama’s appointees to Senate-confirmed positions had served in the Clinton administration. See Peter Baker, *It’s Not About Bill*, N.Y. TIMES, May 31, 2009, § 6 (Magazine), at 47.

<sup>20</sup> See WARD & WEIDEN, *supra* note 5, at 107–08 (describing the allocation of clerks by ideology across the Justices’ chambers).

William Shakespeare.<sup>21</sup> Much of this earlier work is based on close textual,<sup>22</sup> but ultimately subjective, impressions of his writing. In the last fifty years, scholars have turned to statistical analysis. Perhaps most famously, Frederick Mosteller and David L. Wallace, in their seminal 1964 study of the unsigned Federalist Papers, concluded that James Madison, not Alexander Hamilton, was the likely author.<sup>23</sup>

Recent scholarship has examined questions of authorship in judicial opinions. A study comparing the first-draft opinions of Justice Lewis Powell and Justice Marshall found that Justice Marshall's clerks' writing styles were more identifiable than those of Justice Powell's clerks.<sup>24</sup> In a study of federal appellate judges, legal scholars using judges' citations to their own earlier opinions as a measure of their own writing found a modest relationship between the two.<sup>25</sup> While these studies advance our understanding of the judiciary, neither offer a convincing approach to evaluating judicial authorship. Comparing initial drafts may reflect more the collaboration *across* clerks than the Justices' reliance on them, while self-citations are a weak proxy for judges writing their own opinions if their own clerks are also more inclined to cite their judges' opinions.

Our approach differs from these earlier attempts because we explore judicial authorship based on a comprehensive evaluation of the jurists' writing style.<sup>26</sup> The central intuition here is that the more participants in the opinion-writing process, the more heterogeneous the writing style of the Justice's opinions. Justices who write their own

---

<sup>21</sup> See, e.g., JOHN MICHELL, *WHO WROTE SHAKESPEARE?* (1996) (providing a survey of arguments for alternative authors); S. SCHOENBAUM, *SHAKESPEARE'S LIVES* (1993) (same); James D.A. Boyle, *The Search for an Author: Shakespeare and the Framers*, 37 AM. U. L. REV. 625, 628–29 (1988) (noting debates on authorship of Shakespeare's works in connection with changing conceptions of authorship throughout history).

<sup>22</sup> See Thomas Regnier, Comment, *Could Shakespeare Think Like a Lawyer? How Inheritance Law Issues in Hamlet May Shed Light on the Authorship Question*, 57 U. MIAMI L. REV. 377, 378 (2003) (noting that Shakespeare's frequent use of legal terms has authorship implications).

<sup>23</sup> See FREDERICK MOSTELLER & DAVID L. WALLACE, *INFERENCE AND DISPUTED AUTHORSHIP: THE FEDERALIST 263* (1964). During this same period, scholars also used statistical analyses to examine literature. See generally ALVAR ELLEGARD, *A STATISTICAL METHOD FOR DETERMINING AUTHORSHIP: THE JUNIUS LETTERS, 1769–1772*, at 7 (1962); LOUIS TONKO MILIC, *A QUANTITATIVE APPROACH TO THE STYLE OF JONATHAN SWIFT* 16–19 (1967); A.Q. MORTON & JAMES MCLEMAN, *PAUL, THE MAN AND THE MYTH: A STUDY IN THE AUTHORSHIP OF GREEK PROSE* 13–14 (1966) (asserting, based on statistical analysis, that Paul did not author more than five of the fourteen letters attributed to him in the New Testament).

<sup>24</sup> See Wahlbeck et al., *supra* note 12, at 179, 182.

<sup>25</sup> See Stephen J. Choi & G. Mitu Gulati, *Which Judges Write Their Opinions (and Should We Care)?*, 32 FLA. ST. U. L. REV. 1077, 1111–16 (2005) (reasoning that judges who write their own opinions are more likely to cite their own opinions).

<sup>26</sup> Stephen J. Choi and G. Mitu Gulati briefly note that they attempt to analyze opinions through the judges' repetition of certain phrases but find the results inconsistent with their a priori knowledge of certain judges' reputation for writing their own opinions. See *id.* at 1107–10.

opinions would presumptively possess less variable writing styles than Justices who relied heavily on their law clerks.

The Court's institutional design also provides a compelling identification strategy for our hypothesis. Supreme Court clerkships are typically for a single term,<sup>27</sup> running from October through August. Justices who rely more on their clerks to write opinions would likely have a more variable writing style both within and across years than their less reliant colleagues. In addition, historical accounts of the Court suggest that the responsibilities of clerks have grown over time: from stenographer at their inception in the late nineteenth century to legal assistant in the 1920s to law firm associate beginning in the 1950s.<sup>28</sup> If true, we should expect increasingly variable writing styles over time.

In this Article, we analyze the text of majority opinions of all Supreme Court Justices. Using a parsimonious model based on the Justices' use of common function words (e.g., *the*, *also*, *her*), we construct a variability measure for writing style. In most instances, Justices have variability scores that are distinguishable from one another. Moreover, even Justices with comparable or identical variability scores exhibit distinct writing styles based on their different use of function words. Our analysis then allows us to accurately predict authorship in pairwise comparisons of Justices.

More importantly, our model also allows us to evaluate how Justices vary in their writing over their tenure on the Court and in comparison with other Justices. We find that current and recent Justices report significantly higher variability scores than their predecessors,<sup>29</sup> supporting the anecdotal evidence that Justices on average are relying more on their clerks in the opinion-writing process.<sup>30</sup> Moreover, some Justices—most notably Justice Sandra Day O'Connor and Justice Anthony Kennedy—produce variability scores that are not only higher<sup>31</sup> but also vary considerably from year to year,<sup>32</sup> suggesting an even greater reliance on clerks. We test the validity of our model by analyzing the opinions of two judges known to write their own opin-

---

<sup>27</sup> See WARD & WEIDEN, *supra* note 5, at 46 (describing the strong norm with few exceptions of law clerks spending only one year on the Court).

<sup>28</sup> See PEPPERS, *supra* note 3, at 45–46, 83–84, 145 (detailing the changing responsibilities of law clerks from 1886 to the present).

<sup>29</sup> See *infra* Table 2 and accompanying text.

<sup>30</sup> See, e.g., J. Daniel Mahoney, Foreword, *Law Clerks: For Better or For Worse?*, 54 BROOK. L. REV. 321, 339 (1988) (noting that “it is widely known that law clerks play a substantial role in opinion writing” and that the “clerks are often responsible for a judge’s first draft”).

<sup>31</sup> See *infra* Table 2.

<sup>32</sup> See *infra* Figure 2.

ions, Richard A. Posner and Frank Easterbrook.<sup>33</sup> The variability of Judge Posner's and Judge Easterbrook's writing were markedly lower and more consistent than any current or recent Justices on the Court.<sup>34</sup>

This Article proceeds as follows. Part I describes our statistical methodology: our construction of function words and the chi-squared approach to evaluate variability in writing style. In Part II, we briefly describe the data used in our analysis. In Part III, we report our results: the variability scores, both across and within Justices, and our ability to use function words to predict authorship. In Part IV, we discuss the implications of our results and how statistical textual analysis can advance future research of the Court and legal scholarship more generally.

## I METHODOLOGY

Two factors influence what words authors use in their writing. One is subject matter,<sup>35</sup> which can be specific to the author's topic in a particular work or more general to a substantive area.<sup>36</sup> The other is writing style: the sentence structure and patterns of word choice, which are commonly referred to, respectively, as *syntax* and *diction*.<sup>37</sup> Our focus is on writing style. The challenge is to tailor our analysis so that it discerns differences in writing style and not subject matter.

Our project falls within the broader discipline of stylometry, the statistical analysis of texts. Scholars have attempted to discern authorship in many notable works, including biblical documents,<sup>38</sup> Shakes-

---

<sup>33</sup> See Choi & Gulati, *supra* note 25, at 1080 n.6 (noting that information from judges and former law clerks "confirm[s] that Judges Posner and Easterbrook author all of their own opinions").

<sup>34</sup> See *infra* Table 4.

<sup>35</sup> See Choi & Gulati, *supra* note 25, at 1111.

<sup>36</sup> For example, while each of Paul Krugman's columns in the *New York Times* typically focuses on the economic consequences of a specific government policy, the column's general subject matter is economic policy. See Archive of Columns by Paul Krugman, N.Y. TIMES, <http://topics.nytimes.com/top/opinion/editorialsandoped/oped/columnists/paulkrugman/index.html> (last visited July 19, 2011).

<sup>37</sup> See Choi & Gulati, *supra* note 25, at 1100–01.

<sup>38</sup> See Kevin Burns, *Bayesian Inference in Disputed Authorship: A Case Study of Cognitive Errors and a New System for Decision Support*, 176 INFO. SCI. 1570, 1572–74 (2006).



pearean plays,<sup>39</sup> *The Federalist Papers*,<sup>40</sup> and President Ronald Reagan's radio addresses.<sup>41</sup>

Our analysis relies upon function words, which are common words—for example, *all, have, not, than*—whose usage frequencies are largely independent of subject matter.<sup>42</sup> In the words of one statistician, this approach is “topic-free in the sense that the relative frequency with which an author uses, for example, ‘with,’ should be the same regardless of whether the author is describing cooking recipes or the latest news about the oil futures market.”<sup>43</sup>

We considered other approaches, including such larger-scale features as sentence length, paragraph length, or the frequency of multiword phrases and such smaller-scale features as frequency of specific punctuation (e.g., comma, semicolon) or of particular letters. These approaches, however, did not meaningfully improve our results. Therefore, to situate our analysis within a unified methodology, we present results only from the use of function words.

In constructing our list of function words, we began by looking at Mosteller and Wallace's seminal study of *The Federalist Papers*, which had a list of seventy function words.<sup>44</sup> We adopted this list except for seven words—*every, my, shall, should, upon, will, you*—each of which appeared in fewer than 0.1% of Court majority opinions. Table 1 reports our list of function words.

TABLE 1:  
STYLOMETRY OF THE SUPREME COURT  
SIXTY-THREE FUNCTION WORDS

(1–12) *a, all, also, an, and, any, are, as, at, be, been, but,*  
 (13–24) *by, can, do, down, even, for, from, had, has, have, her, his,*  
 (25–36) *if, in, into, is, it, its, may, more, must, no, not, now,*  
 (37–48) *of, on, one, only, or, our, so, some, such, than, that, the,*  
 (49–60) *their, then, there, things, this, to, up, was, were, what, when, which,*  
 (61–63) *who, with, would*

<sup>39</sup> See Oleg Seletsky et al., *The Shakespeare Authorship Question 1* (Dec. 12, 2007) (unpublished manuscript) (on file with authors), available at <http://www.cs.dartmouth.edu/~datamining/Final.pdf>.

<sup>40</sup> See generally MOSTELLER & WALLACE, *supra* note 23 (applying statistical methods to authorship of the disputed Federalist papers).

<sup>41</sup> See Edoardo M. Airoidi et al., *Whose Ideas? Whose Words? Authorship of Ronald Reagan's Radio Addresses*, 40 POL. SCI. & POL. 501, 501–02 (2007); Edoardo M. Airoidi et al., *Who Wrote Ronald Reagan's Radio Addresses?*, 1 BAYESIAN ANALYSIS 289, 291–92 (2006).

<sup>42</sup> For a technical explanation of our approach, see Jeffrey S. Rosenthal & Albert H. Yoon, *Detecting Multiple Authorship of United States Supreme Court Legal Decisions Using Function Words*, 5 ANNALS APPLIED STAT. 283, 288–90 (2011), available at <http://projecteuclid.org/euclid.aoas/1300715191>.

<sup>43</sup> David Madigan et al., *Author Identification on the Large Scale 2–3* (2005) (unpublished manuscript), available at <http://www.stat.columbia.edu/~madigan/PAPERS/>.

<sup>44</sup> See MOSTELLER & WALLACE, *supra* note 23, at 38–39.

Our goal is to construct for each Justice a variability measure of opinions. As stated in the Introduction, the intuition behind this measure is that the greater the variability in writing style, the greater the likelihood that the law clerk has received at least part of the Justice's writing responsibilities.

