Determinants of Writing Style on the United States Circuit Courts of Appeals

JEFFREY BUDZIAK, Western Kentucky University
MATTHEW P. HITT, Colorado State University
DANIEL LEMPERT, SUNY Potsdam

ABSTRACT

A rapidly burgeoning literature in judicial politics concerns the variation in elements of writing style such as reading difficulty, cognitive complexity, affective language, and informality in judicial opinions. Some of these studies argue that judges strategically alter their writing style in anticipation of reactions from other actors. Others indicate that writing style is a function of judge characteristics as well as case-related factors. We investigate the correlates of writing style in US Circuit Courts of Appeals by analyzing a stratified random sample consisting of 11,771 opinions. Construing style broadly to encompass several dimensions suggested by prior work, we find that case and judge characteristics explain substantially more variance in writing style than do strategic considerations.

Political scientists frequently examine how choices made by actors shape the outputs of the political process. Empirically minded scholars tend to focus on actors' choices rather than the content of the resulting outputs: discrete choices lend themselves to systematic classification and are thus easily amenable to examination using quantitative techniques (e.g., Poole and Rosenthal 1997). This does not suggest that empirically minded scholars are uninterested in the substantive content of outputs. However, systematically classifying content raises many difficult methodological questions. Classifying the choice made by a legislator on a roll-call vote is relatively simple; systematically classifying the content of the bill itself is more difficult. Yet comprehensive explanation of the political process requires an understanding of both the choices made by actors and the content of those choices.

We thank Larry Baum, Julian Brooke, Greg Caldeira, Morgan Hazelton, and the reviewers and editor for helpful comments and suggestions. We also thank Brook Spurlock for excellent research assistance. Daniel Lempert acknowledges support from a New York State/United University Professions Individual Development Award. Contact the corresponding author, Daniel Lempert, at lemperds@potsdam.edu.

Journal of Law and Courts (Spring 2019) 2164-6570/2019/0701-00XX$10.00. Electronically published December 14, 2018.

Nowhere has this distinction emerged more prominently than in scholarly examinations of the judiciary. The most important developments in the field of judicial politics have emerged through investigations of the choices made by judges (Pritchett 1948; Rohde and Spaeth 1976; Baum 1997; Epstein and Knight 1998; Segal and Spaeth 2002). While these examinations have advanced the collective understanding of judicial politics in a number of ways, they are not without limitations. Of greatest practical concern, the federal judicial hierarchy is not governed by dispositional votes. The substantive content of judges' choices—taking the form of legal policies articulated in written opinions—not only dictates the appropriate disposition of the case but shapes both the dispositional choices and substantive opinion content generated by judges when confronting similar cases. This has led to calls for a shift in focus from judicial votes to the content of opinions (Maltzman, Spriggs, and Wahlbeck 2000, 154).

Greater scholarly attention to the study of opinion content is unquestionably a laudable goal. Unfortunately, the challenges of systematically classifying that content have severely constrained empirical analyses of opinions. Recent technological developments in automated textual (text) analysis have permitted scholars to overcome some of these constraints. These tools have been employed most readily to examine the text characteristics of the language used in opinions, which we refer to collectively as "writing style." Scholars can now test many hypotheses involving opinion content that would have previously been impossible using large samples (Coleman and Phong 2010; Owens and Wedeking 2011; Owens, Wedeking, and Wohlfarth 2013; Goelzhauser and Cann 2014; Hansford and Coe 2014; Johnson 2014; Black et al. 2016, 2016b; Bryan and Ringsmuth 2016).

This line of inquiry helps break down the barriers to the systematic analysis of opinion content. The valuable insights to be gained from these advances, and the potential for systematic classifications of opinion content using larger samples than previously possible, have led to a rapid proliferation of studies using automated text analysis. The primary focus of these applications has been to examine how opinion writing style varies based on the preferences of other actors. For example, these studies examine whether Supreme Court justices write in different styles depending on the preferences of Congress (Owens et al. 2013), lower federal judges (Corley and Wedeking 2014; Black et al. 2016b), and the public (Black et al. 2016a, 2016b). While the conditions under which justices employ specific writing styles are beginning to crystallize in this still-nascent literature, many questions concerning writing style as a concept have been overlooked.

Here we focus on one such question: What are the determinants of judicial writing style? As discussed below, scholarship examining the effects of writing style generally assumes that it is a function of purposeful choices made by judges. From this perspective, judges strategically use writing style to further specific goals. We suggest that while strategic behavior is one potential determinant of writing style, there may be other, less studied factors that have a more substantial role in shaping judicial writing style. Specifically, we propose that characteristics of the judge and of the case have the potential to affect the writing style used in opinions, at least as much as strategic choice.

To investigate the determinants of writing style, we conduct several analyses using a stratified random sample of 11,771 courts of appeals opinions from 1950 to 2002. Each analysis follows the same general research design. First, we identify sets of covariates associated respectively with strategic behavior, judicial characteristics, and case-specific considerations to create groups of strategic, judge, and case variables. Second, we scale opinions on a large number of text characteristics, including those elements of writing style most commonly cited in the judicial politics literature (namely, reading difficulty, cognitive complexity, and affective language). Third, we regress each set of variables on each text characteristic score to determine the proportion of the variance explained by each set of variables.

Our first analysis provides an example of this research design by investigating the variance of one text characteristic: reading difficulty, that is, (a lack of) "clarity," the most frequently examined Supreme Court opinion text characteristic (e.g., Coleman and Phong 2010; Owens et al. 2013; Goelzhauser and Cann 2014; Hansford and Coe 2014; Johnson 2014; Black et al. 2016b). We present a set of regression models to assess what proportion of the variance in the reading difficulty of courts of appeals opinions is explained by strategic, judge, and case variables. The results suggest that variables associated with judges and cases, respectively, explain a much larger proportion of the variance than those associated with strategic considerations.

Our second analysis extends this application beyond reading difficulty to 97 additional possible text characteristics. The results suggest that while no set of variables explains a substantial percentage of the variance in writing style, those related to the case explain the greatest portion of variance for nearly all characteristics. By contrast, strategic variables explain very little variance.

Our final analysis investigates the role of case variables in greater detail. Here we restrict our research design only to opinions accompanied by separate opinions (i.e., concurrences and dissents). If case variables are the primary determinants of writing style, we should expect the writing styles of majority and separate opinions to be similar. The results confirm this expectation. Ceteris paribus, separate opinions' text characteristics statistically predict the text characteristics of corresponding majority opinions for nearly every text characteristic. Moreover, for approximately half of all text characteristics, separate opinions' text characteristics explain a higher percentage of variance in the corresponding majority opinions' text characteristics than all other case, judge, and strategic variables combined. We conclude by discussing the implications of these results for scholarship examining opinion content using automated text analysis and identify several other important questions regarding writing style beyond the scope of our analyses.

DETERMINANTS OF WRITING STYLE

Many variables likely determine the style used by judges to craft opinions. We organize these plausible determinants into three distinct categories: strategic, judge, and case. We give thorough attention to strategic variables, as scholarship using automated text analysis

has primarily approached questions about writing style from this perspective. But, as we argue below, there are good reasons to expect that judge or case variables have the potential to influence writing style at least as much as strategic variables.

Strategic Determinants

Most scholarship using empirical measures of writing style hypothesizes that judges use writing styles to shape the behavior of other actors in an effort to achieve their goals. This expectation is consistent with the strategic approach to judicial behavior, which argues that judges recognize that their ability to achieve their goals depends on several considerations, including the preferences of other actors, the choices those actors are expected to make, and the institutional context in which they operate (Epstein and Knight 1998, 9). In fact, many analyses explicitly discuss variation in writing style in terms of "strategic opinion writing" (e.g., Black et al. 2016a, 705–10; 2016b, 12–14) or "strategic behavior" (e.g., Owens et al. 2013, 37–41). For scholars operating in this strategic framework, the primary question of interest centers on the effectiveness of writing style in shaping the behavior of other actors.

Several analyses have produced results that are consistent with a strategic account of opinion writing. Owens et al. (2013) find that Supreme Court opinions are less "clear" (as measured by reading difficulty) when the Court's majority is ideologically distant from relevant pivot points in Congress. They conclude that this pattern is evidence of strategic manipulation of writing style designed to avoid congressional review. Black et al. (2016b) find that the Court produces opinions with lower levels of reading difficulty in several different circumstances: when the lower courts are ideologically scattered and distant from the Court, when the Court rules against administrative agencies with poor performance records, when the Court rules against states with less professionalized legislatures, and when the Court rules against prevailing public opinion. This last finding is replicated by Black et al. (2016a). Corley and Wedeking (2014) and Black et al. (2016b) demonstrate, respectively, that higher levels of certainty and lower levels of reading difficulty in Supreme Court opinions are associated with more positive treatment by lower federal courts. The authors interpret their results as evidence of the relationship between writing style and compliance.

