Writing Style and Persuasion on the U.S. Circuit Courts of Appeals*

Jeffrey Budziak† Daniel Lempert‡

2022 APSA Meeting Version

Abstract

Judges regularly try to shape legal policy by influencing the behavior of others. The primary way they can do so is through their legal opinions. We investigate whether an opinion's writing style influences a judge's subsequent decision to treat the opinion. We theorize that opinions written in a style conforming to elite norms should be more influential. To test this expectation, we consider the universe of federal appellate court treatments of the 2160 cases in the Update to the Appeals Courts Database, creating a dataset of over 9 million treatable-treating dyads. The results of our analyses suggest judges' ability to influence others through writing style is limited. Only one element of writing style—certainty—is consistently related to treatment decisions, and this effect is limited to in-circuit treatments. In conclusion, we discuss the consequences of these findings for our understanding of the importance of opinion writing style.

*We thank Matt Hitt for several important contributions.

†Associate Professor of Political Science, Western Kentucky. email: jeffrey.budziak@wku.edu

‡Associate Professor of Politics, SUNY Potsdam. email: dalempert@gmail.com.

Judges' preferences over legal policy lead them to care about both the outcomes of cases and the substantive policies announced in opinions. Whether motivated by law (Bailey and Maltzman 2008; Baum 1997, 58-59) or policy (Segal and Spaeth 2002), judges want to see their preferences enacted broadly. To realize this, judges have to consider how their actions resolve more than a single dispute between two parties. As Hansford and Spriggs (2006) notes when discussing the Supreme Court, "justices do not merely seek to hand down decisions and set precedents that are consistent with their policy preferences. Instead, they endeavor to create legal policy that will actually influence legal and extralegal outcomes in the intended manner" (17).

What affects judges' ability to influence others to achieve their goals? The most important tool judges have at their disposal is their opinion. Written opinions are the dominant way judges communicate (Schauer 1995; Wald 1995). They are the key output of the decision making process (Clark and Lauderdale 2010, 871), and the practice of announcing decisions through written opinions allows judges to communicate with future courts through a medium that guarantees some degree of circulation and longevity of their ideas (Baum 2009, 34). Given their centrality in the legal process, scholars recognize that "opinions represent [judges'] most important tool for influencing social, political, and economic outcomes" (Hansford and Spriggs 2006, 18).

Here we examine how judges' choices regarding organizational structure, language, and grammar, which we refer to collectively as "writing style," affect how influential their opinions are on future courts. Methodological advancements in text analyses have allowed judicial politics researchers to use empirically-reliable measures of writing style elements in their scholarship (Black, Owens, Wedeking and Wolfarth 2016b; Budziak, Lempert and Hitt 2019; Corley and Wedeking 2014; Feldman 2016; Goelzhauser and Cann 2014; Hansford and Coe 2019; Owens and Wedeking 2011; Owens, Wedeking and Wohlfarth 2013). Many have done so to examine whether judges on higher courts can write in a style that influences the behavior of lower courts. The results broadly suggest they can (e.g., Black et al. 2016b; Corley and Wedeking 2014). However, because this one type of influence, frequently characterized as "compliance," is complicated by expectations of stare decisis (see Gruhl 1980; Songer 1987;

Songer and Sheehan 1990, among others), this scholarship cannot comprehensively assess the relationship between writing style and influence. We seek to do so here by examining both whether and how opinions of the United States Courts of Appeals are treated by future courts.

The manuscript proceeds as follows. First, we theorize that opinions conforming to elite norms with respect to writing style are more likely to be influential. Judges disproportionately care about legal elites, including other judges (Devins and Baum 2019). To influence this audience, judges should conform to elite expectations regarding judicial behavior. We argue that legal elites establish norms regarding writing style, and as a result, opinions written in a style that conforms to those norms are more likely to be influential. Second, using legal scholarship and writing style guides produced by legal elites, we identify several elements of writing style—clarity, certainty, affective language, and (in)formality—for which broadly agreed-upon elite norms have emerged. Third, we examine whether opinions incorporating those elements are more influential. Using *Shepard's Citations* ("*Shepard's*"), we test both whether these opinions are more likely to be treated and whether they are more likely to be treated positively once they are treated. Fourth, we present our results. There is limited evidence that writing in a style that conforms to elite expectations consistently affects either the decision to treat opinion or the type of substantive treatment assigned. Only one element of writing style—certainty—appears to consistently influence treatment decisions, and this effect is limited to in-circuit treatments. We conclude by discussing the consequences of these findings for the literature on opinion writing style.

1 Elite Norms and Judicial Behavior

Our research question asks whether judges can use writing style to affect the treatment of their opinions. Treatment decisions can generally be understood as capturing "influence." Klein and Morrisroe (1999) defines influence as "the extent to which the actions of one person have an effect on the views or behavior of others" (372). From this perspective, both the decision to treat an opinion at all and the decision of how to treat an opinion (i.e., positively, negatively, or neutrally) are suggestive of influence. An opinion can be influential by shaping what is discussed by future courts. For example, if an opinion recommends a novel approach to a legal

question that future courts engage with when resolving similar cases, the opinion has been influential on the decision making process. This is true whether it affects the direction of the decision or not. An opinion can also be influential by affecting the direction of the decision in future cases. Building on the previous example, the same opinion could recommend an approach to a legal question that tends to favor one party (e.g., the government) over another (e.g., a criminal defendant). Here, an opinion is influential because it shaped what is discussed by future courts, but it is also influential because it affected the direction of the decision in future cases. Because "influence" is inclusive of many related-but-distinct forms such as these, scholars have regularly adopted the term to describe how judges' opinions broadly affect the behavior of future courts, particularly with respect to whether and how future courts treat opinions (Hansford and Spriggs 2006; Hume 2009; Szmer, Christensen and Grubbs 2020, see footnote 1 for discussion).

There are a number of reasons to think judges wish to influence others. The desire to shape the world around one is a basic human motivation. There is no reason to believe judges are immune from it (Devins and Federspeil 2010, 91). If anything, the decision to forego a (frequently) more lucrative career as an attorney and submit oneself to a grueling confirmation process suggests that people who serve as judges might be more strongly motivated by the desire to influence than others (Baum 2008, 32).[1] From this perspective, judges may desire to influence others as an end in itself. In addition, judges may seek to influence others as a means of advancing their goals. Influencing other judges is the best way to align the outcomes of cases decided by other courts with their own preferences.

Scholarship suggests that some individual-level judge characteristics affect the ability to influence others. Klein and Morrsiroe (1999) shows that legal rules announced in opinions written by judges perceived to be "prestigious" are more likely to be adopted by other circuits in three areas of law (386-387). Szmer, Christensen and Grubbs (2020) finds that an opinion author's age and caseload affect the number of citations to the opinion by future court of appeals panels (61). Nelson and Hinkle (2018) examines the total number of citations, the

[1]Individuals may do this for a variety of other reasons, including the desire for prestige or achievement. As Baum (2010) notes, these motivations are related, but likely distinct, concepts. The decision to become a judge is undoubtedly shaped by all of these considerations.

total number of positive treatments, and the total number of negative treatments to search and seizure cases. The authors find that ideological extremity is correlated with total citations and negative treatments, and subject matter expertise is correlated with all three. However, the result of the latter is counterintuitive; opinions written by judges who are former prosecutors (experts in search and seizure cases by assumption) are less likely to be cited and are less likely to be treated, both positively and negatively, by future courts. These results suggest that some individual-level factors do affect the potential for influence.

