Motivated Reasoning in the Field: Polarization of Prose, Precedent, and Policy in U.S. Circuit Courts, 1930-2013

Elliott Ash, Daniel L. Chen, Wei Lu*

March 20, 2019

Abstract

We document motivated reasoning among U.S. Circuit Court judges. We employ a supervised learning approach to measure partisan influences on prose (writing style), precedent (citations to previous cases), and policy (dissenting votes). We find persistent but low partisanship of language overall, with the notable exception of Civil Rights and First Amendment jurisprudence. Citations display a significant level of partisanship and increase over time. Voting along party lines (dissenting against judges from opposing party) and strategic retirement (retiring while one's own party controls the presidency) have also increased.

*Elliott Ash, ashe@ethz.ch, ETH Zurich; Daniel L. Chen, daniel.chen@iast.fr, Toulouse School of Economics, Institute for Advanced Study in Toulouse, University of Toulouse Capitole, Toulouse, France; Wei Lu, wei.lu17@rotman.utoronto.ca, University of Toronto, Rotman School of Management. First draft: June 2017. Current draft: February 2019. Latest version at: http://nber.org/∼dlchen/papers/Motivated_Reasoning_in_the_Field.pdf. Work on this project was conducted while Daniel Chen received financial support from the European Research Council (Grant No. 614708), Swiss National Science Foundation (Grant Nos. 100018-152678 and 106014-150820), and Agence Nationale de la Recherche.

"'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, New York Times, Sep. 11, 2017).*

1 Introduction

Motivated reasoning is the well-documented tendency where individuals actively seek out confirmatory information. When presented with information, individuals have an easier time absorbing facts supporting what they want to be true than facts supporting what they want to be false (Epley and Gilovich 2016 [19]).[1] In the lab, motivation is 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 (Tour-Tillery and Fishbach 2014 [38]).[2]

In prior studies of motivating reasoning in law, law student subjects are exogenously provided precedents (reasons) (Braman 2006 [11]; Braman & Nelson 2007 [12]). The experiments fix the set of precedents to choose, such that differences in choices cannot be due to knowledge. In vignettes involving interpreting facts, those scoring highest in cognitive reflection were the most likely to display ideologically motivated cognition (Kahan 2013 [27]). But whether these studies on law students are externally valid to judges or other policymakers is still an open question (Sood 2013 [34]).

The task of this paper is to analyze motivated reasoning among real-world judges, using as a natural laboratory the U.S. federal courts – a high-stakes common-law space. Circuit judges can introduce new legal theories,[3] shift standards or thresholds,[4] and rule on the constitutionality of federal and state statutes. Circuit judges

[1] The mechanism is said to be implicit emotion regulation – the brain converges on judgments that minimize negative affective states associated with threat or maximize positive affective states associated with attainment of motives.

[2] Typically, studies only measure the final decision, rather than reasoning (Zeigarnik 1927 [40]; Förster, Liberman, and Higgins 2005 [20]), though recent studies have employed vignettes to measure ex post justification in moral dilemmas (Haidt 2000 [24]). In economics, several recent models and experiments on motivated reasoning are summarized in Benabou and Tirole (2016).

[3] 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

[4] (e.g., shift from reasonable person standard to reasonable woman standard for what constitutes

provide the final decision on tend 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 the majority of what law students are reading). If there is motivated reasoning among these judges, that could have important legal and policy impacts.

There are a handful of key features of Circuit Courts 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),[5] meaning that 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.[6] 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 decisions (almost a million judge votes) for the years 1930-2013. Circuit Court judges are appointed by the U.S. president (Democrat or Republican) with life tenure. The measures of judicial reasoning are constructed from the judges' votes, the texts of the opinions, and the citations between opinions. These outcome measures are linked to judicial biographical features, and in particular the party of the judge's nominating president.

The partisanship measure is based on Gentzkow, Shapiro, and Taddy (2018, [23], henceforth GST). It is defined as the accuracy with which we can predict judge political affiliation based on the content of their decisions. The model uses penalized estimation of parameters to avoid finite-sample bias. Earlier efforts to measure polarization in text include Jensen et al (2012) [26], whose non-penalized measure might over-estimate polarization in early years. Ash et al (2017) [6] use a similar measure to analyze variation in polarized language in the Senate over the electoral cycle. Outside of Congress, Jelveh et al. (2017) [25] use the text of academic articles to predict political donations by economists. An important difference between our context and these previous papers is that congressmen (and economists) have discretion over the topics they address, while judges are assigned topics randomly.

A paralle lliterature has looked at polarization of citizens rather than policymakers. Bertrand et al. (2018) show that partisan affiliations are most associated with social attitudes (rather than consumption and time use).

Like GST, we first predict judge party using text – prose. We represent judicial

sexual harassment, or waive need to prove emotional harm in court to a jury.)

[5]This randomness has been used in a growing set of economics papers ([28, 7, 8, 17, 30, 3]).

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

prose as N-grams and use those as predictors in the GST model. We find that average prose partisanship for judges has remained low, unlike Congress. There is no clear trend over time, either upwards or downwards. By analyzing particular topics, we found that Democrats are more partisan (more easily predicted) on civil rights cases, while Republicans are more partisan on First Amendment cases.

A new contribution is to look at polarization of precedent. We use counts over citations to previous decisions, rather than N-grams, to predict partisan affiliation. Unlike the case of prose, we find high and significant levels of precedent partisanship which ebbs and swells over time. We can say that in our context, judges tend to express ideological differences through citations instead of choice of phrases. These results complement the previous work by Choi and Gulati (2008) [16] showing that circuit judges tend to cite judges from the same party, and that of Niblett and Yoon (2016) [31] showing that circuit judges tend to cite Supreme Court cases authored by judges from the same party.

Finally, we look at polarization of policy. To do this, we look at dissents along party lines. We find that vote polarization has increased: Dissents are increasingly cast by judges sitting in the minority position (with two judges appointed by the opposing party). Democrats are more likely to issue minority dissents on civil rights cases, while Republicans tend to issue minority dissents on First Amendment cases. These findings are related to Sennewald et al. (2015 [**?**]), who hand-coded a large sample of district court cases as liberal or conservative and showed that in recent decades Republican district court judges vote more conservatively, and Democrat district court judges vote more liberally. The results for judges complement the large literature on polarization of votes between political parties in Congress (e.g. McCarty et al 2006 [29]; Andris et al 2015 [1]).

Further evidence of increasing partisan policy interest is increasing prevalence of politically strategic retirement. Using two centuries of data, we see roughly 13% of retirements and 36% of resignations following political cycles, and an increase in retirement partisanship over time. Previous large-scale quantitative studies of the relationship between politics and judicial exits in the U.S. Courts of Appeals have not found electoral cycles in judicial turnover rates using a research design conducted at the yearly level ([39]; [35]; [36]; [41]).

In the appendix we analyze variation across judges in motivated reasoning. Motivated reasoning is more pronounced for Republican-appointed judges. It grows with judicial experience (but not age). For circuit judges who later ascend to the Supreme Court, the motivated-reasoning measure is predictive of future Supreme Court voting.

This paper adds to the emerging empirical literature in economics using text data (Gentzkow, Kelly, and Taddy, 2017, [21]), and the emerging theory literature that is interested in narratives (Benabou et al. 2018 [9]). Our findings are also related to Berdejó and Chen (2017) [10] and Chen (2017) [14], which show the ideological bias of judges shifts before Presidential elections through an increase in dissents and partisan voting, and that this electoral dissent has been growing over time. Ash, Chen, and Naidu (2017) [5] use phrases from court opinions and journal articles to construct measures for the influence of law and economics in the federal judiciary. Ash and Chen (2018[4]) show that document embeddings are not predictive of ideology.

2 Data

The dataset uses the universe of opinions published by U.S. Circuit Courts for the years 1891 through 2013. Only majority opinions written by a specific author (excluding opinions labeled *per curiam*, which are authored by the whole panel without designating a specific author) are included in the analysis. For judge characteristics, we link the opinions to the dataset of Berdejó and Chen (2017) [10], and use variables like circuit, year, gender, education, and years of service.

In addition, the cases are divided into 8 topics that are used in the Songer database [33]. The eight topics are criminal, civil rights, First Amendment, due process, labor relations, and economic activity/regulation.[7] These Songer topics are meant to capture *the nature of the conflict between the litigants* (Songer, 2008 [33]).

2.1 Prose

For each opinion, we use Python scripts to tokenize paragraphs and sentences. After removing capitalization and punctuation, we follow a standard approach for text featurization and represent the cases as frequency distributions over phrases. To construct the vocabulary, we start with the set of observed bigrams (two-word phrases, as opposed to unigrams, trigrams, etc) in the corpus. This is a standard featurization choice in information extraction or document classification. In addition, in previous work we found that most legal concepts (memes) that are shared betewen cases are encoded in bigrams (Chen et al. 2016 [15]).

To achieve a tractable sample of bigrams, we take the following filtering steps. First, we exclude phrases that did not appear in at least 10 different opinions in 10 separate years. Second, we apply a parts-of-speech tagger and include only noun and verb phrases, which tend to be relatively informative and familiar (Denny et al 2015 [18]; Ash 2018 [2]). Finally, we normalize the features using a Porter stemmer [32] to strip suffixes. This process leaves a total of 32,945 phrases.

[7] We drop the privacy and miscellaneous categories for lack of cases.

We then sum the phrase counts to judge-year level, and use these as our observations for estimation.

2.2 Precedent

For citations, we include citations to prior Circuit Court cases that appeared in at least 2 distinct cases in 2 separate years. The total number of legal precedents that satisfy this requirement is 105,948.

2.3 Policy

We have data on all votes case by judges on the panels for each case. Partisan policy is constructed using data on judge party affiliation and judge dissenting votes. to look at strategic retirement, we use data from 1800 to 2004 from the Multi-User Data Base on the Attributes of U.S. Appeals Court Judges to sum up the number of retirements or resignations per month.

3 Measuring polarization from text and citations

3.1 Bayesian interpretation of partisanship

Suppose we have a neutral observer who is reading a written opinion from the Circuit Court. The observer knows little about the writer, and simply reading the opinion does not provide much information about the writer's political affiliation. This is true in theory because judges are asked that they "*not be swayed by partisan interests*",[8] so opinions are not supposed to convey explicit partisan information. If judges follow this edict, then a reasonable guess on party affiliation for the naive reader is 0.5 – i.e., the probabilities of the writer being a Republican or a Democrat are the same.