Because we are counting the appearance of function words, we adopt a chi-squared approach. The chi-squared statistic allows us to test the distribution of the observed count of words against a theoretical or expected distribution. In our case, we will compare the observed count with the null hypothesis that the variability follows a chi-squared distribution; that is, the total count of each function word is equally likely to occur in any of the total number of opinions for each Justice.

The following paragraphs more formally describe our approach. The sixty-three function words are numbered from  $j = 1$  to  $j = 63$ . Suppose a given Justice has written opinions numbered from  $i = 1$  to  $i = K$ . Let  $w_i$  be the total number of words in judgment  $i$  and  $c_{ij}$  be the number of times that function word  $j$  appears in judgment  $i$ . We further define

$$e_{ij} = w_i \left( \frac{c_{1j} + c_{2j} + \dots + c_{Kj}}{w_1 + w_2 + \dots + w_K} \right),$$

where  $e_{ij}$  is the expected number of times that function word  $j$  would have appeared in judgment  $i$ . The null hypothesis is that the total number of  $c_{1j} + c_{2j} + \dots + c_{Kj}$  appearances of  $j$  were each equally likely to occur in any of the total number of  $w_1 + w_2 + \dots + w_K$  words in all of Justice  $K$ 's combined judgments.

The chi-squared statistic is the following:

$$chisq = \sum_{i=1}^K \sum_{j=0}^{63} \frac{(c_{ij} - e_{ij})^2}{e_{ij}}$$

Under the null hypothesis, *chisq* should follow a chi-squared distribution with  $(63 + 1 - 1)(K - 1) = 63(K - 1)$  degrees of freedom, and therefore mean  $63(K - 1)$ . This summation also includes  $c_{i0} = w_i - c_{i1} - \dots - c_{iK}$ , the number of words in judgment  $i$  that are *not* function words.<sup>45</sup>

Our variability measure is accordingly:

$$V \text{ score} = \frac{chisq}{df} = \frac{chisq}{63(K - 1)},$$

which should approximate one under the null hypothesis and exceed one for opinions that collectively exhibit greater variability in writing style.

---

<sup>45</sup> The "+1" arises from the  $c_{i0}$  terms.

There are, of course, other ways of constructing variability scores, many of which we also tried but ultimately decided against using. Because opinions vary in length, using raw counts of  $c_{ij}$  by themselves would not be informative. One alternative would be to evaluate the fraction of words in judgment  $i$  that use function word  $j$ , that is,  $f_{ij} = c_{ij}/w_i$ . Under this approach, a given Justice has a fixed unknown propensity  $p_j$  for using  $j$ , independently for each word of each opinion. Accordingly, the distribution of  $c_{ij}$  of  $j$  in  $i$  is binomial  $(w_i, p_j)$ , so that  $f_{ij}$  has mean  $p_j$  and variance  $p_j(1 - p_j)/w_i$ , which would depend on individual propensities  $p_j$  and  $w_i$ .<sup>46</sup> We decided against this analysis because it is dependent on  $w$  and  $p$ , meaning the variability score will be biased downward for Justices who write shorter opinions. Similarly, while it is possible to modify the measure such that it is independent of  $w$  and  $p$ ,<sup>47</sup> it would still be based on an imperfect estimate if the propensities  $p_j$ .<sup>48</sup>

It is worth noting that the chi-squared values may be less stable and therefore less meaningful when many of the expected cell counts for given words approximate zero. We control for this in part by excluding function words that have a very low frequency in majority opinions, as well as opinions shorter than 250 words. A small fraction of the cell counts—both expected and observed—have a count less than one. It is possible to correct for this through, for example, the use of a Yates correction, but it is typically used for two-by-two tables. As a check, we recomputed our variability (V) score omitting all cells with a very small expected cell count; the scores reduced slightly in a consistent format, but the bootstrap tests of significance below were unchanged. Accordingly, we do not change our definition of V score.

For these reasons, we ultimately chose the chi-squared approach, which serves as the foundation for our analysis.<sup>49</sup> The first stage determines whether the Justices' variability scores are distinguishable from our null hypothesis. Establishing that they are, we then use a boot-

<sup>46</sup> The variability measure would then be the sum of sample standard deviations, that is  $V = \sum_{j=1}^{63} sd(f_{1j}, f_{2j}, f_{3j}, \dots, f_{Kj})$ , where larger standard deviations reflect a more variable writing style.

<sup>47</sup> Because  $f_{ij}$  has variance  $p_j \frac{(1 - p_j)}{w_i}$ , this is also true for  $f_{ij} - \mu_j$ , where  $\mu_j = \frac{c_{1j} + c_{2j} + \dots + c_{Kj}}{w_1 + w_2 + \dots + w_K}$  is our best estimate of  $p_j$ . The quantity  $r_{ij} = w_i^{1/2}(f_{ij} - \mu_j)$  has mean zero and variance  $p_j(1 - p_j)$ , resulting in  $V = \sum_{j=1}^{63} sd(r_{1j}, r_{2j}, r_{3j}, \dots, r_{Kj})$

<sup>48</sup> We can eliminate formal dependence on both  $w$  and  $p$  by creating  $V = \sum_{j=1}^{63} sd(q_{1j}, q_{2j}, q_{3j}, \dots, q_{Kj})$ , where  $q_{ij} = \frac{w_i^{1/2}(f_{ij} - \mu_j)}{(\mu_j(1 - \mu_j))^{1/2}}$ . The term  $\mu$  generates uncertainty where  $p$  is close to zero (in the most extreme case, where  $\mu_j = 0$ ,  $q_{ij}$  is undefined).

<sup>49</sup> For a discussion of the chi-squared approach, see Rosenthal & Yoon, *supra* note 42, at 8–9.

strap approach to determine whether their variability scores are statistically distinguishable from one another. Third, we construct a linear classifier to test whether our model is able to accurately predict authorship, allowing us to construct variability scores for each Justice, over their entire tenure as well as particular periods. It also allows us to directly test variability in pairwise comparisons of Justices using bootstrap simulations. Finally, we test that accuracy of our model in predicting authorship through *leave-one-out cross-validation*. We describe these latter analytic approaches in detail in Part IV.

## II

### DATA

Our data consists of the written opinions from the Supreme Court, which we obtained from Justia.com (Justia), a website that provides free access to federal government documents. Specifically, Justia includes all published Supreme Court opinions from 1791 through the current term.<sup>50</sup>

In order to statistically analyze the opinions, we wrote software in C and Unix<sup>51</sup> that downloaded the decisions directly from Justia. Among other things, our program converts the HTML pages into plain text and culls from the opinion(s) any text not written by a Justice, such as headnotes, synopses, and other notes. In the process, the program separates majority from concurring and dissenting opinions, a surprisingly challenging task attributable to the changing conventions by which the Court demarcates various types of opinions.

For both methodological and substantive reasons we focus on majority opinions. Although more common in recent years, dissents were historically relatively rare until 1941.<sup>52</sup> Concurrences then as well as now are less common than dissents.<sup>53</sup> Compared with majority opinions, dissents and concurrences are also typically shorter and often create significant instability in the textual analysis. In addition, we exclude any unsigned opinions, which typically arose in the form

---

<sup>50</sup> See *US Supreme Court Cases & Opinions*, JUSTIA.COM, <http://supreme.justia.com/us/year/> (last visited Apr. 4, 2011).

<sup>51</sup> On Justia, for most volumes, the majority opinion and any dissenting or concurring opinions for each case are contained within a single HTML file. Our software program identifies the beginning and end of each majority opinion. We have made the software for downloading and analyzing these texts available. For a full description of the software program, see Jeffrey S. Rosenthal, *Explanation of the Software*, PROBABILITY.CA, <http://probability.ca/usscj/README> (last visited July 24, 2011).

<sup>52</sup> Prior to 1941, approximately 10% of decisions included a dissent. Subsequently, in most Court terms, over 50% of decisions include a dissent. See LEE EPSTEIN ET AL., *THE SUPREME COURT COMPENDIUM: DATA, DECISIONS, AND DEVELOPMENTS* 211–15 tbl.3-2 (3d ed. 2003).

<sup>53</sup> See *id.* at 216–20 tbl.3-3.

of per curiam opinions, court orders, decrees, or motions. We also exclude any opinions shorter than 250 words.

We analyzed opinions spanning the entire history of the Court, which includes 111 Justices.<sup>54</sup> Given this sizable list, we report only the most recent forty Justices in our analysis of individual Justices based on their years of departure from the Court.

### III RESULTS

In this Part we describe the V scores for each Justice. V scores allow us to determine the distinctiveness of the Justices' writing style generally and as compared to one another. First, we establish whether these scores are statistically distinguishable from the null hypothesis and then from one another. Afterwards, we test our model's ability to accurately predict authorship in pairwise comparisons of Justices.

#### A. Variability Scores

Table 2 lists the forty most recent Justices that left the Court after 1940 and provides brief demographic information and the V scores for each one. The V scores listed represent each Justice's overall variability score based on all of his or her majority opinions.<sup>55</sup> For reasons described earlier in Part II, we exclude dissenting and concurring opinions as well as any majority opinion of length less than 250 words.

The interpretation of the V score is as follows: the higher a Justice's V score, the more variable her writing style, based on the use of function words in Table 1. A higher V score is consistent with the view that the Justice relies more on her law clerks in writing the opinions. It bears repeating that our analysis cannot prove this result and is open to competing explanations. For example, even in a world where every Justice writes her own opinions, one would expect that some Justices simply have more variable writing styles than others. This heterogeneity is also likely true in a world where every Justice relies on law clerks to assist in writing opinions. The variability in V scores may also reflect differences across law clerks—namely, their ability to mimic their Justices' writing style.<sup>56</sup>

What is the purpose of a V score? The first is to test our construction of the null hypothesis that that the Justices writing style follows a

---

<sup>54</sup> Justice Kagan, the 112th Justice in the Supreme Court's history, is not included in the analysis.

<sup>55</sup> We provide a list of Justices departing prior to 1941 in Table A1 of the Appendix.

<sup>56</sup> Robert O'Neil, one of Justice Brennan's clerks early in his tenure, commented that Justice Brennan allowed him and O'Neil's co-clerk Richard Posner to draft opinions throughout their clerkship because of their ability to "'mimic' Brennan's style or voice." PEPPERS, *supra* note 3, at 158.

uniform and random distribution of function words. To test this, we randomly generate two hundred pseudodocuments each consisting of 2000 independently and randomly generated words. Each word was chosen to be a nonfunction word with a probability of 70%, and it was uniformly selected from the list of function words in Table 1 with a probability of 30%. The null is a V score that approximates one. Repeating this experiment ten times, we produced a mean V score of 1.004622 with a standard deviation of 0.001702, which is consistent with our null hypothesis having a true mean equal to one.