There is also evidence that case- or institution-specific factors shape the ability of judges to influence others. Szmer, Christensen and Grubbs (2020) finds no evidence that panel- or circuit-level variables affect the total number of future citations, but it does find evidence that a number of case-level variables, such as case complexity, issue area, and the presence of an amicus brief, can affect how frequently a case is cited. Hume (2009), examining administrative cases in the courts of appeals, finds variables like the procedural basis for the case, the type of appeal, and whether the case was decided en banc likewise affects how frequently an opinion is treated outside of its originating circuit. In addition, Nelson and Hinkle (2018) finds that a court's baseline citation rate affects the number of future citations as well as the number of positive and negative treatments. These results suggest that some contextual factors do matter. However, even if some individual and contextual factors do shape degrees of influence, none of these factors are under judges' control, particularly in the near term.[2]

Opinions, which are under judges' control, have the potential to be influential because judicial decision making and legal opinion writing are time-consuming, cognitively-demanding activities. Sometimes, these are only moderate burdens. When judges have strong preferences about an issue presented in a case, the difficulties inherent to decision making and, to a lesser degree, opinion writing, are reduced. In those circumstances, judges already have a clear idea of how they will proceed. However, even if judges have preferences, they are not necessarily fixed (see Martin and Quinn 2007 for discussion of Supreme Court justices or Kaheny, Haire and Benesh 2008 for discussion of court of appeals judges), and in many cases, judges confront novel and difficult legal questions over which they lack preferences (see Baum

[2]Even if some factors like "prestige" can be cultivated by judges over a long period, it is not a factor judges can manipulate in the short term to exert influence.

2017, 7-10). In those circumstances, the task of judging is more difficult. When individuals are confronted with difficult tasks, many look for resources to help make the tasks easier. There is no reason to believe judges would be unusual in this regard, and opinions written by other judges are excellent resources to make the decision making and opinion writing processes more manageable.

Opinions are more likely to reduce the burdens of judging when they conform to shared professional expectations, best understood as norms. Scholars have demonstrated the usefulness of thinking about judicial behavior as being directed at audiences (Baum 2009; Black et al. 2016b), and different audiences are likely to have different expectations regarding judicial behavior. Devins and Baum (2019) theorizes that Supreme Court justice[3] behavior is disproportionately directed at an audience of political, social, and professional elites. Colleagues on other courts are an important part of that elite class (Devins and Baum 2019, 40-46), and judges may view them as an important audience for both personal and instrumental reasons (Baum 2008, 50-60; Micli and Cosgel 1994). When discussing the importance of elite audiences, Devins and Baum (2019) focuses primarily on the personal desire of judges to be well-regarded. To cultivate esteem from legal elites, justices conform to behavioral norms established by them.

A judge seeking to influence others should also be motivated to conform to elite norms. The goals of securing esteem and influencing others are at least overlapping, if not reinforcing—courts who hold judges in high esteem are more likely to be influenced by them (see Devins and Baum 2019, 54; Klein and Morrisroe 1999). Therefore, failure to conform to elite norms would lead judges to "lose both the respect of people with whom they interact intensively *and their capacity to influence fellow judges to achieve the results they favor*" (Devins and Baum 2019, 54, emphasis added). The authors identify two such norms that shape the substantive content of judges' opinions: collegiality and legal orientation (53-57). Professional legal training puts emphasis on both collegiality (e.g., Edwards 1998) and law-oriented decision-making (Bailey and Maltzman 2008; Martinek 2010), and professional/personal socialization reinforces their importance in the legal community.

[3]While Devins and Baum (2019) focuses exclusively on Supreme Court justices, it is reasonable to believe other federal judges would be similarly motivated.

Judges have substantial discretion to shape their opinions in a way that signals their fidelity to elite norms. There is ample evidence judges can promote collegiality by making substantive accommodations of their colleagues during the opinion drafting process (e.g., Maltzman, Spriggs and Wahlbeck 2000). These accommodations frequently produce publicly-visible evidence that is at least consistent with collegial decision making, such as unanimous votes or a lack of separate opinion writing (see Corely, Steigerwalt and Ward 2013, chapters 3 and 4, respectively). The ability of opinion authors to signal their adherence to law-oriented decision making is even easier.[4] By writing opinions that engage in legal reasoning, cite existing legal authorities, and purport to rely on existing precedent or doctrine to reach conclusions, judges can signal they have engaged in law-oriented behavior.

Empirical scholarship offers limited support for the hypothesized relationship between collegiality and influence. Hansford and Spriggs (2006) demonstrates that the presence of concurring opinions increases the probability of a negative interpretation of Supreme Court precedent by the future Court, while larger voting margins decrease the probability of a negative interpretation. However, the authors do not find any evidence that the presence of concurring opinions or the size of vote margins affects the treatment of Supreme Court precedent by lower federal courts. In the context of the courts of appeals, Szmer, Christensen, and Grubbs (2020) finds that the presence of a dissenting opinion reduces the number of citations to an opinion by future panels from within the same judicial circuit. However, the authors do not find a similar effect for concurrences or dissents—neither affect citations to an opinion by future panels from outside the opinion's originating circuit. Further, Hume (2009), using a sample of administrative law opinions, does not find evidence that unanimous court of appeals opinions are more likely to be treated more positively by future out-of-circuit panels.

There is stronger evidence suggesting law-oriented decision making shapes how opinions influence future courts. Cross, Spriggs, Johnson, and Wahlbeck (2010) finds subsequent Supreme Court and lower federal courts are more likely to positively treat Supreme Court opinions that

[4] Audiences could conclude that judges have interpreted the law incorrectly or have taken law-looking actions (such as citing precedents) to disguise preference-based decision making. However, the mere act of engaging with legal factors could be viewed as promoting the norm of law-oriented decision making. It is difficult to imagine that legal elites would react to an opinion which engages in legal reasoning in a similar manner to an opinion that made no such efforts, even if it was widely believed that personal preferences determined the outcome in both cases.

include greater numbers of citations to precedents. Hansford and Spriggs (2006) examines the effects of citing particular types of Supreme Court precedent: those with high levels of vitality. The authors conclude that citing high-vitality cases induces more positive treatments among future courts in a variety of contexts. Szmer, Christensen, and Grubbs (2020) finds a similar result in the courts of appeals: opinions that include a higher number of citations per page are more likely to be treated positively both within and outside an opinion's originating circuit. Nelson and Hinkle (2018) offers the only countervailing evidence to this pattern. The authors find that increases in non-binding citations lead to more negative future treatments, while increases in binding-citations decreases the number of positive future treatments. However, Nelson and Hinkle (2018) focuses exclusively on search and seizure cases and looks at treatments in both federal and state courts, so their contrary results may be a function of these differences.

2 Elite Norms and Elements of Writing Style

There are also likely to be elite norms regarding writing style. The norms of collegiality and law-orientation emerge from a variety of sources, with judges' shared educational and professional experiences playing an integral role in promoting these as common expectations (Devins and Baum 2019, 56). Norms regarding writing style are likely to develop for similar reasons. A substantial portion of an aspiring lawyer's legal education is devoted to learning to write effectively, and the legal profession puts considerable weight on writing when evaluating its members. Shared norms about writing style are particularly likely in the judicial profession, where judges primarily communicate their ideas through writing.

Opinions written in a style that conforms to shared norms are more likely to be influential because judges' favorable reactions to style may shape their reactions to substantive content. As discussed above, one way judges can try to influence others is by appealing to the norm of law-oriented decision making. This norm emerges from shared educational and professional experiences and shapes expectations concerning *what* judges should rely on in their opinions. This primarily includes legal factors like formal legal arguments, citations to precedent, and use of proper legal terminology. Those shared experiences may also lead to expectations

concerning *how* these legal factors should be presented. Formal legal arguments frequently take a common structure, and many opinions adopt similar language, grammar, and tone when discussing legal factors. Failure to write in this way could disappoint, frustrate, or even confuse judges, leading them to react negatively to the opinion. This negative reaction concerning style could color their response to the substantive content of the opinion. Conversely, by successfully employing the expected style, judges put their arguments where readers expect to see them and use language readers expect to read. Meeting these expectations is likely to produce a positive reaction in most judges, which could likewise lead them to view the substantive content of the opinion in a positive light.