The evidence generated by this line of scholarship suggests that any explanation of writing style would be incomplete without a consideration of strategic behavior. However, it would be premature to conclude that writing style for all judges is exclusively, or even primarily, a function of strategy. It is clear that justices can manipulate their styles in an attempt to achieve certain policy goals, but the circumstances under which judges might engage in these forms of strategic behavior are frequently limited. For example, justices may seek to avoid congressional review of their decisions, but, as Owens et al. (2013, 44) note, "justices need not always worry about the preferences of Congress," since the Supreme Court is often unconstrained, given the preferences of other actors, especially in constitutional cases.

Similar considerations apply to the US Circuit Courts of Appeals. Congressional review—let alone override—of lower-court decisions is particularly rare (Lindquist and Yalof 2001). Also, lower-court compliance is unlikely to be a concern for appeals judges when authoring nonprecedential opinions (a majority of courts of appeals decisions). Still, there is potential for Supreme Court and en banc review, which provides opportunity for strategic behavior.

Judge-Related Determinants

Characteristics of the author offer one plausible alternative explanation for variance in writing styles. Scholars in the fields of information science and technology have recently relied on author-based patterns to attribute authorship to unsigned documents (e.g., Zheng et al. 2006, 379). Work in this field demonstrates that writing styles can accurately predict author characteristics like gender (Koppel, Argamon, and Shimoni 2002), educational experiences (Corney et al. 2002), and cultural backgrounds, among others, offering support for the expectation that writing styles are shaped by these traits.

Systematic efforts to assess the effect of judge characteristics on writing style by judicial politics scholarship are lacking. Previous analyses tend to compare only justices themselves and not judicial characteristics. Owens and Wedeking (2011, 1043–44) provide a descriptive comparison of opinion cognitive complexity by 17 Supreme Court justices. The results offer evidence that cognitive complexity does vary by justice, but no explanation as to why is developed (other than the observation that cognitive complexity is not correlated with ideology). Li et al. (2013) use algorithmic attribution techniques to determine the authorship of unsigned Supreme Court opinions. The accuracy of the method is noteworthy, but questions about why judges possess certain writing styles are not considered.

The lack of systematic consideration of judge characteristics, particularly in light of scholarship suggesting that author traits affect writing style, leaves our understanding of opinion writing style incomplete. The absence of such scholarship is likely due to the near-singular focus of scholarship examining writing style on US Supreme Court opinions. Given the small number of authors available for analysis (and the lack of diversity among them), there is simply not enough variation (at the Supreme Court) to systematically assess how judge characteristics shape writing styles. We attempt to overcome that limitation here.

Case-Related Determinants

Case variables offer a second alternative explanation for variance in writing styles. All cases vary in terms of characteristics largely outside of the litigants' or judges' control. The substantive content of judicial opinions is inherently constrained by case characteristics; judges presumably must address the facts of the case, the legal question, potentially relevant precedents, and so on. While writing style is conceptually distinct from substantive content, it is not likely to be totally independent. Judges writing opinions that interpret statutes or

apply relevant precedents crafted in particular writing styles may have no choice but to match (at least to some degree) the style of the source to achieve consistency and precision in their opinions. A related pattern may exist for the facts of a case itself; some factual patterns may not lend themselves to explanation or discussion using a particular writing style, thus leading judges to adopt different styles than they might select if completely unconstrained by case facts.

Empirical scholarship on the effect of case variables on writing style is minimal. Owens and Wedeking (2011) provide a descriptive comparison of opinion cognitive complexity using one example of a case variable: issue areas. They find that while justices are reasonably consistent in cognitive complexity, there is appreciable within-author variation in cognitive complexity across issue areas. While this provides a preliminary investigation of how writing style is shaped by case variables, greater systematic attention is clearly warranted.

Determinants of writing style can generally be organized by strategic, judge, and case variables. Although the overwhelming majority of scholarship examining writing style has focused on variables associated with strategic behavior, the circumstances under which such behavior is expected or warranted may be limited. However, it is probably more intuitive to see how opinions could be affected by characteristics of the judge and the case, suggesting that these variables likely explain a relatively significant amount of variance in writing styles. But exactly how much variance is explained by these three sets of variables? To determine the answer, we first must identify characteristics associated with each.

IDENTIFYING VARIABLES

We adopt an intentionally inclusive approach to defining the variables associated with each potential influence on writing style. We are not primarily interested in testing hypotheses about associations between individual covariates and elements of writing style. Rather, we seek to determine how much variance in writing style is associated with a given set of variables. To do so, it is important to include as many plausible covariates as possible; failing to include a potentially relevant variable could lead our analyses to underestimate the amount of variance explained by a plausible influence. Our research design should therefore be understood as seeking to identify the maximum amount of variance explained by each influence. While this approach means that the theoretical justification for including certain variables may be reasonably contested, we believe it is the most appropriate design to address our research question.

Consistent with this inclusive approach, we examine courts of appeals opinions to assess the influences on writing style. Prior empirical examinations of writing style have focused exclusively on opinions of the US Supreme Court. The Court frequently warrants disproportionate attention because of its unique position atop the federal judiciary. However, focusing exclusively on the Court would dramatically limit our ability to assess the influences on writing style—particularly judge-related variables. There is simply not enough variance in the characteristics of Supreme Court opinion authors to make a sensible de-

termination of how judge variables affect writing style. The diversity of jurists serving on the courts of appeals helps overcome this concern.

Strategic Variables

The characteristics we identify as being consistent with strategic opinion writing emerge from three different audiences toward which an author could behave strategically: other members of the courts of appeals, members of the Supreme Court, and members of Congress. We address each of these audiences individually.

The most immediate audience judges must consider when engaging in strategic behavior are other members of the courts of appeals. Of greatest importance for an author are the other panel members. To achieve a majority, an author must convince at least one other member of the three-judge panel to agree to the opinion. If an author were to craft an opinion inconsistent with the preferences of the other panel members, they could find themselves in the minority. To achieve a favored policy outcome, authors must consider the preferences of other panel members.

They must also consider the preferences of a circuit. When panels make errors or fail to comply with the preferences of the circuit as a whole, the circuit can review their decisions en banc (George 1999). If an author or panel were to create a legal policy sufficiently inconsistent with the preferences of the circuit, the other members could intervene through en banc review and move the policy away from the preferences of the author or panel and to the preferences of the circuit. Scholarship confirms this intuition: voting behavior on panels is conditioned on the preferences of the full circuit (Kim 2008; Kastellec 2011).

To account for the opportunities for strategic behavior within the courts of appeals, we include three variables. First, Author-Panel Distance is the ideological distance between the opinion author and the median member of the decision-making panel. Second, Author-Circuit Distance is the ideological distance between the author and the median member of the circuit. Finally, Panel-Circuit Distance is the ideological distance between the median member of the decision-making panel and the median member of the circuit. The inclusion of both Author-Circuit Distance and Panel-Circuit Distance reflects the potential that both the opinion author and panel median may affect the content of the legal policies articulated in the opinion of the court. To create these measures (and all measures involving ideology), we rely on Judicial Common Space scores (Epstein et al. 2007). These scores are particularly well suited for analyses such as ours, as they place lower-court judges, Supreme Court justices, and members of Congress into a common policy space.

Strategic courts of appeals judges must also consider the preferences of the Supreme Court. Just as the decisions of courts of appeals panels can be reviewed by the entire circuit en banc, the Supreme Court can review panel decisions through its writ of certiorari. If an author or panel were to create a legal policy sufficiently inconsistent with the preferences of the Supreme Court, members of the Supreme Court could intervene and move the policy away from the preferences of the author or panel and to the preferences of Court (see Clark [2009] for a discussion of the relationship between en banc and Supreme Court review).

To account for this opportunity to engage in strategic behavior, we include two additional variables that mirror those previously defined: Author-Court Distance is the ideological distance between the author and the median member of the Supreme Court.[1] Panel-Court Distance is the ideological distance between the median member of the decision-making panel and the median member of the Court. Again, both measures are included to reflect the potential influence of both the opinion author and the panel median.

Finally, strategic courts of appeals judges may also consider the preferences of Congress. Just as the circuit as a whole or the Supreme Court can review the decision of a panel, so too (theoretically) can Congress. Owens et al. (2013) offer some support for this possibility: they demonstrate that Supreme Court opinions are written less clearly when the ideological composition of Congress heightens the potential for review. While congressional review of courts of appeals decisions is rare (Lindquist and Yalof 2001), in an effort to be inclusive, we include four variables: Author-House Distance, which is the ideological distance between the author and the median member of the House;[2] Author-Senate Distance, which is the ideological distance between the author and the median member of the Senate; Panel-House Distance, which is the ideological distance between the median member of the panel and the median member of the House; and Panel-Senate Distance, which is the ideological distance between the median member of the panel and the median member of the Senate.

