Measuring Judicial Ideology Through Text

Jake S. Truscott
University of Florida

Michael K. Romano
Shenandoah University

Abstract

Explorations of ideology retain special significance in contemporary studies of judicial politics. While some existing methodologies draw on voting patterns and coalition alignments to map a jurist's latent features, many are otherwise reliant on supplemental proxies – often directly from adjacent actors or via assessments from various prognosticators. We propose an alternative that not only leverages observable judicial behavior, but does so through jurists' articulations on the law. In particular, we adapt a hierarchical factor model to demonstrate how latent ideological preferences emerge through the written text of opinions. Relying on opinion content from Justices of the Supreme Court, we observe a discernible correlation between linguistic choices and latent expressions of ideology irrespective of known ideological preferences or voting patterns. Testing our method against validated and commonly used measures of judicial ideology, we find that our approach strongly correlates with existing measures. We conclude by discussing the intuitive power of text as a feature of judicial ideology, as well as how this process can extend to judicial actors and institutions beyond the Supreme Court.

Keywords: Supreme Court; Judicial Ideology; Text-as-Data

Introduction

Terminology is an expression of language and language is an expression of ideology (Thompson, 1987; Woolard and Schieffelin, 1994). The choices made by judges regarding how they express decisions through written opinions is thus an expression of preferences, which are shaped by both personal inclinations (Segal and Spaeth, 2002), strategic machinations (Bailey and Maltzman, 2011; Bonneau et al., 2007), and considerations of the perceived audience (Baum, 2009; Romano and Curry, 2019). The ramifications of these choices are consequential for the interpretation of the law, as ideological differences ostensibly result in lexical variation employed to delineate facts and legal aspects pertinent to a decision. Put more simply: the words judges choose matter and are a reflection of their own ideological beliefs, which facilitates a considerable impact on how we know and speak about the law. An illustrative example is the contrasting use of terms such as "healthcare provider" – a neutral descriptor for medical professionals – and "abortionist" – a perceptively derogatory term endowed with legal significance through rulings like *Dobbs v. Jackson Women's Health Organization* (2022). Despite serving the same lexical purpose, these terms carry distinct ideological meanings.

There remains a notion, particularly among some members of the legal academy and especially among Article III judges, that jurists are otherwise precluded from exhibiting ideology as a component of their decision-making. This belief retains special significance in American government, insofar as deeply rooted principals in the separation of powers system imply that judicial actors are insulated from the political influences ingrained within elected branches. However, studies of judicial decision-making routinely find support for voting behaviors being emblematic of distinctive ideological preferences. Even if we presume that judges resist actively considering their own ideologies when making decisions, former Chief Justice Rehnquist's (1986) proclamation that judging does not occur in a vacuum seems indicative of the fact that judging may be a matter of perspective.[1] Perspective not only informs judges how to approach jurisprudence and adjudication but inevitably how they shape the law, which "by extension, has significant social and economic consequences for individual litigants and society" (Bonica and Sen, 2021, 97). As Romano and Curry (2019) note with respect to state supreme courts, decisions on the merits can often tell us the baseline directionality of the law – i.e., whether a court swung liberal or conservative on a particular issue or remained neutral, but it cannot explain *why*. A deeper examination of the language judges choose when crafting opinions reveals not just what the law *means*, but also reflects on how judges *think*, and provides a richer tapestry of ideological cues that might be missed by simply looking at the votes (Bailey and Maltzman, 2011; Hinkle, 2015; Songer and

[1] Justice Rehnquist's exact quote was specific to the influence of public opinion and outside audiences on judges, however it is not a far-step to say these external influences also influence the "philosophy" or ideology of a judge when making a decision. See Rehnquist's comments on Judicial Isolation and Constitutional Law: https://www.nytimes.com/1986/04/17/us/required-reading.html

Lindquist, 1996).

Here, we build on these measures by introducing a process that leverages the wealth of data obtained from written text. Using a `Wordshoal` approach, which has proven to be a reliable tool for hierarchical estimation of latent preferences in legislative speech (Lauderdale and Herzog, 2016), we demonstrate how ideology emerges in legal text. We thus propose an alternative strategy for estimating spatial models of judicial ideology by relying on written opinions as a form of speech and expression of belief, aiming to demonstrate how lexical variation in these opinions can be used to estimate latent preferences. To showcase this, we examine non-unanimous cases – including majority opinions, (special) concurrences, and dissents – decided by the Supreme Court between the 2005 and 2022 terms. This time period, which corresponds with John Roberts' tenure as Chief Justice, has been highlighted as having an increased level of partisanship and polarization develop between what is commonly believed to be a collegial court (Salamone, 2018).[2] As such, we believe that the Roberts Court represents a good "first test" for whether judicial voice, through the opinion, correlates with ideology. We contend that Justices form distinctive ideological voices through the creation and adherence to specific forms of precedent across their careers. Measuring judicial opinions in this way will offer better clarity to the creation of judicial regimes and the development of clearer concepts of judicial philosophy and how ideology impacts votes.

Measurement of the ideologies of judicial actors remains an important prospect. However, scholars and observers alike continue to debate the merits of quantitatively representing these features. From rudimentary labeling of jurists using proxies to comprehensive statistical measures using Bayesian processes and machine learning, several mechanisms exist to estimate these observed and latent preferences. These efforts have culminated in a rich body of methods for estimating judicial ideology. We argue that `Wordshoal` provides an additional step in the right direction toward furthering our collective ability to open the "black box" of judicial decision-making. In doing so, we believe that that our measures represent an important contribution in furthering our understanding how judges' latent ideologies become activated beyond mere voting. By considering how language reveals dormant aspects of ideology used to justify legal decision-making, we argue that our methodology better situates judicial ideology into a broader policy (or issue) space and engages with the nuances of judicial behavior. In doing so, we not only take the law seriously, but further showcase how ideology is an important determinant of what becomes law.

[2] *Note*: We provide an extended discussion in the Supplemental Appendix regarding our selection of the Roberts Court as our principal observation period, as well as provide summary statistical information regarding the data (opinion-level observations) employed in this analysis.

Understanding Judicial Motives: The Search for Judicial Ideology

Since Pritchett (1942) and Murphy (1964), judicial scholarship has placed pre-eminent importance on reliability and accuracy in measuring the ideological positions of judges.[3] It is generally accepted within political science that judges' policy preferences play an important role in judicial behavior and decision-making, both in the United States and comparatively abroad. However, measuring judicial preferences and the extent to which they are influenced by both internal and external factors has been the dominant point of contention for expanding our understanding of judicial behavior, as well as critiquing the limits of judicial ideology to explain decision making.[4]

Researchers have been implementing various measures of judicial ideology for several decades. Bonica and Sen (2021) provide a comprehensive overview of this literature, which – as they describe – has exhibited considerable change in the breadth of methodological rigor. Understandably, the bulk of this research concerns the Supreme Court,[5] though scholars are often keen to warn against over-extrapolation of inferences from the Justices' behaviors because it represents a small sample of decision-making and "places considerable importance on nine idiosyncratic individuals who are relatively unconstrained in their position atop the American judicial hierarchy (Bailey, 2007)" (Bonica and Sen, 2021, 98). Nonetheless, ideology – particularly when measured through judge-level features of partisanship – has proven to be reliable predictors of the Justices' decision-making. We provide an overview for some of these various methodologies in Table 1, which include both the source of observable data, as well as the approaches' strengths and weaknesses posited by Bonica and Sen (2021).[6]

While political science has embraced the role of ideology in judicial decision making, some legal scholarship and other observers remain unconvinced. As noted by Fischman and Law (2009), "there is little reason to expect those who practice or teach the craft of legal argument to embrace a body of research that questions the extent to which judicial decision-making is actually driven by legal

[3] While we touch briefly on the various ways scholars have measured judicial ideology, our intention here is not to relitigate which measurement strategy is (arguably) best, most accurate, or the most sincere representation. For an excellent and comprehensive discussion of measurement strategies, *see* Bonica and Sen (2021).

[4] The list of citations here would arguably be immeasurable and surely incomplete. We agree with the sentiment of Hughes et al. (2023, 1) that "judicial ideology is a cornerstone of public law" – one that all research must touch in some way.

[5] "This is no surprise: the US Supreme Court is the most important court in the country and the final stopping point for many politically sensitive issues. Also, from a research standpoint, the Supreme Court lends itself relatively well to ideological measurement. First, unlike most other courts in the United States, all nine members of the Court hear and vote on cases together. Second, a small and tractable docket makes it possible to subjectively hand-code cases in order to estimate judicial ideology" (Bonica and Sen, 2021, 98).

[6] We would note, however, that Table 1 is by no means fully encompassing. For additional information, including those related to measures of judicial ideology at the state level, *see* Brace et al. (2000), Hughes et al. (2023), and Windett et al. (2015), among others.

Table 1: Overview of Existing Measures of Judicial Ideology (Bonica and Sen, 2021)

Type	Data	Strengths	Critiques
Segal-Cover[1]	Pre-confirmation newspaper coverage from Conservative and Liberal news editorials.	- Uses external information to predict ideology. - Draws on expertise of Court observers (and reporters) to characterize a nominee's partisanship. - Strong correlation between Justices' scores and votes (Segal and Cover, 1989, 561).	- Scores are static, whereas judicial ideology is known to 'drift' over time (Owens and Wedeking, 2012; Epstein et al., 2007). - Requires an extensive volume of pre-confirmation information. - Relies on subjective assessment of ideology.
Martin-Quinn[2]	Observed votes in cases.	- Estimates relative measures of judicial ideology with uncertainty. - Proven reliable across literatures, especially Congressional ideology (Clinton et al., 2004). - Relies on observable behaviors (votes) rather than inferred ideology.	- Only feasible if cases are decided together as a single body; not easily transferable across courts. - Endogeneity concern: random walk prior from previous term affects prediction; voting behavior used as both DV and IV. - Unidimensional and assumes the "distribution of case characteristics is constant over time" (Ho and Quinn, 2010).
Appointment Indicators[3]	Party or ideal-point proxies.	- Presidential appointments tend to align with ideological preferences, subject to Senate constraints. - Adequate correlation for politically divisive issues generally aligned along Liberal–Conservative dimensions.	- Assumes Democratic appointees are more liberal and Republicans more conservative. - Assumes Senate and appointment constraints are sufficient to predict ideological variance. - Substantial variation exists among judges appointed by the same president.
DIME[4]	Campaign contributions by federal judges or law clerks (proxy measure).	- Approximately 80% of federal appeals judges contributed to campaigns between 2001 and 2016 (Bonica and Sen, 2017). - Draws on dynamic behaviors of clerks, updated annually. - Captures ideological variance: stronger partisanship reflected accordingly.	- Indirect measure for judges who do not contribute, especially at the Supreme Court level. - Assumes judges hire ideologically similar clerks (though mitigated by averaging scores across clerks).