But this 50/50 guess might not be the best in practice. If the expressed reasoning of judges is motivated by partisan views, then the choice of language and citations might be informative. This motivated reasoning would alter the way judges interpret and evaluate information from briefs and precedents, which would lead to differences in their expressed arguments. Therefore the Bayesian observer would update their beliefs toward more informative posteriors.

Following Gentzkow, Shapiro, and Taddy (2018) [22] (GST), we adopt an intuitive approach to demonstrate how political considerations may lead to partisanship of language and citations in the written opinions of U.S. Circuit Courts. Consider each court opinion as lists of phrases and citations. Each of these lists can be viewed as a realization of a probabilistic token generating process. To study the properties of

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

this process, we want to learn the probability that each token appears in an opinion. Our theoretical model follows the Naive Bayes approach that these probabilities are conditional on individual characteristics of the judges. Therefore the token counts follow a certain distribution characterized by personal traits and total number of phrases written.

In particular, judges from different parties may use different phrases and citations in their opinions. Once we know their choice of tokens, we may infer party membership of a judge better than guessing the prior. Our measure of partisanship, then, is a function of posterior probabilities of the observer. After we read their opinions, the posterior probabilities of a judge being a Republican or Democrat is updated to to $\boldsymbol{q}^R$ and $\boldsymbol{q}^D = 1 - \boldsymbol{q}^R$, respectively. Partisanship can be understood as the difference between these posterior beliefs. If prose and precedent are more predictive of partisan affilitation, then we can say that judicial outputs are more polarized.

3.2 Theoretical foundation for measuring partisanship

We assume the following model, based on GST (2018) [22]. Let $\boldsymbol{c}_{it}$ be the vector of phrase counts for judge i in year t, with each member c_{ijt} representing the frequency of a certain phrase j. Let $m_{it} = \sum_j c_{ijt}$ be the total number of phrases written (verbosity), and let l be the list of phrases written. The probability of writing phrases 1, ..., l a number c_{i1t}, ..., c_{ilt} times, with the total number of phrases equal to m_{it}, is given by a multinomial distribution. Thus $\boldsymbol{c}_{it}$ draws from

$$\boldsymbol{c}_{it} \sim \mathrm{MN}(m_{it}, \boldsymbol{q}_t^{P(i)}(\boldsymbol{x}_{it})), \tag{1}$$

where $\boldsymbol{q}_t^{P(i)}(\boldsymbol{x}_{it}) \in [0, 1)$ is the l-vector of probabilities of writing each phrase at time t conditional on judge characteristics $\boldsymbol{x}_{it}$ and judge party $P(i) \in \{R, D\}$. This probability is unknown to the observer, and it is recovered by the machine learning algorithm described below. Given $\boldsymbol{c}_{it}$ and m_{it}, our goal is to estimate $\boldsymbol{q}_t^{P(i)}(\boldsymbol{x}_{it})$ in order to compute the partisanship measure described below.

As in GST, we have a simple discrete choice model that can micro-found the probabilistic model. The utility function for judge i at year t is

$$u_{it} = \begin{cases} \delta y_t + (1-\delta)(\boldsymbol{\alpha}_t' + \boldsymbol{x}_{it}'\boldsymbol{\gamma}_t)\boldsymbol{c}_{it}, \ i \in R; \\ -\delta y_t + (1-\delta)(\boldsymbol{\alpha}_t' + \boldsymbol{x}_{it}'\boldsymbol{\gamma}_t)\boldsymbol{c}_{it}, \ i \in D, \end{cases} \tag{2}$$

where $y_t = \boldsymbol{\varphi}' \sum_i \boldsymbol{c}_{it}$ is a measure of peer or public opinion, $\boldsymbol{\alpha}_t$ is a vector of time intercepts, $R = \{i : P(i) = R, m_{it} > 0\}$, $D = \{i : P(i) = D, m_{it} > 0\}$ give the sets of Republican and Democrat judges, and δ is a utility weight that disappears in the following transformation.

To simplify the model, assume that every judge chooses his or her phrases to maximize utility u_{it} with respect to a choice-specific i.i.d. type 1 extreme value shock. Then we can reduce the above equation into the following:

$$u_{ijt} = \widetilde{\alpha}_{jt} + \boldsymbol{x}'_{it}\widetilde{\boldsymbol{\gamma}}_t + \widetilde{\varphi}_j \mathbf{1}_{i\in R}, \tag{3}$$

where $\widetilde{\boldsymbol{\varphi}} = 2\delta\boldsymbol{\varphi}$, $\widetilde{\boldsymbol{\gamma}}_t = \boldsymbol{\gamma}_t(1-\delta)$ and $\widetilde{\boldsymbol{\alpha}}_t = (1-\delta)\boldsymbol{\alpha}_t - \delta\boldsymbol{\varphi}$. The probability of judge i writing phrase j at year t is:

$$\Pr(\text{Phrase} = j|\text{Party} = P(i), \text{Time} = t, \boldsymbol{x}_{it}) = q_{jt}^{P(i)}(\boldsymbol{x}_{it}) = \frac{e^{u_{ijt}}}{\sum_l e^{u_{ilt}}} \tag{4}$$

where e is Euler's number.

We would like to measure partisanship by judge and by year. To get partisanship for a judge we first have to aggregate over phrases. To this end, partisanship $\pi_t(\boldsymbol{x})$ is defined as the distance between $\boldsymbol{q}_t^R(\boldsymbol{x})$ and $\boldsymbol{q}_t^D(\boldsymbol{x})$, the vectors of $\boldsymbol{x}$-conditioned phrase probabilities for Republicans and Democrats. As in GST, we select the following distance metric:

$$\pi_t(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{q}_t^R(\boldsymbol{x})\boldsymbol{\rho}_t(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{q}_t^D(\boldsymbol{x})(1-\boldsymbol{\rho}_t(\boldsymbol{x})) \tag{5}$$

where $\boldsymbol{\rho}_t$ is an l-vector over phrases with entries $\rho_{jt}(\boldsymbol{x}) = \Pr(\text{Party} = R|\text{Phrase} = j, \text{Time} = t, \boldsymbol{x}_{it}) = \frac{q_{jt}^R(\boldsymbol{x})}{q_{jt}^R(\boldsymbol{x})+q_{jt}^D(\boldsymbol{x})}$; that is, $\boldsymbol{\rho}_t(\boldsymbol{x})$ is the posterior belief that a neutral observer with a neutral prior assigns to a judge being Republican if he writes phrase j in year t and has characteristics $\boldsymbol{x}$. Therefore we have the following elegant Bayesian interpretation: $\pi_t(\boldsymbol{x})$ gives the posterior probability that an observer with a neutral prior would expect to assign to the speaker's true party (Republican or Democrat) after reading all tokens, given judge characteristics $\boldsymbol{x}$.

To compute partisanship in the data at year t, we then have to aggregate over judges. This is just the average over all the judges at t,

$$\bar{\pi}_t = \frac{1}{|R\cup D|}\sum_{i\in R\cup D}\pi_t(\boldsymbol{x}_{it}).$$

This partisanship measure is the average of the posterior across all judges and phrases at year t, and lies between 0.5 and 1. It is the measure we use to analyze partisanship in the time series below.

3.3 Model Estimation

Now that we have our matrix of frequencies, we describe the estimation procedure. We learn the parameters $(\widetilde{\alpha}_{jt}, \widetilde{\boldsymbol{\gamma}}_t, \widetilde{\varphi}_j)$ to minimize

$$\sum_j \left\{ \sum_t \sum_i [m_{it} \exp(\widetilde{\alpha}_{jt} + \boldsymbol{x}'_{it}\widetilde{\boldsymbol{\gamma}}_t + \widetilde{\varphi}_j \mathbf{1}_{i\in R}) - c_{ijt}(\widetilde{\alpha}_{jt} + \boldsymbol{x}'_{it}\widetilde{\boldsymbol{\gamma}}_t + \widetilde{\varphi}_j \mathbf{1}_{i\in R}) + \lambda_j |\widetilde{\varphi}_j|] \right\} \quad (6)$$

where $\mathbf{1}_{i\in R}$ is a vector of indicator variables equaling one for Republican judges, and λ_j is a phrase-level penalty term for the political party affiliation of judges.

The characteristics $\boldsymbol{x}_{it}$ include circuit, gender, years of service, and type of undergraduate degree (public or private). These enter as controls in the multinomial inverse regression. That is to say, we compute the phrase partisanship conditional on individual characteristics, but the posterior inference on partisanship is made after controlling for those characteristics. For instance, if all Democrat judges were women and all Republican judges men, but text was completely random, our posterior update on partisanship using text is still 0.

Taddy (2015) [37] proves that minimizing (6) is equivalent to minimizing the multinomial logit model, where the the multinomial logit likelihood is approximated with a Poisson model likelihood. This method allows us to estimate each phrase separately in parallel, with the phrase count of each individual modeled as

$$c_{ijt} \sim \text{Pois}(\exp[\log(m_{it}) + u_{ijt}]).$$

This estimation procedure would be infeasible without parallelized computation across phrases. We impose an L_1 penalty on the party membership loadings, which results in sparsity and shrinkage toward zero. The optimal value of λ_j is chosen by minimizing a corrected Bayesian Information Criterion.[9]

We also include circuit times year fixed effects, which ensures random assignment of cases. These fixed effects, and the permutation inference, are two methodological differences from Choi and Gulati (2008) and Niblett and Yoon (2016) [31].

[9] To implement these methods, we use the R packages `distrom` and `textir`, developed by Taddy (2015) [37]. In the latest version of Gentzkow, Shapiro and Taddy (2018) [22], the authors add a constant penalty on intercepts and other covariates. This feature is not currently available in the current versions of distrom and textir. We skip this step; but as noted in their article, this is not necessary for most datasets. Also, instead of penalizing whether a judge is of a certain party *in a certain year*, we are just using the penality term for the party affiliation only. We rented an Amazon Web Service EC2 instance m4.16xlarge with 64 vCPU (virtual CPUs) and 256 GB RAM (later we transferred to Microsoft Azure virtual machine with similar specifications).