Judge Variables

The lack of scholarship on the relationship between judge characteristics and writing style makes our selection of those characteristics particularly challenging. In our efforts to be inclusive, we have identified as many plausible characteristics for which we could find consistent and reliable data. This is not to suggest that we expect each of these characteristics to individually affect how judges craft opinions; rather, we aim to develop as complete a picture as possible of the cumulative effects of judge characteristics on writing style. Most of the measures created to capture judge characteristics are derived from the Attributes of US Appeals Court Judges Database (Gryski and Zuk 2008). We outline these measures in table 1.

Case Variables

The lack of scholarship examining the effect of case characteristics on writing style likewise limits the strength of theoretical justification available for the inclusion of certain

1. We opt to use the Court median for the sake of consistency with our other measures. We do not take a position on the debate concerning what members control the policies created by the Court (see Anderson and Tahk [2007] and Carrubba et al. [2012] for discussion).

2. There are many different potential pivotal players in Congress (see Krehbiel 2010). We rely on the House and Senate medians for the sake of consistency with our other measures. We again take no position on scholarly debates about the US lawmaking process.

Table 1. Judge Factor Variables

Variable	Description	Min	Mean	Max
Appeals Judge	Court of appeals judge	0	.95	1
Ideology	Judicial Common Space score	−.70	−.02	.64
Republican	Identified as Republican	0	.51	1
Democratic President	Appointed by Democratic president	0	.47	1
Male	Male	0	.93	1
Nonwhite	Racial minority	0	.06	1
Age	Age	31	62	94
Experience	Years of experience	0	10.2	49
Undergraduate Ivy	Ivy legal undergraduate education	0	.23	1
Elite Law School	Elite law school education	0	.53	1
Graduate Degree	Possess a graduate degree	0	.14	1
Law Professor	Formerly a law professor	0	.22	1
Catholic	Identified as Catholic	0	.24	1
Jewish	Identified as Jewish	0	.14	1
Federal Prosecutor	Formerly a federal prosecutor	0	.31	1
State Prosecutor	Formerly a state prosecutor	0	.24	1
Private Practice	Formerly in private practice	0	.93	1
Region	US Census region of residence at appointment	. . .	. . .	. . .

measures. Owens and Wedeking (2011) provide some guidance by demonstrating that opinion cognitive complexity varies consistently across some legal issue areas. To account for this potential variance, we include dummy variables for eight general issue areas identified in the US Appeals Courts Database (Songer 2008). These include Criminal Case, Civil Rights Case, First Amendment Case, Due Process Case, Privacy Case, Economic Activity and Regulation Case, and Miscellaneous Case. Labor relations serves as the excluded category.

In addition to issue areas, we theorize that the circuit in which the case originated may affect writing styles. The types of legal disputes conceivably raised in the courts of appeals are not randomly distributed across the country; certain types of cases are undoubtedly more common in some judicial circuits than others. To account for this systematic variation, we include fixed effects for each circuit (save one).

Finally, we adopt a similar strategy when it comes to the time period in which the case originated. Just as the types of legal disputes raised in the courts of appeals are not randomly distributed across geographic location, they are also not randomly distributed over time. Different issues emerge on legal agendas, and changes in judicial interpretation or statutory law have the potential to shape the types of disputes courts of appeals judges are resolving. To account for this variation, we include fixed effects for each year in our data set (again, save one). While the strategy of including fixed effects for both circuit and year is an admittedly broad one for capturing case characteristics, it is consistent with our decision to be as inclusive as possible when considering variables that make up our influences on writing style. These fixed effects for time and circuit naturally absorb a great

deal of unmodeled heterogeneity. While one could argue that these fixed effects do account for properties of cases not covered by judge characteristics or strategic considerations, we describe below an alternative approach that excludes these fixed effects, to ameliorate concerns about our strategy.

Identifying Text Characteristics

Just as we adopt an intentionally inclusive approach to defining the potential influences on writing style, we do the same when identifying potentially relevant text characteristics. The literature to date incorporates several features. Prominently, many scholars (Coleman and Phong 2010; Owens et al. 2013; Goelzhauser and Cann 2014; Hansford and Coe 2014; Johnson 2014; Black et al. 2016b) focus on the reading difficulty of judicial texts, often termed "clarity." While specific formulas for calculating readability vary (the Flesch-Kincaid Grade Level and Coleman-Liau Index being among the most common), all these metrics focus on some combination of character, syllable, word, and sentence counts. Therefore, a text that on average contains longer sentences, more multisyllabic words, and a high ratio of characters to words will be considered more difficult to read. We use the Flesch-Kincaid Grade Level (FKGL) as a metric of reading difficulty in the analysis to follow.[3]

In addition to reading difficulty, scholars have analyzed other stylistic features of judicial texts. Technological advances, notably Linguistic Inquiry and Word Count (LIWC) software (e.g., Tausczik and Pennebaker 2010), allow for the consideration of dozens if not hundreds of theoretically distinct linguistic features. Most of these stylistic features involve some form of dictionary-based classification. A predetermined list of words assumed to adequately encompass some characteristic is compared to a text to determine the relative prevalence of that characteristic in the text. LIWC 2015 includes dictionaries for dozens of characteristics, including concepts such as affective processes, certainty, insight, and causation (Pennebaker et al. 2015).

Many scholars apply LIWC categories to judicial texts (Owens and Wedeking 2011; Bryan and Ringsmuth 2016). These scholars have variously considered categories such as affective (emotional) language (Black et al. 2016; Bryan and Ringsmuth 2016), negative language (Wedeking and Zilis 2015), certainty (Corley and Wedeking 2014), and, combining several categories, cognitive complexity (Owens and Wedeking 2011). For inclusivity, we analyze all of LIWC's categories. Generally speaking, these variables are the percentage of words in the preprocessed documents that fall in one of the predefined LIWC categories.

Additionally, for comparability with prior work (Owens and Wedeking 2011), we construct an index of cognitive complexity by factor analyzing 10 LIWC categories: causation, insight, discrepancy, inhibition, tentativeness, certainty, inclusiveness, exclusive-

3. The Flesch-Kincaid Grade Level (Kincaid et al. 1975) is calculated as FKGL = 0.39(total words / total sentences) + 11.8(total syllables / total words) − 15.59.

ness, negations, and percentage of words containing six or more letters. The first eigenvalue of this factor analysis was well above 1, while no other factors presented an eigenvalue above 1, indicating that a single factor captures this underlying construct well (Kaiser 1960). We use the 2007 LIWC dictionaries for this index for direct comparability with prior published work (i.e., Owens and Wedeking 2011).

We suggest that another textual characteristic, informality, deserves consideration as a textual feature as well.[4] (In)formality is a stylistic element that practitioners (e.g., Scalia and Garner 2008) emphasize. Writing about linguistic variation in general, Heylighen and Dewaele (1999, 2) state that the "most frequently mentioned" aspect of linguistic style is formality. Formality manifests itself in a variety of observable ways in written language, and scholars have taken several approaches toward quantifying a text's formality. Two approaches from the linguistics literature for measuring formality deserve brief mention. First, Brooke and Hirst (2014) take a vocabulary-based approach, presenting a method of ranking words by formality. A key practical advantage of the method is the low degree of supervision required: the analyst need not specify the absolute location of training words on a formality scale but only give the relative formality ranking of words in training-word pairs.[5]

Sheika and Inkpen (2012) alternatively propose several elements of formal and informal language that we also incorporate into our measure of informality. Informal style, in contrast to formal style, uses sentences written in active voice, intensifiers, phrasal verbs, contractions, emotional language, and idiomatic language. Elements of formal style include impersonal style (third person and passive voice), complex words and sentences, technical vocabularies, polite words and formulas, and objective style. Formal style avoids abbreviations, contractions, colloquialisms, vague expressions, and slang words (Sheika and Inkpen 2012, 6–8).

For this litany of features, we constructed regular expressions in the R programming language (R Core Team 2016) to count the instances of the passive voice, intensifying

4. Formal and informal categories exist in LIWC as well; we prefer our approach, as it allows us analytic control over the words and phrases included in the dictionaries. Additionally, we include in our index of informality several features that LIWC does not include, such as intensifying adverbs and phrasal verbs.

5. At least two resources—Hayakawa (1994) and Brooke and Hirst (2013)—provide the required pairwise rankings. Technically, Brooke and Hirst (2014) applies SVM^{rank} (Joachims 2002), which is a Support Vector Machine approach to ranking objects when only relative rankings of a subset of the objects are available (the original application was to rank web pages based on click-through data). Roughly speaking, the Brooke and Hirst (2014) approach is as follows. Select a set of *profile words*—words that appear in a "moderate proportion" of the texts in a corpus. Then, for each word in the training set, construct a *co-occurrence profile*—a vector where coordinate i is a normalized probability that the word occurs in the same text as the ith profile word. Last, find the vector of weights that maximizes the number of training-word pairs for which the inner product of the co-occurrence profile and the weight vector is greater for the more (in)formal word in the pair. All words in the corpus can then be ranked based on the inner product of this weight vector and the word's co-occurrence profile.

adverbs (such as "very"), phrasal verbs (based on the list of the 100 most common such verbs in Gardner and Davies [2007]), contractions, and idiomatic language (based on Spears 2008).[6] We then scaled these counts by the total number of words in an opinion and standardized these values across our sample. We measured emotional language as the absolute value of an opinion's sentimentality score, as calculated by the polarity function (which uses a voluminous dictionary of positive and negative words) in the qdap package (Rinker 2013) in the R programming language. We then factor analyzed all of these metrics to construct a single composite index, Informality.

