Motivated Reasoning in the Field: Polarization of Prose, Precedent, and Policy in U.S. Circuit Courts, 1891-2013

Wei Lu* Daniel L. Chen†

February 20th, 2024

Abstract

This study explores politically motivated reasoning among U.S. Circuit Court judges over the past 120 years, examining their writing style and use of previous case citations in judicial opinions. Employing natural language processing and supervised machine learning, we scrutinize how judges' language choices and legal citations reflect partisan slant. Our findings reveal a consistent, albeit modest, polarization in citation practices. More notably, there's a significant increase in polarization within the textual content of opinions, indicating a stronger presence of motivated reasoning in their prose. We also examine the impact of heightened scrutiny on judicial reasoning. On divided panels, judges show a decrease in polarization in both writing and citation practices. Furthermore, our study explores polarization dynamics among judges who are potential candidates for Supreme Court promotion. We observe that judges on the shortlist for Supreme Court vacancies demonstrate greater polarization in their selection of precedents.

*Baruch College, City University of New York, New York, NY 10010. Email: wei.lu@baruch.cuny.edu

†Toulouse School of Economics, Toulouse, Haute-Garonne 31000, France. Email: dlchen@nber.org. Work on this project was conducted while Daniel Chen received financial support from the European Research Council (Grant No. 614708) and Swiss National Science Foundation (Grant Nos. 100018-152678 and 106014-150820).

"'I pay very little attention to legal rules, statutes, constitutional provisions ... The first thing you do is ask yourself — forget about the law — what is a sensible resolution of this dispute? ... See if a recent Supreme Court precedent or some other legal obstacle stood in the way of ruling in favor of that sensible resolution. ... When you have a Supreme Court case or something similar, they're often extremely easy to get around." *(An Exit Interview with Richard Posner*, The New York Times, *Sep. 11, 2017).*

1 Introduction

Can we quantitatively identify when judges have an easier time recruiting evidence supporting what they want to be true than the evidence supporting what they want to be false [23]? This tendency is called motivated reasoning, and several recent models and experiments on motivated reasoning are summarized in [4]. Motivated reasoning is a subject of much policy debate. Does it affect real-world decision-makers? Moreover, what affects motivated reasoning? These are a few questions this paper seeks to address. Motivated reasoning is the well-documented tendency where individuals actively seek out confirmatory information. The mechanism is said to be implicit emotion regulation – the brain converges on judgments that maximize positive affective states associated with the attainment of motives. In the lab, motivation is typically inferred by the degree to which goal-related concepts are accessible in memory: The greater the motivation, the more likely individuals are to remember, notice, or recognize concepts, objects, or persons related to that goal [53].[1] Recently, motivated reasoning has been used to explain polarization. For example, when responding to moral dilemmas, subjects come to snap judgments, and ex-post generate a justification [29]; or when interpreting data on climate change, subjects update their beliefs following their political party, and this was greatest among those scoring highest in cognitive reflection [34].

In prior studies of motivating reasoning in law, law student subjects are exogenously provided precedents (reasons) [11, 12]) to address the issue that differences in reasoning might be due to memory or knowledge. The experiments fix the set of precedents to choose from. Nevertheless, whether these studies on law students are externally valid to judges or other policymakers is still an open question [50]. In another study using a series of experiments on statutory interpretation with the cultural identity of parties involved as the main manipulation, [36] shows that judges and lawyers do not exhibit cultural biases, unlike law students and

[1] The classic studies only measure the final decision, rather than reasoning [55, 25].

the general public. In our case, we show that when judges are making high-stakes decisions in Court, they could exhibit polarization in writing. Building on the framework by [35] on politically motivated reasoning, recent work by [52] provides a new design to assess politically motivated reasoning based on trust in news.

This paper explores motivated reasoning among real-world judges. Our paper sits at the intersection of constitutional law, politics, and judicial legitimacy, examining the hypothesis that judicial decision-making is politically motivated. Grounded in the debate between jurisprudential decisionism and the separation of political interests from legal procedures [45, 46, 47, 6, 31, 38, 22], we employ a quantitative approach to assess the extent to which recent shifts in the U.S. judiciary reflect broader political dynamics. Our analysis contributes to the understanding of how constitutional-legal proceduralism, often seen as a tool for upholding democratic values, may be exploited by political-economic elites to shape legal outcomes. This paper provides empirical insights into the political nature of judicial decision-making, offering a novel perspective in the context of ongoing debates on the judiciary's role in liberal democracies.

Specifically, we aim to analyze motivated reasoning using as a natural laboratory the U.S. federal courts – a high-stakes common-law space. Circuit judges can introduce new legal theories,[2] shift standards or thresholds,[3] and rule on the constitutionality of federal and state statutes. Circuit judges provide the final decision on tens of thousands of cases per year, compared to just a hundred cases or so on the U.S. Supreme Court. Therefore Circuit decisions are the majority of what creates the law in this common-law space (and most of what law students are reading). If there is motivated reasoning among these judges, that could have substantial legal and policy impacts. Existing research on how Democrat and Republican Circuit Court judges behave differently is extensive (see [10] for a comprehensive review), but almost all of them focus on the decisions made by judges. For example, a recent work examines how the ideology of Circuit Court judges can affect case outcomes in a wide range of Circuit cases [17]. Complementing this literature, we look at polarization in reasoning.

Circuit courts have a handful of critical features that make them a desirable context for this empirical work. First, there is random assignment of cases to judges (who sit in panels, without juries),[4] meaning that

[2] E.g., contract duty posits a general obligation to keep promises, vs. a party should be allowed to breach a contract and pay damages if it's more economically efficient than performing, also known as efficient breach theory, articulated by Richard Posner in a 1985 opinion.

[3] (E.g., shift from reasonable person standard to reasonable woman standard for what constitutes sexual harassment, or waive the need to prove emotional harm in court to a jury.)

[4] This randomness has been used in a growing set of economics papers [37, 20, 3, 19, 41, 2].

judges rule on similar legal issues on average. Second, there is an adversarial system where the litigants are responsible for bringing all the reasons (arguments and precedents) to a judge's attention. This means that differences in reasoning are not due to differences in knowledge.[5] In addition, the briefs are filed prior to judicial assignment, so strategic information provision according to judge type is not feasible.

We have data on 300,000 Circuit Court opinions (almost a million judge votes) for the period 1891-2013. Circuit Court judges are appointed by the U.S. president (Democrat or Republican) with life tenure. The measures of judicial reasoning are constructed from texts of the opinions and legal citations of other Circuit Court opinions. These outcome measures are linked to judicial biographical features, particularly the judge's political affiliation.