TABLE 2:  
VARIABILITY SCORE  
CURRENT AND DEPARTED JUSTICES SINCE 1941

Number	Justice	Appointing President	Year Joined Court	Year Left Court	Age at Retirement	Years on Court	Opinions	Average Word Length	V Score
1	Sonia Sotomayor	Obama	2009			1	8	4167	3.12
2	Samuel A. Alito	Bush, G. W.	2006			4	33	4522	3.38
3	John G. Roberts	Bush, G. W.	2005			5	32	4404	3.12
4	Stephen G. Breyer	Clinton	1994			16	125	3750	3.06
5	Ruth Bader Ginsburg	Clinton	1993			17	138	4472	3.57
6	Clarence Thomas	Bush, G. H. W.	1991			19	149	3987	3.70
7	Anthony M. Kennedy	Reagan	1988			22	193	5061	3.73
8	Antonin Scalia	Reagan	1986			24	221	4317	3.08
9	John Paul Stevens	Ford	1975	2010	90	35	369	4231	3.30
10	David H. Souter	Bush, G. H. W.	1990	2009	70	19	149	5302	3.65
11	Sandra Day O'Connor	Reagan	1981	2006	76	25	272	4878	3.85
12	William H. Rehnquist	Reagan	1972	2005	81	33	478	3458	3.01
13	Harry A. Blackmun	Nixon	1970	1994	86	24	306	4353	3.70
14	Byron Raymond White	Kennedy	1962	1993	76	31	463	4833	3.71
15	Thurgood Marshall	Johnson, L.	1967	1991	83	24	323	3520	3.33
16	William J. Brennan	Eisenhower	1956	1990	84	34	449	3594	3.23
17	Lewis F. Powell	Nixon	1972	1987	80	15	252	3776	3.22
18	Warren Earl Burger	Nixon	1969	1986	79	17	252	3708	3.11
19	Potter Stewart	Eisenhower	1958	1981	66	23	315	3361	3.32
20	William Orville Douglas	Roosevelt, F.	1939	1975	77	36	551	2069	2.53
21	John Marshall Harlan	Eisenhower	1955	1971	72	16	169	3517	2.92
22	Hugo Lafayette Black	Roosevelt, F.	1937	1971	85	34	492	2014	2.51
23	Abe Fortas	Johnson, L.	1965	1969	59	4	41	3266	2.68
24	Earl Warren	Eisenhower	1953	1969	78	16	165	3220	3.18
25	Torn Campbell Clark	Truman	1949	1967	68	18	221	2582	2.73
26	Arthur Joseph Goldberg	Kennedy	1962	1965	57	3	37	3682	3.15
27	Charles Evans Whittaker	Eisenhower	1957	1962	61	5	42	2605	3.19
28	Felix Frankfurter	Roosevelt, F.	1939	1962	80	23	266	2858	2.67
29	Harold Hitz Burton	Truman	1945	1958	70	13	96	3028	3.69
30	Stanley Forman Reed	Roosevelt, F.	1938	1957	73	19	236	3229	2.80
31	Sherman Minton	Truman	1949	1956	66	7	65	1587	2.31
32	Robert Houghwout Jackson	Roosevelt, F.	1941	1954	62	13	148	2880	2.54
33	Fred Moore Vinson	Truman	1946	1953	63	7	65	4511	3.42
34	Wiley Blount Rutledge	Roosevelt, F.	1943	1949	55	6	65	4511	3.42
35	Frank Murphy	Roosevelt, F.	1940	1949	59	9	132	2200	2.72
36	Owen Josephus Roberts	Hoover	1930	1945	70	15	292	2315	2.74
37	James Francis Byrnes	Roosevelt, F.	1941	1942	63	1	16	2387	2.55
38	Charles Evans Hughes	Hoover	1930	1941	79	11	391	2842	3.08
39	Harlan Fiske Stone	Coolidge	1925	1941	69	16	447	2518	2.95
40	James Clark McReynolds	Wilson	1914	1941	79	27	480	1273	2.11

Note: Justices are arranged in reverse chronological order, first by year of departure, then by year of appointment. Justices who departed the Court prior to 1941 are listed in the Appendix, Table A1.

The V scores range from a low of 2.11 for Justice James McReynolds to a high of 3.85 for Justice O'Connor. For each Justice, the V score is much larger than it would be under the null hypothesis that

the function words are truly distributed uniformly and randomly. For example, Justice Stephen Breyer has a V score of 3.06, which reflects a *chisq* statistic of  $3.06 \times 63 \times (125 - 1) = 23,904.72$ . The null hypothesis has a chi-squared distribution of  $63 \times 124 = 7812$ . Justice Breyer's V score corresponds to a *p* value less than 0.0000001, which allows us to reject the null hypothesis that Breyer's writing based on the function words follows a uniform and random distribution. We can reject the null hypothesis for all the Justices in Table 2, as well as all Justices throughout the history of the Court for which we could produce a V score.<sup>57</sup>

The results also support anecdotal accounts of Justices' approaches to opinion writing. For example, Justice Douglas has a V score of 2.53, among the lowest of the Justices in the modern, post-1950 era. This score is consistent with scholarly accounts<sup>58</sup> as well as Justice Douglas's claim that he wrote his own opinions.<sup>59</sup> Similarly, Justice Scalia has a reputation for writing his own opinions;<sup>60</sup> his V score of 3.08, while higher than Douglas, is among the lowest among the current Justices. Justice Holmes had the lowest V score of all Justices: 1.78.<sup>61</sup> This comports with his reputation as a Justice who used his clerks sparingly and primarily for minor tasks such as checking citations.<sup>62</sup> Similarly, Justice Benjamin Cardozo, who was particularly renowned for his writing,<sup>63</sup> had a V score of 2.32.<sup>64</sup>

Several Justices' high V scores similarly support anecdotal and historical accounts of their reliance on clerks. Justice Blackmun's V score of 3.70 supports some historians' views that he delegated much of the opinion writing to his clerks and more so than his contemporaries.<sup>65</sup> Interestingly, the Justices with the highest V scores were Justices O'Connor (3.85) and Kennedy (3.73). Though journalists and scholars have written at length about these Justices' influence

---

<sup>57</sup> Because several Justices on the Court prior to 1800 did not write any opinions in which they were attributed authorship, we could not calculate a V score for them. See *infra* Table A1 (noting that the only identified opinion writer in the Court's early years was the Chief Justice).

<sup>58</sup> See WARD & WEIDEN, *supra* note 5, at 205 ("Douglas wrote most of his own opinions and in general kept his clerks at arm's length.").

<sup>59</sup> See *id.* ("I [Justice Douglas] have written all my own opinions." (quoting Transcriptions of Conversations between Justice William O. Douglas and Professor Walter F. Murphy, Cassette No. 3, (Dec. 20, 27 1961), [http://www.princeton.edu/~mudd/finding\\_aids/douglas/douglas3.html](http://www.princeton.edu/~mudd/finding_aids/douglas/douglas3.html))).

<sup>60</sup> See LAZARUS, *supra* note 11, at 271 (distinguishing Justice Scalia and Justice Stevens from the other Justices in the opinion-writing process for their participation in preparing first drafts).

<sup>61</sup> See *infra* Table A1.

<sup>62</sup> See WARD & WEIDEN, *supra* note 5, at 35.

<sup>63</sup> See Richard A. Posner, *Judges' Writing Styles (and Do They Matter?)*, 62 U. CHI. L. REV. 1421, 1432 (1995) (deeming Cardozo "one of the finest judicial writers in our history").

<sup>64</sup> See *infra* Table A1.

<sup>65</sup> See Garrow, *supra* note 14, at 28.

on the Court as pivotal or “swing” voters,<sup>66</sup> they have largely ignored the relationships of these Justices with their clerks. Indeed, whether the higher V scores of these Justices are attributable to their ideological position or their personal approach to opinion writing is beyond the scope of this Article. It is worth noting, however, that Justice Byron White, another noted swing Justice,<sup>67</sup> had a V score of 3.71, the third highest of all the Justices.

Some V scores were surprising. Justice Thurgood Marshall’s score of 3.33 runs counter to accounts that he largely delegated opinion writing to his clerks.<sup>68</sup> They were also comparable to his contemporaries on the Court, most notably Justice William Brennan (3.23). Justice Ginsburg’s score of 3.57 appears unremarkable, falling roughly in the middle of her contemporaries. But it is markedly higher than Justice Breyer (3.06) or Justice Scalia (3.08), which suggests she relies more on her clerks in writing opinions than do some of her colleagues who were also formerly law professors.<sup>69</sup>

We can aggregate the Justices’ individual V scores to show trends of the Court over time, as reported in Figure 1 and Table 3.

---

<sup>66</sup> See, e.g., LAZARUS, *supra* note 11, at 209 (describing Justice O’Connor’s influence as the “crucial swing vote” in *Tison v. Arizona*, 481 U.S. 137 (1987)); Douglas M. Parker, *Justice Kennedy: The Swing Voter and His Critics*, 11 GREEN BAG 2D 317, 317–18 (2008) (describing Justice Kennedy as the successor to Justice O’Connor as the swing voter on the Court); see also Lynn A. Baker, *Interdisciplinary Due Diligence: The Case for Common Sense in the Search for the Swing Justice*, 70 S. CAL. L. REV. 187, 199 n.68, 202–03 (1996) (describing the influence of Justices Kennedy and O’Connor as swing voters).

<sup>67</sup> B. Drummond Ayres Jr., *The ‘Swing’ Justice: Byron Raymond White*, N.Y. TIMES, June 30, 1972, at 16.

<sup>68</sup> See *supra* text accompanying note 12.

<sup>69</sup> See Lee Epstein et al., *The Norm of Prior Judicial Experience and Its Consequences for Career Diversity on the U.S. Supreme Court*, 91 CALIF. L. REV. 903, 935 (2003) (noting that in 2003 the Court consisted of “three former law professors—Justices Breyer, Scalia, and Ginsburg—who toiled for a total of forty-two years at Harvard (Breyer), Chicago (Scalia), and Rutgers and Columbia (Ginsburg)”).



FIGURE 1:  
VARIABILITY SCORE BY DECADE  
(1990–2000)

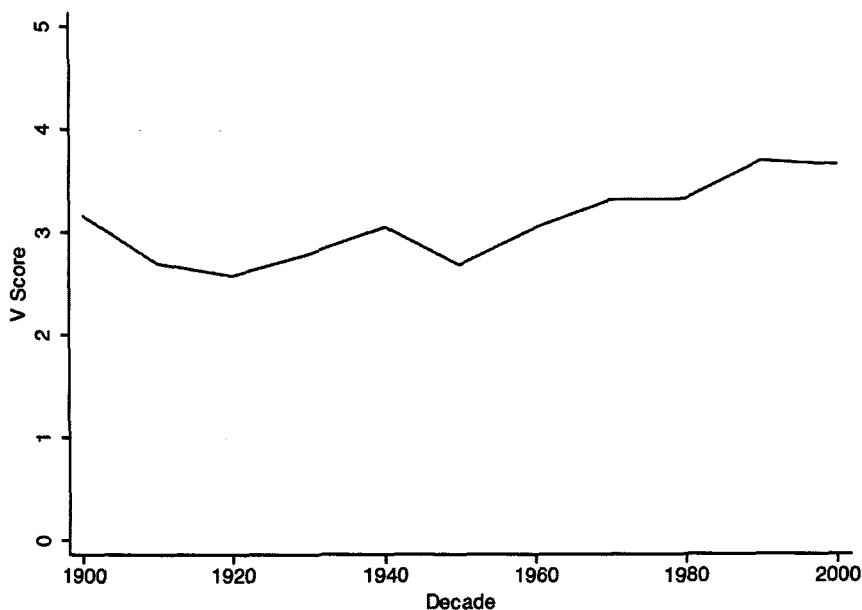


TABLE 3:  
VARIABILITY SCORE  
BY CHIEF JUSTICE (1910–2009)

Number	Chief Justice	Start Yr	End Yr	Start Vol	End Vol	Total Years	Opinions	Average Word Length	V Score
1	Roberts	2005	2009	546	559	5	294	4261	3.59
2	Rehnquist	1986	2005	478	545	19	1685	4384	3.56
3	Burger	1969	1986	395	478	17	2337	3924	3.38
4	Warren	1953	1969	346	395	16	1550	2766	2.92
5	Vinson	1946	1953	329	346	7	756	2713	3.02
6	Stone	1940	1945	314	328	5	719	2730	2.97
7	Hughes	1930	1941	280	313	11	1818	2203	2.81
8	Taft	1921	1930	257	280	9	1602	1851	2.59
9	White	1910	1921	218	256	11	2474	2051	2.66

A couple of patterns emerge. The V scores are lower and generally stable for the period from 1900 to 1950 and steadily increase beginning in 1950. This comports with historical accounts that in the first half of the twentieth century law clerks served a largely administrative role,<sup>70</sup> but that beginning in the 1950s they began to take a more active role in the substantive matters of the Court, including the

<sup>70</sup> See PEPPERS, *supra* note 3, at 83–84, 145 (describing law clerks as stenographers from the 1880s until 1919 and as legal assistants from the 1920s until the 1940s).

drafting of opinions.<sup>71</sup> Not surprisingly, opinion length nearly doubled over this period.<sup>72</sup>

While informative, Justices' aggregate V scores provide an incomplete account of their writing styles. An aggregate score may reflect consistency over time, or it may mask an upward or downward trend or even considerable changes from one year to the next. We examine the Justices' V scores on a year-by-year basis, which allow us to observe time trends.