In addition, opinions written in a style that conforms to shared norms are more likely to be influential because writing style can serve as a heuristic for other judges. There is ample evidence that individuals rely on heuristics when making decisions, particularly when those decisions are difficult. When confronting difficult questions, scholars have demonstrated that individuals will commonly (and subconsciously) substitute a related-but-easier question and answer that (Kahneman 2011, 12). This frequently takes the form of relying on easy-to-evaluate heuristics that are related to, but distinct from, the more difficult question. Judicial politics scholarship suggests judges are not exceptions to this widespread phenomenon. For example, Klein (2002) provides substantial qualitative evidence that judges' identities act as heuristics in the decision making process (93-96). In one representative example, a judge noted "I've been in the business long enough to know many judges. If I know a judge is damn good, well educated, then I would be influenced by an opinion by him" (94). While being "well educated" may be correlated with things like being skilled in argumentation, the interviewee's response implies that judges' ability to influence others is shaped by reputational factors independent of opinion content. This suggests that when judges consider how to treat opinions, they may subconsciously avoid the difficult question "is this a good opinion" by replacing it with the easier "is this opinion written by a good judge?" and answer it. There is also quantitative evidence at least consistent with this explanation. Klein and Morrisroe (1999) finds that legal rules announced in opinions written by judges who are perceived as more "prestigious" are more likely to be adopted by future courts. There are a number of

possible interpretations of this result (see Klein and Morrisroe 1999, 389 for discussion), but the conclusion that judges rely on reputational identifies as heuristics is a plausible one.

Opinion writing style may similarly serve as a heuristic for future courts. When judges consider previous decisions, they have to conduct extensive legal research to decide which opinions to engage with, which opinions to ignore, how much engagement is appropriate, and how to treat relevant opinions.[5] Doing so requires a close reading of a number of opinions, assessing the factual similarities between the various disputes, and weighing the strengths of the arguments presented. Even after all that work is complete, it is still not obvious that a judge will be able to conclusively determine whether to treat an opinion or how to do so (i.e., positively or negatively). Therefore, when evaluating opinions, rather than answering the question "is this a good opinion?", a judge might instead answer the question "is this opinion well-written?" Judges are likely to identify "well-written" opinions as those that conform to shared norms for writing style. If a judge concludes that an opinion is "well-written," that judge may be more likely to adopt substantive arguments of the opinion, independent of their merits.

Legal elites have developed shared norms with respect to several elements of writing style. These norms are recognized in a number of places, including law schools, scholarship of the legal academy, professional conferences, blogs, and social media posts, among others. The recognition of these shared norms is clear in elite-dominated legal institutions whose primary purpose is to educate judges about professional expectations regarding judging. In this world, writing style gets substantial attention. For example, the Federal Judicial Center (FJC), which describes itself as the "research and education agency of the judicial branch," offers courses and resources on the subject of judicial writing, including a well-known Judicial Writing Manual.[6] The National Judicial College (NJC), an entity within the American Bar Association, describes itself as "a national leader in judicial education" and offers similar resources.[7] Dis-

[5] Judges regularly outsource this work to clerks or staff attorneys. However, this does not eliminate the demands of legal research—it simply shifts the burden to another individual. Clerks or staff attorneys could also rely on easier-to-evaluate heuristics to complete their research. Given their comparative lack of experience and the uncertainty inherent in the opinion writing process, clerks may be particularly apt to rely on such heuristics, consciously or not.

[6] "Federal Judicial Center," available at `https://www.fjc.gov/about`. Last accessed 6/1/2022.

[7] "National Judicial College," available at `https://www.unr.edu/judicial-education/national-judicial-college`, Last accessed 6/1/2022.

cussing writing style in its "Judicial Writing Tips" (2017), the NJC teaches that "research shows that different audiences share common preferences about the organization of judicial opinions, use of language, and placement of information. Addressing those preferences leads to audiences' greater understanding of opinion content and enhances parties' and others' positive perceptions of the judicial system." Institutions like these, working in conjunction with law schools and legal scholarship (for one prominent example, see Scalia and Garner 2008) help establish shared norms regarding writing style and disseminate those norms to members of the judiciary.

It is impossible to generate a comprehensive classification of writing style norms. Writing style is made up of a large number of stylistic elements. Some minor elements may simply not warrant enough attention for an elite consensus to develop. In those circumstances, judges are more likely to have unique or idiosyncratic preferences about what is preferable. Even if such a consensus did emerge around these elements, it would be appropriate to question whether minor, non-substantive stylistic choices could reasonably be expected to change how future courts respond to an opinion. For consequential writing style elements, a broad elite consensus is more likely to emerge. However, even for these types of elements, there are good reasons to believe that judges will have some individualized preferences (Budziak and Lempert, forthcoming). Therefore, rather than trying to develop a complete taxonomy of shared norms regarding writing style, we focus our attention on a number of distinct elements for which a broad consensus has most clearly emerged in elite commentary about judicial writing.

All elite commentators stress the importance of writing "clearly." While "clarity" could mean a number of different things (see Owens and Wedeking 2011, 1038) we focus on what is best understood as rhetorical clarity. Many resources stress that readers of opinions expect and prefer a clear rhetorical style. The FJC Judicial Writing Manual advises that "precision and clarity are the main concerns of good writing" and notes that judges interviewed for the manual specifically identified the lack of precision and clarity as a common problem in judicial writing (21). To promote clarity, writing style guides and academic scholarship consistently instruct judges to write using simple and straightforward prose with the goal of producing as readable an opinion as possible. Readable opinions are more likely to use the active voice,

avoid contractions, personal pronouns, intensifying adverbs and slang/idioms (Lebovits and Hildago 2009, 35). The FJC similarly suggests "[writers] should use active voice and... weed out gratuitous adjectives and eliminate unnecessary adverbs" (23). Shapiro et al. (2013) also suggest writers be mindful of their structure: "Writers should strive to keep sentences short, to avoid excessive use of the passive voice, to use active verbs frequently, and to break up long paragraphs" (733). Empirical scholarship suggests that failure to write clearly affects how different audiences respond to opinions. Black et al. (2016b) demonstrates that lower federal courts are more likely to treat clearly-written Supreme Court opinions positively. In the case of the courts of appeals, Nelson and Hinkle (2018) demonstrates that, in the context of search and seizure cases, other courts of appeals and state appellate courts are more likely to engage with and treat an opinion positively when it is written more clearly. For all these reasons, we hypothesize that clearer opinions are more likely to be influential.

Commentators also encourage judges to write definitively (Gardner 1993). The judicial politics literature has commonly described this as writing with "certainty." As described in Corley and Wedeking (2014), commentators argue that judges should write with confidence and assertiveness to enhance the credibility of their messages (Rieke and Stutman 1990). Lebovits and Hildago (2009), as well as *The Judicial Opinion Writing Handbook*, advise clerks and judges to be "definite" and not tentative when writing opinions. Empirical scholarship supports these recommendations. Corley and Wedeking (2014) finds lower federal courts are more likely to treat Supreme Court opinions positively when written with greater certainty. We thus hypothesize that opinions written with greater certainty are more likely to be influential.

In contrast, commentators discourage the use of affective language. Scholarship on legal writing uniformly criticizes the use of emotional, personal, or scandalous material (Lebovits and Hidalgo 2009, 35; Shapiro et al. 2013, 743). Many commentators argue that reliance on emotionally-laden language or personal appeals is indicative of a lack of substantive merit to the arguments being advanced. For example, Scalia and Garner (2008) note that "good judges pride themselves on the rationality of their rulings and the suppression of their personal proclivities, including most especially their emotions" (32). Recent scholarship offers empirical support for the expectation that affective language reduces the influence of legal writing from

the perspective of the Supreme Court. Black et al. (2016a) demonstrate that level of affective language in merits briefs negatively correlates with the probability of success on the merits, controlling for a host of other factors. We therefore hypothesize that opinions employing more affective language are less likely to be influential.

Finally, commentators frequently discuss the importance of writing in a proper tone. One important element of tone is the opinion's level of formality. Writing about style generally, Heylighen and Dewaele (1999, 2) states that the "most frequently mentioned" aspect of linguistic style is formality. Elements of formal style include impersonal style (third person), complex words and sentences, technical vocabularies, polite words and formulae, and objective style (Sheika and Inkpen 2012, 6-8). Judges recognize the importance of formality as an element of legal writing. For example, Scalia and Garner advise that "formality bespeaks dignity" (2008, 118). This suggests that judges have expectations with respect to the level of formality of an opinion. Failure to meet expectations with respect to formality implicate similar concerns associated with the use of affective language; judges may question the legitimacy of opinions written in an inappropriately informal tone. Therefore, we hypothesize that opinions written in a more informal tone are less likely to be influential.