DATA AND RESULTS

To identify a set of courts of appeals opinions, we use the Appeals Court Database (Songer 2008) and the Kuersten and Haire (2007) update.[7] Our sample covers the years 1950–2002 and includes 11,771 signed, published courts of appeals decisions.[8] The majority opinion serves as our unit of analysis. Our approach is straightforward. For every text characteristic we consider (reading difficulty, informality, cognitive complexity, and every characteristic in LIWC), we identify the amount of variance in that characteristic explained by strategic, judge, and case variables.

Our research design does not test the statistical significance of individual coefficients, a departure from prevailing norms in much of quantitative social science. Instead, we analyze how different factors account for some percentage of the total variability in our dependent variable. This strategy recalls Pritchett's (1941) efforts to explain the total variability in Supreme Court dissensus along left-wing/right-wing lines, as one example. Klingman and Lammers's (1984) analysis of how to best explain variance in the "general policy liberalism" of American state politics also fits squarely in the methodological tradition we contribute to here. Our emphasis on substantive explanation also draws inspiration from contemporary advice against overreliance on *p*-values as a sole basis for scientific conclusions (Wasserstein and Lazar 2016, 131).

Example Regression: Reading Difficulty

We provide an example of our approach in table 2, using reading difficulty as an example. Table 2 shows the results of several OLS regression models. Column 1 of table 2 regresses all of the strategic variables on clarity. Column 2 regresses all of our judge characteristics on

6. A phrasal verb is the combination of a verb and another element, usually an adverb or preposition, e.g., "break down" or "see to" (Gardner and Davies 2007).

7. The database is publicly available and is housed by the Judicial Research Initiative at the University of South Carolina. For more information about the coding procedures of the database, as well as a copy of the database itself, see http://www.cas.sc.edu/poli/juri/appct.htm.

8. We drop all majority opinions that are fewer than 1,500 bytes (approximately shorter than a brief paragraph), so as to have enough text in each opinion to meaningfully analyze. We do not include per curiam opinions, since we are interested in author attributes; we also drop a few cases owing to irregularities in the appeals court and attribute databases that we were not able to resolve despite careful cleaning.

Table 2. OLS Regression of Reading Difficulty on Strategic, Judge, and Case Characteristics, All Signed Majority Opinions

	Strategic Variables (1)	Judge Variables (2)	Case Variables (Issue Areas Only) (3)	Case Variables (Including Fixed Effects for Year, Circuit) (4)	All Variables (5)
Author-Panel Distance	−.108	. . .	. . .	. . .	−.057
	(.010)				(.09)
Author-House Distance	−.427	. . .	. . .	. . .	−.558*
	(.230)				(.250)
Author-Senate Distance	.250	. . .	. . .	. . .	.825*
	(.280)				(.350)
Author-Circuit Distance	−.197	. . .	. . .	. . .	−.037
	(.140)				(.140)
Author–Supreme Court Distance	.078	. . .	. . .	. . .	−.346
	(.220)				(.240)
Panel-House Distance	.249	. . .	. . .	. . .	.281
	(.230)				(.230)
Panel-Senate Distance	−.897*	. . .	. . .	. . .	−.583*
	(.280)				(.280)
Panel-Circuit Distance	−.529*	. . .	. . .	. . .	.108
	(.150)				(.140)
Panel–Supreme Court Distance	.822*	. . .	. . .	. . .	.420
	(.22)				(.220)
Appeals Judge	. . .	−1.562*	. . .	. . .	−1.570*
		(.200)			(.190)
Ideology	. . .	−.358*	. . .	. . .	−.135
		(.150)			(.170)
Republican	. . .	.162	. . .	. . .	.248*
		(.110)			(.110)
Nonwhite	. . .	−.046	. . .	. . .	.245*
		(.090)			(.090)
Jewish	. . .	.236*	. . .	. . .	.459*
		(.070)			(.070)
Male	. . .	−.003	. . .	. . .	−.308*
		(.080)			(.080)
Catholic	. . .	−.021	. . .	. . .	.148*
		(.050)			(.060)
Elite Law School	. . .	.072	. . .	. . .	.107*
		(.050)			(.050)
Graduate Degree	. . .	−.170*	. . .	. . .	−.115
		(.070)			(.070)
Law Professor	. . .	.012	. . .	. . .	.029
		(.050)			(.050)
Age	. . .	.008*	. . .	. . .	−.000
		(.000)			(.000)
Experience	. . .	−.024*	. . .	. . .	−.002
		(.000)			(.000)
Federal Prosecutor	. . .	.063	. . .	. . .	.088
		(.080)			(.080)

Table 2 (*Continued*)

	Strategic Variables (1)	Judge Variables (2)	Case Variables (Issue Areas Only) (3)	Case Variables (Including Fixed Effects for Year, Circuit) (4)	All Variables (5)
State Prosecutor	. . .	.095	. . .	. . .	.041
		(.090)			(.090)
Private Practice	. . .	.171*	. . .	. . .	.211*
		(.080)			(.080)
Democratic President	. . .	.081	. . .	. . .	.160
		(.150)			(.160)
Undergraduate Ivy	. . .	.131*	. . .	. . .	.112
		(.060)			(.060)
Northeasterner	. . .	.391*	. . .	. . .	.060
		(.060)			(.160)
Southerner	. . .	.257*	. . .	. . .	.107
		(.060)			(.100)
Westerner	. . .	−.406*	. . .	. . .	−.434*
		(.070)			(.160)
Criminal Case	. . .	. . .	−1.067*	−.950*	−.962*
			(.080)	(.080)	(.070)
Civil Rights Case	. . .	. . .	−.297*	−.114	−.102
			(.080)	(.080)	(.080)
First Amendment Case	. . .	. . .	−.049	.032	.038
			(.120)	(.120)	(.110)
Due Process Case	. . .	. . .	−.017	.079	.118
			(.150)	(.140)	(.140)
Privacy Case	. . .	. . .	−.180	−.002	−.048
			(.250)	(.240)	(.240)
Economic Activity and Regulation Case	. . .	. . .	.070	−.059	−.060
			(.070)	(.070)	(.070)
Miscellaneous Case	. . .	. . .	.211	.287*	.300*
			(.130)	(.130)	(.130)
Constant	14.574*	15.153*	14.796*	16.076*	17.210*
	(.060)	(.330)	(.070)	(.260)	(.450)
Adjusted R^2	.005	.045	.042	.083	.111

Note.—Dependent variable is Flesch-Kincaid Grade Level reading difficulty score. Higher values of the dependent variable indicate a less readable/clear writing style. Robust standard errors in parentheses. Circuit and year fixed effects included but not reported in models 4 and 5 for brevity. $N = 11{,}771$.

* $p < .05$.

clarity. Column 3 regresses our case issue area dummy variables on clarity. Column 4 includes issue area dummies as well as circuit fixed effects and year fixed effects (these latter effects are not presented for brevity). Column 5 includes all of these variables. We focus on adjusted R^2 as our quantity of interest in these five models. That is, taken together, how much total variance in reading difficulty does each category of cases account for?

Because the numbers of variables associated with strategic, judge, and case characteristics differ substantially, we assess the amount of variance explained by each using adjusted

R^2. Letting n be the number of observations and k the number of right-hand-side variables (including the constant), adjusted R^2 is defined as $1 - (1 - R^2)(n - 1)/(n - k)$; since this quantity decreases in k, there is a penalty for the inclusion of extra covariates in a model. This penalty ensures that one set of variables will not receive credit for explaining a higher proportion of the variance simply because of the sheer number of covariates included in the model.[9]

Table 2 shows that we cannot explain even a fifth of the total variation in reading difficulty with all of these covariates combined. Further, when we pay specific attention to the strategic variables, these covariates, taken together, explain an order of magnitude less variation than do case characteristics. Moreover, the individual coefficients on most of these variables fail to attain statistical significance or are signed in the wrong direction.