We report a sample of Justices in Figure 2. Justice Holmes, with an aggregate V score of 1.78, showed remarkable consistency from year to year, with a standard deviation of only 0.22. Justices Cardozo, Douglas, Rehnquist and Scalia also had relatively low variability in V scores, all with standard deviations in the neighborhood of 0.40. By contrast, Justice O'Connor and Justice Kennedy showed greater variability in V scores across years with standard deviations of approximately 0.71 and 0.62, respectively. Justice Stevens and Justice Brennan, with standard deviations of 0.56 and 0.48, respectively, fall between these groups. Their V scores exhibit a different pattern. Justice Stevens's scores have three distinct periods: in his first dozen years (1976–88), Stevens had an average V score of 2.70 with a standard deviation of 0.29; from 1989 to 2000, he had an average V score of 3.29 with a standard deviation of 0.66. From 2001 onward, Justice Stevens's V scores resembled his early years on the Court—averaging 2.79 with a standard deviation of 0.52. Justice Brennan, by contrast, shows an upward trend in V scores by decade with declining standard deviations.

Figure 2 provides additional insight into the writing relationships that Justices have with their law clerks. Because law clerks typically serve for only one year,<sup>73</sup> the annual V score arguably provides a measure of a Justice's reliance on them. Assuming that clerks vary in writing style from one another, it logically follows that Justices who rely more on law clerks would experience greater changes from year to year than those who are known to do write their own opinions. Justices Holmes, Cardozo, and Douglas were all reputed to do their own writing<sup>74</sup>: the year-by-year V scores support this belief. Conversely, by this measure it appears that Justice O'Connor and Justice Kennedy

---

<sup>71</sup> See *id.* at 145 (“[T]he 1950s and 1960s witnessed the transformation of the law clerk into an attorney involved in all aspects of chamber work.”).

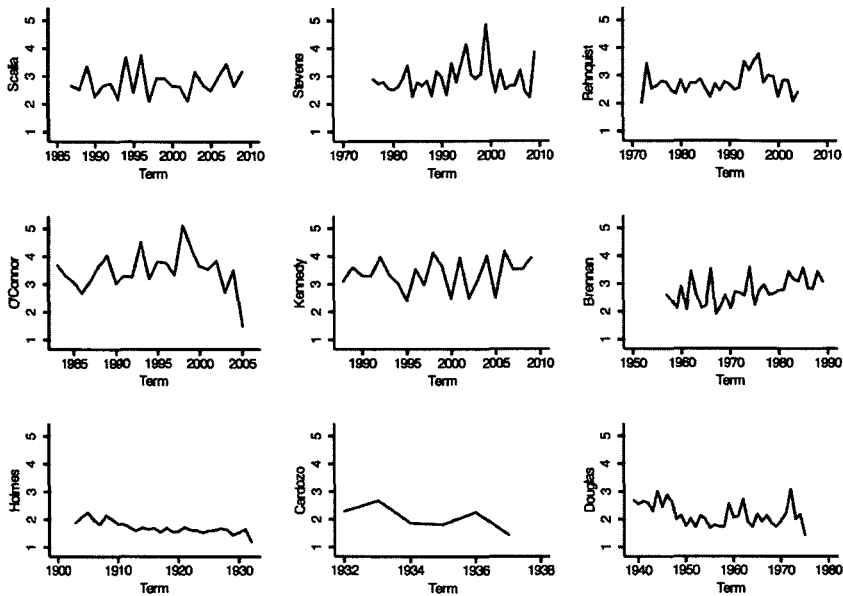
<sup>72</sup> See Ryan C. Black & James F. Spriggs II, *An Empirical Analysis of the Length of U.S. Supreme Court Opinions*, 45 HOUS. L. REV. 621, 634–35, 639–42 (2008) (concluding that although clerks may have contributed to an increase in the length of majority opinions, “they do not appear to have been the driving force behind this change”).

<sup>73</sup> See *supra* note 27.

<sup>74</sup> See PEPPERS, *supra* note 3, at 58, 96–97, 114 (describing the clerkship responsibilities for Justices Holmes, Cardozo, and Douglas).

rely more heavily on their clerks. Their year-to-year scores have the highest standard deviation among the Justices included in Figure 2.

FIGURE 2:  
YEAR-TO-YEAR V SCORES  
SELECT JUSTICES



We recognize that the validity of V scores as a measure of Justices' reliance on their law clerks is ultimately unverifiable. Justices' relationships with their clerks remains shrouded in secrecy,<sup>75</sup> so we cannot definitively separate the Justices' own writing styles from those of their clerks. One recent study has attempted to compare draft opinions to detect the stylistic influence of individual clerks, finding differences in author-identifiability across the Justices based on their respective drafting procedures.<sup>76</sup> Given that only a few Justices' papers are in the public domain, we cannot evaluate this more systematically. More importantly, comparing versions of opinions does not answer questions regarding the authorship process generally.

<sup>75</sup> See, e.g., Paul M. Barrett, *If There Is Blood in an Opinion, We Know Who Wrote It*, WALL ST. J., Oct. 4, 1993, at A8 ("Owing to court secrecy, the public hears little of the clerks . . ."); cf. PEPPERS, *supra* note 3, at 18–20 (discussing the difficulty in obtaining information about clerks due to confidentiality concerns).

<sup>76</sup> See Wahlbeck et al., *supra* note 12, at 174–83 (analyzing variations in Justice Marshall's and Justice Powell's draft opinions during the 1985 term and concluding that Justice Powell's multieditor drafting procedure "had the measurable effect of blurring the fingerprints that individual clerks left on these opinions").

We check the validity of the V scores by comparing them with two jurists known to write their own opinions: Judge Richard A. Posner and Judge Frank Easterbrook, both of the U.S. Court of Appeals for the Seventh Circuit. Judge Posner has described his writing process as one where he writes his own draft opinions and asks his clerks to “make criticisms and do research and try to tie up the loose ends that [he] discover[s] in [his] writing.”<sup>77</sup> He then incorporates their feedback into his revision.<sup>78</sup> Judge Easterbrook has a similar relationship with his clerks, differing in one respect: he writes all of his opinions but may allow each of his clerks to draft one opinion during the term.<sup>79</sup> Together, these judges provide a reasonable gauge by which to determine whether the V score is a reasonable measure of writing variability.

Given that Judge Posner and Judge Easterbrook write their own opinions, one would expect their V scores to be relatively low and stable from one year to the next on the theory that their writing styles are not highly dependent on their law clerks. The summary statistics reported in Table 4 substantiate this view. Both Judge Posner and Judge Easterbrook have V scores that are lower than any of the contemporary Justices on the Court and lower than all but a handful of the Justices in Table 2. It is notable that while their average word length is shorter than most Justices, they have both written many more opinions than any Justice while on the Supreme Court,<sup>80</sup> lending support to our belief that the V scores provide a valid measure of writing variability.

TABLE 4:  
V SCORE  
JUDGE POSNER AND JUDGE EASTERBROOK

Years	Judge	Opinions	Average Word Length	V Score
1981–present	Richard A. Posner	2372	2714	2.60
1984–present	Frank Easterbrook	1651	2400	2.42

Figure 3 shows that the V scores for both Judge Posner and Judge Easterbrook have remained stable during their tenure on the bench, particularly when contrasted with current and recent Justices in Fig-

<sup>77</sup> See Richard Posner, *Diary: Entry 2*, SLATE (Jan. 15, 2002, 11:47 AM), <http://www.slate.com/id/2060621/entry/2060742/>.

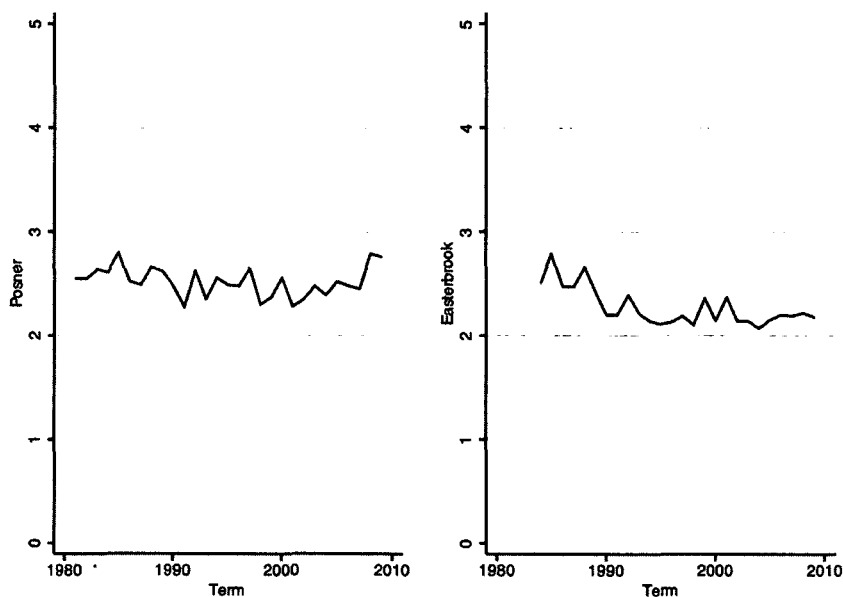
<sup>78</sup> See *id.*

<sup>79</sup> See Choi & Gulati, *supra* note 25, at 1080, 1108. The authors and Judge Easterbrook are in the early stages of a project to discern whether it is possible to identify through textual analysis which opinions Judge Easterbrook allowed his law clerks to draft.

<sup>80</sup> See Tracey E. George & Chris Guthrie, *Remaking the United States Supreme Court in the Courts' of Appeals Image*, 58 DUKE L.J. 1439, 1440 (2009) (quoting sources noting the contraction of the Supreme Court's plenary docket alongside the growth of those of the federal courts of appeals).

ure 2. For Judge Posner, the standard deviation in his V scores across years was 0.14; for Judge Easterbrook, the standard deviation was 0.18. Their standard deviation scores were lower than any of the aforementioned contemporary Justices, which range from 0.41 for Justice Rehnquist to 0.71 for Justice O'Connor. Judge Posner's and Judge Easterbrook's V scores empirically support their reputations for writing their own opinions and strongly suggest by contrast that their contemporaries on the Court rely more on their clerks.

FIGURE 3:  
YEAR-TO-YEAR V SCORES  
JUDGE POSNER AND JUDGE EASTERBROOK



Additionally, we examined a sample of the Justices' dissents. Dissents are typically shorter than majority opinions, although they are still lengthy, averaging over 3000 words. While some Justices write relatively few dissents (e.g., swing Justices), other Justices often write them more often than majority opinions (e.g., ideological Justices of the minority coalition). Looking at dissents of Justices Stevens, Kennedy, and Scalia, we found that their V scores were systematically smaller when compared with those of their majority opinions.<sup>81</sup> This contrast suggests that Justices delegate less in writing dissenting opin-

<sup>81</sup> See Rosenthal & Yoon, *supra* note 42, at 9 (comparing the majority opinions with the dissents for Justices Stevens, Kennedy, and Scalia).

ions than majority opinions, although an intuitive explanation for this escapes us.

### B. Tests of Significance

While the Justices' V scores allow us to convincingly reject the null hypothesis that their writing style follows a uniform and random distribution of function words, they do not establish whether the scores themselves are meaningfully different from one another. There is no straightforward analytic test; because we reject the null hypothesis, the Justices' V scores by definition do not follow a chi-squared distribution. Nor can we analytically determine what type of distribution it is.

It is possible, however, to determine the distribution through a *bootstrap* test. A bootstrap is a procedure of repeated sampling with replacement from a given sample.<sup>82</sup> The intuition behind bootstrapping is that it replaces the unknown distribution with the empirically determined distribution.

With our data, we select, for each Justice, one hundred cases of authored majority opinions uniformly at random, with repetition.<sup>83</sup> For each sample of one hundred cases, we compute the V score in the same manner. We then repeat this process 1000 times for each Justice, which generates 1000 different possible V scores, depending on which one hundred cases we draw.

We can use these values to do *pairwise* comparisons of Justices. This process creates 1,000,000 (one thousand by one thousand) pairs of V scores. We then simply count the fraction of pairs in which the V score for Justice *A* is greater than for Justice *B*, which gives us an estimate of the probability that the V score for Justice *A* is greater than that for Justice *B* for a random selection of judgments. These pairings allow us to estimate the distribution function for the difference of the V score for Justice *A* minus Justice *B*, which we can use to compute a 95% confidence interval for this difference. Accordingly, an entirely positive or negative confidence interval indicates that the differences in V scores between Justice *A* and Justice *B* are statistically meaningful and not due to chance.