3 Data and Measurement

To construct a database appropriate for testing our hypotheses, we began with the Update to the Appeals Court Database (Kuersten and Haire 2007, henceforth "the Update"), a circuit- and year-stratified random sample of 2160 published court of appeals decisions handed down between 1997 and 2002.[8] We then downloaded the text of the opinions associated with these decisions; we refer to these opinions as the "treatable" opinions. Next, we used *Shepard's* to collect data on every published court of appeals opinion that treated any of the treatable opinions, from the day the treatable opinion was published through the end of 2018.[9] There are

[8]The database is publicly available and is housed by the Judicial Research Initiative (JuRI) 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`.

[9]Shepard's is a citation index that links all U.S. court opinions at the state or federal level available in any reporter (i.e., published) since the founding of the United States. Scholars have conducted extensive tests on *Shepard's* data and found that the treatment coding is both reliable and valid (Hansford and Spriggs 2006, 43-54; Spriggs and Hansford 2000).

4466 such opinions, all of which have the potential to treat each "treatable" opinion from the Update. For ease of discussion, we refer to these as "treating" opinions unless it is necessary to call them "potentially treating" for clarity. Our unit of analysis is a treatable-treating dyad, and the initial dataset contains $2160 \times 4466 \approx 9.6$ million observations, although—as we discuss below—the dataset we ultimately analyze is a smaller dataset of some 2.2 million observations.

The dataset we construct consists of the random sample of treatable cases that make up the Update, the population of treating cases that actually treat a given treatable case, and a sample of potentially treating cases that do not in fact treat a given treatable case.[10] Thus, we have what is called, variously, a case-referent, case-control, synthetic retrospective, or choice-based design.

The dataset includes some 9.6 million treatable-treating dyads. However, in practice, some treatable opinions are so irrelevant to a given treating opinion that they will not be cited for that reason (Hinkle 2015; Spriggs and Hansford 2002). To identify treatable-treating dyads that are potentially relevant to each other, we assume that any treating case that in fact treated a treatable case in a given issue area could have treated any other treatable case in that same issue area.[11] Eliminating these "irrelevant" dyads, and dropping observations with missing data, gives a final sample size of about 2.2 million dyads.

The treatable-treating dyad unit of analysis offers the best opportunity to test whether opinions adhering to norms regarding writing style are more influential. Some scholarship examines opinion influence in the courts of appeals using the "treatable" opinion as the unit of analysis (e.g., Hume 2009; Nelson and Hinkle 2018; Szmer, Christensen and Grubbs 2020). However, investigating counts of treatments prevents scholars from assessing important dynamics between the treated and treating courts. Failure to do so represents an important

[10] This is, at worst, an arguably as-if-random sample of non-treating cases. The assumption required is that cases that do not treat *any case in the population*, and are therefore excluded from our sample, fail to treat for the same reasons as the treating cases in our sample—i.e., because of the covariates for which we control. If we are wrong about this, then our estimates do not apply to such "never-treating" cases and apply only to cases in the population that treat at least one case (in the population), which, in our view, is a relatively minor caveat.

[11] We use issue areas based on those in the Update: (1) Criminal, (2) Civil Rights, (3) First Amendment, (4) Due Process, (5) Privacy and Miscellaneous, (6) Labor Relations, and (7) Economic Activity and Regulation. More refined (narrower) categories would also be possible, but do not appear to make much practical difference (Hinkle 2015, 726 fn. 11).

threat to drawing appropriate inferences about the results. As a result, most scholarship on lower court compliance with higher court decisions is organized using a dyad (e.g., Black et al. 2016b; Corley and Wedeking 2014; Westerland, Segal, Epstein, Cameron and Comparato 2010). We adopt a similar approach here.

We use *Shepard's* treatment decisions to measure influence. As defined above, opinions are more influential when they have greater effects on the views or behavior of others. Unlike basic citation analyses, *Shepard's* indicates how courts treat opinions they cite. To determine how an opinion is treated, *Shepard's* asks "What effect, if any, does the citing case have on the cited case?" (see Spriggs and Hansford 2000, 329). If an opinion is influential because it affects the views and behavior of other judges, *Shepard's* is well-designed to measure this effect. *Shepard's* also utilizes a treatment coding scheme well-designed to gauge the degree to which an opinion has affected the views or behavior of the court treating that opinion. This scheme includes codes that are positive ("Followed"), neutral ("Explained," "Harmonized") or negative ("Abrogate, "Supersede," "Distinguishing," "Criticizing," "Limiting," "Questioning," and "Overruling").

The best evidence that a treatable opinion has influenced another court is if that court treated the opinion positively, rather than neutrally or negatively. This occurs when one court "follows" the opinion of another. According to Shepard's, a court follows an opinion only if it "expressly" relied on the opinion by citing it as "controlling authority" (Shepard's 1993, 17, cited in Spriggs and Hansford 2006, 45). "Following" is only applied if the court uses explicit language that extends beyond a "mere going-along," including words like "controlling," "determinative," or "persuasive" (for discussion, see Spriggs and Hansford 2006, 45). Therefore, *Shepard's* does not capture mere citations or passing references to an opinion; for a treatment to be categorized as following, the court must have used specific language demonstrating how the opinion shaped their views or behavior.

To operationalize elements of writing style, we use several measures well-grounded in the judicial politics literature. Clarity, somtimes referred to as rhetorical clarity or readability, has attracted disproportionate attention as both an independent and dependent variable (Black et al. 2016b; Budziak and Lempert forthcoming; Coleman and Phong 2010; Hansford and

Coe 2019; Goelzhauser and Cann 2014; Johnson 2014; Lempert forthcoming; Owens, Wedeking, and Wohlfarth 2013). Linguists characterize readability as the difficulty of reading and understanding written texts (DuBay 2004). Several measures are in common usage. Each—including the Coleman-Liau Index, Flesch-Kincaid Grade Level, Flesch Reading Ease Scale—are calculated from some combination of character, syllable, word, and sentence counts. For each of these measures, longer sentences (i.e., those with more words) and/or longer words (i.e., those with more syllables or characters) indicate that a text is less readable.[12] The particular metric we use in our variable *Clarity* is the Flesch-Kincaid Grade Level (FKGL), a widely-utilized measure of readability based a text's sentence and word length; conveniently, it is scaled to approximate the (U.S.) grade level of education required to understand the text.

To operationalize certainty and affective language, we use Linguistic Inquiry and Word Court (LIWC) software (e.g., Tausczik and Pennebaker 2010). This tool classifies more than 2,000 words into categories that reflect the author's "thinking style" (Owens and Wedeking 2011, 1040). We construct our variables *Certainty* and *Affective Language* by running the preprocessed opinions through LIWC 2015 version 1.1. For Certainty, we follow Corley and Wedeking (2014, 43-45) and use the program's "certainty" category. This category contains 83 words or word stems that capture the use of words implying certainty. Similarly, for Affective Language, we follow Black et al. (2016a, 384-385) and rely on the affective language category. This category contains a total of 919 words or word stems that capture the use of affective language in writing. Certainty and Affective Language takes the value of the percentage of words or word stems from each respective category contained in the opinion.

Our measure of *Informality* combines the approaches of Sheika and Inkpen (2012) and Brooke and Hirst (2014) as detailed in Budziak, Hitt, and Lempert (2019). In short, we create a standardized measure that indicates the prevalence of word choices indicating informal style: intensifiers, phrasal verbs, contractions, idiomatic language, abbreviations, contractions, colloquialisms, vague expressions, slang words (Sheika and Inkpen 2012) and words associated with less-formal synonyms (Brooke and Hirst 2014).[13]

[12]Applying this measure in automated text analysis requires that the text parser accurately identifies syllables, words, and sentences, a non-trivial methodological problem (see Lempert forthcoming). We write a Python script, customized for legal opinions, for this purpose.

[13]We consider emotional (affective) language separately. To ensure that affective language is not driving

In addition to elements of writing style, scholars theorize that elite norms regarding collegiality and law-oriented decision-making affect how influential a judge is likely to be through their opinions (Devins and Baum 2019, 54). We account for the possible effects of these norms on influence specifically by including a number of control variables. To account for shared norms regarding collegiality, we include two variables that capture separate opinion writing. In the courts of appeals, separate opinion writing is unusual (Hettinger, Lindquist and Martinek 2004, 125). Concurring or dissenting opinions are thus easily observable evidence that, at least, the norm of consensus has broken down. Failure to achieve consensus, especially where consensus is so common, could be perceived as evidence of a lack of collegiality. To address this possibility, we include *Concurrence* and *Dissent*, binary variables indicating whether an opinion of that type accompanied a treatable opinion.