[1] *See* Segal and Cover (1989)
[2] *See* Bonica and Sen (2017)
[3] *See* Epstein et al. (2007); Giles et al. (2001); Segal and Spaeth (2002), among others.
[4] *See* Bonica and Sen (2017)

argument" (134).[7] For as long as judicial scholars have amassed evidence that a principal component of judicial behavior is the sincere policy preferences of the judge (or judges), legal scholars have labelled their findings overtly political (Edwards, 1984), "born in a congeries of false beliefs," (Tamanaha, 2009, 687-688), or unfortunately, "innocently ignorant" (Cross, 1997, 251). The primary concern among these and other legal scholars is two-fold: First, either by methodological ignorance or biases against theoretical approaches that would minimize the role of "the law" in judging, legal research often contends that judicial scholars have claimed ideological biases in judges without appropriately measuring or explaining what is meant by "ideology" (Fischman and Law, 2009). To be sure, methodological rigor by judicial scholars attempting to measure judicial ideology often minimize the value of clear conceptualization of the subject, choosing instead to state simply that *Justice X votes the way they do because they are very conservative (liberal)*. Ideology is a "highly flexible conceptual tool" (Gerring, 1997, 957), but often boils down to how actors organize and express their opinions and how those opinions are formed by values. For judicial ideology, this baseline is often used to operationalize ideology by focusing on outward actions such as votes, making it appear that courts are "just another political institution," which has dangerous implications for judicial legitimacy (Gibson and Nelson, 2017).

Second, legal scholarship contends that the measurement of ideology itself misses the point of the law by vastly ignoring the role of precedent and legal opinion in decision making. Part of this, we expect, is born from longstanding notions of how common law principles emerge in judicial behavior. Students of the law are often schooled on the ethics of judicial decision-making, particularly as it relates to how stare decisis applies to synthesizing and formulating jurisprudence. Even in most senior legal institutions like supreme courts, a considerable majority of caseloads will consider routine (or even mundane) applications of the law. In these circumstances, especially when the role of a judge is to simply apply codified standards, elements of ideology and interpretation are effectively restrained. An opinion piece published by *The New York Times* entitled "The Supreme Court is Not as Politicized as You May Think" (Donnelly and Loeb, 2023) draws further on this point by arguing that painting the Supreme Court as invariably political requires observers to ignore the vast majority of statutory litigation reviewed by the Justices, almost all of which end in total (or near) unanimity. Yet, while it may be true that much of the Court's docket lack definitive political salience (as does most dockets across the federal and state judiciaries),[8] this contention itself is seemingly the result of self selecting observations to fit a broader thesis.[9] Unanimous opinions can certainly

[7] There is at least some indication that the legal community is shifting its view in regard to the role of ideology, thanks primarily to recent decision making on the US. Supreme Court. Most recently, *see* Hawaiian State Supreme Court Justice Eddin's concurrence in *City and County of Honolulu v. Sunoco* (2023), in which he states "Enduring law is imperiled. Emerging law is stunted. A Justice's personal values and ideas about the very old days suddenly control the lives of present and future generations."

[8] For discussions on political salience, *see* Bailey et al. (2005) and Clark et al. (2015).

[9] The authors in the aforementioned article are themselves self-selective in which cases

constrain ideological preferences, though often when ideology can be exchanged for some level of legal certainty in the case (Corley et al., 2013). In addition, the Supreme Court's discretionary jurisdictions grants the Justices the ability to think far more about policy goals than law (Baum, 1997; Segal and Spaeth, 2002), as it is "easy for them to find legal justification for whatever position they prefer" (Baum, 1997, 64). And while the vast majority of scholars would agree that personal preference and policy are not the only thing that matters (Hansford and Spriggs, 2006; Richards and Kritzer, 2002; Songer and Lindquist, 1996), ideology often dictates which precedents are recognized and promulgated over time and across courts (Fix and Kassow, 2020; Hinkle, 2015).

Understanding Ideology with Words: Text-Based Methods for Ideological Measurement

As inferred by Table 1, the breadth of measures for judicial ideology are considerably varied, both as it relates to sources of observable data and the inferences that can reliably be drawn from them. However, what remains an emerging and largely unexplored element of measurement is data originating from written (or spoken) text. As Bonica and Sen (2021) observe, research leveraging tools for automated text analysis have emerged as a way of studying ideology across other actors and institutions, and especially as it relates to discerning individual and policy-level positions in elected legislatures like the United States Congress. At their core, these studies assume that language retains special significance as cues of underlying preferences. Given that ideology is understood as a system of belief organization and an attempt at understanding the world around us, language is central to its understanding since language provides us with meaning (Diermeier et al., 2012; Thompson, 1987). "Intuitively," according to Diermeier et al. (2012), "a political ideology specifies which issue positions go together, the 'knowledge of what-goes-with-what'" (31).[10] **While it is correct to say that language can evolve, and with it subtly change the ideological nature of words and their meaning, we presume that this is part of a holistic process of idea generation and refinement over time.** This knowledge helps us better discern variance in the framing of particular issues and their overarching sentiment**, and broadens considerations for understanding how to take language seriously as it changes**, [11] as well as offering a key to un-

deserve the most attention to generalize the Justices' behaviors. They choose to ignore 'constitutional cases' – i.e., those that consider applications of constitutional principles, which are overwhelmingly decided without unanimity, and instead tend to reflect underlying notions of liberal and conservative positions. The key, we argue, is to recognize that any attempts to represent ideology – either empirically or otherwise – require scholars and observers to consider the broadest domain of judicial decision-making because unanimity is hardly an assurance and decisions on the merits of great importance often entail divergent coalitions.

[10] *See* also Poole (2007).

[11] While we do not believe that this is an issue for our analysis here given the time period, researchers should consider how language evolves and the stability of a word's meaning over time. Researchers should exercise caution when extending this and other models relating

derstanding elite discourse and how it is promulgated in the mass public. Indeed, given the progressive development of automated tools and machine learning, a wealth of literature has developed or implemented comprehensive analyses of ideologically oriented Congressional speech, such as `Wordscores` (Laver et al., 2003), `Wordfish` (Slapin and Proksch, 2008), and `Wordshoal` (Lauderdale and Herzog, 2016).

Yet, application of these approaches to the courts, and particularly the Supreme Court, remain few. Until recently, Lauderdale and Clark (2014) remained the most pivotal development in this field, as they were able to use an autoregressive preference model to scale case-level ideology for Justices between 1946 and 2005. Others, including Hausladen et al. (2020), similarly approached case-level positions using data from the Federal Courts of Appeals.[12] Even then, both of these approaches scale ideology as a reflection of the cases themselves, rather than judicial actors. More recently, Cope (2024) developed Jurist-Derived Judicial Ideology Scores (*JuDJIS*), which leverage "information...collected by professional survey firms commissioned by the *Almanac of the Federal Judiciary*, a [triennially] published initiative which surveys a stratified sample of qualified experts for each judge" (2).[13] The result is a dynamic measure of judicial ideology recovered from a hierarchical n-gram analysis conditioned on how lawyers familiar with each judge would review their "ability; demeanor; trial practice/oral argument; settlement/opinion quality; and ideology."

Using Text to Determine Judicial Ideology

To better understand how opinion language acts as a signal of judicial ideology, we must first imagine that the outcome in any case exists within a relatively confined "case space" (Lax, 2011). As part of determining case outcomes, judges write opinions in attempts to justify and persuade others to the "correctness" of their decision (Romano and Curry, 2019). Within each case, there exists a specific number of topics that judges can choose from to frame their argument. Judges choose their words carefully when crafting opinions (Romano and Curry, 2019), but different judges will choose language to explain their argument. Choices in topic selection, framing, and language, are conditioned upon an individual judge's views and beliefs concerning what the right course of action should be in determining the outcome of any case. That is, on their

language to ideology, and be aware of how terms are being used and modified as they become ideologically entrenched. This is particularly important for opinion content, where language is often adopted and borrowed from past precedent, but is also part of a deeper consideration on the politics of language and word choice, and the link between communication and political meaning.

[12] Lauderdale and Clark (2014) focused on incorporating Latent Dirichlet Allocation (LDA) to map ideal points for Supreme Court onto dimensions reflective of the particular issues (topics) discussed in the cases the Justices are deciding. Alternatively, Hausladen et al. (2020) use a supervised machine learning approach to predict the ideological 'direction' of case outcomes from the Courts of Appeals from associated opinion texts.

[13] While this work is surely a core development in the literature, it remains that it is still reliant on proxies (i.e., survey responses from lawyers) to supplement judicial behavior.

ideology.

As we noted in an earlier section, attempts to map legal concepts and broader elements of lexical variation onto ideology remain few in studies of judicial politics. To date, the primary contributions to this literature remain Lauderdale and Clark (2014), Hausladen et al. (2020), and Cope (2024). Even then, none are directly applicable to our efforts. Our goal is to disassociate dichotomous voting behaviors or alternative proxies to instead place the burden of mapping ideology using the text of written opinions on the law. Decisions by the Supreme Court represent a unique case study to accomplish these ends.

However, important questions remain concerning if cues of ideology emerge in legal text, and more importantly how we might go about retrieving it. The answer to the first question must be yes. Irrespective of the underlying facts and merits of each case, opinions authored by the Justices represent policy positions. Particularly when the Court is divided, these positions can manifest across multiple opinions, and the Justices are able to use these opportunities to articulate their perspectives (Brace and Hall, 1993; Hall, 1987; Hettinger et al., 2004; Romano and Curry, 2019; Songer, 1982) Alternatively, when the Court is unified, we can assume they are speaking as one voice and thus conveying some measure of certainty that minimizes ideological cues (Corley et al., 2013). In essence, given that the Justices are unconstrained in providing concurrences and dissents, we can derive comprehensive elements of their perspectives in a dimension that is more robust than whether they voted to affirm or reverse.[14]

The second question is clearly more complicated, though research in ideology as a function of articulated positions in legislative settings provides an intuitive path forward. In particular, we leverage work by Lauderdale and Herzog (2016) by applying `Wordshoal` to opinions by the Supreme Court between the 2005 and 2022 terms. At its core, this methodology employs a two-stage estimation strategy to retrieve ideal points from expressed positions offered during debates in the House of Representatives. The Justices' opinions lend themselves well to this approach, given that the motivations for their decision-making are deterministic of their preferences.[15]

The focus of our work toward yielding inferences from broader lexical variance further lends itself well to this methodology. The `Wordshoal` approach places greater emphasis on how text found within documents can be used to articulate position-taking by its author (or authors). In a legislative setting, this processes addresses relative lexical variance from individual subsets (i.e., issue-specific debates) to draw generalizations of the individuals' (legislators)

[14]This is not to say that they aren't motivated to coalesce towards single opinions – i.e., the notion that unanimity (at least theoretically) prescribes more associative weight to the perceived legitimacy of the Court's decision is not lost on the Justices. However, there exist no institutional constraints on the Justices to author separate concurrences or dissents, if they so choose.