To illustrate, Table 5 provides two examples of bootstrap comparisons. For each, the bootstrap V scores for each of the Justices closely approximate their V scores in Table 3. In the first comparison, Justice O'Connor has a larger bootstrap V score (3.81) than Chief Justice

---

<sup>82</sup> See generally BRADLEY EFRON & ROBERT J. TIBSHIRANI, AN INTRODUCTION TO THE BOOTSTRAP (CRC Press 1998) (1993) (describing bootstrapping and providing applications).

<sup>83</sup> Sampling could occur *without* replacement, but the relatively low number of opinions by some Justices would cause them to drop from our analysis under this approach.

Rehnquist (2.99); the probability that her V score produces a lower score than his is less than 0.0000, a statistically significant difference. In the second example, Justice Stevens produces a lower V score (3.26) than Justice Thomas (3.63), which occurs 0.9241 of the time. This difference, however, is not statistically significant.

TABLE 5:  
SAMPLE PAIRWISE BOOTSTRAP COMPARISONS

<b>Bootstrap V score (O'Connor)</b>	<b>Bootstrap V score (Rehnquist)</b>	<b><i>P</i>(O'Connor &lt; Rehnquist)</b>	<b>95% C.I.: (Rehnquist-Ginsburg)</b>
3.81	2.99	0.0000	(-1.283145, -0.385480)

<b>Bootstrap V score (Stevens)</b>	<b>Bootstrap V score (Thomas)</b>	<b><i>P</i>(Stevens &lt; Thomas)</b>	<b>95% C.I.: (Stevens-Thomas)</b>
3.26	3.63	0.9241	(-0.151552, 0.888814)

Table 6 produces a pairwise comparison of all Justices during the last Rehnquist natural court, a period from 1994 to 2005. The column labeled "V score" provides the V score from Table 2, which in each instance closely approximates the bootstrap V scores.<sup>84</sup> For the remainder of the Table, the cell numbers report the probability that the row Justice has a lower V score than the column Justice. For example, the probability of Justice Ginsburg having a lower V score than Justice Breyer is 0.0109. Across the diagonal, the V scores are mirror images. Accordingly, the probability of Justice Breyer having a lower V score than Justice Ginsburg is 0.9891. A probability greater than 0.95 or less than 0.05 reflects a statistically significant difference in the bootstrap V scores.

Table 6 reveals that in a large fraction of pairings, 44%, the Justices are statistically distinguishable from one another. The larger the difference in V scores, the more likely the bootstrap produces a statistically distinguishable score. Some Justices, such as Justice Kennedy and Justice Thomas, have V scores that are statistically indistinguishable from each other based on our selection of function words. Because their V scores are identical, it is not surprising that the probabilities are very close to 0.50. Conversely, Justice Breyer, with the lowest V score of this cohort, produces bootstrap estimates that are statistically distinguishable from every other Justice during this period, except for Justice Scalia and Chief Justice Rehnquist. Collectively, sixteen out of thirty-six of all the pairwise bootstrap comparisons were statistically significant. As one might expect, Justices with V scores near the median of the Court were less distinguish-

<sup>84</sup> We choose to report the original V score because each bootstrap produces a slightly different V score, all of which are close to the V scores reported in Table 3.

TABLE 6:  
V SCORE BOOTSTRAP—PAIRWISE COMPARISON  
JUSTICES ON LAST REHNQUIST NATURAL COURT (1994–2005)

Row Justice	Column Justice									
	V Score	Stephen G. Breyer	Ruth Bader Ginsburg	Clarence Thomas	Anthony M. Kennedy	Antonin Scalia	John Paul Stevens	David H. Souter	Sandra Day O'Connor	William H. Rehnquist
Stephen G. Breyer	3.06		0.9891	0.9932	0.9995	0.6121	0.8708	0.9868	0.9998	0.4932
Ruth Bader Ginsburg	3.57	0.0109		0.6607	0.7934	0.0095	0.1503	0.6087	0.8803	0.0089
Clarence Thomas	3.70	0.0068	0.3393		0.5824	0.0060	0.0777	0.4702	0.7598	0.0017
Anthony M. Kennedy	3.73	0.0005	0.2066	0.4176		0.0010	0.0361	0.3344	0.6961	0.0003
Antonin Scalia	3.08	0.3879	0.9905	0.9940	0.9990		0.8529	0.9919	0.9999	0.3712
John Paul Stevens	3.30	0.3879	0.8497	0.9223	0.9639	0.1471		0.8934	0.9870	0.1052
David H. Souter	3.65	0.0132	0.8497	0.5298	0.6656	0.0081	0.1066		0.7758	0.0038
Sandra Day O'Connor	3.85	0.0002	0.1197	0.5298	0.3039	0.0001	0.0130	0.2242		0.0000
William H. Rehnquist	3.01	0.5068	0.9911	0.9983	0.3039	0.6288	0.8948	0.9962	1.0000	

Note: V Score column represents V Score for each Justice. Remaining cells represent  $P(\text{Row Justice}) < (\text{Column Justice})$ .

able than those Justices with relatively high or low V scores. Even when the scores were not statistically significant, in 72% of the pairings, the probability was either greater than 0.70 or less than 0.30.

In addition to comparing across Justices, the bootstrap test also allows us to compare V scores *within* Justices. This analysis allows us to examine how an individual Justice may have changed over time. Political scientists have commented how Justices undergo ideological shifts during their tenure on the Court.<sup>85</sup> Although our variability measure—relying on common function words—is nonideological, changes in ideology may coincide with changes in reliance on law clerks.<sup>86</sup>

While there are numerous ways to evaluate change over time, we first examine Justices' writing before and after their sixty-fifth birthday. Sixty-five is an important benchmark for multiple reasons: it rep-

<sup>85</sup> See, e.g., Lee Epstein et al., *Ideological Drift Among Supreme Court Justices: Who, When, and How Important?*, 101 Nw. U. L. Rev. 1483, 1485–87 (2007).

<sup>86</sup> In 1957, Chief Justice Rehnquist famously argued, shortly after he clerked on the Supreme Court, that clerks—and predominantly those from the political Left—influence how Justices make decisions. William H. Rehnquist, *Who Writes Decisions of the Supreme Court?*, U.S. NEWS & WORLD REP., Dec. 13, 1957, at 74, 74–75.



TABLE 7:  
WITHIN-JUSTICE COMPARISON—PRE- AND POST-65  
CURRENT AND DEPARTED JUSTICES SINCE 1941

Justice	Before Age 65			After Age 65			Bootstrapping		
	Opinions	Avg Word Length	V4	Opinions	Avg Word Length	V4	P(Pre65< Post65)	Lower 95 Confidence Interval	Upper 95 Confidence Interval
Sonia Sotomayor									
Samuel A. Alito									
John G. Roberts									
Stephen G. Breyer	68	3807	3.15	53	3646	2.90	0.1451	-0.63	0.18
Ruth Bader Ginsburg	40	4921	3.58	90	4290	3.49	0.44010	-0.45	0.40
Clarence Thomas	147	3967	3.69						
Anthony M. Kennedy	126	4752	3.56	59	5771	3.96	0.9458	-0.07	0.79
Antonin Scalia	143	4473	2.99	71	4003	3.09	0.6767	-0.24	0.38
John Paul Stevens	122	3767	3.02	234	4531	3.33	0.8871	-0.19	0.80
David H. Souter	108	5686	3.75	35	4287	3.11	0.0027	-1.19	-0.17
Sandra Day O'Connor	187	4824	3.61	79	5017	4.33	0.9943	0.14	1.20
William H. Rehnquist	312	3290	2.80	158	3839	3.16	0.9779	0.01	0.68
Harry A. Blackmun	23	4570	4.02	270	4285	3.52	0.1220	-0.92	0.25
Byron Raymond White	273	4930	3.69	176	4462	3.46	0.1899	268.00	236.00
Thurgood Marshall	61	2788	3.07	253	3674	3.28	0.8374	-0.21	0.67
William J. Brennan	174	3071	2.81	268	3970	3.33	1.0000	0.78	1.47
Lewis F. Powell				236	3780	3.23			
Warren Earl Burger	24	5179	3.14	215	3522	3.05	0.1367	-0.70	0.19
Potter Stewart	289	3327	3.27	15	3847	3.22	0.2123	-0.78	0.23
William Orville Douglas	372	2206	2.54	171	1764	2.36	0.2433	-0.56	0.26
John Marshall Harlan	31	2865	2.54	72	4039	3.08	0.9978	0.14	0.96
Hugo Lafayette Black	262	1860	2.48	225	2213	2.46	0.5777	-0.45	0.46
Abe Fortas	41	3266	2.68						
Earl Warren	25	1682	2.39	135	3562	3.29	1.0000	0.54	1.41
Tom Campbell Clark	108	2069	2.48	36	3373	3.11	0.9967	0.16	1.11
Arthur Joseph Goldberg	37	3682	3.15						
Charles Evans Whitaker	42	2605	3.19						
Felix Frankfurter	108	2440	2.48	152	3165	2.77	0.9283	-0.09	0.66
Harold Hitz Burton	51	3461	4.07	38	2653	3.36	0.0768	-1.46	0.24
Stanley Forman Reed	158	3240	2.77	65	3191	2.66	0.3085	-0.43	0.26
Sherman Minton	55	1618	2.36	9	1463	1.96	0.0000	-0.75	-0.35
Robert Houghwout Jackson	148	2880	2.54						
Fred Moore Vinson	75	3331	2.96						
Wiley Blount Rutledge	65	4511	3.42						
Frank Murphy	132	2200	2.72						
Owen Josephus Roberts	198	2148	2.62	85	2736	2.98	0.9284	-0.11	0.78
James Francis Byrnes	16	2387	2.55						
Charles Evans Hughes	140	2843	2.93	251	2841	3.04	0.6548	-0.36	0.55
Harlan Fiske Stone	249	1999	2.63	185	3204	3.30	0.9966	0.18	1.17
James Clark McReynolds	273	1233	2.10	194	1309	2.11	0.5354	-0.34	0.31

Note: We do not analyze Justices who have been on the Court fewer than five years.

resents both the typical retirement age among most workers<sup>87</sup> and the age at which many Justices vest in their judicial pensions.<sup>88</sup> Moreover, some scholars have argued that Justices too often remain on the Court long past their productive years to the point of “mental decrepitude.”<sup>89</sup> A dramatic change in V score before and after age sixty-five

<sup>87</sup> See 42 U.S.C. § 416(l) (2006) (establishing a “normal retirement age” of sixty-five to determine eligibility for social security benefits).

<sup>88</sup> The actual age at which Justices and all other Article III judges vest in their pensions is determined by statute. See 28 U.S.C. § 371 (2006). Under the current requirements, referred to as the *Rule of 80*, the Justice must be at least sixty-five years of age and have served at least ten years, and the age at retirement and years of service combined must be at least eighty. See Albert H. Yoon, *Pensions, Politics, and Judicial Tenure: An Empirical Study of Federal Judges, 1869–2002*, 8 AM. L. & ECON. REV. 143, 147 (2006).

<sup>89</sup> See David J. Garrow, *Mental Decrepitude on the U.S. Supreme Court: The Historical Case for a 28th Amendment*, 67 U. CHI. L. REV. 995, 995 (2000).

may suggest that older Justices are delegating more of their writing to their clerks.

Table 7 reveals no dominant trend comparing Justices pre- and post-sixty-five. A majority of Justices have a higher V score *after* age sixty-five than before: Justice O'Connor, for example, had a pre-sixty-five V score of 3.61 and a post-sixty-five V score of 4.33. Justices Kennedy, Scalia, Stevens, Rehnquist, Marshall, Brennan, John Harlan, Tom Clark, Felix Frankfurter, and Harlan Stone followed this trend. Others, such as Justices Breyer, Ginsburg, Souter, Blackmun, White, Stewart, Douglas, Black, Burton, Reed, Minton, and Chief Justice Burger experienced a reverse trend. In most instances, however, the difference across periods was not statistically significant, suggesting that, at least by this measure, older age does not manifest itself in greater variability in writing style.