The measure of motivated reasoning is defined as the accuracy with which we can predict judges' political affiliations based on the reasoning of their decisions, consistent with previous literature [27, 8]. Earlier efforts to measure polarization in the text include [33], whose non-penalized measure might overestimate polarization in early years. Outside of Congress, [32] use the text of academic articles to predict political donations by economists. An essential difference between our context and previous papers is that members of Congress (and economists) have discretion over the topics they address, while judges are assigned topics randomly. Moreover, political donations are made ex-post, while the political party of appointment is ex-ante. We seek to predict a predetermined measure of ideology prior to reasoning on the case. A parallel literature has looked at the polarization of citizens rather than policymakers. [8] show that partisan affiliations are most associated with social attitudes (rather than consumption and time use). We will find that partisan affiliation is more associated with the text of judgments rather than precedents cited.

We first predict the political party using the text of judicial decisions – i.e., judicial prose. We represent judicial prose as low-dimensional vectors and use those as predictors in a pre-trained large language model for fine-tuning. We find that average prose polarization for judges has remained high and increasing as time goes by. A new contribution is to look at the polarization of precedent, as these are the legal reasons cited to justify a decision. We use a network of citations to previous Circuit Court decisions to predict partisan affiliation. Unlike the case of prose, we find low yet steady levels of precedent polarization over time, indicating that judges tend to express ideological differences through writing instead of choices of

[5] That is, we can distinguish our results from mechanical failures of inference due to bounded rationality or limited attention; in this adversarial setting, briefs bring forward all the citable reasons.

precedents in our context. These results complement previous work with smaller samples by [16] showing that circuit judges tend to cite judges from the same party, and that of [43] showing that circuit judges tend to cite Supreme Court cases authored by judges from the same party.

Finally, we look at how the polarization in prose and precedent changes when judges are under more scrutiny. Specifically, we examine two such scenarios: The first is whether a judge sits on a divided panel of judges from both parties. The second scenario is whether the opinions were filed when the midterm or presidential elections were close. Some research suggests that the threat of actual "whistleblowing" tempers the decisions issued when under scrutiny [18, 5], as reflected by the increase in dissents. Consistent with this interpretation, the polarization in text and citations reduces when under scrutiny. Moreover, we examine how polarization varies when judges have promotion incentives. We find that judges exhibit more polarization in precedent when they are a contender for a Supreme Court vacancy.

2 Measuring motivated reasoning in judicial context

In this paper, we define "motivated reasoning" in the judicial context as the ability to predict a judge's political affiliations based on the way they write opinions and cite precedents, in line with existing literature that used the predictability of texts and other behaviors as measures of polarization and cultural distance [27, 8]. [35] provides a theoretical framework for this measure. They propose that politically motivated reasoning in our setting can be defined as the distortion of how political dispositions (political party) affect the way a judge interprets new evidence (cases) to update his prior beliefs to form a posterior (reflected by texts and citations in opinions). Three features of the institutional setting ensure that the predictability measure we have is not related to varying priors or new evidence, and the predictability of political parties can be attributed to the distortion caused by political dispositions.

Firstly, the style and content of a judge's opinions, along with their chosen citations, offer insights into their formal reasoning processes. Prior literature has shown that judicial fact discretion, how judges believe and interpret the facts presented, can be related to the identity of judges [26]. Since how judges recruit precedents and prose in their opinions constitutes the judicial opinion, we would be observing any slant in the formal reasoning process made explicit in their opinion.

Secondly, the cases are assigned quasi-randomly to judges [6]. The as-if random assignment of cases means that every judge will on average see a similar variety of cases. Notably, cases that should cite certain precedents or refer to certain topics should not systematically differ across judges due to this random assignment process.

Thirdly, absent politically motivated reasoning, the reasoning in the cases should be non-partisan because judges are asked that they "*not be swayed by partisan interests*"[7]. If judges follow this edict, then a reasonable guess on party affiliation based on the opinion is 0.5 – i.e., the probabilities of the writer being a member of Republican Party or a Democratic Party should be the same. However, this might diverge if judges are systematically interpreting the facts and the law in a different manner that is reflective of their political party.

If the expressed reasoning of judges is motivated by partisan views, then the choice of language and citations might be informative of the political party. Motivated reasoning can alter the way judges interpret and evaluate information from briefs and precedents, which would lead to differences in their expressed arguments. If a judge's political affiliations can be predicted based on their writing and citations to legal precedent, it would suggest that their reasoning can be influenced by their political leanings.

2.1 Data

Our dataset includes a collection of 318,474 opinions published by U.S. Circuit Courts from 1891 to 2013 based on [5]. We limit our analysis to opinions written by one judge, excluding opinions labeled *per curiam*, which are authored by the whole panel without designating a specific author, and opinions drafted by multiple judges. Among all opinions, 279,167 are majority opinions and 26,441 are dissent opinions. For each opinion, we observe the full text, legal precedents, as well as all votes cast by judges on the panels for each case. We focus on precedents of previous Circuit Court opinions and the partisan policy is constructed using data on judge dissenting votes.

To study the heterogeneity of motivated reasoning across judicial characteristics, we link the opinions to the United States Courts of Appeals Databases and Attributes of the United States Federal Court Judges

[6] Some research suggests that a few of the courts do not assign cases to judges completely randomly, but the reasons for non-random assignment include workload, scheduling, and professional development[39]. There is no direct evidence that political party is related to the assignment of cases.

[7] http://www.uscourts.gov/judges-judgeships/code-conduct-united-states-judges

from [49], and use variables such as political affiliations of judges, Circuit Court, the political composition of panels, year, quarter to presidential elections for subsequent analysis.

2.2 Classification

Since the predictability of judicial reasoning serves as our measure of motivated reasoning, we conceptualize this measurement problem as equivalent to a binary classification problem in machine learning using high-dimensional text and citation data as inputs and the political affiliations of judges as the outcome variable. In our two prediction tasks, we aim to predict the political affiliation of circuit court judges, specifically whether they belong to the Democratic or Republican party. The affiliation is represented by a binary variable: "1" for Democratic judges and "0" for Republican judges. We use opinion texts and citation embeddings as our predictors. To determine a judge's average stance over a year, we average these embeddings. Our goal is to predict the likelihood that a judge's political affiliation matches their true party affiliation.

We use sample splitting to avoid overfitting the models. In Text Classification, we randomly chose 10% of our dataset, which is about 31,000 opinions as our sample dataset due to computational constraints. Afterward, 30% of the 10% sample is used as the test set and the remaining 70% as the training set. For Citation Classification, we use the full dataset as our sample dataset, with 30% as the test set. The test set is only used after training to assess the performance of the algorithms. The best model (an ensemble of models with best-performing parameters) in each task will be applied to the full sample to generate predictions for all opinions. The remaining strategies and training details are in the SI Appendix.