3.4 Permutation Inference

To analyze polarization over time, we report the main results graphically. To best illustrate the contrast between the actual partisanship and the permutation result, we generate a random assignment of party membership of each judge.[10]

By the definition of partisanship measure, random shuffling of membership should result in a measurement very close to 0.5, as $\boldsymbol{q}_t^R(\boldsymbol{x})$ and $\boldsymbol{q}_t^D(\boldsymbol{x})$ are equal when the other characteristics hold. But if there is small-sample bias, then the permutation placebo could diverge from .5. Therefore in our figures we show the average partisanship computed from the actual data (solid line) and partisanship computed from data with shuffled party affiliation (dotted line). Then these lines are close to each other, that is not evidence of polarization. When the solid line is above the dotted line, that is evidence of polarization.

4 Partisanship in Prose

This section reports the results on partisanship of prose. The results are reported as a series of figures. Blue and red dots give individual partisanship measures for Democrats and Republicans.

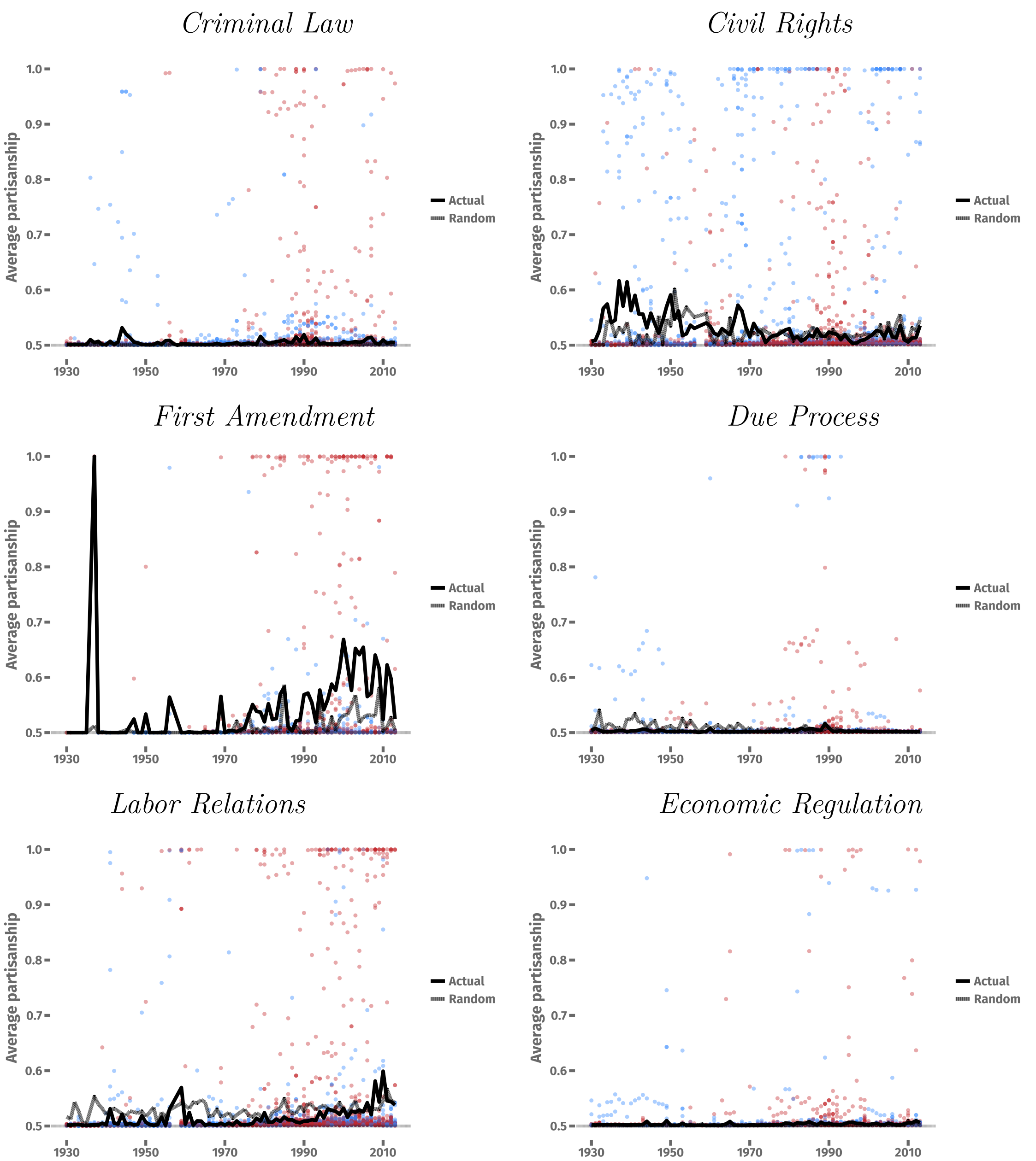
[10]In Gentzkow, Shapiro and Taddy (2018) [22], they use the average share of Republican party during the year which the a particular individual is active for random assignment. In practice, we found that this is similar to just randomly assigning membership with probability of 0.5, so we also use 0.5 for random assignment.

Figure 1: Partisanship in Prose in U.S. Circuit Courts, 1930-2013

Partisanship measures over time in U.S. Circuit Courts, 1930-2013 for phrases. Black solid lines give the average partisanship in the true dataset. Dotted lines give average partisanship in the shuffled dataset (random party affilliations). Red and blue dots show individual observations for Republican and Democrat judges. See text for further discussion of Specification A/B and other technical details.

Figure 1 illustrates the average and individual partisanship across time for two-word phrases used by Circuit judges. The magnitude of average partisanship for phrases of judges is very close to 0.5 for the whole time period. This is evidence of relatively low partisanship in court opinion text. This is different from the main result in GST [22], where the language patterns of congressmen exhibited high partisanship.

Figure 2: Partisanship of phrases by topic



Partisanship measures over time in U.S. Circuit Courts, 1930-2013 for phrases by topic. Black solid lines give the average partisanship in the true dataset. Dotted lines give average partisanship in the shuffled dataset (random party affilliations). Red and blue dots show individual observations for Republican and Democrat judges.

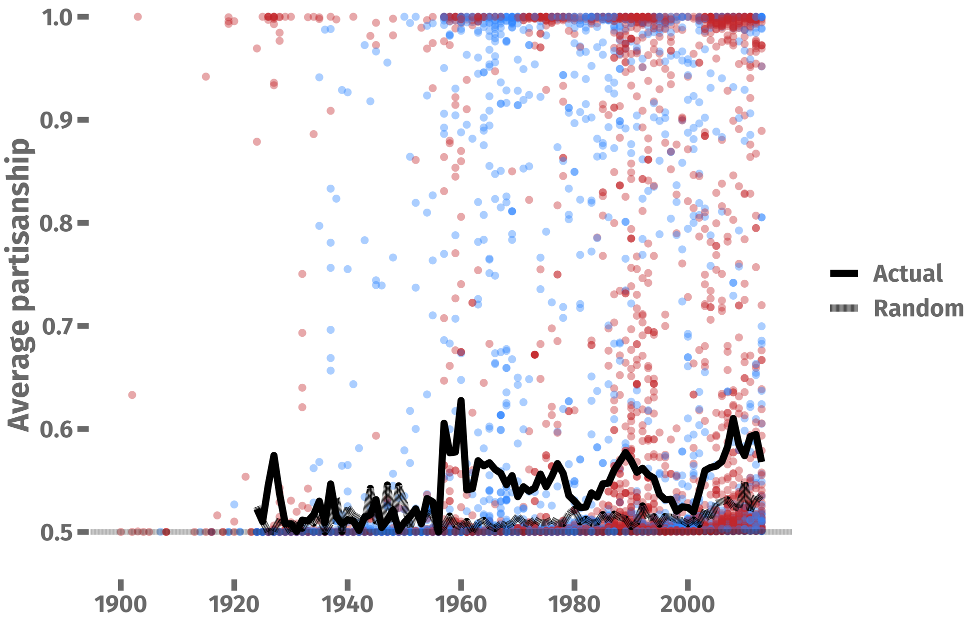
Now we examine how partisanship in phrases varies across topics. These results are reported in Figure 2. Phrase partisanship is consistently low for criminal law, due process, labor, and economic regulation. For Civil Rights, the language was quite polarized during the 1930s and 1940s. For First Amendment, there was significant polarization starting in the 1990s.[11]

[11] The First Amendment's protections have been held to cover flag burning, cross burning, commercial advertising, campaign funding, virtual child pornography, violent video games and DVDs,

5 Polarization of Precedent

In this section we analyze partisan influnces in the precedents that judges select.

