Judges as Party Animals:
Retirement Timing by Federal Judges and
Party Control of Judicial Appointments

Ross M. Stolzenberg\textsuperscript{1,2}
Department of Sociology
University of Chicago
1127 East 59\textsuperscript{th} Street
Chicago, IL 60637
\texttt{r-stolzenberg@uchicago.edu}
https://orcid.org/0000-0003-3034-5600

James Lindgren
Pritzker School of Law
Northwestern University
375 East Chicago Avenue
Chicago, IL 60611
\texttt{jlindgren@law.northwestern.edu}
https://orcid.org/0000-0001-7375-5648
Abstract

Longstanding debate over the Politicized Departure Hypothesis (PDH) asserts that federal judges tend to arrange to retire under presidents of the same political party as the president who first appointed them, thereby giving that party the right to nominate their successor. By timing their departures politically, judges both contribute to the long-term political party orientation of courts and express party agency, even though judges receive no consequent personal benefit. PDH studies inevitably suffer from an absence of data on known and unknown determinants of retirement timing. To avoid these and other problems, we apply 11 sharp regression discontinuity (SRD) analyses to voluntary judicial departures before and after five elections that replace Republican presidents with Democrats, and six that replace Democrats with Republicans, 1920 to 2018. For 10 of 11 analyses, the results of difference tests, difference-in-difference tests, and others are as predicted by PDH for pre-election and post-inauguration observation periods of 270 days. These political effects appear to be stronger for Republican appointees than for Democratic appointees. We also offer a novel explanation of politicized departure based on normative reciprocity rather than simple ideology. The implications of the results from this pseudo-experimental design are considered.

1 Corresponding author.
2 For useful and constructive advice on earlier versions, we thank...
**Introduction.** Much debated hypotheses claim that, as judges appointed under Article III of the U.S. Constitution end their courtroom careers, they seek replacement by others who share their political party orientation as Republicans or Democrats (Yoon 2017; Stolzenberg and Lindgren 2010; Yoon 2006; Zorn and Van Winkle 2000; Spriggs and Wahlbeck 1995). Delivery of appointment rights is straightforward, but sometimes difficult to execute: If judges retire, resign, or accede to senior status early in the four-year term of an incumbent president, or whenever the sitting president’s party controls the Senate, that president has the right to nominate a successor judge, and time for confirmation by a cooperative Senate. By fine-tuning their dates of departure, judges can control assignment of rights to appoint their replacements, barring their sudden death, poor prediction of presidential election outcomes, or unforeseen personal exigencies (Chabot 2019; Erikson and Wlezien 2008; Campbell 2008).

In one version of this politicized departure hypothesis (hereafter, PDH), presidents seek to mold legal decisions by nominating judges who share their political ideology, values, attitudes and opinions. Conversely, when judges leave full time court service, they seek to have their replacements named by presidents who share their judicial ideology, values, attitudes or, at least, opinions. Presidents choose among many candidate judges, but judges choose only between departure under a Democratic or a Republican president, and can only trust that presidents’ ideologies, attitudes, values and opinions correlate with their political party affiliation (e.g., Chabot 2019; Stolzenberg and Lindgren 2010). We call this hypothesis the “enduring ideology” version of PDH, because it relies upon judges’ maintenance of the political and legal attitudes and values that led to their first appointment to the federal bench.

We propose a second version of PDH in which the norm of reciprocity, not enduring ideology, calls on judges to return rights of appointment to presidents of the same party as the presidents who first appointed them. The reciprocity norm is a mainstay of culture-focused social science (see Gouldner 1960 and Molm, Collett, and Schaefer 2007 in sociology; Lubell and Scholz
2001 in political science; Whatley et al. 1999 in social psychology; Kloppenberg 2016 in history; Graeber 2001 in anthropology; Fehr and Schmidt 2006 and Malmendier et al. 2014 in economics). We call this variant of PDH the “reciprocity norm” version.

Both versions of PDH assert that when age, health, family matters, occupational fatigue or anything else induces judges to end full-time judicial service, they tend to delay or accelerate their separations, thereby delivering rights to nominate their successors to presidents of the same party as the presidents who first appointed them. Neither version of PDH excludes the other.