To account for shared norms regarding legal orientation, we include several variables that capture citations to precedent. The easiest way for judges to demonstrate legally-oriented decision making is to rely on precedent. Scholars have demonstrated that opinions citing a greater number of precedents are more likely to be cited (Cross et al. 2010) and treated positively, including by out-of-circuit court of appeals panels (Szmer, Christensen, and Grubbs 2020, but see Nelson and Hinkle 2018). We therefore include a variable, *Citations*, which takes the value of the number of Supreme Court and Circuit Court citations in the treatable opinion.

There are several other covariates likely to affect how opinions are treated. Primary among them is ideological distance between the courts producing the treatable and treating opinions. Judges are more likely to treat an opinion positively when the authoring court is ideologically proximate (Black et al. 2016b; Hansford and Spriggs 2006; Westerland et al. 2010). To control for this possibility, we create a variable, *Ideological Distance*. To construct our measure, we utilize DIME (CF) scores (Bonica 2014), which exploit data on campaign contributions to construct ideal point estimates of political elites, including federal judges.[14] For each dyad, Ideological Distance takes the value of the absolute difference between the DIME scores of the

the results for Informality, we exclude affective language when constructing our measure of informality.

[14]While conventional measures of circuit court judge ideology rely on voting behavior of home state senators and position-taking of appointing presidents (e.g., Giles, Hettinger, and Peppers 2001), the key advantage of DIME scores is that they are primarily a function of *judicial* behavior that is not endogenous to voting (Bonica and Sen 2016).

two opinions' authors.

In addition to ideological distance, scholars have demonstrated that the prestige of an opinion author can shape how the opinion is treated by future courts (Klein and Morrisroe, 1999). To account for this possibility, we generate a prestige measure for all opinion authors in our dataset following Klein and Morrisroe (1999). Briefly, this measure relies on citations to judges *by name*, rather than overall citation counts. To construct the measure, Klein and Morrisroe (1999) count the number of named citations and then adjust them in a variety of ways to account for concerns related to dissenting and concurring opinions, number of opinions written, and other temporal concerns (see Klein and Morrisroe 1999 for relevant details). We largely follow that approach to create the variable *Prestige*.[15]

In addition to our variables of substantive interest, we include a number of additional controls. *Treatable Length* takes the value of the number of words in the treatable opinions. Longer opinions may be more substantive, and are thus more likely to be treated. We also include *Self Treatment*, which takes the value of 1 if the treating opinion author treats an opinion they have previously written. We expect judges are more likely to engage in self-treatment, independent of other opinion characteristics. Finally, we include fixed effects for the treatable opinion's circuit and issue area to account for unmodeled heterogeneity.

4 Models and Results

To determine whether opinions written in a style conforming to shared norms are more influential, we estimate two separate models. The first model examines the decision to treat an opinion (1) or not (0). Because the dependent variable for this model is dichotomous, we estimate a logistic regression model. The second model examines the substantive treatment (positive, neutral, negative) given to opinions that have been treated by future courts. Here,

[15]The measure developed in Klein and Morrisroe (1999) produced a single static measure for courts of appeals judges at a fixed point in time. Because we assign a measure of judicial prestige for each judge in different years, we were forced to decide how to assign the scores. The obvious choice might have been to assign each judge a different prestige score for each observed year. While this is the simplest choice, it is unlikely to be the most theoretically appropriate. There is clear consistency across years, but there is also some variation in year to year scores. This is unlikely how prestige actually operates. There is little reason to believe that judges gain or lose prestige quickly. The variance in the measure is much more likely the result of a degree of randomness inherent to the citation process. To account for this, we average each judge's annual measure to produce a single prestige measure for each judge.

because the dependent variable is unordered and trichotomous, we estimate a multinomial logit model. This two model approach is consistent with the literature (see Hinkle 2015 for a prominent example) and better captures the various ways opinions can be influential. Opinions can influence future courts by changing what is being considered in the decision-making process (Model 1), the outcomes of the decision-making process (Model 2), or both. These are related-but-distinct paths to influence—examining only one path or collapsing the two paths together has the potential to obscure important differences in how opinions are influential. As the results below demonstrate, there are meaningful differences across these two categories.

4.1 Model 1

Recall that our study of the decision to treat is designed as a case-referent study. Mantel (1973) showed that coefficient estimates from a standard logit model are unbiased in this context; however, the estimate of the intercept is biased, which in turn biases the estimated predicted probabilities. Nonetheless, since we can calculate the number of opinions in the population, in each year, that do not treat a given treatable opinion (since we know the total number of published opinions by year), we can obtain unbiased predicted probabilities by using inverse probability weights that account for the under-sampling of cases that do not treat a given treatable case (Manski and Lerman 1977; Stata 17 User's Guide Sec 20.24.3).[16]

Column 1 in Table 1 gives the estimates for the model predicting the decision to treat a treatable case, including both in- and out-of-circuit potential treating opinions. Column 2 gives these estimates for out-of-circuit potential treating opinions only, and Column 3 for in-circuit treating opinions. Because the mean probability of a potential treating opinion treating a given treatable opinion is so low (.00045 for in-circuit treating opinions and .00006 for out-of-circuit treating opinions), we illustrate substantive impact in the discussion below by referring to odds ratios (which, given the low probabilities of a positive outcome, are practically equal to relative risk ratios, i.e., ratios of the probabilities of a positive outcome across the two values of a covariate).

We consider first the model of in- and out-of-circuit treating opinions (Column 1). Only one

[16]We construct our weights to account also for the Update's circuit-year stratified sample of treatable cases.

Covariate	(All)	(Out-of-Circuit)	(In-Circuit)
Ideological Distance	−0.094*	0.011	−0.056
	(0.04)	(0.05)	(0.05)
Clarity	0.013	0.008	0.019
	(0.02)	(0.03)	(0.02)
Informality	−0.046	−0.275*	0.074
	(0.04)	(0.07)	(0.05)
Affective Language	−0.054*	−0.076*	−0.033
	(0.03)	(0.04)	(0.03)
Certainty	0.107	−0.196+	0.305*
	(0.08)	(0.12)	(0.09)
Self Treatment	2.430*		0.493*
	(0.10)		(0.11)
Prestige	0.005	0.001	0.007
	(0.01)	(0.02)	(0.02)
Citations	0.006*	0.005+	0.008*
	(0.00)	(0.00)	(0.00)
Treatable Length (× 1000)	0.093*	0.116*	0.080*
	(0.01)	(0.02)	(0.02)
Dissent	0.204*	0.246+	0.159
	(0.10)	(0.14)	(0.13)
Concurrence	0.165	0.299	0.093
	(0.34)	(0.51)	(0.40)
N	2,184,154	1,996,154	188,000

Table 1. Was a Treatable Opinion Treated? (1=Yes; 0=No). Logit coefficients, standard errors clustered by treatable case in parentheses. Fixed effects and constant omitted from table. $^{+}$ $p < 0.1$ * $p < 0.05$

of the writing style variables, Affective Language, is statistically significant ($p = .04$). Opinions written in a less emotional style are more likely to be treated; the odds ratio associated with going from the 90th to the 10th percentile of Affective Language is 1.18. None of the other style variables (Clarity, Certainty or Informality) are correlated with the decision to treat an opinion. Of the non-style variables, the most dramatic effect is for Self Treatment; the odds ratio of 11.35 indicates that, all else equal, judges are over eleven times more likely to treat an opinion that they wrote themselves, compared to one someone else wrote. Four other covariates are statistically significant: Ideological Distance, Citations, Treatable Length, and Dissent. An opinion written by a judge who is relatively closer to the treating judge is more likely to be cited: going from the 90th percentile to the 10th percentile in Ideological Distance is associated with an odds ratio of 1.17. Longer opinions and those with more citations to circuit court and Supreme Court opinions are more likely to be treated: the odds ratio associated with going from the 10th to the 90th percentile in Length is 1.68; the analogous odds ratio for Citations is 1.21. The odds ratio of 1.23 for Dissent indicates that opinions

with dissents are somewhat more likely to be treated.