Figure 3: Partisanship in Precedents in U.S. Circuit Courts, 1930-2013

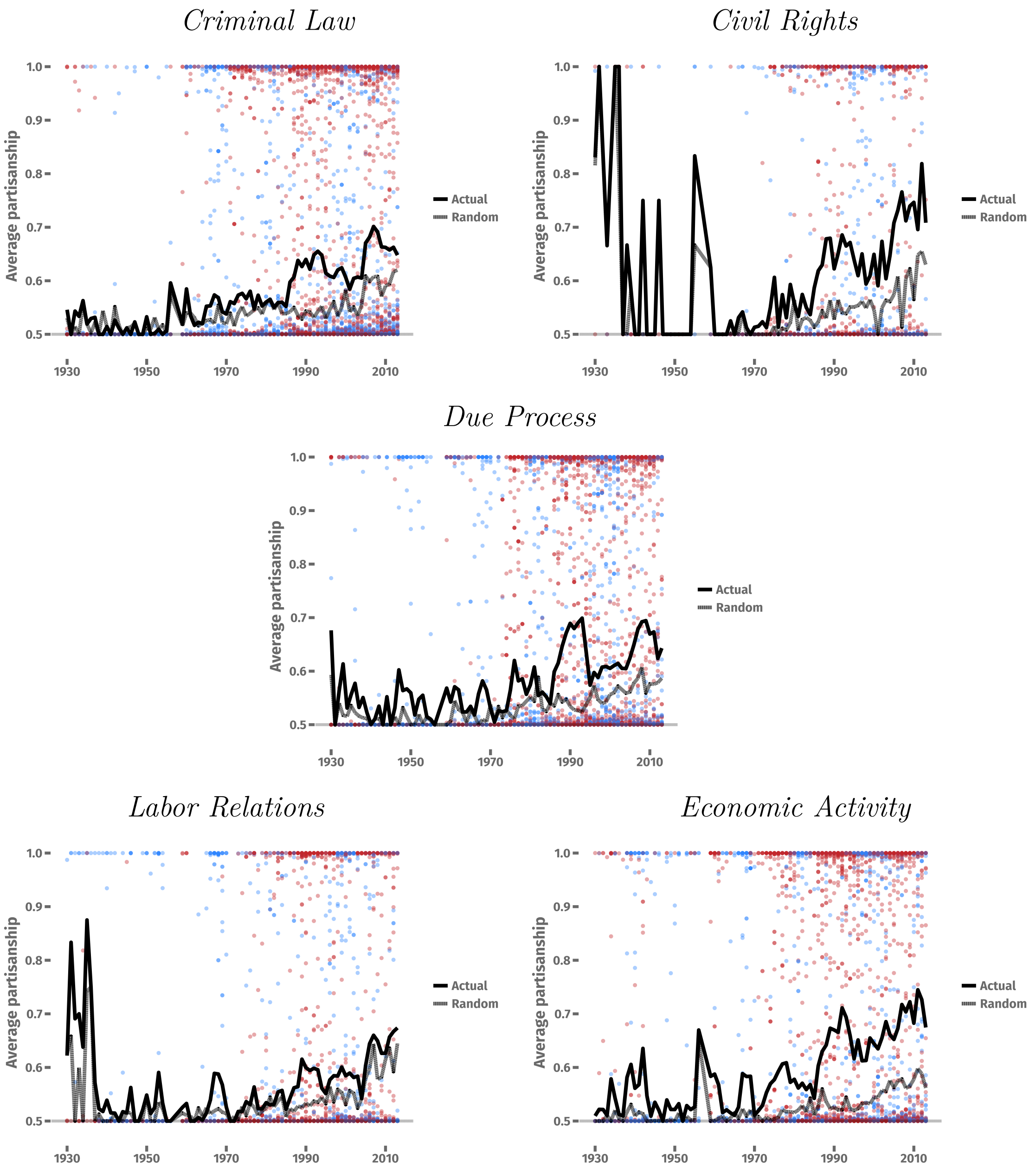


Partisanship measures over time in U.S. Circuit Courts, 1930-2013 for citations. Black solid lines give the average partisanship in the true dataset. Dotted lines give average partisanship in the shuffled dataset (random party affilliations). Red and blue dots show individual observations for Republican and Democrat judges. See text for further discussion of Specification A/B and other technical details.

Figure 3 illustrates the average and individual partisanship across time for citations used by Circuit judges. Unlike the text measures, the citations exhibit relatively high partisanship, at least since the lat 1950s. This measure more comparable to the congressional language in GST. Unlike choice of phrases, it seems that judges tend to cite different cases in line with their preferred outcome. These choices are distinct enough that by looking at their citations one can guess their party affiliation. However, this does not seem to have increased significantly since 1960.

expressive association, protests at military funerals and abortion clinics, false statements of fact, nude dancing, limits to disciplinary measures in public schools, government employment actions, and conditions attached to government benefits.

Figure 4: Partisanship of precedents by topic



Partisanship measures over time in U.S. Circuit Courts, 1930-2013 for citations by topic. Black solid lines give the average partisanship in the true dataset. Dotted lines give average partisanship in the shuffled dataset (random party affilliations). Red and blue dots show individual observations for Republican and Democrat judges.

We look at precedent polarization across topics in Figure 4. Unlike with text, the precedents for criminal law are polarized. With civil rights, there are some small-sample issues in early years but again judges are polarized in recent years. For due process, labor, and economic activity, again we see polarization in precedents, especially in recent years.

6 Partisanship in Policy

In this section we follow up on whether the polarization observed in prose and precedents are reflected in policy choices, as indicated by voting. We also look at whether judges are more politically strategic in their retirement decisions.

6.1 Vote Partisanship

Table 1: Vote Partisanship

	Dissent Vote	
	(1)	(2)
Minority	0.0103**	-0.487**
	(0.000895)	(0.0549)
Minority * Year		0.000251**
		(0.0000280)
N	799180	799180
Case FE	X	X
Judge FE	X	X
Cluster	Judge	Judge

Notes. Effect of being a minority judge (D of DRR or R of RDD) on the likelihood to cast a dissenting vote, controlling for case and judge fixed effects. Standard errors clustered by judge. Sample is cases with judges from both political parties. $+p < .1, *p < 0.05, **p < .01$.

Figure 5: Growth in Vote Partisanship

Notes. Left figure displays decade-by-decade effects of being a minority judge (D of DRR or R of RDD) on the likelihood to cast a dissenting vote, controlling for case and judge fixed effects. Error spikes give 95% confidence intervals. Right figure displays the share of dissenting votes every year. Vertical line is at 1970.

The section looks at partisanship using judge voting. Table 1 shows that minority judges are 1 percentage points more likely to dissent and this is growing over time. Since roughly 2% of votes cast are dissents, the effect size is substantial relative to the average. Figure 5 shows that minority dissent is growing more sharply than any dissent, especially after 1970 when the rate of any dissent has flattened.

Table 2: Vote Partisanship by Topic

	Dissent Vote					
	Criminal	Civil Rght	1st Amend	Due Process	Labor	Econ
Minority	0.0105**	0.0196**	0.0244**	0.0105**	0.0142**	0.00398**
	(0.00124)	(0.00245)	(0.00870)	(0.00118)	(0.00247)	(0.000920)
N	171019	46179	3278	179019	37262	232199
Case FE	X	X	X	X	X	X
Judge FE	X	X	X	X	X	X
Cluster	Judge	Judge	Judge	Judge	Judge	Judge

Notes. Effect of being a minority judge (D of DRR or R of RDD) on the likelihood to cast a dissenting vote for cases of different topics, controlling for case and judge fixed effects. Standard errors clustered by judge. Sample is cases with judges from both political parties. $+p < .1, *p < 0.05, **p < .01$.

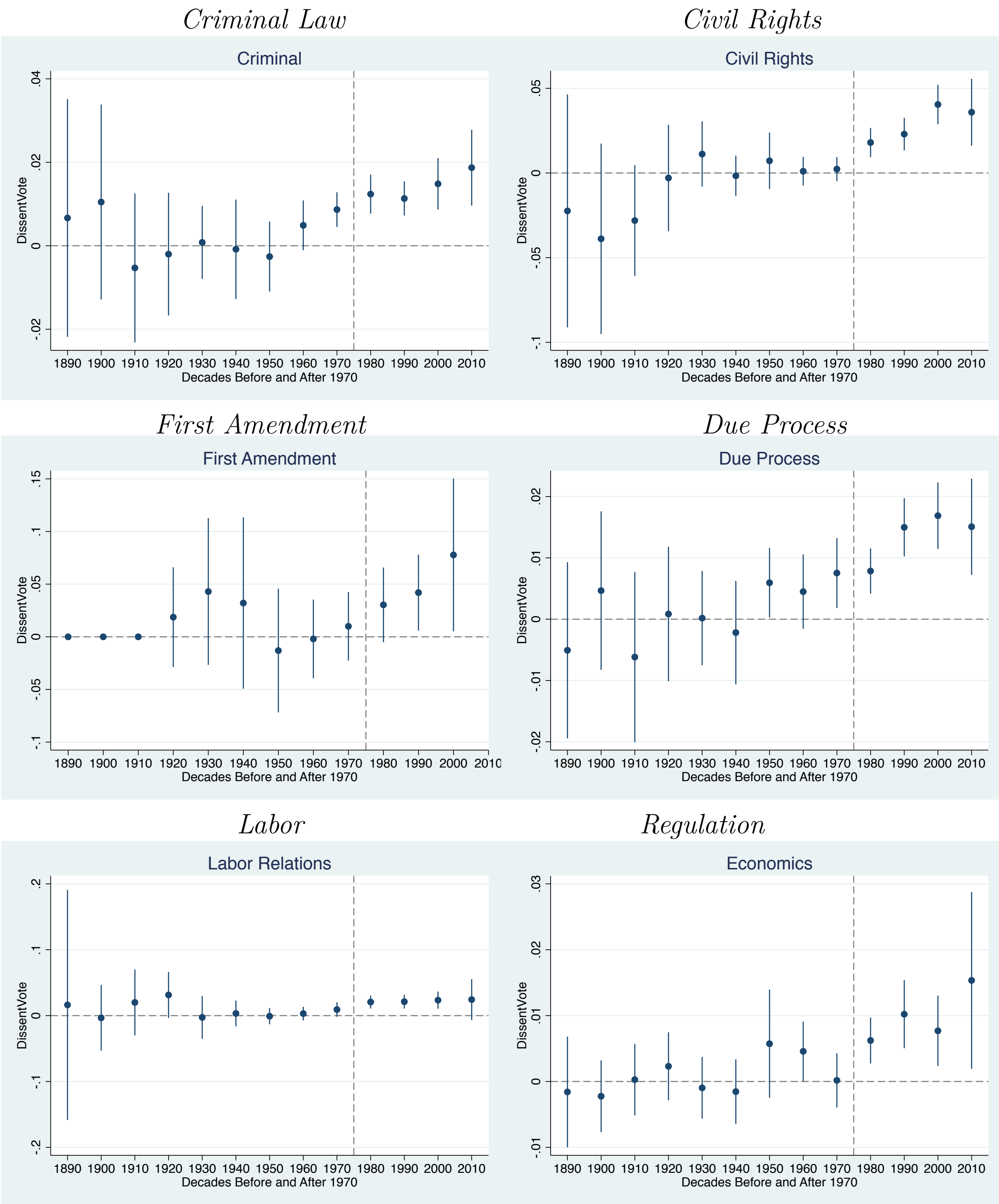
Table 3: Vote Partisanship by Topic and Party

	Dissent Vote					
	Criminal	Civil Rght	1st Amend	Due Process	Labor	Econ
Minority	0.00959**	0.0112*	0.0382+	0.00826**	0.00307	0.00534*
	(0.00254)	(0.00545)	(0.0227)	(0.00255)	(0.00486)	(0.00237)
Minority	0.00285	0.0184+	-0.0267	0.00483	0.0235*	-0.00174
* Dem	(0.00445)	(0.00989)	(0.0408)	(0.00468)	(0.00945)	(0.00474)
N	171019	46179	3278	179019	37262	232199
Case FE	X	X	X	X	X	X
Judge FE	X	X	X	X	X	X
Cluster	Judge	Judge	Judge	Judge	Judge	Judge

Notes. Effect of being a minority judge (D of DRR or R of RDD), fully interacted with party, on the likelihood to cast a dissenting vote for cases of different topics, controlling for case and judge fixed effects. Standard errors clustered by judge. Sample is cases with judges from both political parties. $+p < .1, *p < 0.05, **p < .01$.