If true, PDH is important for its far-reaching implications. Although party politics may be unavoidable for judicial aspirants, PDH suggests that judges themselves act politically, without financial or career advancement incentives, as they end their courtroom careers. If objective data indicates that judges tend to act politically at career end, then it also provides evidence that party politics also tends to influence their behavior throughout their judicial careers, and when evidence of influence is less available.

More abstractly, PDH is important because it describes a self-replicating system shaped by societal norms, supported by judges’ values, attitudes and behavior; facilitated by judicial, presidential and Senatorial organizational structure, practices and procedures; replete with influence of previous judges on the nominations of current jurists, and, barring unforeseen changes in judicial selection, full of promise that current judges will have opportunities to choose the party whose president will nominate their future replacements. If it is indeed institutionalized, as just described, then politicized departure is likely to be durable, with diffuse effects extending beyond the careers and decisions of individual judges, and past the tenures of individual presidents.

Finally, PDH is important because it implies that party politics influence American social stratification through the courts as well as through the elected officials of the legislative and executive branches of government. Competition for market advantage, indicia of social status, and political power (i.e., the entire Weberian stratification paradigm) in the United States is governed by
federal laws, and refereed by federal judges. When disputes over resources, privileges and
competition reach those courts, “Article III judges” interpret laws and admissibility of facts, instruct
juries, decide sentences, make monetary awards, sometimes reach verdicts themselves, and issue
injunctions to halt prohibited behaviors. Thus, judges are arbiters of competition and dispute
involving labor and product markets, public accommodations, schools, housing, voting rights, civil
rights, and intragovernmental conflict. If correct, PDH provides a concise explanation of previous
findings of correlation between the political parties of presidents and the decisions of judges they
appoint (see, e.g. Sunstein et al. 2006; Shepherd 2009; Kang and Shepherd 2011; Kastellec 2011;
Spitzer and Talley 2016). Thus, a proper empirical test of PDH is broadly important for theoretical
and policy-related understanding of significant social, political, economic and legal issues.

Previous studies use judicial career data to consider PDH (see Stolzenberg and Lindgren
2010 for a review through 2010; and Yoon 2017, 2006; Bailey and Yoon 2011; Zorn and Van
Winkle 2000; Spriggs and Wahlbeck 1995; Barrow and Zuk 1990; Choi 2013; Van Tassel 1993
Hansford, Savchak, and Songer 2010; and Nixon and Haskin 2011). Those studies focus on subsets
that constitute a minority of Article III judges (usually Supreme Court justices; occasionally Circuit
Courts of Appeals judges), and therefore heighten current interest in testing PDH in the entire
Article III judiciary.

New PDH tests are also motivated by labor force studies that suggest the need to control for
health, family circumstances, work attitudes, long term career plans and other career characteristics
that are difficult to ascertain for living judges and simply unavailable for most or all of the dead (see
Munnell, Sanzbacher and Rutledge 2018 on the importance of these measures in the general
population; and Greenhouse 1984 regarding difficulties in obtaining such information from judges).
These data difficulties are analogous to problems found in studies of class size effects on school
learning, and minimum wage legislation effects on demand for labor (Angrist and Lavy 1999; Card
and Krueger 1994), and we propose that the same methods used to cope with those problems in
school and employment data can be applied to tests of PDH. Those methods include regression
discontinuity methods and sharp regression discontinuity (hereafter SRD) methods. Application of
SRD to PDH is novel, but SRD is both old and now widely applied social science data (see e.g,
Thistlewaite and Campbell 1960; Holland 1986; Wasserman 2003:251; Morgan and Winship 2014;
Cunningham 2021: Chapter 6).

In short, this paper describes SRD tests of PDH for judges who were appointed under
Article III of the U.S. Constitution, and who terminated full-time judicial service from 1919, when
employment terms of these judges first approximated their current form, to 2018, when we began
this research. To further identify unanswered questions, necessary methods and appropriate data, the
next section reviews previous PDH studies. Then, we describe the SRD method that we apply to
those data. Finally, we describe findings and draw conclusions.

**Previous Research.** This section considers previous PDH research. Stolzenberg and
Lindgren (2010: Table 1) list and briefly describe some 20 previous analyses of departures from the
Supreme Court of the United States (hereafter, SCOTUS). Some of these studies examine only the
statistical distribution of SCOTUS vacancies (Wallis 1936; Callen and Leidecker 1971; Ulmer
1982). Some other judicial career research is not probative of PDH. For example, King (1987) and
Hagel (1993) combine death-in-office with retirement, resignation and partial retirement (“senior
status”), although death-in-office is an involuntary biological consequence of failure to leave the
bench before death, while senior status accession, resignation and retirement are voluntary actions
reserved for the living.