[15]That is, the Justices' decisions – expressed as both their votes and the reasoning articulated in their opinions – are reflective of their inherent beliefs. This assumption draws heavily from the attitudinal nature of Supreme Court decision-making (Segal and Spaeth, 2002), being that the Justices are policy-oriented actors whose decision-making reflect authentic preferences.

broader positions. Further, given how attribution of equal predictive weight to specific words (or phrases) across debates of varying substance is sure to yield dubious results, the implications of particular word choice are sensibly restricted to individual debates. In essence, while certain words or phrases may inform us where to draw distinctions between competing coalitions on specific issues, assuming their influence is constant irrespective of subject matter would bias our results. Instead, `Wordshoal` offers a practical alternative that (1) addresses distinct lexical variance as emblematic for expressing "stated positions" (Lauderdale and Herzog, 2016, 375) within reduced (debate) settings, and subsequently (2) considers how that variance contributes to drawing broader discrimination in a legislator's latent preferences. We discuss our amended implementation of `Wordshoal` below.

Implementing `Wordshoal`

`Wordshoal` is implemented using a two-stage hierarchical process, which uses a Poisson scaling model to retrieve debate-level estimates from `Wordfish` (Slapin and Proksch, 2008) and subsequently aggregate to a general latent position for each individual actor. We review this process below and maintain notation from Lauderdale and Herzog (2016), though we indicate where substitutions are made for the Supreme Court versus their original legislative observations using brackets. In this substitution, our notation follows that the Justices [legislators] are indexed as $i \in 1, 2\ldots, N$, cases [debates] as $j \in 1, 2\ldots, M$, and words as $k \in 1, 2\ldots, K$.

$$\omega_{ijk} \sim \rho(\mu_{ijk})$$
$$\rho(\mu_{ijk}) = \exp(\nu_{ijk} + \lambda_{jk} + \kappa_{jk}\psi_{ij})$$

Being that "the frequency that [Justice] i will use word k in [case] j depends on a general rate parameter ν_{ijk} for [Justice] i's word usage in [case] j, word-[case] usage parameters λ_{jk}, κ_{jk} and the individual's [case]-specific position ψ_{ij}. The ν_{ijk} parameters capture the baseline rate of word usage in a given [opinion], which is simply a function of the length of the [opinion]. The λ_{jk} capture variation in the rate at which certain words are used. The κ_{jk} capture how word usage is correlated with the [Justice]'s [case]-specific position ψ_{ij}" (Adapted from Lauderdale and Herzog, 2016, 377).

With these estimates from case-level dimensions, we can subsequently use `Wordshoal` for aggregating to a (in this instance) single latent dimension with normally distributed error.

$$\psi_{ij} \sim N(\alpha_j + \beta_j\theta_i, \tau_i)$$
$$\theta_i \sim N(0, 1)$$
$$\alpha_j, \beta_j \sim N\left(0, \left(\tfrac{1}{2}\right)^2\right)$$
$$\tau_i \sim \mathcal{G}(1, 1)$$

Again, drawing from Lauderdale and Herzog (2016), "this specification means that the primary dimension of word usage variation in individual [cases] can be

more or less strongly associated with the aggregate latent dimension θ being estimated across all [cases], with either positive or negative polarity for any particular [case]. Essentially, this allows the model to select out those debate-specific dimensions that reflect a common dimension (larger estimates values of β_j), while down-weighting the contribution of debates where the word usage variation across individuals seems to be idiosyncratic ($\beta_j \approx 0$). The priors on θ_i and θ_j allow the model to remain agnostic about the relative polarity of individual [case] dimensions, while constraining the common latent dimension of interest to a standard normal scale" (378).

However, our methodology departs slightly from Lauderdale and Herzog (2016) in how we consider the relative responsibility associated with opinion authorship. Unlike Congressional debates, where we can directly associate speeches with particular legislators, opinion authorship on the Supreme Court dictates that a majority (or at least a plurality) of the Justices will coalesce to a single opinion. Similar concerns emerge when we recognize that Justices frequently join concurrences or dissents. It might be problematic to assume that Justices choosing to join opinions – whether they be majorities, concurrences, or dissents – should be granted the same weight in estimating latent ideal points as those who actually authored them. In one sense, we can infer that given the Justices' capacity to author their own concurrences and dissents, failure to deviate from the language of the opinions they're joining is indicative of it being their own words, as well. However, it clearly lacks the same degree of authenticity that structured assumptions in Lauderdale and Herzog (2016) of directly tying each speech to individual legislators. To account for this, we introduce a procedure that assigns relative weights reflective of opinion authorship. We can assume that in any circumstance, authorship of a decision can be weighed fully as a representation of the writer's voice, while choosing not to join an opinion bears no weight. Alternatively, joining majority opinions, dissents, or (special) concurrences might require additional considerations.[16] Lacking sufficient guidance from existing literature, we implement a combination of weighting arrangements to establish whether degrees of associative responsibility induce significant variance in the estimates.[17] We subsequently amend the second set of equations to include parameter ϕ_{ij},[18] representing the relative responsibility of an opinion for Justice i in case j.[19]

[16] While we recognize additional considerations exist concerning opinion authorship in collegial courts like the Supreme Court – in particular, norms concerning opinion assignment – we do not attempt to model such. While we may choose to remedy this in the future, we must acknowledge that Justices are generally unconstrained to author separate opinions. Given such a dynamic, we are confident their choice to sign-on to opinions or otherwise deviate represent behaviors independent of who is assigned authorship of the Court's majority opinion.

[17] For more information regarding the weighting schemes, please *see* the Supplemental Appendix (Table A1).

[18] Such that $\psi_{ij} \sim N(\alpha_j + \beta_j\theta_i, \tau_i\phi_{ij})$

[19] The established approach for `Wordshoal` used by Lauderdale and Herzog (2016) is currently available in the `quanteda` package for `R` (Benoit et al., 2018). We would ask readers to direct themselves to Lauderdale and Herzog (2016) for a full account of the replication materials, though we will make ours (with adjustments for authorship responsibility weights) available upon publication.

Scaling Supreme Court Justices Between the 2005 and 2022 Terms

To this point, we have introduced an amended application of `Wordshoal` as a methodology to estimate latent ideology as an expression of lexical variance without actually considering whether a Justice voted to affirm or reverse. That is, it draws on the expectation that the variance in word choice signals their relative positions and reflects separability among individuals. Granted, the notion of coalition structures raises potential concerns. Apart from our inability to draw the same assumptions of invariably associating opinions to single authors, it is important to further recognize that the Justices have a tendency to coalesce among likeminded partisans – particularly in response to divisive cases. With respect to our attempts to remove voting and ideology from our procedure, we are cognizant that these behaviors encapsulate aspects of both. However, given our application of different weighting schemes and the capacity for the Justices to voice divergent perspectives to any opinion, we are confident that we took the most appropriate steps to map the Justices' ideal points reflective of the language used in their opinions.

We provide estimates of individual Justices reflective of 1,972 opinions authored in 678 non-unanimous cases between the 2005 and 2022 terms below (Figure 1),[20] which coincides with the Roberts Court (2005-present). Apart from the scheme that only applies to those who authored a particular opinion ('Sole Weight'),[21] we observe noticeable consistency in our estimates. Furthermore, the breadth of error – which represent 95 percent confidence intervals – are shown to only be spacious for those whose presence in the data is more limited.[22] Alternatively, those present for longer periods of observation tended to display more consistency in their estimated positions.

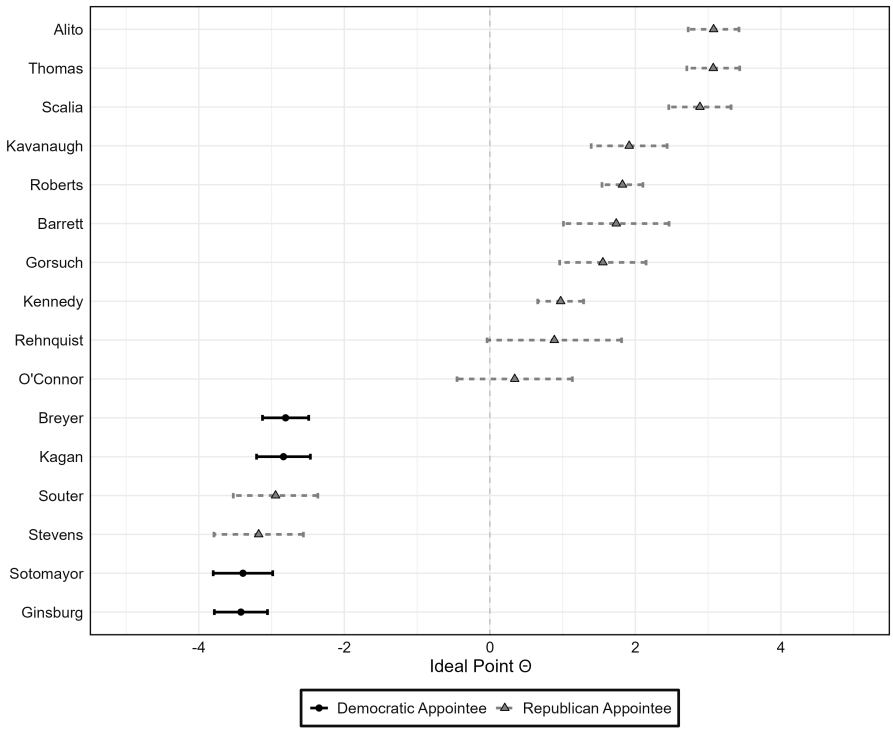
At face value, our estimates retain discernible validity. Insofar as we can provide normative assessments of the Justices' categorical positions on the ideological spectrum, we know that Justices Sotomayor and Ginsburg should populate divergent positions from Thomas and Scalia. Further, the nuance of those we'd expect to populate the center – e.g., Justices Kennedy and O'Connor – is likewise reflected in our estimates. Alternatively, the ordinal rankings of our estimated positions require careful consideration. Notwithstanding measures of potential error, our rankings place Justices Ginsburg and Alito in the most polarized positions. The exact locations of the Justices relative to each other on

[20] Given the emphasis of first-stage `Wordfish` to draw distinctions between positions given divergent word choice on refined subjects (i.e., demarcating language most pivotal to establishing majority and minority coalitions), our estimates do not include unanimous decisions. Please see the Supplemental Appendix for more information regarding data collection and processing.

[21] Where estimates for a particular Justice only consider opinions that they themselves authored.

[22] For example, Justices William Rehnquist (who passed away in 2005), Sandra Day O'Connor (who retired in 2006), David H. Souter (2009), and John Paul Stevens (2010). This also includes newer Justices, such as Justices Amy Coney Barrett and, to a lesser extent, Brett Kavanaugh and Neil Gorsuch.