Table 2 shows that almost all topics display vote partisanship. Table 3 shows that Democrats issue more minority dissents in civil rights and labor cases. These findings echo the greater degree of phrase partisanship among Democrats for civil rights cases.

Figure 6: Growth in Vote Partisanship by Topic



Notes. Figures display decade-by-decade effects of being a minority judge (D of DRR or R of RDD) on the likelihood to cast a dissenting vote for cases of different topics, controlling for case and judge fixed effects. Error spikes give 95% confidence intervals.

Figure 6 shows the changes in vote polarization by topic over time. There is some degree of partisanship increase in criminal cases, starting in the 1960s. Civil rights partisanship appears to have sharply increased in the 1980s and is presently at a high level relative to partisanship in other legal topics. There is some cycling in first amendment partisanship, though the standard errors are wide due to the smaller sample. Due process partisanship increased in the 1950s, a decade prior to criminal case partisanship. The figure indicates some partisanship in labor cases, perhaps cylical as well with high partisanship in the 1920s. There has been a marked increase in partisanship for economics cases since the 1980s.

Together, these findings tend to corroborate findings in Berdejó and Chen (2017) [10] and Chen (2017) [14] that there has been a marked increase in the role of partisan ideology in appellate decisions over the last century.

6.2 Strategic Retirement

In a recent survey, 410 federal judges were asked, "Why do you remain in active service?" ([13]). Two of the political responses were: "I plan to take senior status but am waiting for a different appointing authority (i.e., a different **political** administration) to nominate my successor" and "I intend to retire but am waiting for a different appointing authority (i.e., a different **political** administration) to nominate my successor". Most judges reported a 1 on a Likert scale of 1-7, which is "not at all important or not applicable.[12] The highest-rated responses were, instead, "want to participate" and "want to retain staffing"/"full caseload". Less than 1% of U.S. Federal judges report political motivations for retirement and resignation.

With due respect to the survey answers, we aim to provide empirical evidence on partisan motivations for judicial retirement. Regressions are of the form:

$$\text{Exit}_i = F(t) + \textbf{Proximity}_i'\beta + \epsilon_i \tag{7}$$

where the outcome variable Exit_i is the number of judicial retirements or resignations in month i; the explanatory variable of interest, $\textbf{Proximity}_i$, is a set of quarter-to-election fixed effects;[13] $F(t)$ includes a set of year-specific fixed effects and fixed effects for each quarter of the year (e.g., January through March, April through June, etc.);[14] finally ϵ_i is a mean–zero stochastic error term. In all regressions, the unit of analysis is the month in order to control for seasonality.[15]

[12] To be sure, Likert scales can lead to bias and are not revealed preference measures of true motivations (Cavaille et al. 2018).

[13] We compare to quarter 16, i.e., the quarter immediately following an election, which is the omitted quarter, so the interpretation is akin to a regression discontinuity design.

[14] The set of year-specific fixed effects is intended to capture shocks or trends affecting judicial retirement that are common to all judges in a given year, while the quarterly fixed effects control for seasonal variation in judges' retirement decisions.

[15] In all calculations of statistical significance in this section, robust standard errors are used.

Figure 7: Judicial Exit across the Political Cycle

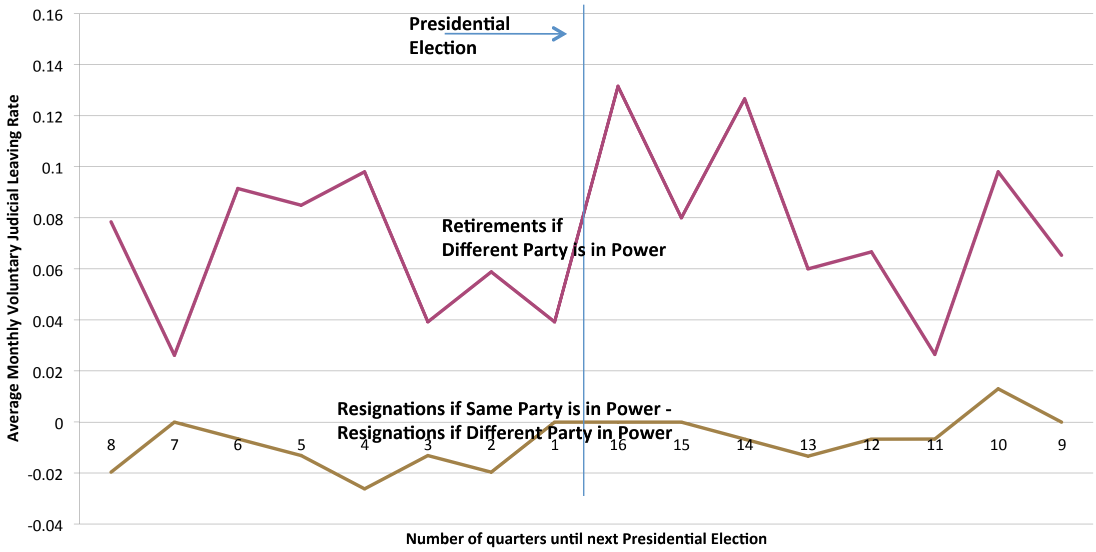


Figure 7 visualizes the results without any covariates. When the party in power is different, retirements dip in a pronounced manner before presidential elections. This would be consistent with judges intending to retire but waiting for a different appointing authority (i.e., a different political administration) to nominate their successor, contra what the survey evidence indicates. We include a number of robustnes checks in the appendix.

To investigate whether political cycles in judicial exits have increased, we divide the dataset into pre- and post-1975 periods. We compare judicial retirements in the three quarters immediately following an election with the three quarters immediately preceding an election (analogizing to a regression discontinuity framework). Formally we estimate:

$$Exit_i = F(t) + \beta_1 After_i * Recent_i + \beta_2 After_i + \beta_3 Recent_i + \varepsilon_i \tag{8}$$

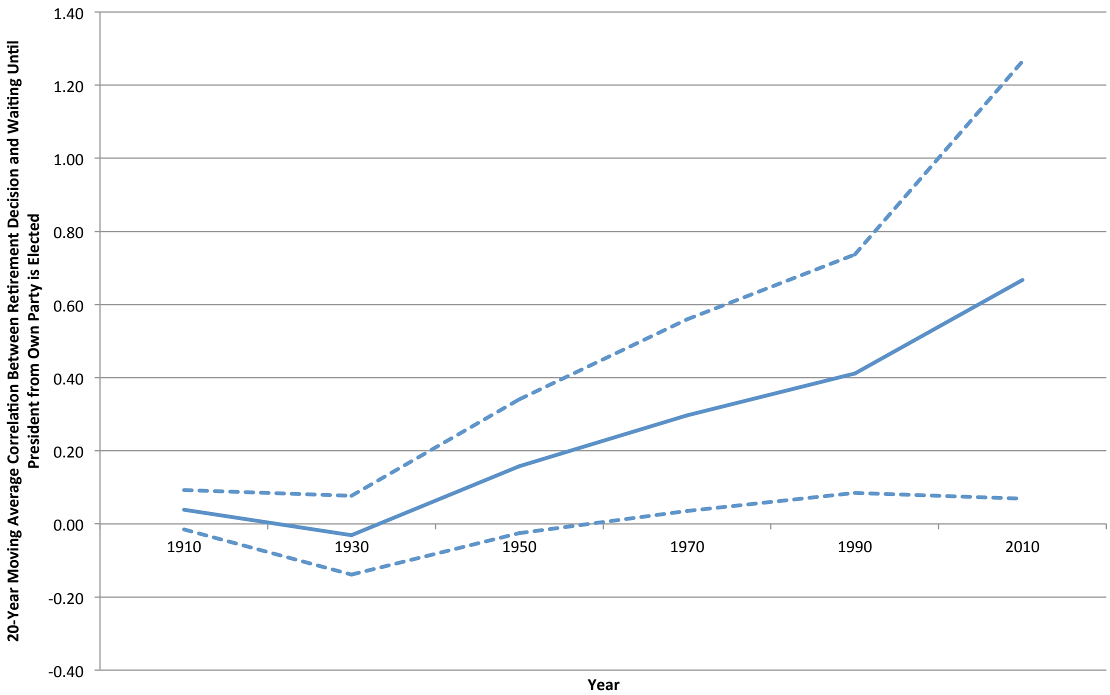
where $F(t)$ are year and quarter fixed effects; $After_i$ is an indicator equal to 1 for the three quarters immediately following a Presidential election; and $Recent_i$ is an indicator variable equal to 1 for the period of time after 1975.

Table 4: Political Cycles in Judicial Exits Over Time

	(1)	(2)
	Number of Judicial Retirements	
After Election	-0.0191	-0.0153
	(0.0436)	(0.0374)
After Election * Year > 1975	0.497**	
	(0.223)	
After Election * Year > 1900		0.145*
		(0.0827)
Year FE	Yes	Yes
Season FE	Yes	Yes
Observations	911	911
R-squared	0.328	0.313

Notes: Robust standard errors in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). The explanatory variables of interest are dummy variables indicating whether it is after an election or not (the first three quarters after an election count as "after" while the three quarters before an election count as "before") and whether it is recent (before or after 1975; before or after 1900) and the regressions also include year fixed effects and seasonly quarter fixed effects.

Figure 8: Increase in Electoral Cycles in Judicial Exits Over Time



We can see the result in Table 4. The higher rate of voluntary retirements following an election appears entirely attributable to the post-1975 period. The 20-year moving average correlation between retirement decision and whether it is after an election (Figure 8) suggests that the electoral cycles we observe in judicial retirement decisions may be entirely a recent phenomenon. Political cycles may explain a much larger proportion of judicial exits in recent years, suggesting that judges have stronger partisan policy motivations than they did in earlier years.

The sclerotization of the normal churning of judges to reflect the preferences of the electorate may raise questions about the non-partisanship of the judiciary. If judges wait to have their replacements selected by a President from the same party, and if judges observe others appointed by the opposing party are also waiting, they may choose to wait as well, creating a positive feedback for the judiciary to become more polarized over time. Sclerotization of the natural churning of judges may lead the judiciary to become less reflective of the preferences of the electorate over time. The fact that judges do not admit to political cycles in exits is consistent with motivated reasoning.