Of course, some courts of appeals cases are not "ripe" for strategy. These courts regularly dispose of cases that are routine, noncontroversial, and not especially prone to strategic ideological machinations (Bowie and Songer 2009). However, there are circumstances under which courts of appeals judges should be more likely to behave strategically. Following a similar logic to Owens et al. (2013, 35), which concludes that Supreme Court justices strategically obfuscate the language of majority opinions to avoid unfavorable review from Congress, we theorize that opinion authors will write less clearly when the probability that their decision will be reviewed (en banc, by the Supreme Court, or by Congress) is greatest. As such, table 3 replicates the reading difficulty analysis of table 2 but only on a subset of cases in our data that are more likely to be reviewed. We define cases likely to be reviewed somewhat broadly: any case that features an amicus brief or a dissent or that reverses the trial court is included in this subset. Despite narrowing our focus to cases that more likely contain genuine controversy and disagreement, table 3 shows a strikingly similar pattern as table 2. Our strategic variables, taken together, explain very little variation in reading difficulty, and the individual coefficients (which should all be signed positively under a strategic model of opinion writing) are mostly not statistically significant, and several are signed in the wrong direction. For reading difficulty, at least, strategic calculations do not influence writing style in the circuits in the same way observed by other scholars in the Supreme Court (Black et al. 2016b).

All Signed Majority Opinions

We continue with an analysis of all majority opinions in the data set, applying our approach exemplified above to nearly 100 different stylistic features. Table 4 shows the mean and median adjusted R^2 from the 98 sets of regressions—one set for each of our 98 text features, with each set including a regression of a given text characteristic on strategic, judge, and case variables, respectively. The results demonstrate that most of the variance

9. The results are substantively robust to the use of conventional R^2 (which does not penalize for additional covariates) to assess the quantity of variance explained.

Table 3. OLS Regression of Reading Difficulty on Strategic, Judge, and Case Characteristics, Strategy-Ripe Majority Opinions Only

	Strategic Variables	Judge Variables	Case Variables (Issue Areas Only)	Case Variables (Including Fixed Effects for Year, Circuit)	All Variables
Author-Panel Distance	−.258*	. . .	. . .	. . .	−.236
	(.130)				(.120)
Author-House Distance	−.448	. . .	. . .	. . .	−.470
	(.300)				(.330)
Author-Senate Distance	.293	. . .	. . .	. . .	.654
	(.380)				(.480)
Author-Circuit Distance	−.262	. . .	. . .	. . .	.403*
	(.190)				(.180)
Author–Supreme Court Distance	.017	. . .	. . .	. . .	−.341
	(.300)				(.340)
Panel-House Distance	.159	. . .	. . .	. . .	.347
	(.310)				(.310)
Panel-Senate Distance	−.695	. . .	. . .	. . .	−.564
	(.390)				(.390)
Panel-Circuit Distance	−.816*	. . .	. . .	. . .	−.421*
	(.200)				(.190)
Panel–Supreme Court Distance	.902*	. . .	. . .	. . .	.699*
	(.290)				(.280)
Appeals Judge	. . .	−1.120*	. . .	. . .	−1.115*
		(.290)			(.280)
Ideology	. . .	−.351	. . .	. . .	−.076
		(.220)			(.250)
Republican	. . .	.034	. . .	. . .	.127
		(.140)			(.130)
Nonwhite	. . .	−.001	. . .	. . .	.226*
		(.120)			(.110)
Jewish	. . .	.147	. . .	. . .	.380*
		(.090)			(.090)
Male	. . .	.014	. . .	. . .	−.311*
		(.120)			(.090)
Catholic	. . .	−.025	. . .	. . .	.137
		(.080)			(.080)
Elite Law School	. . .	.087	. . .	. . .	.111
		(.070)			(.070)
Graduate Degree	. . .	−.194*	. . .	. . .	−.139
		(.090)			(.090)
Law Professor	. . .	.090	. . .	. . .	.098
		(.070)			(.070)
Age	. . .	.007	. . .	. . .	−.002
		(.010)			(.010)
Experience	. . .	−.020*	. . .	. . .	.002
		(.010)			(.010)
Federal Prosecutor	. . .	.063	. . .	. . .	.083
		(.110)			(.120)

Table 3 (*Continued*)

	Strategic Variables	Judge Variables	Case Variables (Issue Areas Only)	Case Variables (Including Fixed Effects for Year, Circuit)	All Variables
State Prosecutor	. . .	.123	. . .	. . .	.092
		(.130)			(.013)
Private Practice	. . .	.138	. . .	. . .	.190
		(.110)			(.110)
Democratic President	. . .	−.036	. . .	. . .	.037
		(.210)			(.200)
Undergraduate Ivy	. . .	.132	. . .	. . .	.113
		(.080)			(.080)
Northeasterner	. . .	.288*	. . .	. . .	−.128
		(.090)			(.210)
Southerner	. . .	.149	. . .	. . .	−.028
		(.090)			(.140)
Westerner	. . .	−.458*	. . .	. . .	−.449*
		(.100)			(.210)
Criminal Case	. . .	. . .	−1.043*	−.926*	−.938*
			(.100)	(.100)	(.100)
Civil Rights Case	. . .	. . .	−.361*	−.169	−.156
			(.110)	(.100)	(.100)
First Amendment Case	. . .	. . .	.036	.128	.144
			(.140)	(.130)	(.130)
Due Process Case	. . .	. . .	.031	.060	.060
			(.220)	(.220)	(.220)
Privacy Case	. . .	. . .	.206	.317	.252
			(.350)	(.330)	(.330)
Economic Activity and Regulation Case	. . .	. . .	−.104	−.082	−.084
			(.100)	(.100)	(.100)
Miscellaneous Case	. . .	. . .	.332	.404*	.418*
			(.190)	(.190)	(.190)
Constant	14.585*	15.008*	14.836*	16.027*	16.876*
	(.080)	(.470)	(.090)	(.370)	(.630)
Adjusted R^2	.004	.029	.035	.072	.086

Note.—Dependent variable is Flesch-Kincaid Grade Level reading difficulty score. Higher values of the dependent variable indicate a less readable/clear writing style. Robust standard errors in parentheses. Circuit and year fixed effects included but not reported in models 4 and 5 for brevity. $N = 5{,}748$.

* $p < .05$.

in writing style remains unexplained despite our intentionally inclusive variable selection process. No set of variables explains a substantial amount of the variance in writing style. In fact, the collective impact of all covariates is limited; variables related to the case, judge, and strategic opportunities together explain approximately one-tenth of the total variance in writing style. These preliminary results suggest that despite the many advances in measuring writing style, most of the variance in the actual content of legal opinions remains unexplained.

Table 4. Average Adjusted R^2 (across 98 Regressions) for All Text Characteristics, Signed Majority Opinions ($N = 11{,}771$)

	Judge Variables	Case Variables	Strategic Variables
Mean	.016	.083	.003
Median	.008	.058	.001
SD	.017	.073	.003

Despite the limited variance that our covariates collectively explain, there are meaningful differences in the relative amount of variance that our sets of covariates explain, as shown in table 4. The largest substantive gap is between the variance explained by case variables and all other potential explanations. On average, case variables explain just over 8% of the variance in writing style. While not large in absolute magnitude, this represents a substantial increase compared to judge and strategic variables. Judge variables explain a little over 1% of the variation in writing style, while strategic variables explain less than one-half of 1% of the variation in writing style. Even measured crudely by issue area, year, and circuit, case variables explain about 25 times as much variance in writing style, on average, compared to variables associated with strategic behavior.

Table 4 indicates only a general pattern across dozens of textual features. Naturally, not all of these features are of equal import. Thus, figure 1 shows the adjusted R^2 associated with regressing the most theoretically relevant features in the literature—reading difficulty, cognitive complexity, informality, and affective language—on judge, case, and strategic variables, respectively. This more focused exercise explores whether the patterns in table 4 hold for the text features that are particularly meaningful theoretically.

The general pattern observed in table 4 is largely replicated in figure 1. Case variables explain the most variance for three of the four text characteristics. Judge characteristics explain a higher proportion of the variance in cognitive complexity. This pattern may not be entirely surprising—cognitive complexity is theoretically distinct from other elements of writing style, as it seeks to measure the complexity of ideas employed (Owens and Wedeking 2011, 1038). For the three elements associated exclusively with writing style—affective language, informality, and reading difficulty—variables related to the case explain substantially more variance. For each of these characteristics, case variables explain more than twice as much variance as variables associated with the judge or strategic considerations. For all four text characteristics (including cognitive complexity), case variables explain more variance than strategic considerations by an order of magnitude or more.