Labor force-wide studies find that the probability of voluntary employment termination
varies inversely with workers’ health or “vitality” (French 2005; Bound 1991; Dwyer and Mitchell
1999; Parsons 1982). Virtually all previous historical narrative studies of SCOTUS voluntary
terminations consider the retirement effects of declining vitality, or, in Garrow’s (2000) sensational
wording, “decrepitude.” In statistical analyses, Squire (1988) includes a measure of poor health
which is criticized by Hagle (1993: 35) and Zorn and van Winkle (2000: 162). For dead justices, Stolzenberg and Lindgren (2010) use years-left-to-live at a time before death to indicate health at that time. However, remaining lifetime is more reliable for measuring population average health than individual health. Zorn and van Winkle (2000) use justices’ written opinion production to measure physical health, but the many determinants of productivity raises questions about the validity of this measure (see Green and Baker 1991). Finally, we suggest that judges’ career and employment decisions seem likely to be less affected by actual health and future longevity than by judges’ unobservable perceptions of those things: Sick judges may refuse retirement if they think themselves healthy; healthy people may be more likely to retire if they think themselves ill. Moreover, Hagle (1993:46) asserts that SCOTUS justices are flagrantly dishonest and willfully misleading about their health. Thus, controlling for health in judicial career studies requires methods that do not require direct health measurement or candid self-reporting by judges. We return to this issue below, after describing previous efforts to distinguish effects of political party identity, which would be central to the reciprocity norm version of PDH, from political ideology effects, which would be central to the enduring ideology version of PDH.

“Conceptually, differences between ideology and party are stark, because parties are organizations of people, and ideologies are complexes of values, attitudes, ideas and perceptions (see J.L Martin 2015). Conceptual differences notwithstanding, empirical observations of ideologies and party affiliations of individuals can be correlated empirically, even to the point that effects of one are difficult or impossible to distinguish from effects of the other. In the general population, party identification and ideology of individuals are regularly measured by survey questions. For judges, party identification is conveniently defined and observed as the party of the president who first appointed them to the Article III bench. But judicial custom and ethics makes measurement of ideology has been more involved. Pinello (1999: 219) reviews and exhaustively meta-analyzes 84 prior studies, then concludes, “party is a dependable yardstick for ideology.” Thus, Pinello implies
that ideology and reciprocity versions of PDH are empirically indistinguishable, even if their conceptual dissimilarity suggests otherwise.

Judicial ideology measurement has grown considerably since Pinello (1999). Martin and Quinn (2002) (hereafter MQ) show that ideology can be measured without reference to party, by Item Response Theory (IRT) scaling of SCOTUS justices’ votes in court decisions. In a computational tour de force, MQ calculate annual ideal point IRT ideology scores for SCOTUS justices, starting in 1937, based on voting in case decisions. Whatever their advantages, MQ methods cannot be applied to judges of district courts, because they do not cast votes on panels, as do SCOTUS and appellate court judges. Judicial Common Space (JCS) scales combine MQ scores with other data for SCOTUS justices. For judges of Circuit Courts of Appeals, the JCS scale confounds party and ideology, which is inferred from the political parties of the appointing president and senators from the judge’s home state. In a recent, novel, indirect measurement strategy applied to judges of all Article III courts, Bonica et al (2019) use political donations of money by law clerks of Article III judges to indicate political ideologies of the judges for whom they work.

In short, techniques for measuring judicial ideology have developed considerably since Pinello’s analysis. So one can no longer rely on the implication of his study that empirical measures of party identity and judicial ideology are not different enough to allow separate tests of the ideological endurance and reciprocity norm versions of PDH. To make the necessary update, we examine the empirical congruence of party and ideology measures by principal components factor analysis of data on the 31 SCOTUS justices who served at any time from 1960 to 2018. We focus on SCOTUS justices because they are the only judges for whom there exist ideology measures that are not at least partially based on party identity – i.e. MQ scores. We focus on 1960 to 2018 to include other ideology scores that are available only after 1960. We end observations in 2018, because that is the year we began research reported here. Factor analyzed variables include lifetime
averages of Bailey, MQ and JCS scores, plus Rep (= 1 for justices first appointed the Article III
courts by a Republican president; = 0 else). Data and analysis details are given in an appendix.