7 Conclusion

Judges are nominally expected to sit above the partisan fray -- but we find they are divisive in their rhetoric and decisions. What are the doctrinal sources of ideological partisanship in the judiciary? One idea is that the language used, or the authorities cited, might be informative about political views. We find that there is small partisanship across judges in text, but significant partisanship in citations.

This is accompanied by vote partisanship. They tend to dissent along partisan lines–dissenting only against judges appointed by the opposing party's president. In addition, we document retirement partisanship–judges strategically timing their retirement decisions to be replaced by someone ideologically similar.

Comparing results for phrases and citations, it appears that judges are less polarized through their phrases than their citations (legal precedents cited as justifications). Another open question for future research is the causal relationship between polarized phrase usage and polarized precedent.

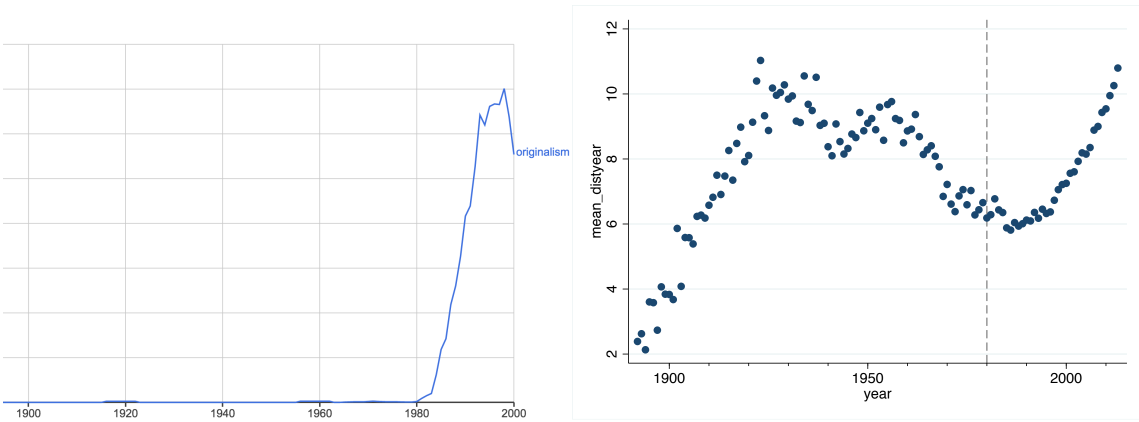
Using all 26 Supreme Court judges (1946-2016) who sat on at least 50 Circuit cases, we find that judge who moves from the most Democrat to the most Republican in precedent and phrase usage is 32 percentage points and 23 percentage points, respectively, more likely to vote conservative. A judge who moves from the lowest to highest rank in vote polarization is 25 percentage points more likely to vote conservative. This last result is noteable as either Democrats or Republicans can cast minority dissents, but this pattern more saliently predicts conservative votes on the Supreme Court.

We also present a number of findings for future research: does prose concentration cause citation concentration, and does it lead to social change? One question is whether motivated reasoning is conscious or unconscious, deliberate or implicit. The fact that partisanship persists at the time of retirement and is expressed by a judge in his exit interview would suggest that judges are deliberately engaging in motivated reasoning. This suggests not all behavioral biases are "Type I" that can be eliminated by a nudge.

Asymmetric partisanship suggests that judges from one party are using more concentrated language whose usage would indicate the party, but non-usage would not indicate membership in either party. For example, Contract with America was a smaller vocabulary that the Republican party used in 1994 that marked the increase in partisanship of congressional speech. The civil rights language concentration by Democrats occurred in the midst of civil rights changes while First Amendment concentration by Republicans occurred in the midst of the originalist movement ("How Conservatives Weaponized the First Amendment", New York Times, 06/30/2018). A future question is whether concatenated vocabulary causes social change.

Originalism itself is an example of neologism, coined by Paul Brest in 1980, when he stated, "By 'originalism' I mean the familiar approach to constitutional adjudication that accords binding authority to the text of the Constitution or the intentions of its adopters." In Google Books, the term "originalism" does not appear until 1980, and around that time, the time distance between the case and the cases it cites has grown. Moreover, citations to the Bill of Rights experienced an inflection in growth around the 1970s, such that currently 30% of cases cite the Bill of Rights. Democrats have said "we are all Originalists now".

Figure 9: Originalism



References

[1] C. Andris, D. Lee, M. J. Hamilton, M. Martino, C. E. Gunning, and J. A. Selden. The rise of partisanship and super-cooperators in the us house of representatives. *PloS one*, 10(4):e0123507, 2015. 1

[2] E. Ash. The political economy of tax laws in the u.s. states. Technical report, 2016. 2.1

[3] E. Ash and D. L. Chen. Religious freedoms, church-state separation, and religiosity: Evidence from randomly assigned judges. 2017. 5

[4] E. Ash and D. L. Chen. Judge vectors: Spatial representations of the law using document embeddings. 2018. 1

[5] E. Ash, D. L. Chen, and S. Naidu. The impact of legal schools of thought. Technical report, 2017. 1

[6] E. Ash, M. Morelli, and R. Van Weelden. Elections and divisiveness: Theory and evidence. *Journal of Politics*, 2017. 1

[7] D. H. Autor, 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. Working Paper 258, Michigan Retirement Research Center, 2011. 5

[8] 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. 5

[9] R. Bénabou, A. Falk, and J. Tirole. Narratives, imperatives and moral reasoning. Technical report, Working Paper, 2018. 1

[10] C. Berdejó and D. L. Chen. Electoral cycles among us courts of appeals judges. Technical report, Institute for Advanced Study in Toulouse (IAST), 2017. 1, 2, 6.1

[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. 1

[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. 1

[13] S. B. Burbank, S. J. Plager, and G. Ablavsky. Leaving the bench, 1970-2009: The choices federal judges make, what influences those choices, and their consequences. *University of Pennsylvania Law Review*, 161(1):1, 2012. 6.2, B

[14] D. L. Chen. Priming ideology: Why presidential elections affect us judges. Technical report, Institute for Advanced Study in Toulouse (IAST), 2017. 1, 6.1

[15] D. L. Chen, A. Parthasarathy, and S. Verma. The Genealogy of Ideology: Identifying Persuasive Memes and Predicting Agreement in the U.S. Courts of Appeals. Technical report, 2016. 2.1

[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. 1

[17] G. B. Dahl, A. R. Kostøl, and M. Mogstad. Family Welfare Cultures. *Quarterly Journal of Economics*, 129(4):1711–1752, 2014. 5

[18] M. J. Denny, B. O'Connor, and H. Wallach. A little bit of nlp goes a long way: Finding meaning in legislative texts with phrase extraction. Microsoft Research NYC, http://dirichlet.net/, April 2015. 2.1

[19] N. Epley and T. Gilovich. The mechanics of motivated reasoning. *Journal of Economic perspectives*, 30(3):133–40, 2016. 1

[20] 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

[21] M. Gentzkow, B. T. Kelly, and M. Taddy. Text as data. Technical report, National Bureau of Economic Research, 2017. 1

[22] M. Gentzkow, J. M. Shapiro, and M. Taddy. Measuring group differences in high-dimensional choices: Method and application to congressional speech. Technical report, National Bureau of Economic Research. 3.1, 3.2, 9, 10, 4

[23] M. Gentzkow, J. M. Shapiro, and M. Taddy. Measuring group differences in high-dimensional choices: Method and application to congressional speech. Technical report, National Bureau of Economic Research, 2018. 1

[24] J. Haidt. The positive emotion of elevation. 2000. 2

[25] Z. Jelveh, B. Kogut, and S. Naidu. Political language in economics. 2015. 1

[26] J. Jensen, E. Kaplan, S. Naidu, and L. Wilse-Samson. Political polarization and the dynamics of political language: Evidence from 130 years of partisan speech. *Brookings Papers on Economic Activity*, 2012(2):1–81, 2012. 1

[27] D. M. Kahan. Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision Making*, 8(4):407–424, 2013. 1

[28] J. R. Kling. Incarceration length, employment, and earnings. *The American Economic Review*, 96(3):863–876, June 2006. 5

[29] N. M. McCarty, K. T. Poole, and H. Rosenthal. *Polarized America: The Dance of Ideology and Unequal Riches*. Walras-Pareto lectures. The MIT Press, 2006. 1

[30] M. Mueller-Smith. The criminal and labor market impacts of incarceration: Identifying mechanisms and estimating household spillovers. Working paper, Columbia University, July 2014. 5

[31] A. Niblett and A. H. Yoon. Friendly precedent. *William and Mary Law Review*, 57(5):1789–1824, 2016. 1, 3.3

[32] M. F. Porter. An algorithm for suffix stripping. *Program*, 14(3):130–137, 1980. 2.1

[33] 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. 2

[34] A. M. Sood. Motivated cognition in legal judgments—an analytic review. *Annual Review of Law and Social Science*, 9:307–325, 2013. 1

[35] J. F. Spriggs ll. and P. J. Wahlbeck. Calling it quits: Strategic retirement on the federal courts of appeals, 1893-1991. *Political Research Quarterly*, 48(3):573–597, 1995. 1

[36] R. M. Stolzenberg and J. Lindgren. Retirement and death in office of u.s. supreme court justices. *Demography*, 47(2):269–298, 2010. 1

[37] M. Taddy. Distributed multinomial regression. *The Annals of Applied Statistics*, 9(3):1394–1414, 2015. 3.3, 9

[38] 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. 1

[39] A. Yoon. Pensions, politics, and judicial tenure: An empirical study of federal judges, 1869-2002. *American Law and Economics Review*, 8(1):143–180, 2006. 1