2.3 Training Algorithms for Texts

We leverage recent advances in natural language processing to classify political affiliations of Circuit Court judges by fine-tuning pre-trained large language models using opinion texts directly as inputs. In recent years, transformer-based pre-trained large language models have been proven to have satisfactory performance on a variety of NLP tasks. Even with a small sample for fine-tuning, pre-trained models can further significantly improve the performance [21]. In this paper, we will use an ensemble of several commonly-used pre-trained transformer models to ensure the robustness of our results by averaging the predictions across models. Before fine-tuning, for each opinion, we use the Microsoft Presidio tool

[40] to detect and replace all names (including judges and any person's names) and locations to the word "PERSON" and "LOCATION". Doing so will prevent the pre-trained models from relying on the name and location information for classification, a problem known as data leakage. After that, the first 512 words of each opinion, which is the maximum length allowed by models, will be used as inputs for pre-trained models to learn.

2.4 Training Algorithms for Citations

For precedents, we combine network representation models with ensemble supervised learning for classification. In the first step, we construct and transform the citation network into dense low-dimensional vectors. Specifically, we create a weighted directed graph of 310,282 nodes, and transform the citation network into citation embeddings of 300 dimensions using the node2vec algorithm by [28]. The node2vec algorithm rests on the idea that a word is represented by its "neighboring words" [24] in natural language processing. It adopts a random walk approach across the network to generate sequences of citations, and by maximizing the probability of neighboring citations, we can have latent vectors that "*maximize the likelihood of preserving network neighborhoods*" of citations. Then, we use an ensemble of commonly used supervised machine learning algorithms as our prediction algorithm, consistent with similar strategies used in previous literature [7].

2.5 Permutation Inference

To ensure that the algorithms are indeed learning from the training set, we generate a random permutation of political parties with an equal probability of two parties for all authors and use this list as the dependent variable for training. If the algorithms are learning correctly from the data, using random series as the dependent variable should result in random predictions. A similar strategy is also implemented by [27], who randomly shuffle the share of Republicans/Democrats in Congress during the year in which a particular congressman. In practice, we train another set of models with the same parameters on the random series for both Text and Citation Classification.

3 Results

3.1 Polarization in Prose and Precedents across time

We begin with an overview of how polarization in prose and precedent evolves over time. Figure 1 illustrates the trend in average polarization levels within opinion texts. The magnitude of average polarization in writing consistently exceeds 0.5 and surged towards 0.8 after 1950. These trends imply an increasing propensity for motivated reasoning among judges when drafting their opinions. Although [27] identified an increase in polarization in congressional speeches after 2000, it is significantly lower compared to the polarization observed in judicial opinions. To put this effect size in perspective, in [27], the polarization varies between 0.5 and 0.515. Notably, the placebo test involving random shuffling series aligns closely with the 0.5 benchmark, validating our models' ability to produce random predictions when analyzing data with randomly permuted party affiliations of judges. The fact that polarization was more present a century ago is consistent with other analyses of partisan behavior in the judiciary [14]. In the analyses that follow, we demonstrate how scrutiny influences this measure of partisanship, taking into account the specific time period in question.

Furthermore, we investigate the presence of motivated reasoning in the selection of precedents by judges. Figure 2 shows that, over the past 120 years, Circuit Court judges have consistently demonstrated a lower level of motivated reasoning in their choice of legal precedents, especially when compared to the more pronounced motivated reasoning observed in texts. This suggests that, unlike the choice of language, judges are more constrained in their choice of precedents. However, the level of polarization is still distinguishable from the placebo random series of 0.5 and higher than the polarization in congressional speech found in [27].

The larger polarization in text may be attributable to the rhetorical style of judicial overstating, a product of cognitive processes and a means to enhance the judiciary's legitimacy [48]. Circuit Court judges may be constrained by precedents as they face the reversal from higher courts if deviating too much from precedent [30].

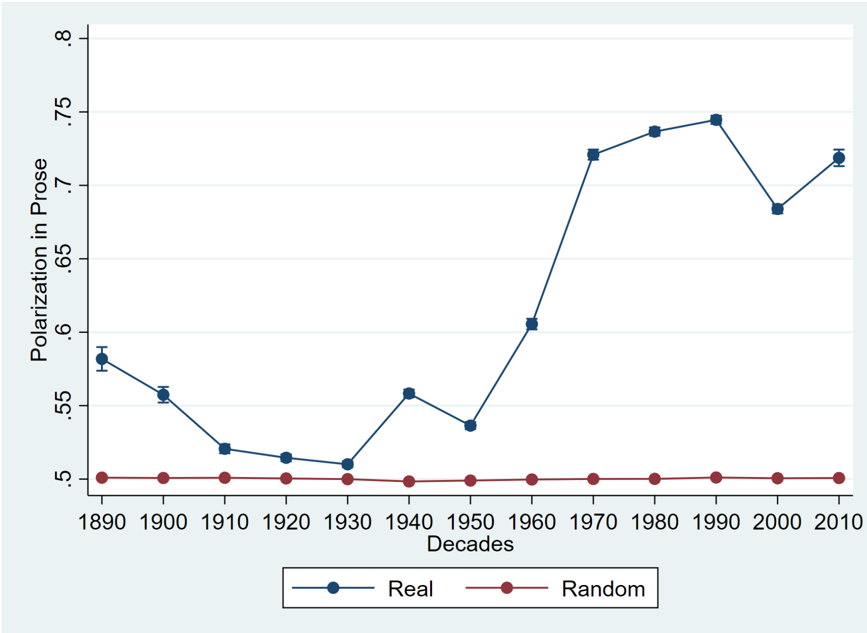

Figure 1: Polarization in Prose in U.S. Circuit Courts

Notes: Polarization measures over time in U.S. Circuit Courts, 1891-2013 for writing. The blue line gives the average polarization in the true dataset. The red line gives the average polarization in the shuffled dataset (random party affiliations). Error bars indicate the 99% confidence interval.