Because pre- and post-sixty-five comparison is a broad measure of change in the Court, it may make sense to look at more narrow time periods. For example, it might make sense to contrast Justices' first five years with their final five years on the Court. The intuition behind examining these periods is that, even for Justices with prior judicial experience, the Court signifies a big adjustment. The Court's caseload and writing requirements differ from other state or federal courts. The Court has a docket that is nearly entirely discretionary and dedicated to addressing new issues of law. These demands likely create intellectual challenges that many Justices had not previously encountered as judges or as lawyers, suggesting that they face a learning curve.

The results, reported in Table 8, reveal greater differences within the Justices. Sixteen out of twenty-eight Justices had higher V scores during their first five years as compared with their last five years. In most of these instances, this change was not statistically significant, but some interesting results emerge. Chief Justice Rehnquist, for example, has shown remarkable consistency; his V score was 2.85 during his first five years and 2.72 during his most recent five years. Justice Thurgood Marshall, whom some speculated to have "relied more heavily on his law clerks" during his first years on the Court,<sup>90</sup> actually had a slightly lower V score during his final years on the Court (2.84) compared with his first five (3.02). Other Justices, by contrast, changed considerably across these periods. Justice Souter had a significantly higher V score (3.69) in his first five years than in his last five years (3.11), while his predecessor, Justice Brennan, had a markedly *lower* V score (2.46) in his first five years than in his last five years (3.63). These results are consistent with the claim that Justice Souter

---

<sup>90</sup> Tushnet, *Brethren*, *supra* note 12, at 2112.

TABLE 8:  
WITHIN-JUSTICE COMPARISON—FIRST FIVE AND LAST FIVE YEARS  
CURRENT AND DEPARTED JUSTICES SINCE 1941

Number	Justice	First 5 Years on Court			Last 5 Years on Court			Bootstrapping		
		Opinions	Avg Word Length	V4	Opinions	Avg Word Length	V4	P(First5<Last5)	Lower 95 Confidence Interval	Upper 95 Confidence Interval
1	Sonia Sotomayor									
2	Samuel A. Alito									
3	John G. Roberts									
4	Stephen G. Breyer	32	3860	2.86	39	3922	3.09	0.9472	-0.05	0.52
5	Ruth Bader Ginsburg	40	4921	3.58	39	3692	3.33	0.1193	-0.64	0.16
6	Clarence Thomas	38	4777	3.73	39	3328	3.86	0.6340	41.00	33.00
7	Anthony M. Kennedy	56	4397	3.37	32	6699	4.00	0.9985	0.20	0.97
8	Antonin Scalia	50	4216	3.04	41	4364	3.19	0.7873	-0.19	0.43
9	John Paul Stevens	54	3607	2.97	33	4280	3.46	0.9582	-0.06	0.93
10	David H. Souter	35	5440	3.69	35	4287	3.11	0.0002	-0.90	-0.24
11	Sandra Day O'Connor	63	5159	3.60	30	3993	4.01	0.9056	-0.16	0.79
12	William H. Rehnquist	85	3787	2.85	37	3409	2.72	0.1732	-0.52	0.16
13	Harry A. Blackmun	52	4635	3.87	48	4563	3.34	0.0402	-1.05	0.06
14	Byron Raymond White	52	2976	2.91	74	4872	3.32	0.9963	63.00	77.00
15	Thurgood Marshall	42	2707	3.02	67	3326	2.84	0.1794	-0.57	0.19
16	William J. Brennan	54	2765	2.46	63	4189	3.63	1.0000	0.78	1.47
17	Lewis F. Powell	85	4026	3.38	77	3279	3.01	0.0589	-0.76	0.08
18	Warren Earl Burger	54	4705	3.29	78	2978	3.01	0.1367	-0.70	0.19
19	Potter Stewart	47	3013	2.84	73	3744	3.32	0.9925	0.09	0.91
20	William Orville Douglas	105	2544	2.70	87	1863	2.60	0.2849	-0.69	0.32
21	John Marshall Harlan	45	2728	2.56	52	4639	3.33	0.9997	0.31	1.17
22	Hugo Lafayette Black	73	1831	2.50	72	3023	2.75	0.9033	-0.12	0.60
23	Abe Fortas									
24	Earl Warren	45	1915	2.34	47	4003	3.54	1.0000	0.65	1.65
25	Tom Campbell Clark	48	2099	2.43	61	3498	3.06	0.9967	0.16	1.11
26	Arthur Joseph Goldberg									
27	Charles Evans Whittaker	42	2605	3.19	39	2726	3.25	0.6372	-0.28	0.39
28	Felix Frankfurter	70	2454	2.50	49	4381	3.18	0.9984	0.18	1.08
29	Harold Hitz Burton	28	4137	4.88	38	2653	3.36	0.0123	-2.12	-0.44
30	Stanley Forman Reed	76	2682	2.53	42	3131	2.57	0.6184	-0.26	0.37
31	Sherman Minton	46	1694	2.38	44	1487	2.19	0.0656	-0.41	0.05
32	Robert Houghwout Jackson	59	2938	2.62	47	2767	2.26	0.0028	-0.60	-0.11
33	Fred Moore Vinson	50	3561	3.01	54	2854	2.87	0.2545	-0.52	0.27
34	Wiley Blount Rutledge	49	4490	3.38	52	4703	3.11	0.0635	-0.67	0.08
35	Frank Murphy	70	1904	2.35	68	2451	2.87	0.9968	0.14	0.80
36	Owen Josephus Roberts	92	1914	2.57	85	2736	2.98	0.9455	-0.08	0.83
37	James Francis Byrnes									
38	Charles Evans Hughes	260	2795	3.03	99	2718	2.98	0.3409	-0.53	0.35
39	Harlan Fiske Stone	109	1917	2.60	95	3530	3.33	0.9995	0.29	1.33
40	James Clark McReynolds	94	1343	2.10	57	1195	1.90	0.0235	-0.41	0.00

Note: We have not analyzed Justices who have been on the Court for fewer than five years. For current Justices, the last five years on the Court are their most recent five years.

relied less on his clerks as his tenure proceeded, and Justice Brennan relied more.

### C. Authorship Identification

Thus far, our writing variability measure has established that the Justices have writing styles that are statistically distinguishable from the null and as shown through our bootstrap test in many instances statistically more variable than other Justices. We now turn our attention to answer the question of practical importance: whether it is possible to accurately predict authorship of judicial opinions.

The consensus from our informal inquiries with constitutional law scholars suggests that while they might be able to identify authorship based on known passages, they could not likely discern author-

ship based on writing style alone. The task is all the more challenging when comparing Justices of similar judicial ideology. In this Part we test whether it is possible to use function words to accurately identify authorship.

As with the bootstrap, we approach this question through a pairwise approach. We consider a particular pair of Justices—for example, Justice *A* and Justice *B*—and the universe of majority opinions in which the author is one of these Justices. Consistent with hypothesis testing and in order to avoid overfitting,<sup>91</sup> we partition the data into two separate parts: a training set and a testing set. We use the training set to develop a model for classifying judgments as being authored by either Justice *A* or Justice *B*. We reserve the testing data to determine the predictive ability of our model.

To test its accuracy, we use leave-one-out cross-validation: for each judgment written by either Justice *A* or *B*, the judgment is the test set, and all other judgments written by either Justice *A* or *B* serve as the training set. We determine whether the model properly classifies the test judgment as belonging to Justice *A* or *B*. We repeat this process for each judgment and count the number of accurate classifications over the number of judgments by Justice *A* and Justice *B*, respectively.

With a linear classifier, we let  $T$  be a training set consisting of all judgments by Justice *A* or *B*, with  $|T| = n$ , where  $n$  are the total number of opinions. We use the following linear regression model:  $Y = x\beta + \varepsilon$ , in which  $\varepsilon$  is an  $n \times 1$  vector of independent errors with mean zero.  $Y$  represents an  $n \times 1$  vector of  $\pm 1$ : we assign a value of  $-1$  for each judgment written by Justice *A* and a  $+1$  for each judgment written by Justice *B*. The term  $x$  is an  $n \times 64$  matrix defined as:

$$x = \begin{pmatrix} 1 & f_{1,1} & f_{1,2} & \cdots & f_{1,63} \\ 1 & f_{2,1} & f_{2,2} & \cdots & f_{2,63} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & f_{n,1} & f_{n,2} & \cdots & f_{n,63} \end{pmatrix},$$

where  $f_{ij}$  are the fraction of words in judgment  $i$  in the training set that are from function word  $j$ .

The least-squares estimate for  $\beta$  corresponds to the maximum likelihood estimate (MLE) if the errors ( $\varepsilon_i$ ) are assumed to be independent and identically distributed (IID) normal random variables, and is defined as:  $\hat{\beta} = (x^T x)^{-1} x^T Y$  where  $\beta = (\beta_0, \beta_1, \dots, \beta_n)$ . Given a test

---

<sup>91</sup> Overfitting of the data is where one constructs a model that fits well on the existing data but fails to effectively predict on new data. See Rosenthal & Yoon, *supra* note 42, at 287 (“[A]dditional [statistical] quantities [beyond function words] did not greatly improve . . . predictive power . . .”).

judgment with function words fractions of  $g_1, g_2, \dots, g_{63}$ , we can calculate the linear fit value  $l = \beta_0 + \sum_{j=1}^{63} \beta_j g_j$ .

We classify Justice *A* as the author of the test judgment if  $l < 0$ , and otherwise assign it to Justice *B*. While there are other possible means of classification (e.g., neural networks or support vector machines), we choose the linear classifier for its parsimony and accuracy.<sup>92</sup>

Table 9 provides two examples of the linear classifier. When comparing Justice Breyer and Justice Ginsburg, the linear classifier accurately predicted the author of Justice Breyer's opinions 94% of the time and Justice Ginsburg's opinions 96% of the time. A comparison of Justice Clarence Thomas and Justice Kennedy produces lower accuracy results, accurately predicting authorship of Justice Thomas's opinions 77% of the time and Justice Kennedy's opinions 84% of the time.

Table 10 produces the results of the pairwise comparisons for the linear classifier for all Justices from Chief Justice Rehnquist's last natural court (1994 to 2005). The lowest prediction rates involved opinions written by Justice Thomas. It predicted only 56% when compared with Justice Stevens, 49% with Chief Justice Rehnquist, and only 40% with Justice O'Connor. By contrast, the model's highest

<sup>92</sup> We also tried a Naïve Bayes classifier, which assumes that (1) the different fractions of function words  $j$  are conditionally independent; (2) on the condition that Justice *A* is the author of record, the conditional distribution of the fraction  $f_j$  of function word  $j$  appearing in the judgment is normal; and (3) the corresponding mean and variance are given by the sample mean and variance by Justice *A* in the training data. See generally Burns, *supra* note 38, at 1574 (discussing use of the Bayes Rule in aggregating probabilities). This classifier produces a log likelihood of Justice *A* having written the opinion,  $\text{loglike}(A) = C - \sum_{j=1}^{63} \left( \frac{1}{2} \log(v_j) + \frac{(f_j - m_j)^2}{2v_j} \right)$  for some constant  $C$ , where  $m_j$  and  $v_j$  are the same mean and variance of the fraction of words which are reference word  $j$ , over all judgments in the training set written by Justice *A*. We compute the  $\text{loglike}(B)$  in the same manner. The Naïve Bayes produces similar, but on average slightly less accurate, predictions to the linear classifier.

TABLE 9:  
LINEAR CLASSIFIER  
SAMPLE COMPARISONS

Justice <i>A</i>	Justice <i>B</i>	Fraction Accurately Predicting	Success Rate Predicting	Fraction Accurately Predicting	Success Rate Predicting
		Justice <i>A</i>	Justice <i>A</i>	Justice <i>B</i>	Justice <i>B</i>
Stephen G. Breyer	Ruth Bader Ginsburg	117/125	0.9360	133/138	0.9638
Clarence Thomas	Anthony M. Kennedy	114/149	0.7651	163/193	0.8446

prediction rates involved Justice Breyer, achieving a minimum accuracy rate of 86% in each of the eight pairings.