The results in table 4 and figure 1 suggest that the most important determinants of writing style are case variables. These variables consist of fixed effects for year, circuit, and issue area. A skeptical reader may question whether year and circuit truly represent characteristics of the case itself; is it possible that the quantity of variance explained by case variables is primarily a function of these two components? To address this concern, refer to table 5. The mean and median adjusted R^2 for the regressions of judge and strategic var-

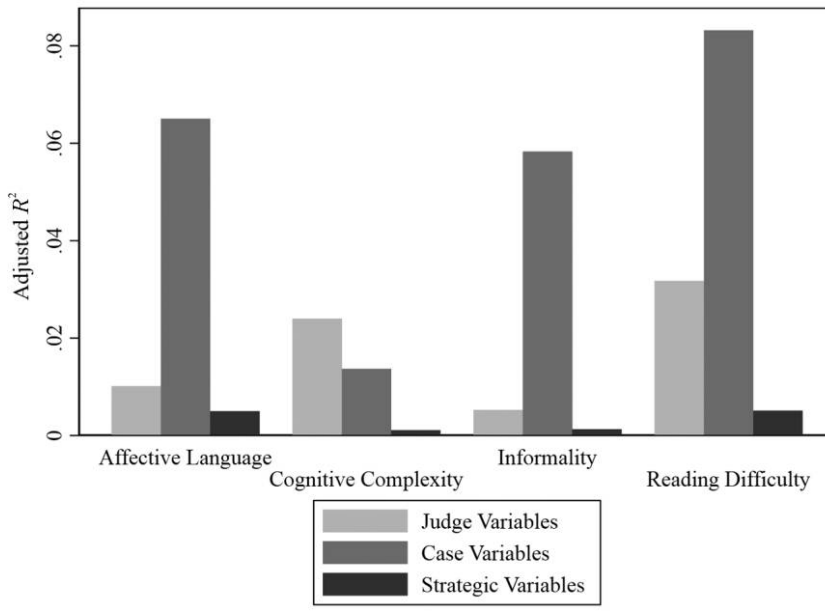

Figure 1. Adjusted R^2 for four common textual features, all signed majority opinions

iables are unchanged. However, the regression for the case variables model now only includes fixed effects for issue area; both year and circuit are excluded from the analysis. While the total quantity of variance explained by issue area alone is lower, on average, it still explains substantially more variance than judge and strategic variables combined. A similar pattern holds for reading difficulty, cognitive complexity, informality, and affective language (not displayed). As with the more expansive definition of case variables, issue areas explain less variance than judge characteristics in cognitive complexity but more variance in informality and reading difficulty. However, judge variables explain more variance in affective language than issue areas alone. All of the results, taken together, suggest that case variables explain the greatest portion of the variance in judicial writing style, although the finding is less definitive if we define case variables narrowly.

Table 5. Average Adjusted R^2 (across 98 Regressions) for All Text Characteristics, Signed Majority Opinions: Issue Area Only ($N = 11{,}771$)

	Judge Variables	Case Variables (Issue Area Only)	Strategic Variables
Mean	.016	.043	.003
Median	.008	.019	.001
SD	.017	.056	.003

Majority Opinions with Separate Opinions

Given the apparent influence of case variables on writing style, we examine their role further through an analysis of separate opinion writing. While fixed effects for year, circuit, and issue area capture many characteristics of cases, there are surely factors idiosyncratic to each individual case or narrow issue area or topic that are out of reach of such measures. To capture these more idiosyncratic factors, we compare writing style within cases by comparing majority and separate opinions authored in a common decision. Adopting this approach is useful for isolating the impact of case variables from other potential motivations and capturing case factors that cannot be captured by the blunter measures that are available.

We note that the motivation for strategic behavior differs substantially for separate opinion writers. Authors of concurrences and dissents do not typically need to worry about whether their legal and policy position is ideologically acceptable to other actors as majority authors might. Separate opinion writing also varies in terms of judge characteristics; separate opinion authors frequently have different individual characteristics than the majority opinion author. However, we know that separate opinions and a majority opinion share a common set of (unobserved) case variables. Comparing these types of opinions thus allows the most direct test of the effect of case variables on writing style.

To assess the magnitude of this effect, we follow a similar approach as outlined above. For all text characteristics, we regress judge variables, case variables, and strategic variables and report the adjusted R^2 values. We differ from our previous approach in two respects. First, we necessarily restrict our analysis to only cases that include at least one separate opinion. While this approach limits our sample size considerably—the bulk of majority opinions in the courts of appeals fail to inspire any accompanying separate opinions—we believe the inferences that can be drawn from this unique subsample of cases warrant narrowing the set of observations analyzed. Second, we include the adjusted R^2 from a fourth regression: one that regresses the separate opinion's value of a text characteristic on that characteristic for the majority opinion. If the text characteristic of separate opinion(s) explains variance in the text characteristics of the majority opinion, then this explained variance can reasonably be attributed to shared idiosyncratic case factors.

Table 6 presents the results for the 1,283 signed opinions accompanied by at least one separate opinion. As was the case for the larger sample that included unanimous opinions,

Table 6. Average Adjusted R^2 (across 98 Regressions) for All Text Characteristics, Signed Majority Opinions with Separate Opinions ($N = 1{,}283$)

	Judge Variables	Case Variables	Strategic Variables	Textual Feature in Separate Opinion(s)
Mean	.018	.112	.005	.220
Median	.015	.088	.003	.173
SD	.016	.080	.007	.175

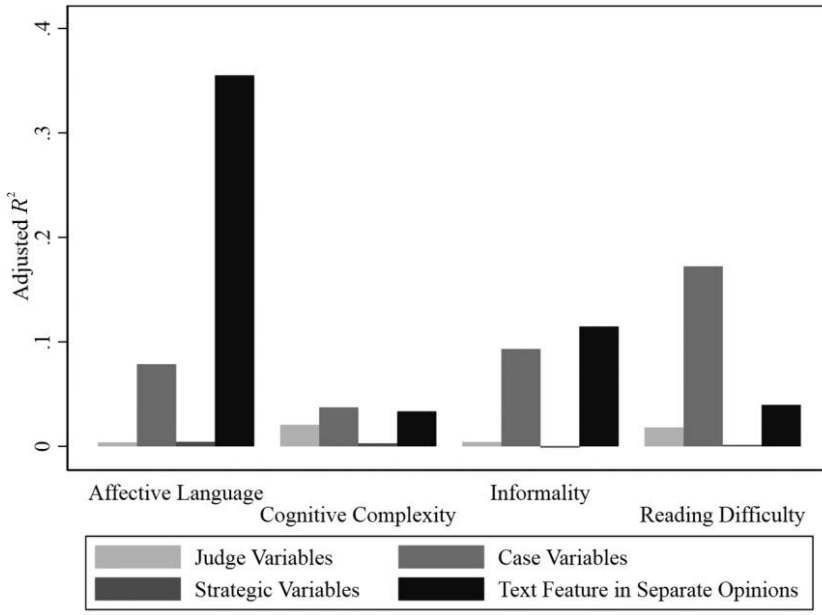

Figure 2. Adjusted R^2 for four common textual features, signed majority opinions with separate opinions.

for majority opinions that attracted at least one separate opinion, case variables (fixed effects for year, circuit, and issue area) explain substantially more variance, on average, than do judge or strategic variables.[10] Most importantly, the text characteristics of separate opinions explain the largest quantity of variance in writing style. The observed magnitude is noteworthy. For example, the text characteristics of the separate opinion(s) explain nearly 17 times the amount of variance in the text characteristic of majority opinions as do judge characteristics and 40 times the amount of variance as do strategic considerations. Additionally, our blunter measure of case variables accounts for only about half as much variance in textual features of majority opinions, compared with the same text characteristic in separate opinions.

To probe the existence of these patterns in the most theoretically relevant text characteristics, figure 2 shows the adjusted R^2 values of regressions of reading difficulty, cognitive complexity, informality, and affective language on judge, case, and strategic variables, as well as the corresponding text characteristic in the case's separate opinion(s), respectively. The results demonstrate that for these four commonly used text characteristics, variables related to the case explain far more variance than those relating to the judge or strategic

10. This result and the other substantive effects discussed below are all robust to the alternative specification of case variables in which fixed effects for issue areas are the only variables included in the model.

behavior. Levels of informality and affective language in separate opinion(s) explain the greatest portion of variance in the informality and affective language in corresponding majority opinions. This is particularly noteworthy for affective language; the level of affective language used in a separate opinion explains approximately 35% of the variance in the levels of affective language in the majority opinion. For both text characteristics, our blunter measure of case variables (i.e., year, circuit, and issue area) explains the second-largest proportion of variance. For cognitive complexity and reading difficulty, the opposite pattern occurs: case variables explain the largest portion of the variance, with the levels of cognitive complexity and reading difficulty, respectively, in the separate opinion accounting for the second-largest proportion of variance. As figure 2 demonstrates, while some differences exist across different text characteristics, a common theme emerges: case variables explain the largest portion of variance in judicial writing style. Judge variables and variables capturing strategic behavior explain comparatively less.

CONCLUSION

Automated text analysis now commands great interest from scholars of judicial politics. This level of interest is unquestionably warranted; the ability to systematically classify the content of judicial opinions creates the opportunity to test hypotheses previously out of reach in large samples. Many scholars have taken advantage of these opportunities, and our collective understanding of judicial behavior has grown as a result (Owens and Wedeking 2011; Corley and Wedeking 2014; Goelzhauser and Cann 2014; Nelson 2014; Black et al. 2016, 2016b; Bryan and Ringsmuth 2016).