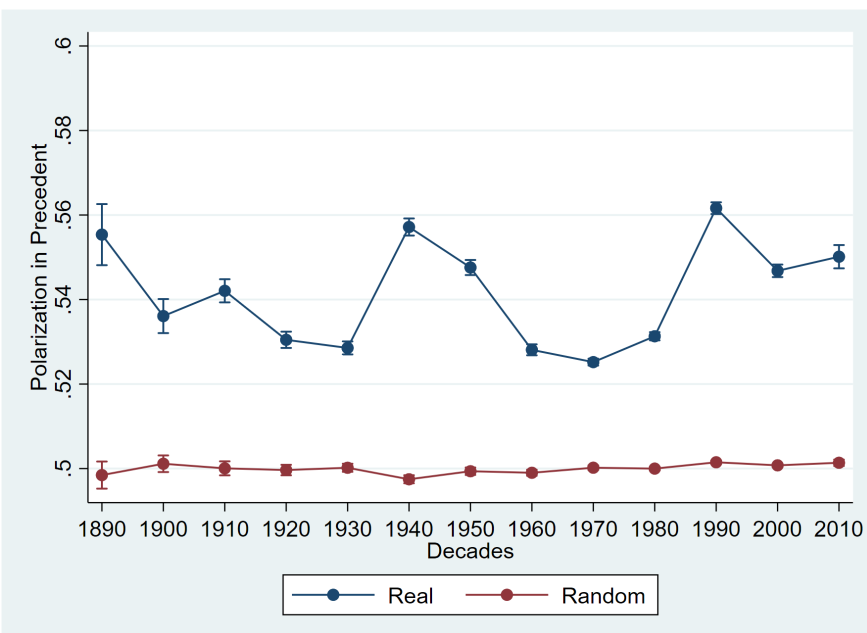

Figure 2: Polarization in Precedents in U.S. Circuit Courts

Notes: Polarization measures over time in U.S. Circuit Courts, 1891-2013 for citations. The blue line gives the average polarization in the true dataset. The red line gives the average polarization in the shuffled dataset (random party affiliations). Error bars indicate the 99% confidence interval.

3.2 Polarization in Reasoning or Decision when judges are under greater scrutiny

In this section, we examine the dynamics of polarization in prose and precedents, along with the polarization in policy, under scenarios of increased judicial scrutiny. We hypothesize that under such scrutiny, judges do not want to appear as politically motivated in their reasoning, even if they come to different conclusions in their decisions.

3.2.1 Divided panels

Our investigation begins with how judicial reasoning and decisions are influenced when judges serve on a three-member panel that consists of Republican and Democrat judges. Previous studies, such as those by [18, 17, 2], indicate that political divisions within a panel creates an opportunity for whistleblowing, through dissenting opinions, to expose disobedient decisionmaking by the majority. In the presence of such a whistleblower, the majority must sometimes capitulate and keep its decision within the confines of doctrine. [51] furthers this idea and suggests that in ideologically divided panels, judges may exhibit ideology-dampening effects. To explore these dynamics, we employ a linear regression model, focusing on polarization in reasoning and the propensity for dissenting votes, with the key independent variable being participation in a politically divided panel. The counterfactual comparison group is participation in a politically homogenous panel. This model controls for Circuit $\times$ Year and legal issues fixed effects.

Tables 1 and 2 show that judges are more likely to cast dissenting votes when they are part of a divided panel or as a minority member within such a panel. Concurrently, our findings indicate a reduced degree of motivated reasoning in their prose and precedent citations, aligning with our hypothesis. Compared to the previous literature on congressional speech [27] where the polarization variation is only 0.015, the dampening effect we observe is large.

3.2.2 Electoral Cycles

We observe a similar pattern in another scenario of scrutiny, namely, the periods leading up to Presidential and midterm elections. With heightened scrutiny, we may expect a decrease in polarization with their reasoning. To investigate this hypothesis, we divided the time into 16 quarters preceding a Presidential election. Our analysis focuses on the electoral cycles and their impact on polarization in judicial reasoning.

Table 1: Polarization in Divided Panels

	(1) Text	(2) Citation	(3) Dissent vote
Divided Panel	-0.032***	-0.038***	0.006***
	(0.005)	(0.004)	(0.001)
Observations	310604	269155	1030343
R^2	0.335	0.125	0.009
Circuit × Year FE	✓	✓	✓
Legal Issue FE	✓	✓	✓

Notes: This table shows how judges on a divided panel would exhibit polarization in prose, precedent, and policy. The unit of observation for Column (1) and (2) is at the opinion level, and Column (3) is at the vote level. Every case has three votes from three judges sitting in a panel and judges are allowed to write concurring or dissent opinions besides the majority opinion for each case. We controlled for Circuit × Year and legal issues fixed effects. Standard errors clustered at judge level in parentheses.* $p < 0.1,^{**}$ $p < 0.05,^{***}$ $p < 0.01$

Table 2: Polarization in Divided Panels

	Text (1)	Citation (2)	Dissent Vote (3)
Minority	-0.020***	-0.030***	0.012***
	(0.006)	(0.004)	(0.001)
Observations	225817	196097	742495
R^2	0.320	0.065	0.012
Circuit × Year FE	✓	✓	✓
Legal Issue FE	✓	✓	✓

Notes: Effect of being a minority judge (D of DRR or R of RDD) on the polarization in texts and citations, and the likelihood to cast a dissenting vote, controlling for Circuit × Year and legal issues fixed effects. The unit of observation for Column (1) and (2) is at the opinion level, and Column (3) is at the vote level. Standard errors clustered at judge level in parentheses. The sample is cases with judges from both political parties. $*p < .1, **p < 0.05, ***p < .01$.

The results of this analysis are detailed in Table 3.

We find that judges reduce polarization in both the texts of opinions and citations preceding midterm elections, indicating a notable shift in judicial behavior during these periods. This trend contrasts with the pattern observed before Presidential elections, where such polarization does not exhibit a significant change. A plausible explanation for this discrepancy can be linked to the heightened partisan political priming associated with Presidential elections. According to [14], this kind of priming is intense during Presidential elections, potentially neutralizing the judges' inclination to reduce polarization under scrutiny.

Furthermore, an intriguing pattern emerges in the context of both midterm and Presidential elections as they draw near: an increase in dissenting votes. This observation, documented in Column 5, aligns with what can be described as a 'whistleblowing effect' similar to that found on politically divided panels. During periods of increased scrutiny, which are common around election times, judges might express dissent more openly, a behavior that is consistent with a whistleblowing response.

To summarize, we investigate two different situations of increased scrutiny, which prior research has suggested can lead to greater dissent and whistleblowing. With this greater scrutiny, we observe a reduction in polarization in prose and citations to precedent. Such a response suggests a complex interplay between the political environment and judicial decision-making.