Next, we turn to treatments by out-of-circuit opinions (Column 2). As before, opinions written in a more emotional style are less likely to be treated; the odds ratio of going from the 90th to the 10th percentile of Affective Language is 1.27. Considering only out-of-circuit treatments, less formal opinions are less likely to be cited: going from the 90th to the 10th percentile of Informality is associated with an odds ratio of 1.67. Certainty and Clarity are again non-significant. While Treatable Length remains significant (odds ratio 1.90), the other control variables are not statistically significant when examining only out-of-circuit treatments.

Turning to in-circuit treatments (Column 3), we find that the only textual variable that is statistically significant is Certainty. Opinions written in a style that is more certain are more likely to be cited: the odds ratio associated with going from the 10th to the 90th percentile in certainty is associated with an odds ratio of 1.34. Statistically significant control variables here include Self (odds ratio: 1.64), Citations (odds ratio: 1.26) and Treatable Length (odds ratio: 1.57).

4.2 Model 2

Table 2 presents the model of treatment type, conditional on the decision to treat. As before, Column 1 includes both in- and out-of-circuit treatments, Column 2 models only out-of-circuit treatments, and Column 3 model in-circuit treatments only. The base (reference) outcome is a negative treatment. In discussing these results, we focus on the probability of a positive treatment, relative to a negative treatment, but we include the (not particularly illuminating) results for neutral relative to negative in the tables as well. We include the same covariates as in the models presented in Table 1.

Covariate	(All)	(Out-of-Circuit)	(In-Circuit)
Outcome: Neutral Treatment			
Ideological Distance	−0.028	−0.065	0.069
	(0.15)	(0.21)	(0.23)
Clarity	0.073	0.178^{+}	−0.066
	(0.07)	(0.09)	(0.11)
Informality	−0.188	−0.045	−0.409^{+}
	(0.16)	(0.18)	(0.21)
Affective Language	−0.033	0.039	−0.084
	(0.08)	(0.11)	(0.11)
Certainty	0.018	0.015	−0.026
	(0.27)	(0.33)	(0.44)
Prestige	0.028	0.057	0.003
	(0.04)	(0.05)	(0.05)
Self Treatment	−0.269	0.000	−0.157
	(0.50)	(.)	(0.51)
Citations	0.003	−0.002	0.007
	(0.01)	(0.01)	(0.01)
Treatable Length (× 1000)	0.018	−0.018	0.070
	(0.05)	(0.05)	(0.07)
Dissent	0.075	0.160	−0.086
	(0.26)	(0.32)	(0.41)
Concurrence	−0.467	−0.410	−0.405
	(0.39)	(0.46)	(0.50)
Outcome: Positive Treatment			
Ideological Distance	−0.104	0.023	−0.121
	(0.09)	(0.12)	(0.12)
Clarity	−0.001	0.053	−0.050
	(0.04)	(0.06)	(0.05)
Informality	−0.022	−0.196^{+}	0.014
	(0.08)	(0.11)	(0.11)
Affective Language	0.073	0.117	0.048
	(0.05)	(0.08)	(0.06)
Certainty	0.339*	0.095	0.417*
	(0.15)	(0.22)	(0.19)
Prestige	0.017	0.045	−0.002
	(0.02)	(0.03)	(0.03)
Self Treatment	0.478^{+}	0.000	0.143
	(0.27)	(.)	(0.29)
Citations	0.001	−0.007	0.007
	(0.00)	(0.01)	(0.01)
Treatable Length (× 1000)	0.022	0.016	0.045
	(0.02)	(0.03)	(0.04)
Dissent	−0.237^{+}	−0.281	−0.191
	(0.14)	(0.21)	(0.20)
Concurrence	−0.383^{+}	−0.235	−0.372
	(0.20)	(0.30)	(0.26)
N	3,572	1,395	2,177

Table 2. Type of treatment a treated opinion received. Multinomial logit coefficients, standard errors clustered by treated case in parentheses. Omitted outcome is negative treatment. Fixed effects and constant omitted from table. $^{+}$ $p < 0.1$ * $p < 0.05$.

Predicting the type of treatment an opinion receives proves to be difficult. The only covariate that has a statistically significant correlation with the probability of a positive treatment, relative to a negative treatment is Certainty, in both the combined in- and out-of-circuit model and the in the in-circuit model. In the combined model, moving from the 10th to the 90th percentile in Certainty decreases the probability of a negative treatment from .24 to .19, and increases the probability of positive treatment from .70 to .75; in the in-circuit model, it decreases the probability of a negative treatment from .20 to .15 and increases the probability of a positive treatment from .75 to .81.[17] In the combined model, Self Treatment, Concurrence and Dissent are marginally significant ($.05 < p < .1$), with changes in the probability of a positive treatment of .09 (Self Treatment), $-.06$ (Concurrence) and $-.05$ (Dissent).

5 Discussion and Conclusion

The results of our analyses fail to coalesce into a single, cohesive narrative. Most of the writing style variables are unrelated to both the decision to treat an opinion and the substantive treatment given to opinions that have been treated. This is largely true whether one considers all treatments, only treatments made by in-circuit panels, or only treatments made by out-of-circuit panels. Affective Language correlates with the decision to treat, primarily when looking at out-of-circuit treatments, but fails to affect the type of substantive treatment assigned. Only one writing style variable—Certainty—consistently behaves as expected: opinions written with greater certainty are more likely to be treated and are more likely to be treated positively by in-circuit panels. Below we attempt to draw reasonable conclusions from the results and discuss the consequences of this finding for the broader writing style literature.

We theorized that judges are influenced by elite norms, that elite have developed norms regarding writing style, and that opinions written in that style should be more persuasive. The lack of consistent relationships between writing style variables and treatment decisions suggests this theory, at least as it is framed here, is not supported. There are a number of potential reasons why. First, it could be the case that judges are not influenced by legal elites. Given recent scholarship (e.g., Bonica and Sen 2020; Devins and Baum 2019), this

[17] These probabilities are calculated by setting the values of the other covariates at their in-sample values.

seems unlikely. There are strong theoretical reasons to expect judges to be influenced by legal elites. These reasons may be particularly strong for judges on courts that operate outside the view of anyone except for legal elites. There is also robust empirical scholarship consistent with this theoretical perspective. For example, scholars have consistently demonstrated that lower federal courts generally comply with Supreme Court decisions (Benesh and Reddick 2002; Songer, Segal and Cameron 1994; Westerland et al. 2010). However, as this scholarship makes clear, mechanisms to ensure compliance are limited. As a result, it is not obvious why lower federal courts comply with the Supreme Court at such high rates. One potential explanation is that compliance, much like consensus and legal orientation, represents an elite norm into which judges are educated and socialized. Lower court judges comply with Supreme Court opinions because they have been trained to do so, and failure to conform to that norm could bring social sanction from their peers. While additional empirical scholarship explicitly investigating the relationship between elite norms and judicial behavior would be beneficial, we are skeptical that judges are not influenced by them.

Second, it could be the case that elite norms for writing style exist, but we misunderstand them. For example, elite norms could have developed around other elements of writing style, but not clarity, certainty, affective language, or informality. Again, we are skeptical of this explanation. The elements we identified here stand out in a large volume of legal writing scholarship. No other element of writing style receives comparable attention. Had one element failed to influence treatments, it would be fair to question whether elite norms regarding that unique element had emerged. But the lack of a consistent effect across most of the writing style variables makes this explanation less likely.