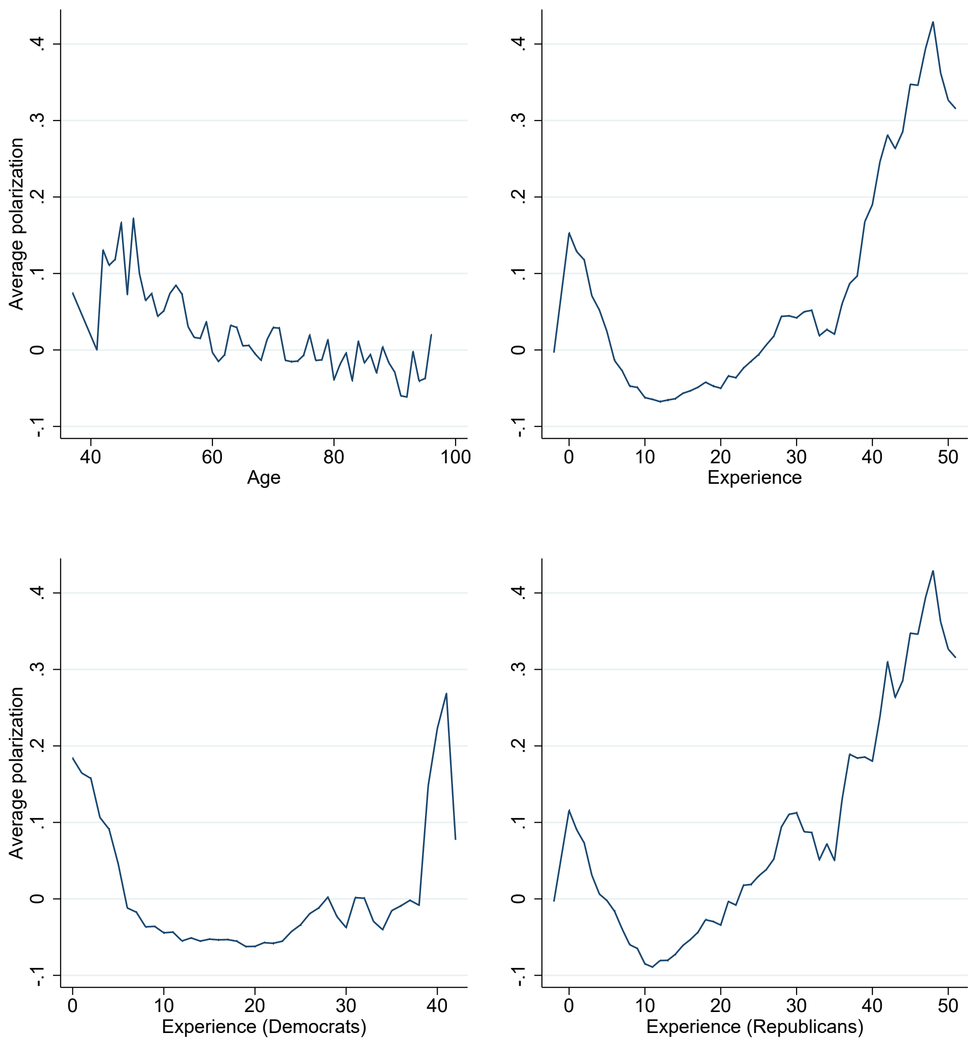
[40] B. Zeigarnik. On the retention of completed and uncompleted activities. *Psychologische Forschung*, 9:1–85, 1927. 2

[41] G. Zuk, G. S. Gryski, and D. J. Barrow. Partisan transformation of the federal judiciary, 1869-1992. *American Politics Quarterly*, 21(4):439–457, 1993. 1

A Variation in Precedent Partisanship with Experience

As an additional appendix result we show that precedent polarization is U-shaped with experience but declines with age. The growth with experience is more pronounced for Republican judges.

Figure 10: Motivated reasoning grows with experience (not age)



Notes. Demeaned by Circuit-year.

B Robustness Checks for Strategic Retirement

In order to calculate the share of judicial exits that are politically motivated, we assume that the benchmark is essentially random retirements or resignations, spread evenly over 16 quarters between elections and evenly without regards to the party of the appointing President. We use as a baseline the fact that on average, 0.14 judges voluntarily leave the bench each month in my sample; of these, 0.12 are retirements and 0.02 are resignations.

In each of the three quarters before a Presidential election, the number of retirements for judges when the party in power is different drops by 0.08-0.10 per month (Table 4 Column 2). This is rather large—summary statistics displayed on the first row of numbers in Table 4 indicate that when the party in power is different from the party of the appointing President of the judge, 0.07 judges retire per month. The magnitudes are invariant to the controls as one might expect from the unconditional visualization. These effects are also statistically significant at the 1% or 5% level and much larger in magnitude than the other quarters.[16] Estimates from the negative binomial model also indicate statistically significant reductions in retirements when the party in power is different (at the 1% or 5% level) for each of the three quarters preceding a Presidential election.

Finally, to interpret the magnitudes, assuming that we should expect $0.124 * 48 = 5.95$ judges to retire every 4 years, a back-of-the-envelope comparison yields the abnormal number of judges not retiring before the election. Regression coefficients in the three quarters (each containing 3 months) prior to election indicates that $(0.079 + 0.076 + 0.107) * 3 = 0.79$ judges are missing, which suggests 13% of judicial retirements are politically motivated.

[16] All significance tests are two-tailed with respect to the null hypothesis of no effect. There is one other quarter that is significant at the 5% level. No statistically significant effects are observed when the party in power is the same (Column 1). One might expect one effect to be significant at the 5% level after 20 tests.

Table 5: Political Cycles in Judicial Exits

	(1)	(2)	(3)	(4)
	Number of Retirements		Number of Resignations	
Party in Power	Same	Different	Same	Different
Mean of dep. var.	0.051	0.073	0.015	0.008
Quartertoelect = 1	0.00741	-0.0793**	-0.00832	-0.00430
	(0.0269)	(0.0365)	(0.0109)	(0.00413)
Quartertoelect = 2	-0.0130	-0.0762**	0.00861	0.00122
	(0.0254)	(0.0386)	(0.0175)	(0.00865)
Quartertoelect = 3	-0.0302	-0.107***	0.00257	0.0117
	(0.0245)	(0.0383)	(0.0184)	(0.00920)
Quartertoelect = 4	0.0270	-0.0101	0.00685	-0.00489
	(0.0508)	(0.0531)	(0.0243)	(0.00587)
Quartertoelect = 5	0.00829	-0.00447	0.00834	-0.00265
	(0.0539)	(0.0614)	(0.0274)	(0.00979)
Quartertoelect = 6	0.0794	-0.0144	-0.00741	-0.00368
	(0.0571)	(0.0623)	(0.0278)	(0.0109)
Quartertoelect = 7	0.0295	-0.0905	-0.0265	-0.00631
	(0.0543)	(0.0585)	(0.0261)	(0.00897)
Quartertoelect = 8	0.0344	-0.0399	0.0235	-0.00979
	(0.0479)	(0.0582)	(0.0253)	(0.0106)
Quartertoelect = 9	0.0222	-0.0538	0.0315	-0.00755
	(0.0489)	(0.0614)	(0.0249)	(0.0135)
Quartertoelect = 10	0.0541	-0.0377	0.0223	0.00450
	(0.0558)	(0.0659)	(0.0252)	(0.0167)
Quartertoelect = 11	0.0106	-0.121**	0.0376	0.00851
	(0.0481)	(0.0612)	(0.0258)	(0.0173)
Quartertoelect = 12	-0.0377	-0.0699	0.0228*	-0.0150
	(0.0408)	(0.0557)	(0.0136)	(0.0152)
Quartertoelect = 13	-0.0337	-0.0709	0.0442***	-0.0127
	(0.0457)	(0.0576)	(0.0164)	(0.0172)
Quartertoelect = 14	-0.0478	-0.0207	0.0350**	-0.00701
	(0.0453)	(0.0617)	(0.0170)	(0.0195)
Quartertoelect = 15	-0.0651	-0.0781	0.0290*	0.00355
	(0.0416)	(0.0595)	(0.0160)	(0.0206)
Year FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	2433	2433	2433	2433
R-squared	0.198	0.282	0.098	0.091

Notes: Robust OLS standard errors in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). The outcome variables are the number judges that retire in a particular month (Columns 1-2) and the number judges that resign in a particular month (Columns 3-4).

We also see a politically motivated pattern for resignations. As noted above, the baseline is 0.02 judges resigning per month. In each of the four quarters after a Presidential election, the number of resignations for judges when the party in power is the same increases by 0.02-0.04 per month (Column 3). These numbers are again large relative to the mean—when the party in power is the same as the party of the appointing President of the judge, 0.015 judges resign per month (and when the party in power is different from the party of the appointing President of the judge, 0.008 judges resign per month). These effects are statistically significant at the 1, 5, or 10% level and much larger in magnitude than the other quarters. The effects are larger in Column 4, which display the regression for the number of resignations when the party in power is different. Estimates from the negative binomial model also indicate increases in resignations when the party in power is the same, increases that are statistically significant at the 1% level for the significant quarters in the linear model.

To interpret the magnitudes, assuming that we should expect $0.023 * 48 = 1.10$ judges to resign every four years, the missing $(0.023+0.044+0.035+0.029)*3 = 0.39$ judges calculated by summing the four quarters-to-election coefficients suggests that 36% of judicial resignations follow political cycles.

It is clear these rates for retirements and resignations fluctuate across the political cycle. It is important to note that quarter 16, which contains parts of November, December, January, and part of February is the omitted quarter, which has a coefficient of 0. Thus, for instance, in Column 3, the coefficients on quarters 12-15 are estimated to be significant relative to the quarter right after, not relative to the election date. When we omit quarter 1 instead of quarter 16, so that the omitted quarter is the one immediately preceding the election, the coefficients on quarters 12-15 are still statistically significant and slightly larger in magnitude.

The patterns are robust to alternative measures of electoral proximity (i.e., linear quarters-to-next election rather than with quarter-to-election dummies) and dropping one Circuit at a time.[17]

[17] The results are also robust to a specification that employs disaggregated data using the number of retirements per Circuit-month, including Circuit fixed effects, and clustering the standard errors at the Circuit level.