3.3 Polarization and Promotion Incentives

In this section, we analyze an institutional factor that is likely to influence political polarization in the courts: promotion incentives. We concentrate on the nomination process for Supreme Court of the United States (SCOTUS) justices, where Circuit Court judges are potential candidates for elevation to the highest court by presidential and senate appointment. This scenario raises a question: do judges demonstrate increased partisan polarization in their reasoning and decision-making as a strategy to secure a nomination? Drawing on the findings of [9], who observed that judges on the president's "shortlist" are more likely to write dissent opinions and vote in line with the presidents, our analysis seeks to understand if politically motivated reasoning might change with SCOTUS vacancies using a much larger sample of judges and years. Detailed methodology and data processing information for this analysis are provided in the SI appendix.

In Table 4, we present our results using the same specification as in Table 1. The results indicates no

Table 3: Electoral Cycles in Text and Citation

	(1)	(2)	(3)	(4)	(5)
	Prose		Precedents		Dissent Vote
	All Op	Dis Op	All Op	Dis Op	
Quarter to election=1	0.004	-0.019	-0.002	0.007	0.005**
	(0.005)	(0.013)	(0.003)	(0.006)	(0.002)
Quarter to election=2	0.005	-0.021	0.003	0.001	0.003*
	(0.004)	(0.013)	(0.002)	(0.006)	(0.002)
Quarter to election=3	0.009**	-0.003	0.004**	0.002	0.003*
	(0.004)	(0.011)	(0.002)	(0.006)	(0.002)
Quarter to election=4	0.010**	-0.019	0.001	-0.014**	-0.001
	(0.004)	(0.013)	(0.002)	(0.006)	(0.002)
Quarter to election=5	0.010*	0.004	-0.001	-0.010	0.002
	(0.006)	(0.017)	(0.003)	(0.008)	(0.002)
Quarter to election=6	0.005	-0.006	0.000	-0.011	0.001
	(0.006)	(0.017)	(0.003)	(0.008)	(0.002)
Quarter to election=7	0.006	0.008	-0.001	-0.008	-0.001
	(0.006)	(0.017)	(0.002)	(0.008)	(0.002)
Quarter to election=8	-0.008*	-0.025	-0.002	-0.019**	0.001
	(0.005)	(0.015)	(0.002)	(0.007)	(0.002)
Quarter to election=9	-0.012**	0.001	-0.002	-0.018**	0.005*
	(0.006)	(0.019)	(0.003)	(0.009)	(0.002)
Quarter to election=10	-0.012**	-0.009	-0.001	-0.021**	0.003
	(0.006)	(0.019)	(0.003)	(0.009)	(0.002)
Quarter to election=11	-0.006	-0.010	0.000	-0.016*	0.003
	(0.005)	(0.019)	(0.003)	(0.009)	(0.002)
Quarter to election=12	-0.011***	-0.015	-0.003	-0.010	-0.001
	(0.004)	(0.013)	(0.002)	(0.006)	(0.002)
Quarter to election=13	-0.003	-0.021	0.000	-0.004	0.000
	(0.005)	(0.015)	(0.002)	(0.007)	(0.002)
Quarter to election=14	-0.009*	-0.031**	0.000	-0.000	-0.002
	(0.005)	(0.015)	(0.002)	(0.007)	(0.002)
Quarter to election=15	-0.002	-0.021	0.002	0.003	-0.002
	(0.004)	(0.013)	(0.002)	(0.006)	(0.002)
Observations	190135	17110	178609	13494	606999
R^2	0.243	0.137	0.097	0.086	0.008
Circuit × Year FE	✓	✓	✓	✓	✓
Season FE	✓	✓	✓	✓	✓
Legal Issue FE	✓	✓	✓	✓	✓

Notes: The unit of observation for Column (1) to (4) is at the opinion level, and Column (5) is at the vote level. Standard errors clustered at judge level in parentheses. The base period is 16 quarters to Presidential Elections. The sample is cases published after 1975. * $p < 0.1$,** $p < 0.05$,*** $p < 0.01$

systematic differences between judges on the presidential shortlist and their non-contender counterparts; moreover, a Supreme Court vacancy does not result in significant changes in behavior across the entire judicial spectrum in Circuit Courts. However, during a Supreme Court vacancy, contender judges demonstrate noticeably more polarization in their selection of legal precedents. Nevertheless, we observe no significant extension of this trend to their writing style or voting patterns. From a theoretical standpoint, it is remarkable that contender judges choose to standout to a potential nominating president through their citations to precedent, which we previously documented to be less polarizing in general. Presidents may look to nominate partisan/ideological allies rather than individuals that are politically ambiguous or moderate in their behavior in how they follow precedents.

Table 4: Polarization in SCOTUS Vacancies

	Text	Citation	Dissent Vote
	(1)	(2)	(3)
Vacancy	-0.001	-0.001	0.000
	(0.003)	(0.001)	(0.001)
Contenders	-0.041	-0.005	0.003
	(0.039)	(0.021)	(0.005)
Vacancy × Contenders	0.016	0.017**	-0.001
	(0.019)	(0.008)	(0.006)
Observations	49,711	46,759	153,672
R^2	0.257	0.100	0.008
Circuit × Year FE	✓	✓	✓
Legal Issue FE	✓	✓	✓

Standard errors in parentheses
* $p < 0.1$,** $p < 0.05$,*** $p < 0.01$

Notes: Effect of being a SCOTUS vacancy contender on the polarization in texts and citations, and the likelihood to cast a dissenting vote, controlling for Circuit × Year and legal issues fixed effects. The unit of observation for Column (1) and (2) is at opinion level, and Column (3) at the vote level. Standard errors clustered at judge level in parentheses. Sample is cases with judges from both political parties after 1975. $*p < .1, **p < 0.05, ***p < .01$.

4 Conclusion

Judges are nominally expected to sit above the partisan fray. However, we find they are divisive in their rhetoric and citations to legal precedent. We find that both text and citations display polarization, with text being even more polarized. In addition, judges display less polarization in reasoning when under greater scrutiny, sitting on divided panels, or before elections.

Lifetime-appointed judges assert that their decisions are not influenced by politics. However, their voting trends and the intense partisan struggles during confirmation processes suggest otherwise. Our findings reveal the political nature of judicial reasoning measured in their rhetoric and their citations to precedent. If judges cherry-pick their precedents, this casts a shadow over the fairness of their decisions. A diminished sense of legitimacy can lead to decreased compliance with the law, which can have social and economic implications. Trust has been shown to have impacts, see [1]'s recent paper documenting this link causally. They show that enhanced trust spurs reliance on formal institutions. Reliance on formal institutions can, in turn, propel economic development, investments, and entrepreneurial undertakings. While our paper may not directly quantify these effects, it seeks to underscore their significance.