In the pursuit of testing these novel hypotheses, it is easy to overlook foundational questions central to this burgeoning literature. In this analysis, we address what is among the most fundamental questions to consider: What determines the writing styles used to craft judicial opinions? Most judicial scholarship using automated text analysis examines strategic choices in writing style. Our results suggest that much of the variance observed in writing styles remains unexplained, even when accounting for an inclusive array of potential explanatory variables. The explained variance in stylistic differences does not appear to be driven by strategic variables, nor is it a function of judge characteristics. Rather, it appears that variables associated with the case are the most substantial influence. This result does not call into question the results of previous scholarly endeavors; we instead hope to have provided a greater understanding of both the potential promise and limitations of text analysis of judicial opinions.

In that spirit, we close by identifying several questions important to the application and interpretation of text analysis in the context of judicial opinions. We also offer some preliminary reactions to these questions in the hope of continuing an important dialogue in what is clearly a fruitful avenue for future research by judicial politics scholars.

First, what is the substantive impact of judicial writing style? Previous scholarship examining text characteristics has found statistically significant associations between these

characteristics and a variety of behaviors associated with strategic calculations. However, what is the substantive impact of these relationships? Our results here suggest that the substantive differences in style explained by strategic variables are likely small. Scholars working in this area might consider paying special attention to the substantive, rather than statistical, significance of these relationships. Specifically, one might present changes in predicted outcomes or probabilities over plausible ranges of a text variable, or one might present measures of model fit from specifications that include only the text variable predictor(s). If strategic calculations explain less than one-half of 1% of a document's reading difficulty or affective tone, can such a difference be perceived by the judge's audience? If not, what does this say about the likelihood that these differences arise because of a judge's strategic decision? To advance the field of judicial politics with respect to text analysis of judicial opinions, these questions require substantial attention.

Second, could excessive noise in measures of writing style account for the relatively low variance we see explained? If so, there are several standard recommendations to address such concerns. Scholars should pause before adopting measures of writing style developed in contexts other than legal opinions and consider whether modification is appropriate before application to judicial texts. Relatedly, one ought to consider whether meaningful systematic variation in the measure is likely to exist across legal opinions, which, we suspect, occupy a relatively narrow stylistic range across some attributes, given authors' shared training. When possible, use multiple indicators for a single construct, and ensure that the individual measures load onto a single factor with an eigenvalue above 1 (Kaiser 1960). Preprocessing decisions may also affect the amount of noise present in the measure (Denny and Spirling 2018). Ensure that the measures chosen specifically match the context being analyzed (Grimmer and Stewart 2013; Benoit, Munger, and Spirling 2017). Move beyond emphasizing statistical significance in reporting results, examining model fit and effect size carefully (Gelman and Stern 2006). Last, avoid small samples when possible: in this context, noisy measures can lead to substantively large (and statistically significant) apparent effect sizes even when the true effect is tiny (Loken and Gelman 2017).

Researchers should not necessarily abandon the search to explain variability in judicial writing style. Style may truly be largely stochastic. However, it may also be the case that better measurement of key covariates could reveal a greater systematic component to writing style. For instance, our measurement of legal issue area is admittedly coarse. Future scholars could consider refining this variable: beyond a broad, substantive issue area, how do particular legal topics, doctrinal areas, and case facts relate to writing style? Alternatively, perhaps judicial writing style does emanate from a judge's innate skill. Better measuring the training, experience, and talent of a judge could potentially explain a great deal more of the variability in style. Future work should entertain these possibilities.

Third, how does the use law clerks affect the measurement of writing style? Many judges use the assistance of law clerks when drafting opinions (Ward and Weiden 2006, 200–237; Bowie, Songer, and Szmer 2014, 92–93). While the use of law clerks could take a variety of forms, evidence suggests that clerks have played an increasingly prominent

role in the actual writing of opinions over time. At the Supreme Court, according to Ward and Weiden (2006, 212), "the most common form [of opinion writing process] is *delegation*, which involves a justice assigning the first draft of an opinion to a clerk and then revising the clerk's original draft." A similar pattern appears to occur in the courts of appeals. According to Bowie, Songer, and Szmer (2014, 92), "Many judges ask their clerk to write a first draft of the opinion, after providing guidance about the direction the opinion should take."

This working arrangement leads judges to cede some control over the style in which the opinion is written to clerks. As the author of the first draft, clerks necessarily play an important role in the selection of language and grammatical structure used in judicial opinions. Even if a judge "heavily edits" the initial draft of opinions (Bowie, Songer, and Szmer 2014, 93), it is important to determine whether those edits are directed toward style, substantive content, or both. Surely these edits must include concerns about substantive content; it is unlikely that the delegation of the early opinion drafting would lead to the dissemination of opinions containing substantive content that is inconsistent with judges' preferences. But do judges instruct clerks as to writing styles? Are the judges' edits heavy enough to change not only the substantive content of the opinion but the style in which it was originally drafted? Future examinations of judicial writing style should consider the role that law clerks play in the opinion writing process.

Next, how broadly should scholars conceive of judicial writing style? Several distinct elements of writing style have already emerged in the literature, with more to surely come as the quality of automated text analysis improves. Scholars have already examined affective language (Bryan and Ringsmuth 2016), certainty (Corley and Wedeking 2014), cognitive complexity (Owens and Wedeking 2011), and reading difficulty (Owens et al. 2013; Black et al. 2016a, 2016b), the last of which has received a disproportionate share of scholarly attention. While specific hypotheses may dictate which element of writing style warrants investigation in a specific instance, the literature is largely silent in offering a broader justification for prioritizing certain elements of writing style over others. For example, is reading difficulty a more important element of writing style than certainty, affective language, or informality, as the current literature seems to imply? A more complete theoretical development of how these elements fit together, and what the relationship between these elements means for judicial opinion writing, is warranted.

Finally, what are the costs of writing in specific styles? Strategic explanations of judicial opinion writing suggest that judges vary their writing style to achieve specific goals. In many analyses, judges are hypothesized to alter an element of writing style, under certain conditions, to affect the behavior of others. If, for example, decreased reading difficulty increases the likelihood of lower-court compliance with Supreme Court opinions (Black et al. 2016b, 141–55), why would justices not always minimize the reading difficulty of an opinion to help achieve compliance? The most apparent explanation must be that there is a cost associated with changing writing styles. Currently, this answer is largely assumed rather than tested. Further examinations of writing style should explore

the cost of writing in particular styles in greater detail to see whether opinions written in a particular style are in fact more costly.[11] If changing writing styles is costless, explanations as to why judges do not always maximize the strategic impact of their writing style are needed.

REFERENCES

Anderson, Robert, and Alexander Tahk. 2007. "Institutions and Equilibrium in the United States Supreme Court." *American Political Science Review* 101 (4): 811–25.

Baum, Lawrence. 1997. *The Puzzle of Judicial Behavior*. Ann Arbor: University of Michigan Press.

Benoit, Ken, Kevin Munger, and Arthur Spirling. 2017. "Measuring and Explaining Political Sophistication through Textual Complexity." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3062061.

Black, Ryan C., Matthew E. K. Hall, Ryan J. Owens, and Eve M. Ringsmuth. 2016. "The Role of Emotional Language in Briefs before the US Supreme Court." *Journal of Law and Courts* 4 (2): 377–407.

Black, Ryan C., Ryan J. Owens, Justin Wedeking, and Patrick C. Wohlfarth. 2016a. "The Influence of Public Sentiment on Supreme Court Opinion Clarity." *Law and Society of Review* 50 (3): 703–32.

——— 2016b. *U.S. Supreme Court Opinions and Their Audiences*. Cambridge: Cambridge University Press.

Bowie, Jennifer Barnes, and Donald R. Songer. 2009. "Assessing the Applicability of Strategic Theory to Explain Decision Making on the Courts of Appeals." *Political Research Quarterly* 62 (2): 393–407.

Bowie, Jennifer Barnes, Donald R. Songer, and John Szmer. 2014. *The View from the Bench and Chambers: Examining Judicial Process and Decision Making on the U.S. Courts of Appeals*. Charlottesville: University of Virginia Press.

Brooke, Julian, and Graeme Hirst. 2013. "Hybrid Models for Lexical Acquisition of Correlated Styles." Paper presented at the 6th International Joint Conference on Natural Language Processing, Nagoya, Japan, October 14–18. http://aclweb.org/anthology/I13-1000.

———. 2014. "Supervised Ranking of Co-occurrence Profiles for Acquisition of Continuous Lexical Attributes." Paper presented at the 25th International Conference on Computational Linguistics, Dublin, August 23–29. http://www.aclweb.org/anthology/C14-1000.

Bryan, Amanda C., and Eve M. Ringsmuth. 2016. "Jeremiad or Weapon of Words? The Power of Emotive Language in Supreme Court Dissents." *Journal of Law and Courts* 4 (1): 159–85.

Carrubba, Cliff, Barry Friedman, Andrew D. Martin, and Georg Vanberg. 2012. "Who Controls the Content of Supreme Court Opinions?" *American Journal of Political Science* 56 (2): 400–412.

Clark, Tom S. 2009. "A Principal-Agent Theory of En Banc Review." *Journal of Law, Economics, and Organization* 25 (1): 55–79.