The third, and we think most likely, explanation is that some elite norms regarding writing style exist, but are weaker than we have theorized here. Thinking about the strength of elite norms as a continuum, on one end, our theoretical perspective posits relatively strong norms. We argued that elites have coalesced around clear expectations regarding writing style, that those expectations well-known, and that the resulting norms are consequential enough to influence judicial behavior. On the other end of this continuum, it could be the case that no elite norms exist, or the norms are sufficiently weak to have no effect on behavior.[18] Our

[18]While it is beyond the scope of our discussion to parse the definition, it is fair to question whether a norm

results suggest that elite norms are, at least, not at the strongest end of the continuum. This could be true for many reasons, but one likely possibility is conceptual confusion. To take one example, legal writing scholarship consistently promotes the importance of clarity, but clarity can mean a number of different things (see Owens and Wedeking 2011, 1038-1039 for discussion). If commentators are conflating rhetorical clarity with cognitive clarity or doctrinal clarity, the resulting ambiguity makes it more difficult to operationalize the concept. But even if commentators successfully cabin their discussion to rhetorical clarity, disagreement about what constitutes a clear style could weaken the influence of this norm. Some commentators may believe that short, declarative sentences are the clearest. Others might believe that brevity obscures nuance, potentially making it more difficult to understand the subtleties of what is being communicated. While this is just one example, it is suggestive of the possibility that elite norms are relatively weak because the concepts underlying them are more nuanced than much commentary recognizes.

However, our results also reject the conclusion that writing style norms never affect treatment decisions. Certainty is significantly correlated with the decision of in-circuit panels to treat an opinion and the decision of those panels to treat it positively. Even if the strongest version of our theory is largely unsupported, this finding may help illuminate the strength of elite norms by clarifying how and under what conditions writing style influences treatments. We identified two potential mechanisms that could link elite norms to treatment decisions. First, judges might be influenced by opinions conforming to elite norms because their positive reactions toward style bleed over into their reaction toward substance. Second, judges might use style as a heuristic for substantive content. The lack of any consistent relationship between writing style and out-of-circuit treatments, coupled with the specific effect of certainty on in-circuit treatments, can help distinguish between these mechanisms and provide a better sense of the strength of elite norms.

For Affective Language and Informality, there is no plausible mechanism, aside from the two identified above, linking these elements of writing style to treatment decisions. The lack of a consistent relationship between these elements and treatment decisions suggests neither

exists if it fails to affect behavior. A norm this weak may simply not be a norm at all.

identified mechanism operates as we expected. While Affective Language does correlate with the decision to treat (particularly in the out-of-circuit context), its failure to influence the substantive direction of treatments in any context cautions against strong conclusions of a meaningful effect. For Clarity and Certainty, there are alternative potential mechanisms linking these elements to treatment decisions, particularly when judges are expected to comply (as is this case for in-circuit treatments). Focusing on clarity, Black et al. 2016b identifies three. Clear opinions remove the discretion of (among others) lower court judges to evade opinions, help actors monitor and identify cases of defiance, and provide instructions to actors inclined to follow decisions but lacking the resources to do so (Black et al. 2016b, 11). Focusing on certainty, Corley and Wedeking (2014) argues that writing with greater certainty likewises reduces discretion under expectations of compliance. The authors argue that a key strategy for judges is to "create the impression that the judge has no choice" but to follow the opinion (Gardner 1993, quoted in Corley and Wedeking 2014, p. 39).

Our empirical results help distinguish between these alternative mechanisms. In the context of the court of appeals, reducing discretion seems to be the most relevant. Court of appeals judges presumably do not need help monitoring or understanding the decisions of their own circuit. The significant effect of Certainty, combined with the nonsignificant effect of Clarity, suggests that writing with greater certainty may be more effective in reducing discretion. Writing with certainty frequently involves using language that is closely related to the goal of reducing or eliminating discretion; words like "must," "never," or "always" connote definitive writing. Clarity, as operationalized by readability, is well suited to make opinions easier to read. However, more readable opinions do not necessarily reduce judges' discretion. Therefore, our results suggest that if judges wish to enhance compliance with their opinions, writing with certainty is more important than writing clearly. However, a context wherein the expectation of compliance operates appears to be a necessary predicate to the effect of certainty.

What does this mean for the strength of elite norms? It appears that norms are limited and contextual. Our results suggest that one element of writing style influences behavior, but only under conditions of compliance. While this finding is noteworthy, standing alone, it is

not sufficient evidence that writing style norms broadly influence behavior. If writing style norms are at the weak end of the continuum described above, writing style is likely a function of judges' individual aesthetic preferences. There are good reasons to believe both that judges have unique aesthetic preferences regarding writing style and that those preferences play some role in shaping their behavior. Budziak and Lempert (forthcoming) details why judges are likely to have aesthetic preferences over writing style and presents empirical evidence that is consistent with the conclusion that these preferences are important determinants of writing style. That result, taken in conjunction with our findings here, suggest that scholars of judicial politics should be careful to recognize the importance of these preferences when investigating questions of writing style. Our results suggest that writing style influences behavior, but its ability to do so is more limited and contextual than the literature tends to suggest.

References

Anderson, Robert. 2011. "Distinguishing Judges: An Empirical Ranking of Judicial Quality in the United States Court of Appeals." *Missouri Law Review* 76(2): 315–83.

Bailey, Michael A. and Forrest Maltzman. 2008. "Does Legal Doctrine Matter? Unpacking Law and Policy Preferences on the U.S. Supreme Court." *American Political Science Review* 102(3): 369–84.

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

Baum, Lawrence. 2009. *Judges and Their Audiences: A Perspective on Judicial Behavior*. Princeton, NJ: Princeton University Press.

Baum, Lawrence. 2010. "Motivation and Behavior: Expanding the Scope of Inquiry" in *The Psychology of Judicial Decision Making*, eds. David Klein and Gregory Mitchell. Chapter 1, pp. 3–27. New York, NY: Oxford University Press.

Baum, Lawrence. 2017. *Ideology in the Supreme Court*. Princeton, NJ: Princeton University Press.

Benesh, Sara C. and Malia Reddick. 2002. "Overruled: An Event History Analysis of Lower Court Reaction to Supreme Court Alteration of Precedent." *The Journal of Politics* 64(2): 534–50.

Black, Ryan C., Matthew E. K. Hall, Ryan J. Owens and Eve M. Ringsmuth. 2016a. "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. 2016b. *U.S. Supreme Court Opinions and Their Audiences*. Cambridge, UK: Cambridge University Press.

Bonica, Adam. 2014. "Mapping the Ideological Marketplace." *American Journal of Political Science* 58(2): 367–87.

Bonica, Adam and Maya Sen. 2016. "A Common-Space Scaling of the American Judiciary and Legal Profession." *Political Analysis* 25(1): 114–21.

Bonica, Adam and Maya Sen. 2020. *The Judicial Tug of War: How Lawyers, Politicians, and Ideological Incentives Shape the American Judiciary*. New York, NY: Cambridge University Press.

Brooke, Julian and Graeme Hirst. 2014. "Hybrid Models for Lexical Acquisition of Correlated Styles." Paper presented at the 25th International Conference on Computational Linguistics, Dublin, Ireland. August 23–29th.

Budziak, Jeffrey, Matthew P. Hitt and Daniel Lempert. 2019. "Determinants of Writing Style on the United States Circuit Courts of Appeals." *Journal of Law and Courts* 7(1): 1–28.

Budziak, Jeffrey and Daniel Lempert. Forthcoming. "Aesthetic Preferences and Policy Preferences as Determinants of U.S. Supreme Court Writing Style." *Journal of Law and Courts*.

Choi, Stephen J. and G. Mitu Gulati. 2004a. "Choosing the Next Supreme Court Justice: An Empirical Ranking of Judge Performance." *Southern California Law Review* 78(1): 23–118.

Choi, Stephen J. and G. Mitu Gulati. 2004b. "A Tournament of Judges? " *California Law Review* 92(1): 299–322.

Choi, Stephen J. and G. Mitu Gulati. 2005. "Which Judges Write Their Opinions (and Should We Care)?" *Florida State University Law Review* 32(4): 1077–1122.

Choi, Stephen J. and G. Mitu Gulati. 2006. "Ranking Judges according to Citation Bias (As a Means to Reduce Bias)." *Notre Dame Law Review* 82(3) 1279–1310.

Choi, Stephen J., G. Mitu Gulati and Eric A. Posner. 2009. "Judicial Evaluations and Information Forcing: Ranking State High Courts and Their Judges." *Duke Law Journal* 58(7): 1314–82.

Clark, Tom S., and Benjamin Lauderdale. "Locating Supreme Court Opinions in Doctrine Space." *American Journal of Political Science* 54(4): 871–90.