We note two interesting points of interpretation from the linear classifier analysis. First, Justices may have similar V scores yet have writing styles that are clearly distinguishable from one another. For example, as noted in Table 6, Justice Thomas and Justice Kennedy have nearly identical V scores of 3.70 and 3.73, respectively. The linear classifier, however, is able to predict with 77% accuracy authorship of Justice Thomas's opinions and 84% accuracy authorship of Justice Kennedy's opinions. The explanation lies in the construction of the V score: similar or even identical V scores can reflect large differences between Justices in their frequency of use of various function words. For example, while Justice Thomas and Justice Kennedy use the word *his* at identical rates (mean = 0.0022), Justice Thomas uses the word *such* (mean = 0.0025) nearly as often as Justice Thomas (mean = 0.0012).

TABLE 10:  
LINEAR CLASSIFIER  
JUSTICES ON LAST REHNQUIST NATURAL COURT (1994–2005)

	V Score	Non-Authoring Justice (Justice B)								
		Stephen G. Breyer	Ruth Bader Ginsburg	Clarence Thomas	Anthony M. Kennedy	Antonin Scalia	John Paul Stevens	David H. Souter	Sandra Day O'Connor	William H. Rehnquist
Stephen G. Breyer	3.06		0.9360	0.8800	0.9680	0.9600	0.9200	0.9280	0.8640	0.9040
Ruth Bader Ginsburg	3.57	0.9638		0.8261	0.9058	0.9565	0.8406	0.8406	0.8043	0.8406
Clarence Thomas	3.70	0.8523	0.7987		0.7651	0.7450	0.5570	0.8188	0.3960	0.4899
Anthony M. Kennedy	3.73	0.9741	0.9119	0.8446		0.9171	0.7720	0.9275	0.7150	0.7513
Antonin Scalia	3.08	0.9819	0.9412	0.8235	0.8869		0.7919	0.9186	0.8235	0.7511
John Paul Stevens	3.30	0.9810	0.8523	0.8916	0.9079	0.9160		0.9458	0.8211	0.7263
David H. Souter	3.65	0.9732	0.8523	0.8389	0.8591	0.8926	0.8121		0.7584	0.7718
Sandra Day O'Connor	3.85	0.9449	0.9338	0.7978	0.8640	0.8750	0.7426	0.9265		0.6691
William H. Rehnquist	3.01	0.9854	0.9519	0.9247	0.9331	0.9226	0.8033	0.9582	0.8096	

Note: V score column represents V score for all opinions. Remaining cells represent fraction that accurately predicts Justice A compared to Justice B.

Second, unlike in Table 6, the values across the diagonal in Table 10 are not necessarily symmetric. In other words, the fraction predicting Justice A's opinions versus Justice B need not be the same for the fraction predicting Justice B's opinions versus Justice A. For example, in a comparison of Justice Thomas and Justice Stevens, the linear clas-

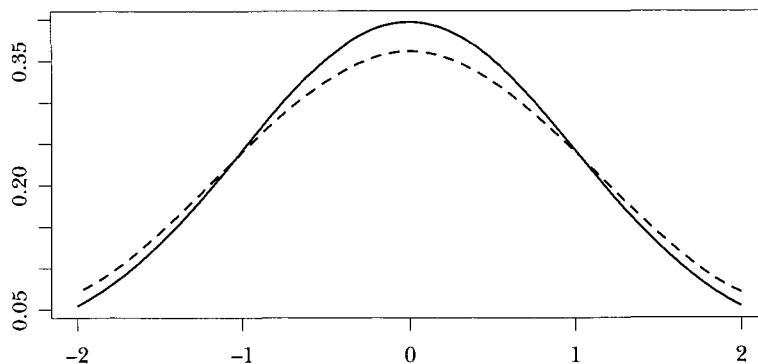
sifier predicts Justice Thomas as the author of his opinions with only 56% accuracy but predicts Justice Stevens as the author with 89% accuracy. This asymmetry is due to the shape of the probability distributions of the two Justices and their degree of overlap. If one distribution largely overlaps with another distribution, it is possible for the prediction to be much higher for one than the other.<sup>93</sup>

Overall, of the seventy-two possible pairings represented in Table 10, the model achieved an accuracy rate of at least 70% in sixty-eight pairings (representing nearly 95% of pairings). In thirty (42%) of the pairings, the accuracy rate exceeded 90%. Comparing these rates to a null hypothesis of authorship being randomly determined (i.e., 50%), our model appears to predict quite well. We achieved these results using function words, and it is possible that a different algorithm—that is, one that is more tailored toward words more common to legal writing—may produce even higher prediction rates.

The question of which Justice authored a Supreme Court opinion is, admittedly, an academic exercise, since most majority opinions reveal the authoring Justice. The purpose of this exercise is to show that the text of Justices' opinions is statistically distinguishable from another, even in instances when their V scores are not statistically significantly more variable.<sup>94</sup> Our analysis shows that statistical analysis can

---

<sup>93</sup> To illustrate how symmetries arise, consider the following figure:



The solid curve is a graph with mean zero and variance one. The dotted curve is a graph with mean zero and variance 1.1. Most of the probability distribution falls between 1 and -1, where the solid curve is larger. In this example, approximately 70% of the points (opinions) chosen from either distribution will be classified as belonging to the solid curve. Accordingly, points from the solid distribution will have the correct classification about 70% of the time, while points from the dotted distribution will have the correct classification about 30% of the time.

<sup>94</sup> There may also be a practical application to the linear classifier when looking at per curiam opinions, in which the Court does not report the authoring Justice. The linear classifier can often help one discern the likely author of these opinions when analyzed in combination with known majority opinions.

meaningfully contribute to our understanding not merely of how Justices vote<sup>95</sup> but also how they write.

#### IV

#### DISCUSSION

This Article had two objectives. First, to show that it is possible to statistically evaluate Justices' writings. Using only function words, we produced measures of Justices' writing variability; in many instances these differences were large and statistically significant. Moreover, our approach allows us to accurately predict authorship, even when Justices have similar or identical V scores. In 94% of pairwise comparisons of the last Rehnquist natural court, our linear classifier model achieves an accuracy rate of at least 70%. In 41% of comparisons, the classifier model exceeded 90%.<sup>96</sup>

Our second and more important objective was to use textual analysis to better understand how Justices produce opinions and how much they rely on their clerks. For the period 1900–50, the Justices' V scores were low, consistent with historical accounts of clerks serving a primarily administrative role. Since 1950, however, their scores have steadily increased.<sup>97</sup> This upward trend supports claims that Justices are increasingly relying on their clerks in the opinion-writing process. In addition, many Justices vary considerably in their V scores from one year to the next.<sup>98</sup>

Moreover, our findings suggest that delegation of writing responsibilities to law clerks is highest among those considered swing Justices. Justices Kennedy, O'Connor, Blackmun, and White reported the highest V scores among all Justices dating back to 1900.<sup>99</sup> Whether being a swing Justice fosters greater variability in writing, or whether the relationship is mere coincidence, the consequences are tangible and warrant closer examination. A large fraction of opinions

---

<sup>95</sup> Most analysis of judicial politics examines the voting patterns of the Justices based on the U.S. Supreme Court Judicial Database, created by Harold J. Spaeth. See THE SUPREME COURT DATABASE, <http://scdb.wustl.edu/> (last visited April 11, 2011) ("The Supreme Court Database is the definitive source for researchers, students, journalists, and citizens interested in the U.S. Supreme Court. The Database contains over two hundred pieces of information about each case . . . [including] the votes of the Justices."); see also *The Genesis of the Database*, THE SUPREME COURT DATABASE, <http://scdb.wustl.edu/about.php> (last visited April 11, 2011).

<sup>96</sup> See *supra* Tables 9 & 10.

<sup>97</sup> See *supra* Figure 1.

<sup>98</sup> See *supra* Figure 2.

<sup>99</sup> See *supra* Table 2 and *infra* Table A1.



each year are 5–4,<sup>100</sup> often the most important cases of the term. Moreover, these cases are often assigned to swing Justices.<sup>101</sup>

We emphasize our earlier caveat that our analysis provides only circumstantial evidence of collaborative authorship. We argue throughout the Article that low V scores suggest the Justice does her own writing. Alternative explanations are that the Justice delegates writing to his clerks but closely edits their writing, or that the clerks effectively mimic their Justices' writing style. Conversely, we contend that high V scores reflect greater delegation of writing responsibilities to the clerk. However, it could be the case that the Justice does her own writing but simply has a naturally higher variability. While these alternative explanations are plausible, our separate analysis of opinions by Judges Posner and Easterbrook, both known to write their own opinions, yields lower V scores with lower variation from one year to the next.<sup>102</sup> These results provide strong support for our analysis.

Given our finding that Justices increasingly delegate to clerks in the opinion-writing process, should we be concerned? Legal scholars have weighed in on the normative and positive implications of judicial reliance on clerks in opinion writing. Judge Posner, not surprisingly, is critical of this trend:

Most judges nowadays, because of heavy caseloads, delegate the writing of their judicial opinions to their clerks. It's a mistake on a number of grounds: [t]he more you write, the faster you write; only the effort to articulate a decision exposes the weak joints in the analysis; and the judge-written opinion provides greater insight into the judge's values and reasoning process and so provides greater information—not least to the judge.<sup>103</sup>

Other scholars have argued that judges who write their own opinions are more influential than judges who rely heavily on clerks.<sup>104</sup> Mark Tushnet, a former law clerk for Justice Marshall,<sup>105</sup> has commented that “an opinion cannot carry the weight of [a] Justice's prior

---

<sup>100</sup> For example, according to the U.S. Supreme Court Database, *supra* note 95, for the period of the 2001–09 terms, a 5–4 majority decided 21% (154 out of 733) of the cases before the Supreme Court.

<sup>101</sup> See Paul H. Edelman & Jim Chen, *The Most Dangerous Justice Rides Again: Revisiting the Power Pageant of the Justices*, 86 MINN. L. REV. 131, 186–88 (2001) (noting that Justice O'Connor and Justice Kennedy write more 5–4 opinions than their seniority would suggest).

<sup>102</sup> See *supra* Table 4 and Figure 3.

<sup>103</sup> See Richard Posner, *Diary: Entry 1*, SLATE, (Jan. 14, 2002, 11:46 AM) <http://www.slate.com/id/2060621/entry/2060676/>.

<sup>104</sup> See, e.g., William M. Landes, Lawrence Lessig & Michael E. Solimine, *Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges*, 27 J. LEGAL STUD. 271, 271–74 (1998).

<sup>105</sup> Tushnet, *Brethren*, *supra* note 12, at 2109 n.\*.

public service when it is written by a recent law school graduate serving as the Justice's law clerk."<sup>106</sup>

Our analysis is silent on whether increased delegation to clerks has improved or harmed the quality of the Court's opinions. Such a claim falls outside this Article's analysis, which at its core is quantitative, not qualitative. Answering this question presents its own challenges on two fronts: first, the nonrandom process by which the Chief Justice (or senior associate Justice if the Chief is not in the majority) assigns opinions; and second, there is no obvious metric for measuring the quality of the Justices' writing. The same opinion often invokes both praise and criticism from legal scholars as well as the public. One point is clear: Justices now write opinions that are longer and more variable along common function words. While technological advances may account for much of this trend, it is not difficult to imagine that Justices would write shorter, less variable opinions without the assistance of their clerks.

Our results also more broadly inform existing debates about the Court. For example, recent scholarship has criticized lifetime tenure for Supreme Court Justices and proposed a constitutional amendment to impose term limits.<sup>107</sup> We do not find that older Justices—that is, those older than sixty-five years old—have systematically larger V scores than younger Justices. Nor do we find strong evidence that Justices rely more on the clerks in their first few years on the Court. If Justices are adopting different approaches to opinion writing over their tenure, it does not manifest itself in the variability of their writing.

Our analysis does raise institutional questions about the Court, which has remained steadfastly lean over time. As in the nineteenth century, today's Court consists of only nine Justices. Their primary support remains their law clerks; the only difference is that each Justice has four rather than two clerks.<sup>108</sup> As the Justices' docket has steadily grown, it is not surprising that law clerks play an increasingly substantive role.

If our main concern is that Justices delegate too much writing to their clerks, an obvious solution is to increase the number of Justices.

---

<sup>106</sup> Mark Tushnet, *Style and the Supreme Court's Educational Role in Government*, 11 CONST. COMMENT. 215, 222 (1994).