Figure 1: `Wordshoal` Estimates (High Weight) by Justice and Weighting Scheme



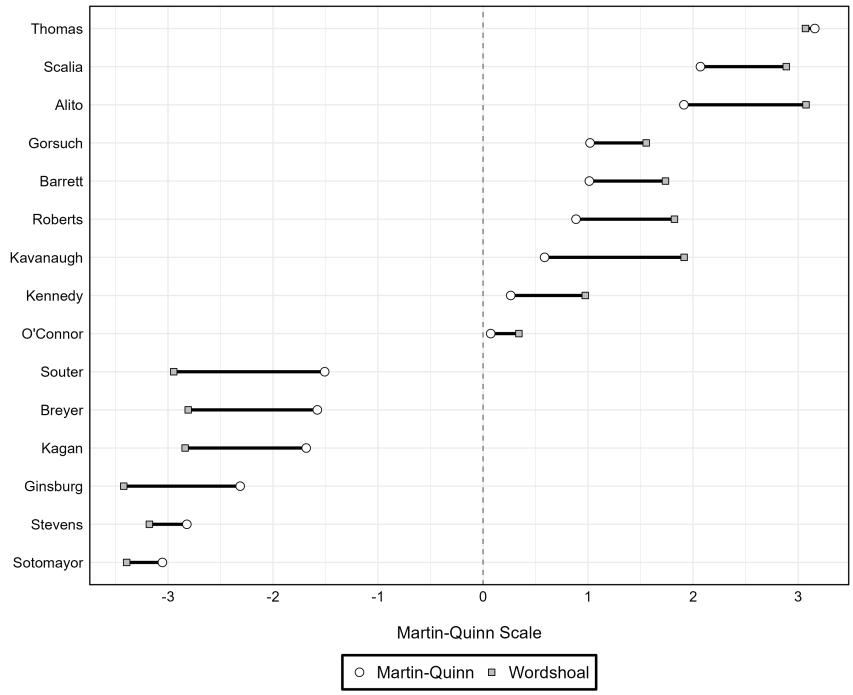
Note: Left-Right axis (θ_i) scaled to represent progressive values of Liberalism to Conservatism. For more information regarding the associated weighting schemes, *see* the Supplemental Appendix (Table A1).

our unbounded scale might lead to some debate concerning whether the ordinal and cardinal distances are, in a sense, the most accurate representation.[23] What we can say is that the relative location of the Justices on this scale lends credibility to its accuracy, particularly given the clustering of objectively likeminded Justices.

Further, another test of our estimates' validity is to compare them to established methodologies. For this, we draw on ideal points estimated by Martin and Quinn (2002), from which most measures of judicial ideology that consider the Justices' voting behaviors draw their inspiration. Given the dynamic nature of their scaling procedure, we plot the average position for each Justice against those recovered using `Wordshoal` (Figure 2). However, while the directionality of both measures is symmetric, their relative magnitudes differ. To account for this, we standardize both using z-score normalization, ensuring they share a common scale with a mean of zero and a standard deviation of one. We then plot the relative variance of Justice-level estimates (Figure 3).

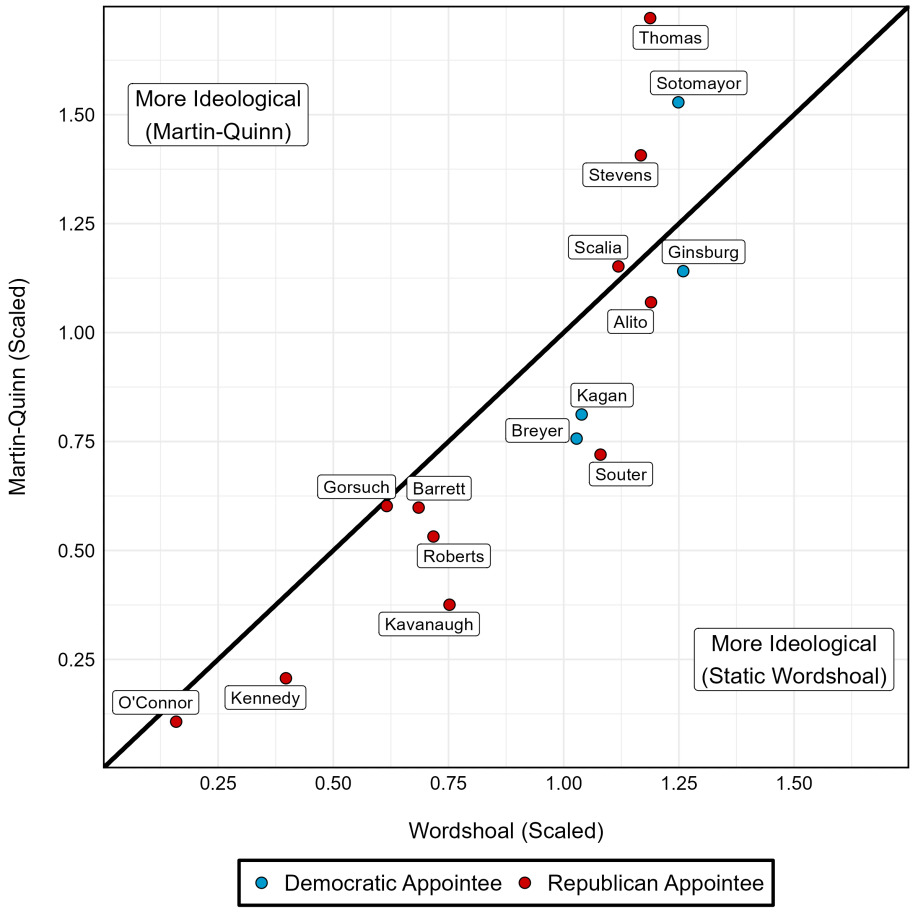
[23] E.g., the ordering of Justices Ginsburg, Sotomayor, Stevens, and Souter on the Left (most liberal), and Alito, Thomas, Scalia, and Kavanaugh on the Right (most conservative).

Figure 2: `Wordshoal` Estimates (High Weight) versus Martin and Quinn (2002)



Note: This figure compares `Wordshoal` estimates using *High Weights* with the dynamic ideal point estimates by Martin and Quinn (2002). Both represent Justice-level means across the 2005-2021 observation terms, where averages for Martin-Quinn were recovered from their `post_mn` variable. While the scales are not equal or normalized, the (left-right) progression of Liberalism to Conservatism is observed in both. Correlation = 96.6%.

Figure 3: Comparison of Static `Wordshoal` (High Weight) versus Martin-Quinn



Note: Both axes represent the absolute values of each Justice's estimated ideal point using Martin-Quinn and Static `Wordshoal`, where both scales are standardized using z-score normalization. Values for Martin-Quinn (Static `Wordshoal`) are measured using the average of `post_mn` (θ_i) across the observation period. Points nearest to the diagonal segment indicate greater correlation between the relative placement of a Justice. Alternatively, values above (below) the diagonal indicate greater relative ideological placement in Martin-Quinn (Static `Wordshoal`). Correlation = 86.84%.

Notwithstanding the difference in observable behaviors used to estimate latent ideal points, we observe a strong correlation between our measure and Martin-Quinn. A particular difference is the relative position of Clarence Thomas, who is estimated to be discernibly more conservative in Martin-Quinn, as well as subtle variance in the Justices' ordinal rankings. However, we are not of the mind that our measure is failing to discern what is arguably the "correct"

ordering or placement of each Justice. Instead, it is important to recognize the incomplete nature of the data. Being that, while Martin-Quinn maps the full extent of each Justice's tenure on the Court, ours only represents the full service of seven.[24] Accounting for those whose tenures (as of the 2022 term) are fully represented in the data, we observe strong similarity with Martin-Quinn's ordinal rankings. We imagine that incorporating the absent data for the remaining eight Justices whose tenures are not fully represented could remediate any potential shortcomings in (ordinal) placement. Even then, the fact that we observe such high degrees of correlation with an established and respected measure of judicial ideology underscores the robustness of our methodology. The Justices are not only keen to express ideology as a principal feature of their authored opinions, but our methodology is able to map these distinct behaviors.

Discussion

Ideology serves as a principal element of judicial decision-making. Research that maps latent features of ideology among judicial actors has drawn from a multitude of direct and indirect observational behaviors. Those with the most robust methodologies – particularly Martin and Quinn (2002) – center on estimates rooted in the Justices' dichotomous voting behaviors. However, much as theirs and others' approaches drew considerably from advancements in studies of legislative behavior,[25] we suggest a similar adoption strategy. Namely, our work adapted a two-stage `Wordshoal` procedure from Lauderdale and Herzog (2016) to instead represent Supreme Court Justices in a space reflective of their opinion writing behaviors. Our goal is not to suggest that this approach is invariably the best and should be adopted by researchers without any hesitation. Instead, we aimed to demonstrate that latent features of ideology are not only prevalent in judicial text, but that our approach provides a means to explore judicial behaviors through a lens that incorporates articulated perspectives on the law.

Using decisions authored between the 2005 and 2022 terms, we implemented `Wordshoal` to develop latent ideal points for each Justice serving during this observation period. Our efforts yielded impressive degrees of correlation with established measures like Martin and Quinn (2002), largely irrespective of how we chose to weight authorship of majority opinions. While there are subtle differences in ordinal placement between our approach and Martin-Quinn, we expect these issues to be the result of incomplete data – which we plan to rectify in the future. Taken together, our approach provides a robust measure of latent judicial ideology rooted in lexical variance that, for all intents and purposes, mirrors normative assumptions of the Justices' relative ideological positions.

However, like any measure of judicial ideology referenced in Table 1, we must recognize there are potential shortcomings. First, as we noted in an earlier sec-

[24] Chief Justice Roberts (Appointed 2005), as well as Justices Alito (2006), Sotomayor (2009), Kagan (2010), Gorsuch (2017), Kavanaugh (2018), and Barrett (2020).

[25] Particularly Poole and Rosenthal (1985); Clinton et al. (2004), among others.

tion, our desire to fully remove coalition alignments as a predictor is something we are not sufficiently confident can fully be achieved. Unlike the initial application of `Wordshoal` in Congressional settings (Lauderdale and Herzog, 2016), the Court's reliance on coalescing to majority and other opinions often negates our ability to draw 1:1 assessments of Justice-author observations. That is, at minimum, four Justices will coalesce to a single plurality opinion – though a minimum of five to constitute a majority opinions is obviously more common. To account for this, we introduced (1) any and all opinions within the scope of decided cases,[26] and (2) a collection of weighting schemes to induce variance in how particular opinions should be associated with the non-authoring Justices joining majorities, dissents, and (special) concurrences. Apart from circumstances where we restricted responsibility to only those who authored a particular opinion, we observe discernible continuity and reliability in the estimates. Even then, our "Sole Weights" scheme has a greater influence on scaled positions, rather than ordinal rankings. Further, we are not convinced that this final restrictive scheme is the best representation of how we should discern ideological positions, if nothing for the fact that coalescing to a written opinion is a conscious decision that the Justices are by no means forced to make. If the Justices felt it was necessary to deviate from the language of opinions written to reflect the majority (minority) position, they are entirely free to do so. Choosing to (or not) is reflective of their authentic feelings and should be viewed as such to some tangible degree.

Alternatively, a second concern that we don't address fully in earlier sections is the constancy assumption underpinning the estimates. In particular, `Wordshoal` is estimated in two stages; first from a local (case-level) dimension, then subsequently to a common dimension reflective of the predictive weights we can derive from the larger collection of local dimensions. However, the estimates themselves are considered to be constant. As such, our estimates displayed in Figure 1 are static representations of each Justices' latent positions across the observation period. Yet, given the broader literature regarding "ideological drift" (Epstein et al., 2007), we know the Justices' positions are unlikely to be static across their tenures. With this in mind, we explore a dynamic specification in the Supplemental Appendix by allowing for variance in θ_i across successive terms. The preliminary results yield interesting inferences, particularly as it relates to improving the relative correlation to Martin and Quinn (2002) using normalized scales to approximately 96 percent.