References

[1] D. Acemoglu, A. Cheema, A. I. Khwaja, and J. A. Robinson. Trust in state and nonstate actors: Evidence from dispute resolution in pakistan. *Journal of Political Economy*, 128(8):3090–3147, 2020. 16

[2] D. Beim, T. S. Clark, and B. E. Lauderdale. Republican-majority appellate panels increase execution rates for capital defendants. *The Journal of Politics*, 83(3):1163–1167, 2021. 3, 11

[3] A. Belloni, D. Chen, V. Chernozhukov, and C. Hansen. Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica*, 80(6):2369–2429, 2012. 3

[4] R. Bénabou and J. Tirole. Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3):141–64, 2016. 2

[5] C. Berdejó and D. L. Chen. Electoral cycles among us courts of appeals judges. *The Journal of Law and Economics*, 60(3):479–496, 2017. 5, 6

[6] R. Berger. *Government by judiciary*. Liberty Fund, 1997. 3

[7] M. Bertrand and E. Kamenica. Coming apart? Cultural Distances in the United States over Time. Working Paper 24771, National Bureau of Economic Research, June 2018. 8

[8] M. Bertrand, F. Kramarz, A. Schoar, and D. Thesmar. The cost of political connections. *Review of Finance*, 22(3):849–876, 2018. 4, 5, 23

[9] R. C. Black and R. J. Owens. Courting the president: how circuit court judges alter their behavior for promotion to the supreme court. *American Journal of Political Science*, 60(1):30–43, 2016. 13, 25

[10] A. Bonica and M. Sen. Estimating judicial ideology. *Journal of Economic Perspectives*, 35(1):97–118, 2021. 3

[11] E. Braman. Reasoning on the threshold: Testing the separability of preferences in legal decision making. *The Journal of Politics*, 68(2):308–321, 2006. 2

[12] E. Braman and T. E. Nelson. Mechanism of motivated reasoning? analogical perception in discrimination disputes. *American Journal of Political Science*, 51(4):940–956, 2007. 2

[13] J. Camacho-collados, K. Rezaee, T. Riahi, A. Ushio, D. Loureiro, D. Antypas, J. Boisson, L. Espinosa Anke, F. Liu, E. Martínez Cámara, et al. TweetNLP: Cutting-edge natural language processing

for social media. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–49, Abu Dhabi, UAE, Dec. 2022. Association for Computational Linguistics. 22

[14] D. L. Chen. Priming ideology: Why presidential elections affect us judges. Technical report, Institute for Advanced Study in Toulouse (IAST), 2017. 9, 13

[15] T. Chen and C. Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794. ACM, 2016. 23

[16] S. J. Choi and G. M. Gulati. Bias in judicial citations: A window into the behavior of judges? *The Journal of Legal Studies*, 37(1):87–129, 2008. 5

[17] A. Cohen. The pervasive influence of ideology at the federal circuit courts. Technical report, National Bureau of Economic Research, 2023. 3, 11

[18] F. B. Cross and E. H. Tiller. Judicial partisanship and obedience to legal doctrine: Whistleblowing on the federal courts of appeals. *Yale LJ*, 107:2155, 1997. 5, 11

[19] G. B. Dahl, A. R. Kostøl, and M. Mogstad. Family welfare cultures. *The Quarterly Journal of Economics*, 129(4):1711–1752, 2014. 3

[20] H. David, N. Maestas, K. J. Mullen, and A. Strand. Does delay cause decay? the effect of administrative decision time on the labor force participation and earnings of disability applicants. Technical report, National Bureau of Economic Research, 2015. 3

[21] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. 7, 22

[22] R. Doerfler and S. Moyn. Democratizing the supreme court. Technical report, 2020. 3

[23] N. Epley and T. Gilovich. The mechanics of motivated reasoning. *Journal of Economic perspectives*, 30(3):133–40, 2016. 2

[24] J. R. Firth. A synopsis of linguistic theory, 1930-1955. *Studies in linguistic analysis*, 1957. 8

[25] J. Förster, N. Liberman, and E. T. Higgins. Accessibility from active and fulfilled goals. *Journal of Experimental Social Psychology*, 41(3):220–239, 2005. 2

[26] N. Gennaioli and A. Shleifer. Judicial fact discretion. *The Journal of Legal Studies*, 37(1):1–35, 2008.

5

[27] M. Gentzkow, J. M. Shapiro, and M. Taddy. Measuring group differences in high-dimensional choices: Method and application to congressional speech. *Econometrica*, 87(4):1307–1340, 2019. 4, 5, 8, 9, 11

[28] A. Grover and J. Leskovec. node2vec: Scalable feature learning for networks. In *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 855–864, 2016. 8

[29] J. Haidt. The positive emotion of elevation. *Prevention & Treatment*, 3(1), 2000. 2

[30] R. L. Hasen. Polarization and the judiciary. *Annual Review of Political Science*, 22:261–276, 2019. 9

[31] R. Hirschl. *Towards juristocracy: the origins and consequences of the new constitutionalism*. Harvard University Press, 2009. 3

[32] Z. Jelveh, B. Kogut, and S. Naidu. Political language in economics. Technical Report 14-57, 2018. 4

[33] J. Jensen, S. Naidu, E. Kaplan, L. Wilse-Samson, D. Gergen, M. Zuckerman, and A. Spirling. Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech [with comments and discussion]. *Brookings Papers on Economic Activity*, pages 1–81, 2012. 4

[34] D. M. Kahan. Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 8(4):407–424, 2013. 2

[35] D. M. Kahan. The politically motivated reasoning paradigm, part 1: What politically motivated reasoning is and how to measure it. *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource*, pages 1–16, 2016. 3, 5

[36] D. M. Kahan, D. Hoffman, D. Evans, N. Devins, E. Lucci, and K. Cheng. Ideology or situation sense: an experimental investigation of motivated reasoning and professional judgment. *U. Pa. L. Rev.*, 164:349, 2015. 2

[37] J. R. Kling. Incarceration length, employment, and earnings. *American Economic Review*, 96(3):863–876, 2006. 3

[38] S. Levinson. *Our undemocratic Constitution: Where the Constitution goes wrong (and how we the people can correct it)*. Oxford University Press, 2006. 3

[39] M. K. Levy. Panel assignment in the federal courts of appeals. *Cornell L. Rev.*, 103:65, 2017. 6

[40] O. Mendels, C. Peled, N. Vaisman Levy, T. Rosenthal, L. Lahiani, et al. Microsoft Presidio: Context

aware, pluggable and customizable pii anonymization service for text and images, 2018. 8

[41] M. Mueller-Smith. The criminal and labor market impacts of incarceration. Technical report, 2015. 3