<sup>107</sup> See generally Steven G. Calabresi & James Lindgren, *Term Limits for the Supreme Court: Life Tenure Reconsidered*, 29 HARV. J.L. & PUB. POL'Y 769, 771–72 (2006) (“[T]he American constitutional rule granting life tenure to Supreme Court Justices is fundamentally flawed . . . .” (footnote omitted)); Garrow, *supra* note 89 (discussing historical attempts to place term limits on Supreme Court Justices).

<sup>108</sup> See WARD & WEIDEN, *supra* note 6, at 36 (“The second major transformation occurred in 1941 when the number of clerks was doubled to two per justice.”); Wahlbeck et al., *supra* note 12, at 169 (“Each justice on the Supreme Court is currently authorized to employ four law clerks . . . .”).

Court expansion would increase the number of Justices to hear cases and write opinions. Article III of the Constitution is silent on the number of Justices, so Court expansion would not require a constitutional amendment.<sup>109</sup> This expansion, however, may be politically infeasible: intense confirmation hearings of any Court nominee appear to be the norm,<sup>110</sup> and Congress and the public vehemently opposed the last attempt to expand the Court.<sup>111</sup>

Another approach, recently proposed by scholars, is for the Court to hear cases in smaller panels (e.g., three Justices) rather than en banc.<sup>112</sup> If we assume that the docket and the number of cases granted cert would remain the same, each Justice could presumably prepare fewer cases for oral argument and dedicate more time to writing opinions. The Court could make this change on its own, without any congressional approval.<sup>113</sup> Other countries have taken this exact approach, such as the Canadian Supreme Court.<sup>114</sup> It is unclear, however, whether the current Court would be willing to do so.

If, however, our concern is less about delegation per se and more about the implications of delegating to relatively inexperienced staff, then the simplest solution is for the Court to modify its clerk selection process. Rather than draw primarily from recent law graduates, the Court could hire more law clerks with greater legal experience. The intuition here is that if there are positive returns to experience in the study of law implicit in the selection criteria of Justices, then it logically follows that these criteria are desirable attributes among law clerks. Some state supreme courts have incorporated this idea by creating permanent clerks.<sup>115</sup>

---

<sup>109</sup> See U.S. CONST., art. III, § 1 (specifying only that Supreme Court Justices “shall hold their offices during good behaviour, and shall, at stated times, receive for their services, a compensation, which shall not be diminished during their continuance in office”).

<sup>110</sup> Cf. Elena Kagan, *Confirmation Messes, Old and New*, 62 U. CHI. L. REV. 919, 920–30 (1995) (reviewing STEPHEN L. CARTER, *THE CONFIRMATION MESS* (1994)) (arguing that the less acrimonious Senate confirmation hearings of the present also run the risk of “ceas[ing] to engage nominees in meaningful discussion of legal issues”).

<sup>111</sup> See WILLIAM E. LEUCHTENBURG, *THE SUPREME COURT REBORN: THE CONSTITUTIONAL REVOLUTION IN THE AGE OF ROOSEVELT 132–62* (1995).

<sup>112</sup> See, e.g., George & Guthrie, *supra* note 80, at 1442, 1458–65; Tracey E. George & Chris Guthrie, “*The Threes*”: *Re-Imagining Supreme Court Decisionmaking*, 61 VAND. L. REV. 1825, 1827 (2008).

<sup>113</sup> See 28 U.S.C. § 2072(a) (2006) (“The Supreme Court shall have the power to prescribe general rules of practice and procedure . . . in the United States district courts . . . and courts of appeals.”).

<sup>114</sup> See Benjamin R.D. Alarie et al., *Is Bigger Always Better? On Optimal Panel Size, with Evidence from the Supreme Court of Canada 2* (Univ. Toronto Legal Studies Research, Paper No. 08-15, 2011), available at [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1152322](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1152322) (“[A]lthough the Supreme Court of Canada is composed of nine justices . . . it routinely sits in panels of five, seven, or nine justices.”).

<sup>115</sup> For example, the California Supreme Court and lower courts now have permanent law clerks. See Paul L. McKaskle, *The European Court of Human Rights: What It Is, How It*

## CONCLUSION

This Article provides a statistical analysis of judicial opinions aimed at improving our understanding of judicial authorship. We show that it is possible to statistically evaluate the content of Supreme Court Justices' opinions, and our results offer strong evidence that Justices are increasingly relying on their clerks when writing opinions. Whether this trend is desirable is a separate, more involved discussion that goes beyond the scope of this Article. This discussion is worth having, given the import that judges, litigants, and scholars give to the Justices' chosen words.

More broadly, this Article also provides what we hope is a step toward bridging research between social science and traditional legal scholarship. The former focuses almost exclusively on how judges and Justices vote on individual cases and has relatively little to say substantively about legal doctrine, while the latter has much to say about legal doctrine but less about how systematically it develops or its subsequent effect. While this Article examines only judicial authorship, statistical textual analysis provides a powerful means to improve our substantive understanding of the law.

---

*Works, and Its Future*, 40 U.S.F. L. REV. 1, 27 n.164 (2005); see also THE SUPREME COURT OF CALIFORNIA: CONTAINING THE INTERNAL OPERATING PRACTICES AND PROCEDURES OF THE CALIFORNIA SUPREME COURT 4 (2007), available at <http://www.courtinfo.ca.gov/courts/supreme/iopp.htm> (noting that "a judicial assistant and five staff attorneys" support each Justice).

## APPENDIX

TABLE A1:  
CURRENT AND DEPARTED JUSTICES—PRE-1941

Number	Justice	Appointing President	Year Joined Court	Year Left Court	Age at Retirement	Years on Court	Opinions	Average Word Length	V Score
41	Pierce Butler	Harding	1923	1939	73	16	323	1909	2.69
42	Louis Dembitz Brandeis	Wilson	1916	1939	83	23	453	1745	2.35
43	Benjamin Nathan Cardozo	Hoover	1932	1938	68	6	126	2718	2.32
44	George Sutherland	Harding	1922	1938	76	16	280	2290	2.36
45	Willis Van Devanter	Taft	1911	1937	78	26	344	2243	2.50
46	Oliver Wendell Holmes	Roosevelt, T.	1902	1932	91	30	874	1161	1.78
47	Edward Terry Sanford	Harding	1923	1930	65	7	128	1936	2.72
48	William Howard Taft	Harding	1921	1930	73	9	253	2786	2.81
49	Joseph McKenna	McKinley	1898	1925	82	27	631	2242	2.61
50	John Hessin Clarke	Wilson	1916	1922	65	6	129	1980	2.47
51	Mahlon Pitney	Taft	1912	1922	64	10	248	2943	2.91
52	William Rufus Day	Roosevelt, T.	1903	1922	73	19	423	2543	2.94
53	Edward Douglas White	Taft	1894	1921	76	27	649	2995	2.95
54	Joseph Rucker Lamar	Taft	1911	1916	59	5	108	1954	2.25
55	Charles Evans Hughes	Taft	1910	1916	54	6	391	2842	3.08
56	Horace Harmon Lurton	Taft	1910	1914	70	4	91	3028	2.87
57	John Marshall Harlan	Hayes	1877	1911	78	34	720	3425	3.48
58	William Henry Moody	Roosevelt, T.	1906	1910	57	4	62	2350	2.60
59	David Josiah Brewer	Harrison	1890	1910	73	20	520	2432	2.58
60	Melville Weston Fuller	Cleveland	1888	1910	77	22	748	2204	2.84
61	Rufus Wheeler Peckham	Cleveland	1896	1909	71	13	300	3227	3.33
62	Henry Billings Brown	Harrison	1891	1906	70	15	441	2918	3.30
63	George Shiras	Harrison	1892	1903	71	11	243	3097	3.16
64	Horace Gray	Arthur	1882	1902	74	20	436	2572	2.97
65	Stephen Johnson Field	Lincoln	1863	1897	81	34	541	2206	2.69
66	Howell Edmunds Jackson	Harrison	1893	1895	63	2	45	3592	3.58
67	Lucius Quintus C. Lamar	Cleveland	1888	1893	36	5	98	2968	3.36
68	Samuel Blatchford	Arthur	1882	1893	73	11	417	3393	4.70
69	Joseph P. Bradley	Grant	1870	1892	79	22	380	3132	3.24
70	Samuel Freeman Miller	Lincoln	1862	1890	74	28	607	2197	2.52
71	Stanley Matthews	Garfield	1881	1889	65	8	228	3309	3.30
72	Morrison Remick Waite	Grant	1874	1888	72	14	766	1315	2.28
73	William Burnham Woods	Hayes	1881	1887	63	6	161	2229	2.93
74	Ward Hunt	Grant	1873	1882	72	9	143	1666	2.67
75	Noah Haynes Swayne	Lincoln	1862	1881	77	19	329	1722	2.19
76	Nathan Clifford	Buchanan	1858	1881	77	23	392	3231	3.25
77	William Strong	Grant	1870	1880	72	10	239	2469	3.23
78	David Davis	Lincoln	1862	1877	62	15	191	1414	2.00
79	Salmon Portland Chase	Lincoln	1864	1873	65	9	114	1385	1.95
80	Samuel Nelson	Tyler	1845	1872	83	27	294	1742	2.31

TABLE A1 (CONTINUED):  
CURRENT AND DEPARTED JUSTICES—PRE-1941

Number	Justice	Appointing President	Year Joined Court	Year Left Court	Age at Retirement	Years on Court	Opinions	Average Word Length	V Score
81	Robert Cooper Grier	Polk	1846	1870	61	24	195	1658	2.21
82	James Moore Wayne	Jackson	1835	1867	90	32	130	4077	3.88
83	John Catron	Jackson	1837	1865	85	28	146	2303	2.34
84	Roger Brooke Taney	Jackson	1836	1864	81	28	241	2324	2.22
85	John Archibald Campbell	Pierce	1853	1861	58	8	90	1740	2.07
86	John McLean	Jackson	1830	1861	81	31	228	2397	2.63
87	Peter Vivian Daniel	Van Buren	1842	1860	68	18	80	3375	2.93
88	Benjamin Robbins Curtis	Fillmore	1851	1857	46	6	47	3353	2.72
89	John McKinley	Van Buren	1838	1852	68	14	12	3352	3.02
90	Levi Woodbury	Polk	1845	1851	57	6	42	3175	2.62
91	Joseph Story	Madison	1812	1845	77	33	228	2815	2.63
92	Henry Baldwin	Jackson	1830	1844	54	14	40	3896	2.71
93	Smith Thompson	Monroe	1823	1843	67	20	77	3105	2.79
94	Philip Pendleton Barbour	Jackson	1836	1841	55	5	16	3442	2.12
95	Gabriel Duvall	Madison	1811	1835	56	24	7	1367	2.28
96	John Marshall	Adams, John	1801	1835	80	34	449	2351	2.80
97	William Johnson	Jefferson	1804	1834	63	30	76	2355	2.20
98	Bushrod Washington	Adams, John	1799	1829	67	30	38	3307	2.68
99	Robert Trimble	Adams, J. Q.	1826	1828	43	2	15	2249	2.25
100	Thomas Todd	Jefferson	1807	1826	61	19	7	1247	2.10
101	Henry Brockholst Livingston	Jefferson	1807	1823	71	16	17	1792	1.98
102	Samuel Chase	Washington	1796	1811	70	15			
103	William Cushing	Washington	1790	1810	78	20	1	314	
104	William Paterson	Washington	1793	1806	61	13	1	440	
105	Alfred Moore	Adams, John	1800	1804	49	4			
106	Oliver Ellsworth	Washington	1796	1800	55	4	4	976	1.87
107	James Iredell	Washington	1790	1799	48	9			
108	James Wilson	Washington	1789	1798	56	9			
109	John Blair	Washington	1790	1795	63	5			
110	John Jay	Washington	1789	1795	50	6	3	600	1.20
111	Thomas Johnson	Washington	1792	1793	61	1			
112	John Rutledge	Washington	1790	1791	52	1	1	885	

Note: In the early years on the Court (1790–1820), the identity of the author of an opinion was either the Chief Justice or commonly not revealed. Accordingly, we could not produce a V score for many of these early Justices.