Future efforts will surely lead us to continue our efforts, particularly as it relates to overcoming these established shortcomings and continuing to extrapolate on the dynamic specification. However, the underlying motivation of this research is not to fill some longstanding gap in judicial politics literature or relitigate contentious debates that often arise with respect to measures of judicial ideology, but rather to demonstrate what we believe to be a strategy for estimating judicial ideology using an observation strategy that leverages articu-

[26] This provides additional robustness to the observation data, insofar as the inclusion of majority opinions, (special) concurrences, and dissents offers a broader accounting of justice-level perspectives, rather than simply addressing whether a Justice coalesced with the majority.

lated perspectives on the law. We retain that our long-term goal is to measure ideology among jurists across both state and federal courts.[27] We anticipate two primary obstacles. First, a sizeable and sufficient repository of published opinions from each of the 52 state courts of last resort, as well as from the federal district, appellate, and Supreme courts. As we demonstrate in Figure 1, the greatest Justice-level variance appears to emerge in response to insufficient data. Those whose tenures are most fully encapsulated in our observation period demonstrated much greater consistency in their estimated positions in the first dimensional space. Second, while retrieving an assortment of estimates spanning state and federal judicial institutions would surely be a worthwhile exercise, we are devoted to developing a strategy that would make them comparable across courts and hierarchy.

References

Bailey, M. A. (2007). Comparable preference estimates across time and institutions for the court, congress, and presidency. *American Journal of Political Science 51*(3), 433–448.

Bailey, M. A., B. Kamoie, and F. Maltzman (2005). Signals from the tenth justice: The political role of the solicitor general in supreme court decision making. *American Journal of Political Science 49*(1), 72–85.

Bailey, M. A. and F. Maltzman (2011). *The constrained court: Law, politics, and the decisions justices make*. Princeton University Press.

Baum, L. (1997). The puzzle of judicial behavior.

Baum, L. (2009). *Judges and their audiences: A perspective on judicial behavior*. Princeton University Press.

Benoit, K., K. Watanabe, H. Wang, P. Nulty, A. Obeng, S. Müller, and A. Matsuo (2018). quanteda: An r package for the quantitative analysis of textual data. *Journal of Open Source Software 3*(30), 774–774.

Bonica, A. and M. Sen (2017). A common-space scaling of the american judiciary and legal profession. *Political Analysis 25*(1), 114–121.

Bonica, A. and M. Sen (2021). Estimating judicial ideology. *Journal of Economic Perspectives 35*(1), 97–118.

Bonica, A. and M. J. Woodruff (2015). A common-space measure of state supreme court ideology. *The Journal of Law, Economics, and Organization 31*(3), 472–498.

[27] To date, the most reliable measure offering a means to bridge judges across the state and federal hierarchy has been Bonica and Woodruff's (2015) judicial "CFscore" methodology. Alternatively, the most reliable alternatives for estimating ideology among state-level jurists is often viewed to be Party-Adjusted Judge Ideology (PAJID) scores (Brace et al., 2000; Hughes et al., 2023) and Windett et al.'s (2015) scores.

Bonneau, C. W., T. H. Hammond, F. Maltzman, and P. J. Wahlbeck (2007). Agenda control, the median justice, and the majority opinion on the us supreme court. *American Journal of Political Science 51*(4), 890–905.

Brace, P. and M. G. Hall (1993). Integrated models of judicial dissent. *The Journal of Politics 55*(4), 914–935.

Brace, P., L. Langer, and M. G. Hall (2000). Measuring the preferences of state supreme court judges. *The Journal of Politics 62*(2), 387–413.

Clark, T. S., J. R. Lax, and D. Rice (2015). Measuring the political salience of supreme court cases. *Journal of Law and Courts 3*(1), 37–65.

Clinton, J., S. Jackman, and D. Rivers (2004). The statistical analysis of roll call data. *American Political Science Review 98*(2), 355–370.

Cope, K. L. (2024). An expert-sourced measure of judicial ideology. *Available at SSRN 4742254*.

Corley, P. C., A. Steigerwalt, and A. Ward (2013). *The Puzzle of Unanimity: Consensus on the United States Supreme Court*. Stanford University Press.

Cross, F. B. (1997). Political science and the new legal realism: A case of unfortunate interdisciplinary ignorance. *Nw. UL Rev. 92*, 251.

Diermeier, D., J.-F. Godbout, B. Yu, and S. Kaufmann (2012). Language and ideology in congress. *British Journal of Political Science 42*(1), 31–55.

Donnelly, N. and E. Loeb (2023, October). The supreme court is not as politicized as you may think. *The New York Times*. [Opinion].

Edwards, H. T. (1984). Public misperceptions concerning the politics of judging: Dispelling some myths about the dc circuit. *U. Colo. L. Rev. 56*, 619.

Epstein, L., A. D. Martin, K. M. Quinn, and J. A. Segal (2007). Ideological drift among supreme court justices: Who, when, and how important. *Nw. UL Rev. 101*, 1483.

Fischman, J. B. and D. S. Law (2009). What is judicial ideology, and how should we measure it. *Wash. UJL & Pol'y 29*, 133.

Fix, M. P. and B. J. Kassow (2020). *us supreme court doctrine in the state high courts*. Cambridge University Press.

Gerring, J. (1997). Ideology: A definitional analysis. *Political research quarterly 50*(4), 957–994.

Gibson, J. L. and M. J. Nelson (2017). Reconsidering positivity theory: What roles do politicization, ideological disagreement, and legal realism play in shaping us supreme court legitimacy? *Journal of Empirical Legal Studies 14*(3), 592–617.

Giles, M. W., V. A. Hettinger, and T. Peppers (2001). Picking federal judges: A note on policy and partisan selection agendas. *Political Research Quarterly 54*(3), 623–641.

Hall, M. G. (1987). Constituent influence in state supreme courts: Conceptual notes and a case study. *The Journal of Politics 49*(4), 1117–1124.

Hansford, T. G. and J. F. Spriggs (2006). *The politics of precedent on the US Supreme Court*. Princeton University Press.

Hausladen, C. I., M. H. Schubert, and E. Ash (2020). Text classification of ideological direction in judicial opinions. *International Review of Law and Economics 62*, 105903.

Hettinger, V. A., S. A. Lindquist, and W. L. Martinek (2004). Comparing attitudinal and strategic accounts of dissenting behavior on the us courts of appeals. *American Journal of Political Science 48*(1), 123–137.

Hinkle, R. K. (2015). Legal constraint in the us courts of appeals. *The Journal of Politics 77*(3), 721–735.

Ho, D. E. and K. M. Quinn (2010). Did a switch in time save nine? *Journal of Legal Analysis 2*(1), 69–113.

Hughes, D. A., T. Wilhelm, and X. Wang (2023). Updating pajid scores for state supreme court justices (1970–2019). *State Politics & Policy Quarterly 23*(4), 463–470.

Lauderdale, B. E. and T. S. Clark (2014). Scaling politically meaningful dimensions using texts and votes. *American Journal of Political Science 58*(3), 754–771.

Lauderdale, B. E. and A. Herzog (2016). Measuring political positions from legislative speech. *Political Analysis 24*(3), 374–394.

Laver, M., K. Benoit, and J. Garry (2003). Extracting policy positions from political texts using words as data. *American political science review 97*(2), 311–331.

Lax, J. R. (2011). The new judicial politics of legal doctrine. *Annual Review of Political Science 14*(1), 131–157.

Martin, A. D. and K. M. Quinn (2002). Dynamic ideal point estimation via markov chain monte carlo for the us supreme court, 1953–1999. *Political analysis 10*(2), 134–153.

Murphy, W. (1964). *ELEMENTS OF JUDICIAL STRATEGY*. Univesity of Chicago.

Owens, R. J. and J. Wedeking (2012). Predicting drift on politically insulated institutions: A study of ideological drift on the united states supreme court. *The Journal of Politics 74*(2), 487–500.

Poole, K. T. (2007). Changing minds? not in congress! *Public Choice 131*, 435–451.

Poole, K. T. and H. Rosenthal (1985). A spatial model for legislative roll call analysis. *American journal of political science*, 357–384.

Poole, K. T. and H. Rosenthal (1997). *Congress: A Political-economic History of Roll Call Voting*. Oxford University Press, USA.

Pritchett, C. H. (1942). The voting behavior of the supreme court, 1941-42. *The Journal of Politics 4*(4), 491–506.

Richards, M. J. and H. M. Kritzer (2002). Jurisprudential regimes in supreme court decision making. *American Political Science Review 96*(2), 305–320.

Romano, M. K. and T. A. Curry (2019). *Creating the law: State Supreme Court opinions and the effect of audiences*. Routledge.

Salamone, M. F. (2018). *Perceptions of a Polarized Court: How Division Among Justices Shapes the Supreme Court's Public Image*. Temple University Press.

Segal, J. A. and A. D. Cover (1989). Ideological values and the votes of us supreme court justices. *American Political Science Review 83*(2), 557–565.

Segal, J. A. and H. J. Spaeth (2002). *The Supreme Court and the attitudinal model revisited*. Cambridge University Press.

Slapin, J. B. and S.-O. Proksch (2008). A scaling model for estimating time-series party positions from texts. *American Journal of Political Science 52*(3), 705–722.

Songer, D. R. (1982). Consensual and nonconsensual decisions in unanimous opinions of the united states courts of appeals. *American Journal of Political Science*, 225–239.

Songer, D. R. and S. A. Lindquist (1996). Not the whole story: The impact of justices' values on supreme court decision making. *American Journal of Political Science 40*(4), 1049–1063.

Tamanaha, B. Z. (2009). The distorting slant in quantitative studies of judging. *BCL Rev. 50*, 685.

Thompson, J. B. (1987). Language and ideology: A framework for analysis. *The Sociological Review 35*(3), 516–536.

Windett, J. H., J. J. Harden, and M. E. Hall (2015). Estimating dynamic ideal points for state supreme courts. *Political Analysis 23*(3), 461–469.

Woolard, K. and B. Schieffelin (1994). Language ideology. annual reviews 23.

Supplemental Appendix

The following are the Supplemental Materials referenced in the manuscript *Measuring Judicial Ideology Through Text*. Replication materials will be provided upon request and will be deposited to the `Harvard Dataverse` upon acceptance and publication.[28]

Data Overview

Our data incorporates written opinions (Majority, Concurrences, and Dissents) in non-unanimous cases decided between the Court's 2005 and 2022 terms. Our choice of observation period was two-fold. First, given the broad consensus that the Roberts Court (2005-present) represents a period dominated by seminal rulings that often draw on ideologically divergent coalitions, we anticipated a rich volume of observational data. Second, while we recognize that we are omitting the most recent term (OT2023), much of our ability to demonstrate reliability and robustness required direct comparisons with the ideal points estimated by Martin and Quinn (2002). Notwithstanding the robustness and potential for our amended `Wordshoal` methodology, we recognize that Martin and Quinn (2002) – much like DW-Nominate (Poole and Rosenthal, 1985, 1997; Poole, 2007) – still assumes prominent status as the figurative gold standard of representing the liberalism of Supreme Court Justices. However, their most recent updates lapse with the 2022 term (as of January 2025). Given such, we made the conscious choice to ensure complete overlap between our analysis of the Roberts Court and the extent of comparable analyses using Martin-Quinn.