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.

Corely, Pamela C., Amy Steigerwalt and Artemus Ward. 2013. *The Puzzle of Unanimity: Consensus on the United States Supreme Court*. Stanford, CA: Stanford University Press.

Cross, Frank B., James F. Spriggs II, Timothy R. Johnson, and Paul J. Wahlbeck. 2010. "Citations in the U.S. Supreme Court: An Empirical Study of their Use and Significance." *University of Illinois Law Review*, (2): 489–576.

Devins, Neal and Lawrence Baum. 2019. *The Company They Keep: How Partisan Divisions Came to the Supreme Court*. New York, NY: Oxford University Press.

Devins, Neal and Will Federspiel. 2010. "The Supreme Court, Social Psychology, and Group Formation" in *The Psychology of Judicial Decision Making*, eds. David Klein and Gregory Mitchell. Chapter 6, pp. 85–102. New York, NY: Oxford University Press.

DuBay, William H. 2004. *The Principles of Readability*. Cost Mesa, CA: Impact Information.

Edwards, Harry. 1998. Collegiality and Decision Making on the D.C. Circuit. *Virginia Law Review* 84(7): 1335–70.

Federal Judicial Center. 2013. *Judicial Writing Manual: A Pocket Guide for Judges.* 2nd Edition.

Feldman, Adam. 2016. "Counting on Quality: The Effects of Merits Brief Quality on Supreme Court Opinion Content." *Denver Law Review* 94(1): 43–70.

Gardner, James A. 1993. *Legal Argument: The Structure and Language of Effective Advocacy.* Charlottesville, VA: The Michie Company.

George, Joyce J. 2007. *Judicial Opinion Writing Handbook*. Buffalo, NY: William S. Hein & Co., Inc.

Giles, Micheal W., Virginia A. Hettinger, and Todd Peppers.2001. "Picking Federal Judges:: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54(3): 623–41.

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.

Gruhl, John. 1980. "The Supreme Court's Impact on the Law of Libel: Compliance by Lower Federal Courts." *Western Political Quarterly* 33(4): 509–19.

Hansford, Thomas G. and Chelsea Coe. 2019. "Linguistic Complexity, Information Processing, and Public Acceptance of Supreme Court Decisions." *Political Psychology* 40(2): 395—412.

Hansford, Thomas G. and James F. Spriggs II. 2006. *The Politics of Precedent on the U.S. Supreme Court*. Princeton, NJ: Princeton University Press.

Hettinger, Virginia A., Stephanie A. Lindquist, and Wendy L. Martinek. 2004. "Comparing the Attitudinal and Strategic Accounts of Dissenting Behavior on the U.S. Courts of Appeals." *American Journal of Political Science* 48(1): 123–37.

Heylighen, Francis and Jean–Marc Dewaele. 1999. Formality of Language: Definition, Measurement and Behavioral Developments. Center Leo Apostel, Free University of Brussels: Internal Report.

Hinkle, Rachael K. 2015. "Legal Constraint in the U.S. Courts of Appeals." *Journal of Politics* 77(3): 721–35.

Hume, Robert J. 2009. "The Impact of Judicial Opinion Language on the Transmission of Federal Circuit Court Precedents." *Law and Society Review* 43(1): 127–50.

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

"Judicial Writings Tips." 2017. The National Judicial College. `http://www.judges.org/judicial-writings-tips/` (December 15, 2017).

Kaheny, Erin B., Susan Brodie Haire and Sara C. Benesh. 2008. "Change over Tenure: Voting, Variance, and Decision Making on the U.S. Courts of Appeals." *American Journal of Political Science* 52(3): 490–503.

Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York, NY: Macmillian Publishers.

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.

Klein, David E. 2002. *Making Law in the United States Courts of Appeals*. New York, NY: Cambridge University Press.

Klein, David and Darby Morrisroe. 1999. "The Prestige and Influence of Individual Judges on the U.S. Courts of Appeals." *The Journal of Legal Studies* 28(2): 371–91.

Kuersten, Ashlyn and Susan B. Haire 2007. Songer Database Update. Available at `http://artsandsciences.sc.edu/poli/juri/appct.htm`.

Landes, William M., Lawrence Lessig and Michael E. Solimine. 1998. "Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges." *The Journal of Legal Studies* 27(2): 271–332.

Landes, William M. and Richard A. Posner. 1993. "The Influence of Economics on Law: A Quantitative Study." *Journal of Law and Economics* 36(1): 385–424.

Lebovits, Gerald and Lucero Ramirez Hidalgo. 2009. "Advice to Law Clerks: How to Draft Your First Judicial Opinion." *Westchester Bar Journal* 36(1): 29–37.

Lempert, Daniel. Forthcoming. "On Separation of Powers and Obfuscation in U.S. Supreme Court Opinions." *Journal of Law and Courts*.

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

Manski, Charles F. and Steven R. Lerman. "The Estimation of Choice Probabilities from Choice Based Samples." *Econometrica: Journal of Econometric Society*. 45(8): 1977–88.

Mantel, Nathan. 1973. "Synthetic Retrospective Studies and Related Topics." *Biometrics* 29(3): 479–86.

Martin, Andrew D. and Kevin M. Quinn. 2007. "Assessing Preference Change on the US Supreme Court." *Journal of Law, Economics and Organization* 23(2): 365–85.

Martinek, Wendy. 2010. "Judges as Members of Small Groups" in *The Psychology of Judicial Decision Making*, eds. David Klein and Gregory Mitchell. Chapter 6, pp. 85–102. New York, NY: Oxford University Press.

Miceli, Thomas J. and Metin M. Cosgel. "Reputation and Judicial Decision–Making." *Journal of Economic Behavior and Organization* 23(1): 31–51.

Nelson, Michael J. and Rachel K. Hinkle. 2018. "Crafting the Law: How Opinion Content Influences Legal Development." *Justice System Journal* 39(2): 97–122.

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.

Rieke, Richard D. and Randall K. Stutman. 1990. *Communication in Legal Advocacy*. Columbia S.C.: University of South Carolina Press.

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

Schauer, Frederick. 1995. "Opinions as Rules." *University of Chicago Law Review* 62(4): 1455–1477.

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

Shapiro, Stephen M., Kenneth S. Geller, Timothy S. Bishop, Edward A. Hartnett and Dan Himmelfarb. 2013. *Supreme Court Practice*. 10th ed. Arlington, VA: Bloomberg BNA.

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.

Shepard's Company. 1993. Shepard's Citations In-House Training Manual. Photocopy.

Songer, Donald R. 1987. "The Impact of the Supreme Court on Trends in Economic Policy Making in the United States Courts of Appeals." *The Journal of Politics* 49(3): 830–41.

Songer, Donald R. and Reginald S. Sheehand. 1990. "Integrating Alternative Approaches to the Study of Judicial Voting: Obscenity Cases in the U.S. Courts of Appeals." *American Journal of Political Science* 36(4): 963–82.

Songer, Donald R., Jeffrey A. Segal and Charles M. Cameron. 1994. "The Hierarchy of Justice: Testing a Principal-Agent Model of Supreme Court-Circuit Court Interactions." *American Journal of Political Science* 38(3): 637–96.

Spriggs, James F. II and Thomas G. Hansford. 2000. "Measuring Legal Change: The Reliability and Validity of Shepard's Citations." *Political Research Quarterly* 53(2): 327–41.

Spriggs, James F., and Thomas G. Hansford. "The U.S. Supreme Court's Incorporation and Interpretation of Precedent." *Law and Society Review* 36(1): 139–160.

Szmer, John, Robert K. Christensen and Samuels Grubbs. 2020. "What Influences the Influence of U.S. Courts of Appeals Decisions?" *European Journal of Law and Economics* 49(1): 55–81.

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.

Wald, Patricia. 1995. "A Reply to Judge Posner." *University of Chicago Law Review* 62(4): 1451–1455.

Westerland, Chad, Jeffrey A. Segal, Lee Epstein, Charles M. Cameron and Scott Comparato. 2010. "Strategic Defiance and Compliance in the U.S. Courts of Appeals." *American Journal of Political Science* 54(4): 891–905.