[42] C. L. Nemacheck. *Strategic Selection: Presidential Nomination of Supreme Court Justices from Herbert Hoover through George W. Bush*. University of Virginia Press, 2007. 24, 25

[43] A. Niblett and A. H. Yoon. Friendly precedent. *William and Mary Law Review*, 57(5):1789–1824, 2016. 5

[44] V. Sanh, L. Debut, J. Chaumond, and T. Wolf. Distilbert, a distilled version of bert: smaller, faster, cheaper and lighter. *arXiv preprint arXiv:1910.01108*, 2019. 22

[45] C. Schmitt. *Gesetz und Urteil: Eine Untersuchung zum Problem der Rechtspraxis*. Beck, 1969. 3

[46] C. Schmitt. The crisis of parliamentary democracy [1923]. *Trans. E. Kennedy*, 1985. 3

[47] C. Schmitt. *Political theology: Four chapters on the concept of sovereignty*. University of Chicago Press, 2005. 3

[48] D. Simon and N. Scurich. Judicial overstating. In *Published in Special Symposium: The Supreme Court and the American Public*, volume 88, pages 14–36, 2014. 9

[49] D. R. Songer. *The United States Courts of Appeals Data Base Documentation for Phase 1*. http://artsandsciences.sc.edu/poli/juri/cta96_codebook.pdf, Last visited June 20, 2018. 7

[50] A. M. Sood. Motivated cognition in legal judgments—an analytic review. *Annual Review of Law and Social Science*, 9:307–325, 2013. 2

[51] C. R. Sunstein, D. Schkade, and L. M. Ellman. Ideological voting on federal courts of appeals: A preliminary investigation. *Virginia Law Review*, pages 301–354, 2004. 11

[52] M. Thaler. The fake news effect: Experimentally identifying motivated reasoning using trust in news. *arXiv preprint arXiv:2012.01663*, 2020. 3

[53] M. Touré-Tillery and A. Fishbach. How to measure motivation: A guide for the experimental social psychologist. *Social and Personality Psychology Compass*, 8(7):328–341, 2014. 2

[54] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le. Xlnet: Generalized autoregressive pretraining for language understanding. *Advances in neural information processing systems*, 32, 2019. 22

[55] B. Zeigarnik. On the retention of completed and uncompleted activities. *Psychologische Forschung*,

9:1–85, 1927. 2

A Training

A.1 Text Classification

We implement fine-tuning on three popular transformer-based pre-trained models and use a simple average ensemble of predictions as the final predictions of texts on political affiliations of judges. The first model we use is DistilBERT [44], a smaller version of the BERT model designed to overcome the slow training problem of BERT [21] due to the large model size while obtaining similar performance as BERT. Secondly, we use two improved version of BERT, XLnet [54] and twitter-RoBERTa [13] that are trained on larger corpus and with improved architecture than the original BERT model.

For fine-tuning, we used the Python package `transformer` and accessed pre-trained models from Huggingface.co, a collaborative open-source platform for model sharing. The distilbert-base-uncased-finetuned-sst-2-english model was fine-tuned using default parameters over five epochs on 70% of a 10% sample (comprising 22,922 opinions), with the remainder serving as the test set. The xlnet-base-cased and twitter-roberta-base-sentiment-latest models were trained on 70% of a 5% sample for five epochs with a learning rate of 2e-5, other parameters being default, due to computational limitations. Post fine-tuning, these models were applied to the entire sample for political party predictions.

Overall, three models exhibited comparable results, consistently achieving a prediction accuracy around 0.7, shown in Table 5. Altering the number of epochs from two to eight did not significantly impact the outcomes, as we consistently employed the best model for predictions.

Table 5: Model Performance Metrics, Text Classification

Model	Training Loss	Validation Loss	Accuracy	N
DistilBERT	0.5013	0.5481	0.707588	22,292
twitter-RoBERTa	0.4897	0.5960	0.700084	11,146
XLnet	0.4740	0.5926	0.698619	11,146

A.2 Citation Classification

We first use a grid search method with K-fold cross validation to tune the parameters used in different algorithms (a list of commonly used algorithms) in order to maximize the evaluation metric of that algorithm (here we used the AUC score). Then we use a voting ensemble method based on the best estimator of each model to average results obtained from the set of algorithms. The analysis is done using Python packages `scikit-learn` and `xgboost`. After training, we apply the ensemble model on full sample.

For each algorithm, we allow the algorithm to search among a set of possible parameters to optimize the prediction, as in [8]:

- Elastic Net. A10-fold cross validation is added to the algorithm to choose the optimal mixing parameter of LASSO and ridge regression among a vector of possible choices: $[0.1, 0.15, 0.5, 0.7, 0.95, 0.99, 1]$.
- Decision Tree. We use a 10-fold cross validation to choose the optimal minimal samples per leaf among a vector of possible choices: $[1, 5, 10, 20, 50, 100, 150, 500, 1000]$.
- Random Forest. We use a 10-fold cross validation to choose the optimal minimal samples per leaf among a vector of possible choices: $[5, 10, 20, 50, 100, 200, 500, 1000]$.
- XGBoost, by [15]. We use a 10-fold cross validation to choose the optimal maximum number of leafs among a vector of possible choices: $[3, 5, 10, 20, 50, 100, 200, 500, 1000]$.
- K-Nearest Neighbors. We use a 10-fold cross validation to choose the optimal number of neighbors among a vector of possible choices: $[20, 50, 100, 200, 300, 500]$.

Overall, the voting ensemble is as good as every individual algorithms, and the accuracy is around 0.60.

Table 6: Model Performance Metrics, Citation Classification

Algorithm	F1 Score	Accuracy	N
Elastic Net	0.5621	0.5821	192,758
Regression Tree	0.5651	0.5764	192,758
Random Forest	0.5865	0.5974	192,758
XGBoost	0.5763	0.5868	192,758
K-Nearest Neighbors	0.5682	0.5884	192,758
Voting Ensemble	0.5797	0.5946	192,758

B Polarization across Time

In this section, we re-examined the patterns presented in Figure 1 of the main paper, employing a linear regression model with fixed effects for Circuit Court and Legal Issue. This analysis aimed to assess polarization across three dimensions: prose, precedent, and policy. As shown in Figure 3, a marked increase in textual polarization is observed starting from the 1970s, indicating a shift towards more politically charged language in judicial opinions. In contrast, precedent polarization does not show a significant change, reinforcing the notion that language, rather than legal precedents, has become a primary medium for expressing politically motivated reasoning. Furthermore, dissent rates along party lines have been on the rise since the 1970s, suggesting an increasing tendency for judges to vote in accordance with their political affiliations.