Using data just described, principal components factor analysis finds only one factor with an
eigenvalue greater than 1, and it explains 99.37 percent of the variance among Rep, Bailey, MQ and
JCS scores. Factor loadings are all above .65 and average .90. Although the small N for the
analysis, and its restriction to SCOTUS justices from 1960 to 1988, calls for restraint, findings are
bolstered by their consistency with Pinello’s summary of previous studies. In short, in spite of new
methods and resurgent interest in distinguishing ideology effects on PDH from party identity
effects, the factor analysis suggests that for SCOTUS justices, party identity and the ideology scales
analyzed here are all indicators of the same underlying factor.\textsuperscript{v \textsuperscript{vi}}

For the present purpose of testing PDH in the entire Article III judiciary, the implications of
past research, and the factor analysis results just presented can be summarized briefly in three
points:

\textbf{First}, a disproportionate share of prior PDH research focuses on SCOTUS justices, who are
a small segment of the Article III judiciary. Thus, testing PDH in the entire Article III judiciary
remains an important task, and is the focus of analyses presented here.

\textbf{Second}, PDH can be hypothesized both as an effect of judges’ ideology and as an effect of
their political party identity. Conceptual differences between political ideologies and political
parties are clear, but factor analysis finds that SCOTUS ideology and party measures are indicators
of the same underlying latent factor. For lower court judges, the ideology measures are computed
from party identity variables. Consequently, our analyses of politicized departure focus on the role
of party identity, which is readily and reliably available for all judges. However, without controlling
for ideology, effects of party identity on politicized departure would be a mixture of both ideology
and political identity effects.
Third, consideration of career and retirement studies in the general labor force suggests that unobservable personal characteristics and circumstances of judges affect their ability to adjust the timing of their retirements and resignations from full time judicial service. The next section describes a strategy to hold unobservable characteristics of judges constant, and to test PDH for the entire Article III judiciary.

Analytic Strategy. We re-state the PDH as follows: When judges are ready to end their full-time federal judicial service, those who were first appointed by a Republican president are more likely to end full-time service when the incumbent president is a Republican than when the president is a Democrat, all else equal. Similarly, when Democratic appointees decide to end their full-time judicial service, they are more likely to do so when the incumbent president is a Democrat, all else equal. PDH is an hypothesis that when judges retire or resign, they tend to adjust the timing of their retirements so that the president who nominates their successors is of the same party as the president who first appointed them to the Article III bench.

Our strategy is to test these hypotheses by selecting pairs of time periods in which all determinants of termination probability, except the political party of the incumbent president, may be regarded as identical, or nearly so, for every judge who terminates full-time service in either period. If pre-election and post-inauguration periods are adjacent and sufficiently short, judges’ attitudes, values, health, family characteristics, finances and other retirement-related characteristics can be considered to be the same in both periods, leaving the political party of the sitting president as the only retirement-related characteristic that changes with the inauguration of a new president. Consequently, any difference between the termination probability after inauguration and the probability before the election is attributed to the change in presidential party.

Circumstances just described occur naturally but irregularly, shortly before “regime-changing” elections (here defined as elections and inaugurations that replace Democratic presidents with Republicans, or vice versa) and after the inaugurations that follow them. For example, consider
the 270 days (about 9 months) before the presidential election of 2008 and the equal period after the
inauguration in 2009. The 2008 pre-election president was Republican; the 2009 post-inauguration
president was Democratic. We assume that retirement-related characteristics of judges do not
differ meaningfully between adjacent pre-election and post-inauguration periods. If this assumption
is tenable, then the average treatment effect of a Democratic president on departures from full time
judicial service of Democratically-appointed judges is the difference between the proportion of
Democratically-appointed judges who retire in the 2009 post-inauguration period and the proportion
of Democratically appointed judges who retire in the 2008 pre-election period. The PDH hypothesis
can be expressed as a positive after-before difference in the number of terminations, a positive after-
before difference in the rate of terminations, an after/before ratio greater than one, or an after/before
odds-ratio greater than one, depending on statistical preferences.