Table 6: Political Cycles in Judicial Exits - Robustness Checks

	(1)	(2)
	Number of Retirements	
	Each coefficient represents a separate regression	
		Drop 1 Circuit at a time
Quarters to Election	0.00563**	
	(0.00286)	
After Election		0.0741*
(Entire Sample)		(0.0425)
After Election		0.0686*
(Drop Circuit 1)		(0.0401)
After Election		0.0784*
(Drop Circuit 2)		(0.0426)
After Election		0.0828*
(Drop Circuit 3)		(0.0427)
After Election		0.0828*
(Drop Circuit 4)		(0.0427)
After Election		0.0414
(Drop Circuit 5)		(0.0375)
After Election		0.0828*
(Drop Circuit 6)		(0.0423)
After Election		0.0752*
(Drop Circuit 7)		(0.0415)
After Election		0.0686*
(Drop Circuit 8)		(0.0415)
After Election		0.0501
(Drop Circuit 9)		(0.0333)
After Election		0.0686*
(Drop Circuit 10)		(0.0400)
After Election		0.0643
(Drop Circuit 11)		(0.0426)
After Election		0.0512
(Drop Circuit 12)		(0.0397)

Notes: Robust standard errors in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). Voluntary judicial leavings are the number of judges that retire or resign in a particular month. The explanatory variables of interest are dummy variables indicating whether it is after an election or not (the first three quarters after an election count as "after" while the three quarters before an election count as "before"). Each coefficient represents a separate regression. The regression also includes year fixed effects and seasonly quarter fixed effects (only in Column (1)).

These patterns are slightly more pronounced for Republican appointees.

Table 7: Who Does Political Cycles in Judicial Exits - Party of Appointment

	(1)	(2)	(3)	(4)
	Number of Judicial Retirements		Number of Judicial Resignations	
	of Democratic Judges	of Republican Judges	of Democratic Judges	of Republican Judges
Quartertoelect = 1	-0.0367	-0.0456	-0.0141	-0.00219
	(0.0307)	(0.0315)	(0.0105)	(0.00772)
Quartertoelect = 2	-0.0390	-0.0578*	-0.00943	0.0291*
	(0.0319)	(0.0350)	(0.0127)	(0.0155)
Quartertoelect = 3	-0.0599**	-0.0683*	-0.00556	0.0203
	(0.0295)	(0.0354)	(0.0142)	(0.0151)
Quartertoelect = 4	0.00730	0.0217	-0.0000577	0.00739
	(0.0427)	(0.0573)	(0.0177)	(0.0170)
Quartertoelect = 5	-0.0359	0.0316	-0.00110	0.00847
	(0.0491)	(0.0636)	(0.0212)	(0.0211)
Quartertoelect = 6	-0.00556	0.0652	-0.0160	0.0202
	(0.0522)	(0.0645)	(0.0190)	(0.0217)
Quartertoelect = 7	-0.0265	-0.0238	-0.0187	-0.00828
	(0.0511)	(0.0591)	(0.0184)	(0.0183)
Quartertoelect = 8	0.0146	-0.0253	-0.000115	0.0246
	(0.0483)	(0.0575)	(0.0183)	(0.0182)
Quartertoelect = 9	-0.0351	-0.0349	0.00211	0.0289
	(0.0517)	(0.0605)	(0.0190)	(0.0179)
Quartertoelect = 10	0.00174	-0.00791	0.0133	0.0341*
	(0.0552)	(0.0680)	(0.0211)	(0.0175)
Quartertoelect = 11	-0.0590	-0.0578	-0.00243	0.0518***
	(0.0502)	(0.0619)	(0.0184)	(0.0195)
Quartertoelect = 12	-0.0375	-0.0727	-0.0102	0.0122
	(0.0432)	(0.0556)	(0.0170)	(0.0106)
Quartertoelect = 13	-0.0381	-0.0892	0.00554	0.0167
	(0.0455)	(0.0630)	(0.0200)	(0.0117)
Quartertoelect = 14	-0.00707	-0.0814	-0.00298	0.0352**
	(0.0498)	(0.0647)	(0.0208)	(0.0145)
Quartertoelect = 15	-0.0152	-0.132**	-0.00565	0.0330**
	(0.0464)	(0.0631)	(0.0214)	(0.0142)
Year FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	2433	2433	2433	2433
R-squared	0.185	0.226	0.088	0.115

Notes: Robust standard errors in parentheses (* significant at 10%; ** significant at 5%; *** significant at 1%). Outcome variables are the number of judges that retire or resign in a particular month. The explanatory variables of interest are dummy variables indicating the number of quarters remaining before the upcoming presidential election (16 quarters to the election is the omitted dummy variable). The regression also includes year fixed effects and seasonly quarter fixed effects.

Including only judicial exits after 1975 would match the sample relevant to the [13] survey. Table 8 exhibits substantially larger political cycles in judicial retirements. The baseline is now 0.51 judges retiring per month. Assuming that we should expect $0.51 * 48 = 24$ judges to retire every 4 years, a back-of-the-envelope comparison with the regression coefficients in the three quarters prior to election suggests that an abnormal number of judges are not retiring before the election—the missing $(0.38+0.48+0.75)*3 = 4.8$ judges who are not retiring would render 20% of judicial retirements to be politically motivated.

Table 8: Political Cycles in Judicial Exits After 1975

	(1)	(2)	(3)	(4)
	Number of Retirements		Number of Resignations	
Party in Power	Same	Different	Same	Different
Mean of dep. var.	0.164	0.351	0.017	0.014
Quartertoelect = 1	-0.0315	-0.379*	0.0165	-0.00274
	(0.131)	(0.207)	(0.0210)	(0.0171)
Quartertoelect = 2	-0.108	-0.478**	0.00506	0.00589
	(0.127)	(0.204)	(0.0269)	(0.0262)
Quartertoelect = 3	-0.203*	-0.747***	0.0295	0.0921*
	(0.122)	(0.187)	(0.0574)	(0.0523)
Quartertoelect = 4	0.155	-0.219	0.0153	0.000417
	(0.286)	(0.273)	(0.0199)	(0.0141)
Quartertoelect = 5	0.205	-0.300	-0.0385	-0.00251
	(0.297)	(0.339)	(0.0461)	(0.0202)
Quartertoelect = 6	0.313	-0.298	-0.0499	0.0537
	(0.296)	(0.351)	(0.0529)	(0.0517)
Quartertoelect = 7	0.0698	-0.556*	-0.0671	0.00899
	(0.281)	(0.328)	(0.0550)	(0.0245)
Quartertoelect = 8	0.341	-0.241	-0.0374	0.000278
	(0.247)	(0.350)	(0.0393)	(0.0147)
Quartertoelect = 9	0.296	-0.275	-0.0198	0.0450
	(0.245)	(0.372)	(0.0333)	(0.0522)
Quartertoelect = 10	0.547*	-0.272	-0.0312	0.00598
	(0.311)	(0.395)	(0.0407)	(0.0284)
Quartertoelect = 11	0.209	-0.673*	-0.000780	0.0565
	(0.230)	(0.365)	(0.0569)	(0.0631)
Quartertoelect = 12	-0.0437	-0.406	-0.0187	0.000139
	(0.178)	(0.340)	(0.0251)	(0.00761)
Quartertoelect = 13	-0.136	-0.488	-0.00109	-0.00278
	(0.218)	(0.347)	(0.0243)	(0.0135)
Quartertoelect = 14	-0.171	-0.295	0.0351	0.00584
	(0.200)	(0.365)	(0.0602)	(0.0239)
Quartertoelect = 15	-0.319*	-0.600*	0.0655	0.00871
	(0.174)	(0.359)	(0.0633)	(0.0193)
Year FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	348	348	348	348
R-squared	0.188	0.248	0.116	0.113

Notes: Robust OLS standard errors in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). The outcome variables are the number judges that retire in a particular month (Columns 1-2) and the number judges that resign in a particular month (Columns 3-4).

Table 6 presents the results excluding exits after 1975. Here, only political cycles in resignations are observed, suggesting that retirements has only recently become a political tool.

Table 9: Political Cycles in Judicial Exits Before 1975

	(1)	(2)	(3)	(4)
	Number of Retirements		Number of Resignations	
Party in Power	Same	Different	Same	Different
Mean of dep. var.	0.032	0.027	0.015	0.007
Quartertoelect = 1	0.0157	-0.0223	-0.0129	-0.00458
	(0.0212)	(0.0186)	(0.0125)	(0.00370)
Quartertoelect = 2	0.00529	-0.00299	0.00924	0.000529
	(0.0197)	(0.0222)	(0.0202)	(0.00912)
Quartertoelect = 3	0.00167	0.00967	-0.00272	-0.00312
	(0.0184)	(0.0226)	(0.0193)	(0.00521)
Quartertoelect = 4	0.00286	0.0261	0.00518	-0.00577
	(0.0245)	(0.0348)	(0.0285)	(0.00634)
Quartertoelect = 5	-0.0279	0.0464	0.0156	-0.00260
	(0.0306)	(0.0408)	(0.0316)	(0.0110)
Quartertoelect = 6	0.0392	0.0270	-0.00108	-0.0130
	(0.0385)	(0.0389)	(0.0318)	(0.00959)
Quartertoelect = 7	0.0200	-0.0146	-0.0208	-0.00889
	(0.0362)	(0.0355)	(0.0296)	(0.00963)
Quartertoelect = 8	-0.0175	-0.00587	0.0336	-0.0115
	(0.0321)	(0.0307)	(0.0291)	(0.0121)
Quartertoelect = 9	-0.0249	-0.0163	0.0401	-0.0161
	(0.0348)	(0.0329)	(0.0287)	(0.0134)
Quartertoelect = 10	-0.0274	0.00269	0.0312	0.00401
	(0.0360)	(0.0373)	(0.0289)	(0.0189)
Quartertoelect = 11	-0.0234	-0.0309	0.0441	0.000615
	(0.0368)	(0.0323)	(0.0288)	(0.0174)
Quartertoelect = 12	-0.0377	-0.0145	0.0297*	-0.0175
	(0.0333)	(0.0237)	(0.0154)	(0.0176)
Quartertoelect = 13	-0.0181	-0.00195	0.0517***	-0.0143
	(0.0358)	(0.0266)	(0.0188)	(0.0199)
Quartertoelect = 14	-0.0286	0.0251	0.0351**	-0.00921
	(0.0377)	(0.0315)	(0.0173)	(0.0224)
Quartertoelect = 15	-0.0244	0.00679	0.0231	0.00265
	(0.0349)	(0.0264)	(0.0157)	(0.0238)
Year FE	Yes	Yes	Yes	Yes
Season FE	Yes	Yes	Yes	Yes
Observations	2073	2073	2073	2073
R-squared	0.211	0.141	0.102	0.097

Notes: Robust OLS standard errors in parentheses (* $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$). The outcome variables are the number judges that retire in a particular month (Columns 1-2) and the number judges that resign in a particular month (Columns 3-4).