Data Collection

Opinions authored by Justices between the Court's 2005 and 2022 terms were retrieved principally from mining Justia – a legal repository service offering "free case law, codes, opinion summaries, and other basic legal texts," including opinions by the United States Supreme Court in an indexed, HTML format.[29] Supplemental and other meta data drew from the Supreme Court Database (SCDB).[30]

Data Summary Statistics

Below we provide summaries of Term (Figure A1) and Justice-level (Figure A2) opinion data. Given the need to draw distinctions in ideology (θ_i) given vari-

[28] *Note*: All replication materials, including data retrieval and processing, estimation, and figure (table) compilation was compiled using the `R` programming language and its associated suite `RStudio`.

[29] For more information regarding Justia and its repository of legal documents, please visit Here.

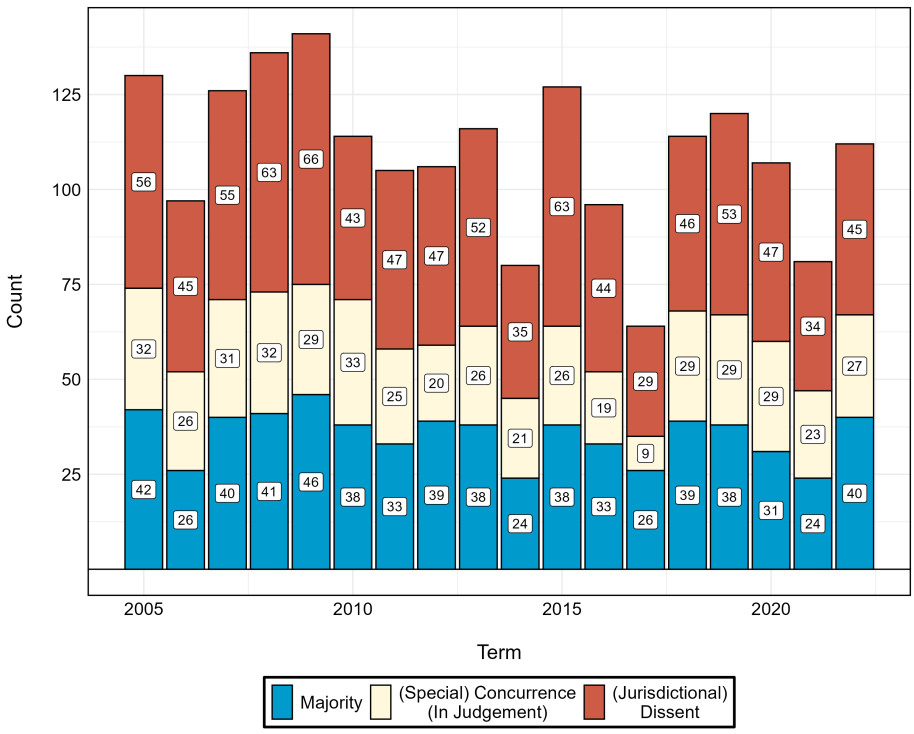
[30] For more information on the SCDB, visit Here. However, as of February 2025, the repository is currently in the processes of relocating from Washington University, St. Louis (MO) to Pennsylvania State University (University Park, PA). For more information, please visit Here.

ance in opinion language, only those cases with an associated concurring or dissenting opinion were included in this analysis.[31] As was conditioned in the `Wordshoal` estimation procedure, Majority, Concurring, and Dissenting Opinions were stratified and organized by their respective typology. For consistency, we observed the following coding rules:

- Majority Opinions: Any opinions representing the majority coalition. These also included opinions representing a *plurality* of the Court, but only if no majority opinion was otherwise rendered.
- Concurring Opinions: Any opinion not reflecting the majority coalition but was nonetheless offered by a member of the non-dissenting (majority) coalition. These included regular concurrences and concurrences *In Judgment*, as well as *special* concurrences (*In-Part*) where the authoring Justice did not fully join the minority.
- Dissenting Opinion: Opinions representing the perspective(s) of the minority (dissenting) coalition(s). These included both regular and *Jurisdictional* dissents.

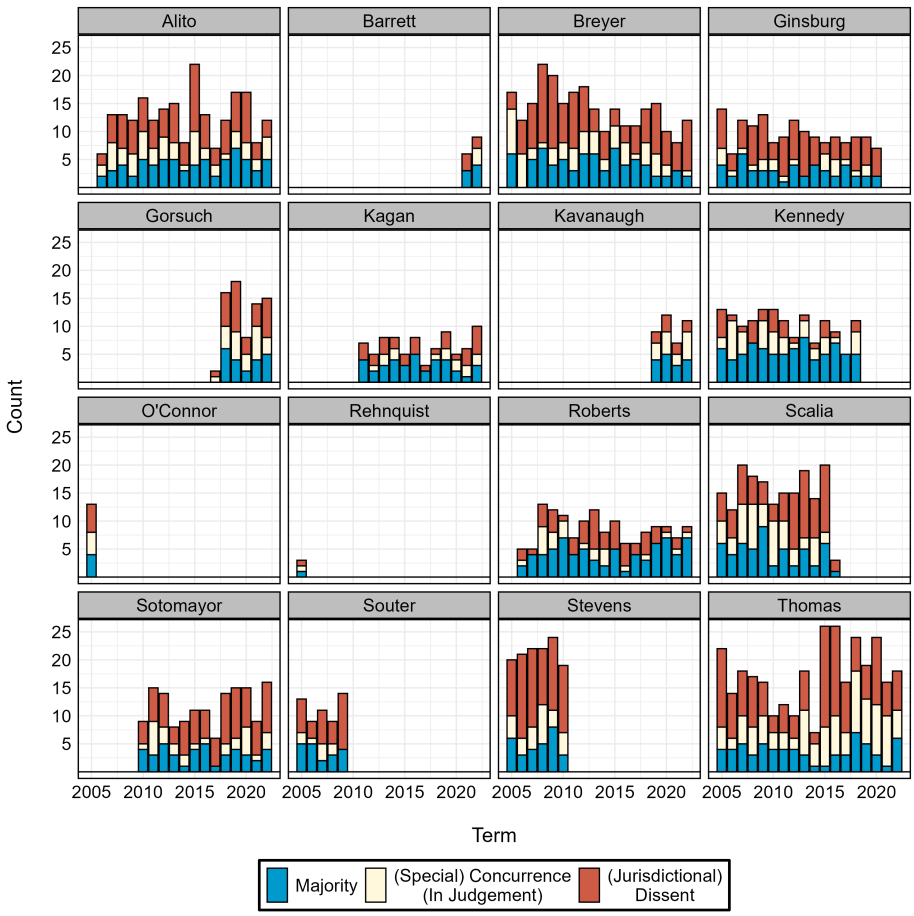
[31] In essence, cases decided without an accompanying concurring or dissenting opinion(s) were omitted. However, cases with an accompanying concurrence, even if the coalitions would otherwise reflect a unanimous agreement of the majority position, were still included. Alternatively, unanimous decisions with no concurring opinions, those decided *per curiam*, or otherwise representing an equally-divided Court with no plurality opinion were omitted.

Figure A1: Distribution of Opinion Data by Term (2005-2022 Terms) Included in `Wordshoal` Estimation.



Note: Bars represent volume of individual opinion types included in the estimation by term. The total observations across terms is 1,972 – while the term-level average is 109 with a median of 113. However, given unforeseen circumstances – e.g., lack of archived opinions on Justia or lack of convergence for at leas two cases (e.g., Docket Numbers: 06-1195 and 08-810), it does not represent a full accounting of non-unanimous opinions between the 2005 and 2022 terms. However, we consider these omissions as missing at random and do not appear to represent any significant inhibition to trusting the validity and robustness of the estimation and subsequent inferences.

Figure A2: Distribution of Opinion Data by Term and Justice (2005-2022 Terms) Included in `Wordshoal` Estimation.



Note: Bars represent the volume of authored opinions (Majority, Concurrences, and Dissents) at the Justice-Term level. For additional information regarding the coding rules to classify opinion types, please *see* the preceding discussion in Section 1.2. Alternatively, please *see* Figure A2 for a special note regarding omissions of observations due to lack of archived opinions on Justia or failure to converge in the estimation procedure.

Testing Weighting of Opinion Ownership

As described in the manuscript, we give special attention to the relative weight(s) associated with opinion authorship. In the original application of `Wordshoal`, Lauderdale and Herzog (2016) are able to identify legislator-specific associations between documents (*speeches*) within group-specific *debates*. In essence, each document-author contribution existed as a single observation, where word-debate usage parameters λ_{jk} and κ_{jk} were unique to legislator i in debate j. Given that much of the members in the chamber (United States Senate or Irish Dáil) do not participate in a given debate, many observations will be missing. However, Lauderdale and Herzog (2016) articulate that in assuming "the positions legislators express are unrelated to their decisions to participate in a debate ..., the measures [they] recover should be interpreted as summaries of the positions actually taken by legislators, relative to their peers, in the debates they participated in" (p. 377). In short, missing observation at random is not detrimental to deriving valid inferences.

However, conditions unique to circumstances where decisions are rendered by panel (and coalitions) – such as the Supreme Court – raise important distinctions. Chief among them being that (1) Instances of Justices failing to participate in the proceedings of any case are fairly uncommon, and (2) as opposed to rendering individual opinions as *seratim*, signed opinions as majority and minority coalitions are the status quo. For all intents and purposes, every $\text{case}_{j..n.}$ will be decided with a majority (plurality or *per curiam*) opinion alongside (perhaps) a concurrence or minority dissent. As only one Justice will author an opinion, those remaining who constitute a majority (minority) coalition will join to varying degrees. Meaning that, unlike Lauderdale and Herzog (2016), each Justice i (generally) participates in every case j and there should be very few missing ψ_{ij}.

As such, the most important question is how to assess relative responsibility of an opinion to a non-authoring Justice. We explore this question in the manuscript and justify our decision to prescribe various weighting schemes as such:

1. Opinion writers, by virtue of being the ascribed author, should be given full weight of association. It is effectively the same as Lauderdale and Herzog (2016) drawing direct association between legislators and their respective speeches. If a distinction must be made to determine which opinions are most reflective of a Justice's latent ideological preferences, the opinions they author should invariably be given the most weight.

2. Notwithstanding implicit dynamics that motivate the Justices towards coalescence, there are no structural barriers to individual (separate) opinion authorship as concurring or dissent. That is, while the holding represents only the opinion associated with a majority (plurality) coalition, there is nothing that inhibits Justices from authoring separate concurrences or dissents. Failure to do so should, at least to some extent, be viewed as a non-authoring Justices assenting to the holding, justification, and other

jurisprudential elements of the opinion they are joining.