11. Scholars have recognized potential trade-offs in the sense that changes in writing style may require judges to pursue the expense of one goal at the expense of another (see Owens et al. [2013] for a discussion about the trade-off between lower-court compliance and avoiding congressional review). However, there has been little discussion as to why judges would vary their writing style if they are pursuing only one goal in the opinion (e.g., lower-court compliance without considering the probability of congressional review).

Coleman, Brady, and Quy Phong. 2010. "The Language of Supreme Court Briefs: A Large Scale Quantitative Investigation." *Journal of Appellate Practice and Process* 11 (1): 75–103.

Corley, Pamela, and Justin Wedeking. 2014. "The (Dis)advantage of Certainty: The Importance of Certainty in Language." *Law and Society Review* 48 (1): 35–62.

Corney, Malcolm, Olivier de Vel, Alison Anderson, and George Mohay. 2002. "Gender-Preferential Text Mining of E-mail Discourse." Paper presented at the 2002 Annual Computer Security Applications Conference, Las Vegas, December.

Denny, Matthew J., and Arthur Spirling. 2018. "Text Preprocessing for Unsupervised Learning: Why It Matters, When It Misleads, and What to Do about It." *Political Analysis* 26 (2): 168–89.

Epstein, Lee, and Jack Knight. 1998. *The Choices Justices Make*. Washington, DC: Congressional Quarterly.

Epstein, Lee, Andrew D. Martin, Jeffrey A. Segal, and Chad Westerland. 2007. "The Judicial Common Space." *Journal of Law, Economics, and Organization* 23 (2): 303–25.

Gardner, Dee, and Mark Davies. 2007. "Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis." *TESOL Quarterly* 41 (2): 339–59.

Gelman, Andrew, and Hal Stern. 2006. "The Difference between 'Significant' and 'Not Significant' Is Not Itself Statistically Significant." *American Statistician* 60 (4): 328–31.

George, Tracey E. 1999. "The Dynamics and Determinants of the Decision to Grant En Banc Review." *Washington Law Review* 74 (1): 213–74.

Goelzhauser, Greg, and Damon M. Cann. 2014. "Judicial Independence and Opinion Clarity on State Supreme Courts." *State Politics and Policy Quarterly* 14 (2): 123–41.

Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21 (3): 267–97.

Gryski, Gerald S., and Gary Zuk. 2008. A Multi-User Data Base on the Attributes of U.S. Appeals Court Judges, 1801–2000. http://artsandsciences.sc.edu/poli/juri/attributes.htm.

Hansford, Thomas G., and Chelsea Coe. 2014. "Linguistic Complexity and Public Acceptance of Supreme Court Decisions." Paper presented at the Annual Meeting of the American Political Science Association, Washington, DC, August.

Hayakawa, S. I. 1994. *Choose the Right Word: A Contemporary Guide to Selecting the Precise Word for Every Situation*. 2nd ed. Revised by Eugene Ehrlich. New York: HarperCollins.

Heylighen, Francis, and Jean-Marc Dewaele. 1999. "Formality of Language: Definition, Measurement, and Behavioral Determinants." Internal report, Center "Leo Apostel," Free University of Brussels.

Joachims, Thorsten. 2002. "Optimizing Search Engines Using Clickthrough Data." In *Proceedings of the 9th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*. New York: Association for Computing Machinery.

Johnson, Stephen N. 2014. "The Changing Discourse of the Supreme Court." *University of New Hampshire Law Review* 12 (1): 29–68.

Kaiser, Henry F. 1960. "The Application 0of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20 (1): 141–51.

Kastellec, Jonathon P. 2011. "Hierarchical and Collegial Politics on the U.S. Courts of Appeals." *Journal of Politics* 73 (2): 345–61.

Kim, Pauline T. 2008. "Deliberation and Strategy on the United States Courts of Appeals: An Empirical Exploration of Panel Effects." *University of Pennsylvania Law Review* 157 (5): 1319–81.

Kincaid, J. Peter, Robert P. Fishburne, Richard L. Rogers, and Brad S. Chissom. 1975. "Derivation of New Readability Formulas (Automated Readability Index, Fog Count, and Flesch Reading Ease Formula) for Navy Enlisted Personnel." Defense Technical Information Center Document Research Branch Report 8-75.

Klingman, David, and William W. Lammers. 1984. "The 'General Policy Liberalism' Factor in American State Politics." *American Journal of Political Science* 28 (3): 598–610.

Koppel, Moshe, Shlomo Argamon, and Anat Rachel Shimoni. 2002. "Automatically Categorizing Written Texts by Author Gender." *Literary and Linguistic Computing* 17 (4): 401–12.

Krehbiel, Keith. 2010. *Pivotal Politics: A Theory of U.S. Lawmaking*. Chicago: University of Chicago Press.

Kuersten, Ashlyn, and Susan Haire. 2007. Update to the Appeals Courts Database, 1997–2002. http://artsandsciences.sc.edu/poli/juri/appct.htm.

Li, William, Pablo Azar, David Larochelle, Phil Hill, James Cox, Robert C. Berwick, and Andrew W. Lo. 2013. "Using Algorithmic Attribution Techniques to Determine Authorship in Unsigned Judicial Opinions." *Stanford Technology Law Review* 16 (3): 503–34.

Lindquist, Stefanie A., and David A. Yalof. 2001. "Congressional Responses to Federal Circuit Court Decisions." *Judicature* 85 (1): 61–68.

Loken, Eric, and Andrew Gelman. 2017. "Measurement Error and the Replication Crisis." *Science* 355 (6325): 584–85.

Maltzman, Forrest, James F. Spriggs II, and Paul J. Wahlbeck. 2000. *Crafting Law on the Supreme Court: The Collegial Game*. New York: Cambridge University Press.

Nelson, Michael J. 2014. "Elections and Explanations: Judicial Retention and the Readability of Judicial Opinions." Working paper. http://mjnelson.org/papers/NelsonReadability.pdf.

Owens, Ryan, and Justin Wedeking. 2011. "Justices and Legal Clarity: Analyzing the Complexity of U.S. Supreme Court Opinions." *Law and Society Review* 45 (4): 1027–61.

Owens, Ryan, Justin Wedeking, and Patrick Wohlfarth. 2013. "How the Supreme Court Alters Opinion Language to Evade Congressional Review." *Journal of Law and Courts* 1 (1): 35–59.

Pennebaker, James W., Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015 "The Development and Psychometric Properties of LIWC2015." Texas ScholarWorks. https://repositories.lib.utexas.edu/handle/2152/31333.

Poole, Keith T., and Howard Rosenthal. 1997. *Congress: A Political-Economic History of Roll-Call Voting*. Oxford: Oxford University Press.

Pritchett, C. Herman. 1941. "Divisions of Opinion among Justices of the U. S. Supreme Court." *American Political Science Review* 35 (5): 890–98.

———. 1948. *The Roosevelt Court: A Study in Judicial Politics and Values, 1937–1947*. New York: Macmillan.

R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rinker, Tyler W. 2013. Quantitative Discourse Analysis Package. Buffalo, NY: University at Buffalo/SUNY. Version 2.2.0.

Rohde, David W., and Harold J. Spaeth. 1976. *Supreme Court Decision Making*. San Francisco: Freeman.

Scalia, Antonin, and Bryan A. Garner. 2008. *Making Your Case: The Art of Persuading Judges*. St. Paul, MN: Thomson West.

Segal, Jeffrey A., and Harold J. Spaeth. 2002. *The Supreme Court and the Attitudinal Model Revisited*. New York: Cambridge University Press.

Sheika, Fadi Abu, and Diana Inkpen. 2012. "Learning to Classify Documents according to Formal and Informal Style." *Linguistic Issues in Language Technology* 8 (1): 1–29.

Songer, Donald R. 2008. The Original U.S. Appeals Courts Database, 1925–1996. http://artsandsciences.sc.edu/poli/juri/appct.htm.

Spears, Richard A. 2008. *McGraw-Hill's Essential American Idioms Dictionary*. 2nd ed. Dubuque, IA: McGraw-Hill.

Tausczik, Yla R., and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods." *Journal of Language and Social Psychology* 29 (1): 24–54.

Ward, Artemus, and David L. Weiden. 2006. *Sorcerers' Apprentices: 100 Years of Law Clerks at the United States Supreme Court*. New York: New York University Press.

Wasserstein, Ronald L., and Nicole A. Lazar. 2016. "The ASA's Statement on *p*-Values: Context, Process, and Purpose." *American Statistician* 70 (2): 129–33.

Wedeking, Justin, and Michael Zilis. 2015. "The Strategic Use of Rhetoric: Disagreeable Language in Supreme Court Opinions." https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2628819.

Zheng, Rong, Jiexun Li, Hsinchun Chen, and Zan Huang. 2006. "A Framework for Authorship Identification of Online Messages: Writing-Style Features and Classification Techniques." *Journal of the American Society for Information Science and Technology* 57 (3): 378–93.