C Polarization by Experience

To explore the underlying mechanism of behavioral anomalies, we examined if such anomalies diminish with experience. Specifically, we focused on whether anomalies are driven by Type I thinking, which may erode with experience, unlike Type II thinking, like motivated reasoning, which are more reflective and intentional. Using the same linear regression framework, we analyzed how polarization in reasoning varies with judges' experience. Our findings, presented in Table 7, reveal that polarization in prose remains largely unchanged with experience, except for a notable increase among judges with 15 to 25 years of experience. These results suggest that while judges' experience do not significantly impact polarization in their textual content, their selection of precedents becomes a bit more polarized in the middle of their careers. This finding is particularly striking given the overall increase in textual polarization over the years, suggesting that this trend might not be primarily driven by the accumulation of judicial experience. Further research is necessary to fully understand these dynamics and the factors influencing them. These patterns, where prose polarization is mostly unaffected by experience, suggest that the behavioral anomalies are driven by Type II thinking, being more reflective and intentional in nature.

D Polarization during Vacancies

As noted by [42], since the era of President Eisenhower, there has been a growing trend for presidents to prefer individuals from federal courts as potential Supreme Court candidates. This preference may be attributed to the clearer ideological traceability of federal judges compared to candidates from other backgrounds. Since President Ford's nomination of Justice John G. Roberts, approximately 73% of the nominees have been Circuit Court judges. In light of this trend, our study focuses on all Supreme Court vacancies from 1975 to 2013. We consider the vacancy period, plus the six months preceding it, as our sample timeframe.

Following the approach of [9] for defining vacancies and contenders, we identify the start of a vacancy as the date a justice first informs the president of their intention to step down. The vacancy period ends when the Senate confirms the nomination. We define contenders as judges included in the president's shortlist for each vacancy, based on the criteria established by [42]. Our analytical specification for examining the influence of promotion incentives on judicial polarization is outlined below:

$$Y_{it} = \alpha + \beta Vacancy_t + \gamma Contender_i + \delta Vacancy_t \times Contender_i + \boldsymbol{\eta}' \boldsymbol{Z}_{it} + \varepsilon_{it} \quad (1)$$

where Y_{it} is the polarization outcome (e.g. dissent rate), and $\boldsymbol{Z}_{it}$ are Circuit $\times$ Year and legal-issue fixed effects. We estimate the equation using OLS with robust standard errors clustered by individual judge. The coefficient of primary interest is δ, which measures the average difference in the polarization outcome, accounting for the fixed effects, for contenders during the periods of judicial vacancies.

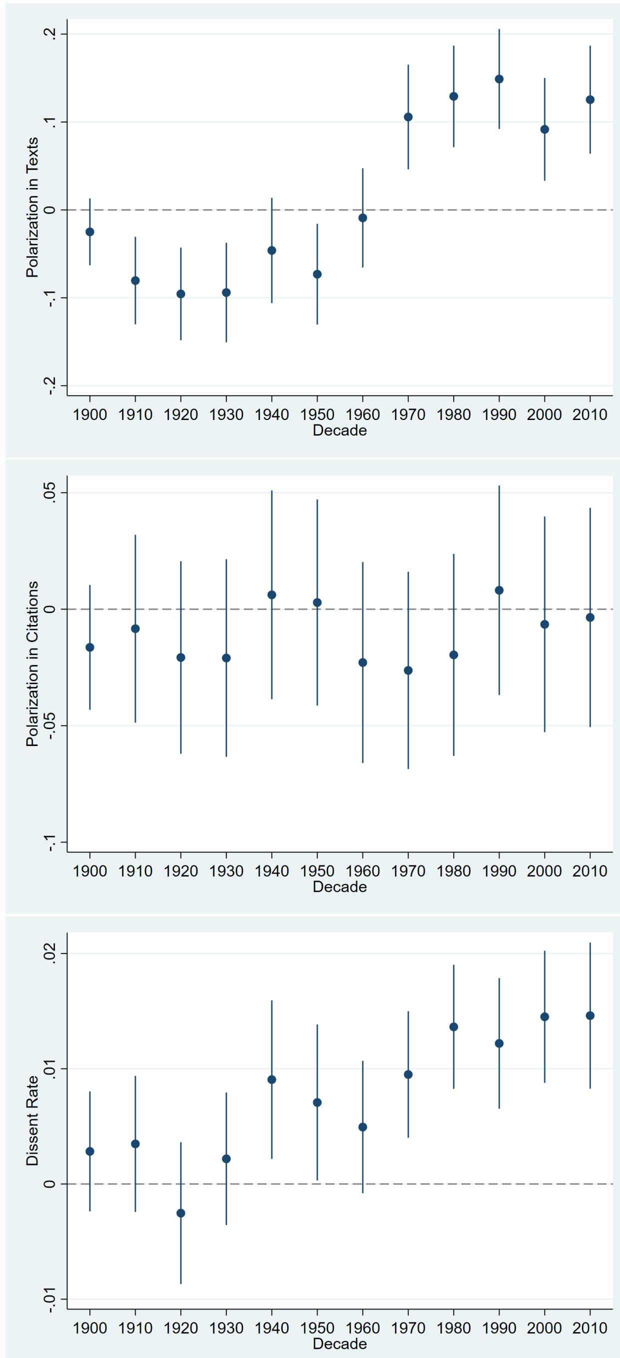

Figure 3: Polarization in prose, Precedent, and Policy across time

Notes: The temporal changes in polarization in texts, citations, and dissent votes. The baseline level is 1890-1900. We control for Circuit and Legal Issue fixed effects. Standard errors clustered at judge level in parentheses.

Table 7: The effect of experience on polarization

	(1) Text	(2) Citation
Age	-0.000	-0.001***
	(0.001)	(0.001)
Experience $\in [5, 10)$	-0.000	0.004
	(0.005)	(0.003)
Experience $\in [10, 15)$	-0.008	-0.001
	(0.010)	(0.006)
Experience $\in [15, 20)$	-0.002	0.017*
	(0.016)	(0.009)
Experience $\in [20, 25)$	-0.005	0.025*
	(0.021)	(0.013)
Experience $\in [25, 30)$	-0.025	0.010
	(0.027)	(0.017)
Experience $\in [30, 35)$	-0.030	-0.007
	(0.033)	(0.022)
Experience $\in [35, 55)$	-0.039	-0.018
	(0.042)	(0.032)
Observations	312930	271059
R^2	0.334	0.117
Circuit × Year FE	✓	✓
Legal Issue FE	✓	✓

Notes: The baseline level is Experience $\in [0, 5)$ years. Standard errors clustered at judge level in parentheses. $*p < .1, **p < 0.05, ***p < .01$.