3. A justifiable argument exists to assert that Justices who author separate opinions but join broader coalitions (e.g., Justices who offer separate concurrences but remain a member of the majority) should be viewed differently with respect to how much we associate them with those opinions they've joined. In short, the decision to offer separate opinions is evident of their desire to articulate positions separate from those expressed in the principal opinion. Surely, these separate opinions should be given full weight to the authoring Justice. Yet, by still joining the majority – rather than coalescing with the minority – this should still be viewed as a signal of assenting to (at the very least) the judgment. The question then is *how much* should majority opinions be viewed as the perspectives of those who author separate opinions? This string of hypotheticals most acutely apply to those who author separate concurrences, but those in the minority face a similar dilemma: For Justices who join a dissent but similarly author their own, how do we prescribe association for those they join when it is clear they are similarly extrapolating in another opinion?

To account for such a dynamic, we estimated a collection of four ideal points in spaces conditioned on varying weighting schemes (Table A1). While opinion authors were always given full associative weight, each subsequent weighting scheme diminished the associated weight for those who simply joined the Majority, or otherwise authored a separate (Concurring or Dissent) opinion. The emphasis of this exercise being to test the volatility of our estimates given divergent schemes. Yet, as we demonstrate in Figure 1 and reproduce below in Table A2, there is effectively no variance in estimates across schemes. The only exception, of course, is found with the most restrictive weighting scheme (*Sole Weight*), where only opinion authors are given any associated weight for their contributions. However, as we express in the manuscript, the capacity for Justices to author separate opinions if they truly disagree with a Majority (Dissenting) Opinion precludes that they at least substantively agree with the findings and inferences of the principal opinions. As such, at least some relative weight from those opinions should be attributed to those who joined – and instigating variance to how those non-zero weights are represented appear to have little-to-no effect in spurring variance in the resulting estimates.

Table A1: Associative Weighting Schemes by Opinion Authorship

Scheme (Responsibility)	Condition	Majority	Concurrence	(Other) Concurrence	Dissent
Full	JM (Exclusively)	1.0			
	JM AC	1.0	1.0		
	JM ASC	1.0		1.0	
	JM JSC	1.0		1.0	
	AD (Exclusively)				1.0
High	JM (Exclusively)	0.9			
	JM AC	0.9	1.0		
	JM ASC	0.75		1.0	
	JM JC	0.75	0.9		
	JM JSC	0.75		0.75	
	AD (Exclusively)				1.0
Low	JM (Exclusively)	0.75			
	JM AC	0.75	1.0		
	JM ASC	0.5		1.0	
	JM JC	0.5	0.75		
	JM JSC	0.5		0.75	
	AD (Exclusively)				1.0
Sole	JM (Exclusively)				
	JM AC		1.0		
	JM ASC			1.0	
	JM JC				
	JM JSC				
	AD (Exclusively)				1.0

Table A2: Static Wordshoal Summary Estimates by Justice (2005-2022 Terms)

Justice	Full	High	Low	Sole
Alito	3.08	3.08	3.08	2.65
Barrett	1.74	1.74	1.74	1.47
Breyer	-2.81	-2.81	-2.81	-2.68
Ginsburg	-3.42	-3.42	-3.42	-3.12
Gorsuch	1.55	1.55	1.55	1.43
Kagan	-2.84	-2.84	-2.84	-2.67
Kavanaugh	1.91	1.91	1.91	1.51
Kennedy	0.97	0.97	0.97	0.61
O'Connor	0.34	0.34	0.34	0.36
Rehnquist	0.89	0.89	0.89	0.83
Roberts	1.82	1.82	1.82	1.4
Scalia	2.89	2.89	2.89	2.76
Sotomayor	-3.39	-3.39	-3.39	-3.13
Souter	-2.95	-2.95	-2.95	-2.65
Stevens	-3.18	-3.18	-3.18	-2.78
Thomas	3.07	3.07	3.07	2.92

Inferences re: Case Importance and Effects on Justice Locations in Ideal Point Space

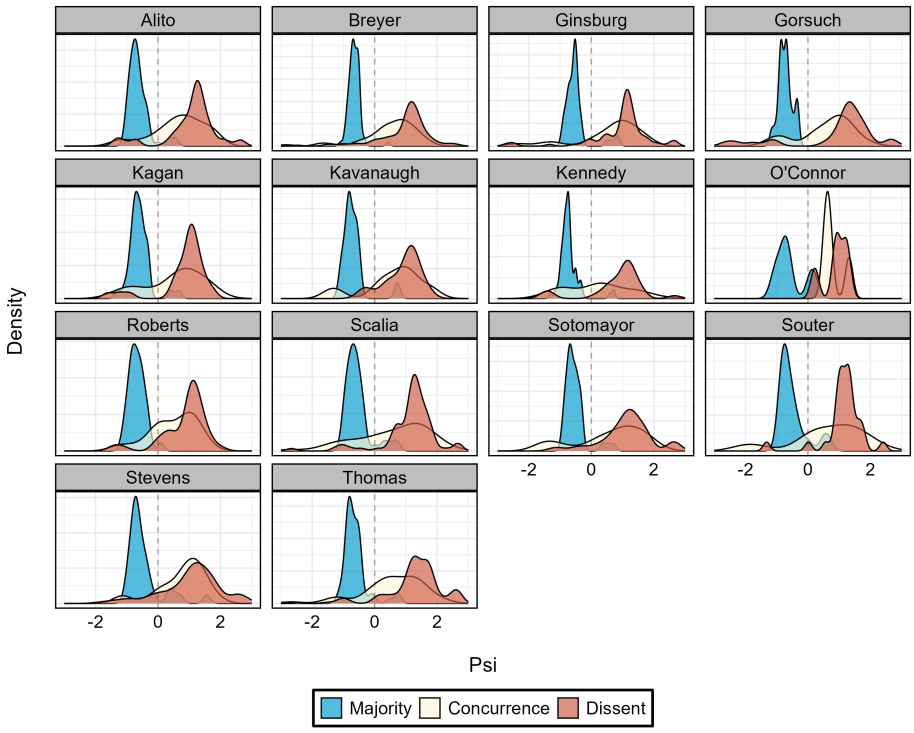
Below we provide summary analyses related to the underlying parameters used to influence each Justice's θ_i. In particular, we produce a sample of the most prominent cases organized by their associated slope (β_j) in Table A3. A similar figure is produced in Table 2 of Lauderdale and Herzog (2016, 12). Alternatively, we produce Justice-level distributions of ψ_{ij} for their authored opinions in Figure 3.

Table A3: Static Wordshoal Summary Outputs by High Beta Estimate

Term	Docket	Case Name	*Abs.* β_j
2005	03-674	Jama v. Immigration and Customs Enforcement	0.351
2005	03-9627	Pace v. DiGuglielmo	0.351
2006	04-1360	Hudson v. Michigan	0.334
2005	04-514	Bell v. Thompson	0.351
2005	04-5286	Dodd v. United States	0.352
2006	04-980	Brown v. Sanders	0.336
2009	07-463	Summers v. Earth Island Institute	0.339
2018	16-1067	Murphy v. Smith	0.335
2018	16-1348	Currier v. Virginia	0.344
2018	16-1362	Encino Motorcars, LLC v. Navarro	0.335
2018	16-1454	Ohio v. American Express Co.	0.335
2018	16-1466	Janus v. AFSCME	0.332
2017	16-373	California Public Employees' Retirement System v. ANZ Securities, Inc.	0.335
2017	16-6219	Davila v. Davis	0.335
2018	16-969	SAS Institute Inc. v. Iancu	0.332
2019	17-1299	Franchise Tax Board of California v. Hyatt	0.332
2019	17-1702	Manhattan Community Access Corp. v. Halleck	0.332
2018	17-530	Wisconsin Central Ltd. v. United States	0.335
2020	18-1109	McKinney v. Arizona	0.332
2020	18-1323	June Medical Services L.L.C. v. Russo	0.34
2019	18-422	Rucho v. Common Cause	0.332
2020	18-725	Barton v. Barr	0.332
2021	19-1257	Brnovich v. Democratic National Committee	0.362
2021	19-438	Pereida v. Wilkinson	0.338
2020	19-7	Seila Law LLC v. Consumer Financial Protection Bureau	0.342
2022	20-1009	Shinn v. Martinez Ramirez	0.355
2021	20-107	Cedar Point Nursery v. Hassid	0.355
2022	20-1088	Carson v. Makin	0.341
2022	20-322	Garland v. Gonzalez	0.364
2022	20-443	United States v. Tsarnaev	0.358
2022	20-826	Brown v. Davenport	0.355
2022	21-12	Federal Election Commission v. Cruz	0.355
2022	21-147	Egbert v. Boule	0.341
2022	21-499	Vega v. Tekoh	0.355

Note: β_j represents the case-specific slope. As Lauderdale and Herzog (2016) note, "this allows the model to select out those debate-specific dimensions that reflect a common dimension (large estimated values of β_j), while down-weighting the contribution of debates where the word usage variation across individuals seems to be idiosyncratic ($\beta_j = 0$)" (p. 5). The cases referenced in this table represent the 95th percentile of absolute values for observed β_j.

Figure A3: Distribution of Opinion-Level Psis by Justice (Static `Wordshoal`)



Note: Figure illustrates the distribution of opinion-level ψ_{ij}, which represents the justice-specific position in a particular case.

Special Note: Modeling Temporal Dependence

Overview

A natural progression of this methodology is the application of dynamic variance in θ_i, which Lauderdale and Herzog (2016) argue can "be achieved from a closely related model that does not change the lower-level model for the texts" (p. 19). They continue by articulating how the application of an existing dynamic strategy similar to Poole and Rosenthal (1997) or Martin and Quinn (2002) would be satisfactory. In the following, we provide an initial attempt at such a strategy, including an updated specification that allows for temporal variance in θ_i as well as updates to figures and tables previously introduced using the original static methodology.

Updated Specification

$$\theta_{i,t} \sim \begin{Bmatrix} N(\theta_{i,t-1}, 1) \text{ If Justice}_i \text{ was present in term}_{t-1} \\ N(0, 1) \text{ If Justice}_i \text{ is newly introduced} \end{Bmatrix}$$

The principal update to the model specification introduced in this manuscript is the amendment to facilitate temporal variance in θ_i. Specifically, for each term, veteran Justices who were present in $\text{term}_{(t-1)}$ will begin the subsequent $\text{term}_{(t)}$ with $\theta_{i,t-1}$. Conversely, newly introduced Justices will assume a standard normal $\theta_{i,t}$. This dynamic structure uses priors from the previous term's θ_i estimates to inform the estimation of the current term's ideal points, capturing the temporal drift of Justice positions. After each iteration, the model refines these estimates until the change in the log-posterior is below a predefined tolerance threshold. Apart from stratifying the observable data to only those observed in term_t, other elements of the model specification remain unchanged.

Updated Tables and Figures Given Dynamic Specification

We replicate the Tables and Figures introduced in the previous sections using the initial attempts at modeling dynamic variance. We further include an additional table (Table A5) to summarize the average ideal points across Martin-Quinn and both `Wordshoal` specifications, as well as two additional figures (Figures A4 and A6) to illustrate Justice-level variance in $\theta_{i,t}$ and the comparative difference given a static or dynamic specification.

Table A4: Dynamic Wordshoal Summary Outputs by High Beta Estimate

Term	Docket	Case Name	*Abs.* β_j
2018	16-285	Epic Systems Corp. v. Lewis	0.564
2009	08-441	Gross v. FBL Financial Services, Inc.	0.565
2010	08-1198	Stolt-Nielsen S. A. v. AnimalFeeds Int'l Corp.	0.567
2010	08-810	Conkright v. Frommert	0.567
2007	06-340	National Assn. of Home Builders v. Defenders of Wildlife	0.568
2007	06-5247	Fry v. Pliler	0.568
2018	16-969	SAS Institute Inc. v. Iancu	0.568
2019	17-1299	Franchise Tax Board of California v. Hyatt	0.57
2019	17-1702	Manhattan Community Access Corp. v. Halleck	0.57
2019	18-422	Rucho v. Common Cause	0.57
2015	14-556	Obergefell v. Hodges	0.572
2010	08-970	Perdue v. Kenny A.	0.574
2020	17-834	Kansas v. Garcia	0.574
2011	09-571	Connick v. Thompson	0.576
2011	09-525	Janus Capital Group, Inc. v. First Derivative Traders	0.577
2011	10-238	Arizona Free Enterprise Club's Freedom Club PAC, et al. v. Bennett, et al; McComish, et al. v. Bennett, et al.	0.577
2011	10-277	Wal-Mart Stores, Inc. v. Dukes	0.577
2011	09-987	Arizona Christian School Tuition Organization v. Winn	0.579
2005	04-5286	Dodd v. United States	0.58
2020	17-1712	Thole v. U. S. Bank N. A.	0.583
2006	04-980	Brown v. Sanders	0.585
2020	18-1109	McKinney v. Arizona	0.588
2020	18-725	Barton v. Barr	0.588
2021	20-107	Cedar Point Nursery v. Hassid	0.589
2010	08-1470	Berghuis v. Thompkins	0.594
2010	08-861	Free Enterprise Fund v. Public Company Accounting Oversight Bd.	0.594
2010	09-497	Rent-A-Center, West, Inc. v. Jackson	0.594
2020	17-1678	Hernandez v. Mesa	0.594
2020	18-1323	June Medical Services L.L.C. v. Russo	0.608
2006	05-493	Ayers v. Belmontes	0.612
2006	04-473	Garcetti v. Ceballos	0.613
2021	19-1257	Brnovich v. Democratic National Committee	0.615
2006	04-1360	Hudson v. Michigan	0.621
2020	19-7	Seila Law LLC v. Consumer Financial Protection Bureau	0.621

Note: β represents the case-specific slope. As Lauderdale and Herzog (2016) note, "this allows the model to select out those debate-specific dimensions that reflect a common dimension (large estimated values of β_j), while down-weighting the contribution of debates where the word usage variation across individuals seems to be idiosyncratic ($\beta_j = 0$)" (p. 5). The cases referenced in this table represent the 95th percentile of absolute values for observed β_j.

Table A5: Comparison of Static and Dynamic Wordshoal Estimates Versus Martin-Quinn by Justice

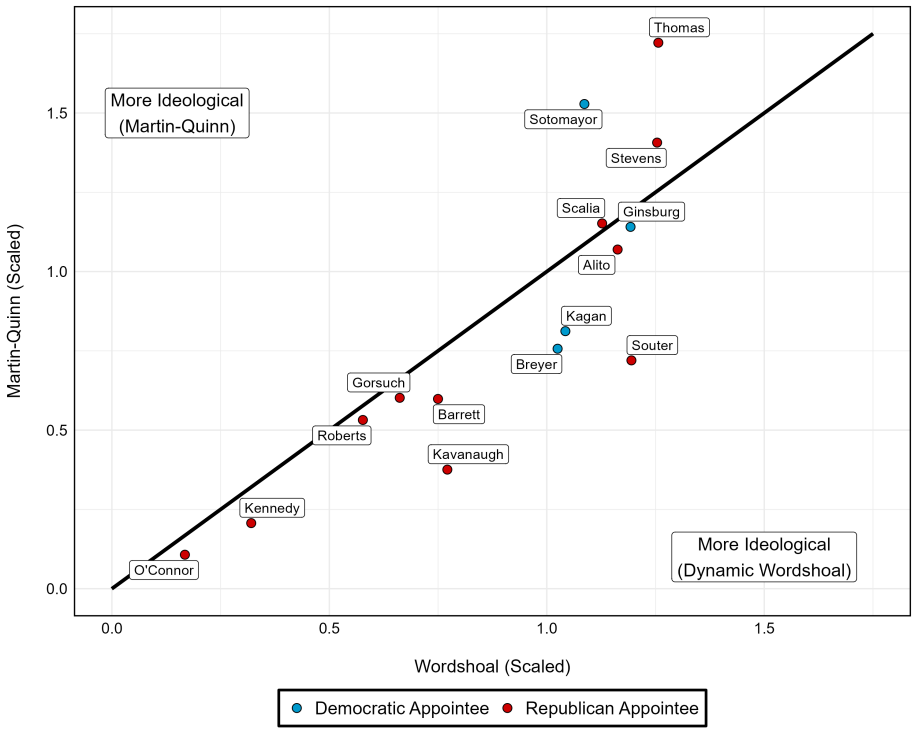
Author	Martin-Quinn	Dynamic Wordshoal	Static Wordshoal
Stevens	-1.81	-1.648	-3.178
Souter	-0.774	-1.571	-2.946
Ginsburg	-1.759	-1.568	-3.422
Sotomayor	-3.127	-1.43	-3.394
Kagan	-1.713	-1.373	-2.837
Breyer	-1.317	-1.35	-2.808
O'Connor	1.016	0.204	0.341
Kennedy	0.678	0.403	0.973
Roberts	1.062	0.737	1.822
Gorsuch	1.027	0.848	1.554
Barrett	0.949	0.962	1.737
Kavanaugh	0.559	0.99	1.915
Rehnquist	2.959	1.07	0.886
Scalia	2.503	1.454	2.887
Alito	1.95	1.5	3.075
Thomas	3.347	1.622	3.07

Note: Values for Martin-Quinn derived using each Justice's respective means from `post_mn` for observations during 2005 to 2021 terms, while Dynamic Wordshoal values represent the same methodology by retrieving the each Justice's average across $\theta_{i,t}$. Both Wordshoal estimates represent values retrieved using the `High Weight` scheme.

Figure A4: Term-Level Variance in Dynamic Ideal Points by Justice

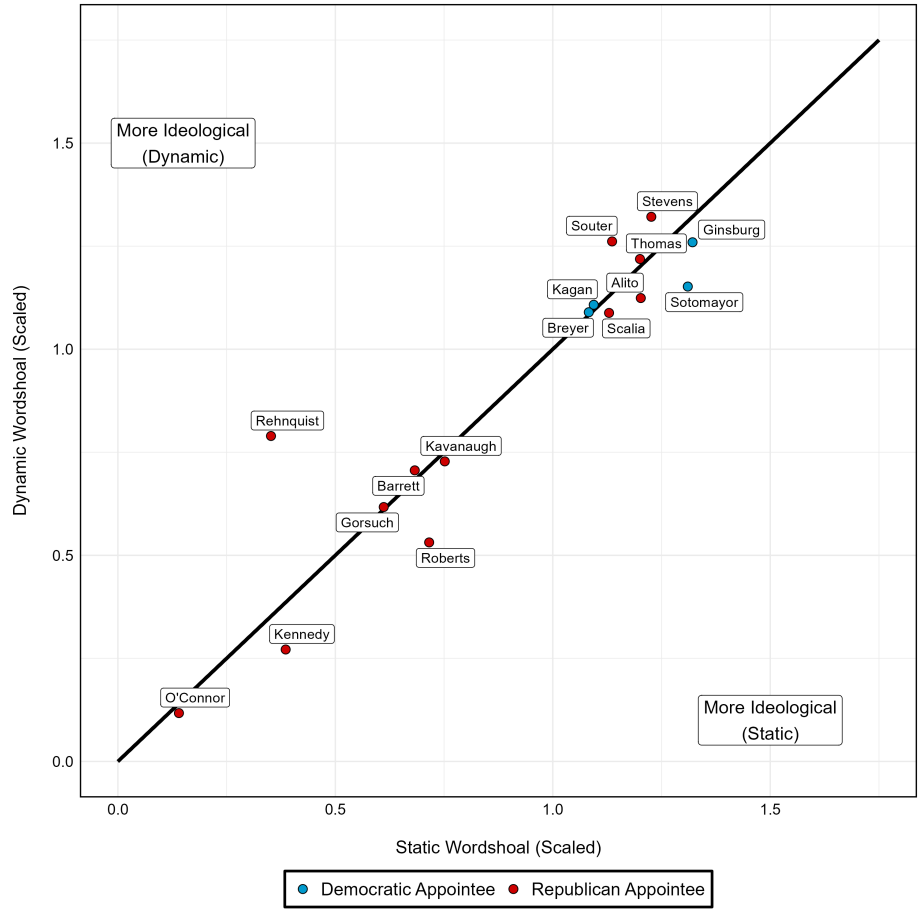
Note: Points represent the estimated value of θ_t for Justice$_i$, while bands represent 95% confidence intervals. With rare exception, both the point estimates and associated bands align with the individual Justice's expected ideological leanings. Circumstances where bands unexpectedly and appreciably surpass zero – e.g., Justices Sotomayor (2009, 2016), Roberts (2005), and Scalia (2016) – are likely (some combination of) freshman effects or, in the case of Scalia, unforeseen death significantly reducing the scope of available data in term$_t$. Otherwise, point estimates and associated margins of error align with expected ideological leanings and placement.

Figure A5: Comparison of Dynamic `Wordshoal` versus Martin-Quinn



Note: Both axes represent the absolute values of each Justice's estimated ideal point using Martin-Quinn and Dynamic `Wordshoal`, where both scales are standardized using z-score normalization. Values for Martin-Quinn (Dynamic `Wordshoal`) measured using the average of `post_mn` (θ) across the observation period. Points nearest to the diagonal segment indicate greater correlation between the relative placement of Justice$_i$ across both methodologies. Alternatively, values above (below) the diagonal indicate greater relative ideological placement in Martin-Quinn (Dynamic `Wordshoal`). Correlation = 96.4%.

Figure A6: Comparison of Static v. Dynamic `Wordshoal` Means (High Weight)



Note: Figure represents the comparative ideal points of each Justice using a static or dynamic `Wordshoal` specification. The scope of variance between static and dynamic point estimates must be contextualized given the relative difference in their associated scales – i.e., the margins of the unbounded scale with the dynamic specification is discernibly larger (-3.42, 3.07) than the static specification (-1.6, 1.6). To account for this, we conduct a similar strategy as Figure A5 to take the absolute values of each Justice's estimate given the normalization of the static and dynamic scales. Estimates nearest the diagonal line indicate lesser variance between the static and dynamic specifications, while those above (below) the diagonal indicate greater relative ideological placement from the dynamic (static) specification. Correlation = 92.8%.