Regime changing elections and inaugurations occur 11 times from 1920 to 2017 (i.e. elections of 1920, 1932, 1952, 1960, 1968, 1976, 1980, 1992, 2000, 2008, and 2016). By starting these analyses in 1920, we evade statistical consequences of a small judiciary in earlier years (the entire Article III judiciary does not exceed 200 active duty judges consistently until 1919), and escape problems of comparing terminations of full-time judicial service before and after the 1919 modifications of Article III judicial employment regulations, which create the option of Senior Service for long-serving, sub-SCOTUS judges. Accession to Senior Service facilitates terminations from full time service by permitting judges a reduced caseload, or no case load at all, without loss of honorific status, income or other perquisites.

As an additional control for confounding and spuriousness due to unobserved variables, we also calculate the same after-before voluntary termination probability difference for judges first appointed by a president of the same party as the presidential election loser, and subtract it from the difference obtained from judges appointed by presidents of the same party as the election winner. This is the “difference-in-differences” (“diff in diff” or, hereafter, DiD) statistic. Again, depending
on statistical preferences, DiD can be expressed as a difference between rates, a ratio or an odds ratio. PDH predicts a positive value for DiD based on differences between rates, or ratios greater than unity, if DiD is based on ratios and odds ratios.

To observe and control effects of historical peculiarities such as time elapsed between regime-changing elections, or the political balance of the Senate, we replicate analyses at each of the 11 regime-changing presidential elections from 1920 to 2016. For example, Eisenhower’s 1952 election was the first regime-changing election after 1932. Perhaps World War II, the Great Depression, or the unusually long, 20-year interval between these regime changes altered career dynamics for politically-influenced federal judges during F.D. Roosevelt’s presidential tenure. Similarly, to control for possible confounding by political party of the incumbent president, we stratify analyses by the party of the winner of the regime-changing election – six Republican and five Democratic regime-changing victories from 1920 to 2016.

We perform all analyses separately for pre-election and post-enumeration periods of 180, 270, 365, 547 and 730 days, or approximately 6, 9, 12, 18 and 24 months before the regime-changing election, and after the subsequent inauguration. Thus, we stratify analyses by length of the pre-election and post-inauguration enumeration periods, to determine if the treatment effect is strongest at the beginnings of presidential terms in office, when incumbent presidents tend to be most popular, have their greatest Senate support, and the maximum time available to negotiate Senate confirmation of nominees.

Finally, we emphasize that the hypothesized presidential party effect on judicial full-time service departures is probabilistic and incomplete (thus neither necessary nor sufficient). For example, judges’ voluntary terminations from full time judicial employment may coincide randomly with White House occupancy by presidents of the same party as the presidents who first appointed them to the federal bench, or fail to coincide despite effort by judges to arrange the contrary. Also, judges’ desires to comply with norms of reciprocity and enduring ideology may be
overwhelmed by their inaccurate predictions of future presidential election outcomes, or by unexpected personal exigencies. As Justice Ginsburg illustrates, inaccurate election predictions and personal exigencies can defeat intentions for politicized departure, thereby reducing the number of politicized departures, biasing Diff and other measures downward, and thereby making tests of PDH more stringent than their significance levels imply. Good luck and accurate predictions neither compel nor motivate politicized departure, and so do not affect tests of it described here.

Research Design And Data. The process just described appears to be a previously unnoticed, naturally occurring example of the sharp regression discontinuity (SRD) research design, with 11 replications (Cattaneo and Vazquez-Bare 2016; Lee and Lemieux 2010; Imbens and Lemieux 2008: 217-19; Thistlewaite and Campbell 1960). The hallmark of SRD is abrupt, exogenous change in the state or value of a treatment.\textsuperscript{viii} We now describe the design of this research in the language of experimentation, focusing on subjects, outcomes and treatments.

Subjects. The units of analysis – the subjects – in analyses presented here are persons who were employed full-time as Article III federal judges for at least 730 days (about two years) prior to a regime-changing presidential election between 1920 and 2016.\textsuperscript{ix} For brevity, we call retirements, resignations and accessions to senior status \textit{“trigger actions,”} because they trigger new presidential nominations to the bench. Prior service of at least 730 days excludes judges who lack a minimal claim to a federal judicial career, rather than a recent posting to a new job. Requiring a year of post-inaugural life avoids the need to distinguish those who take a trigger action in that period from those who might have done so, had they endured. Judges are, or would be, excluded from analysis if they leave office involuntarily due to death, abolition of their appointed court, or Congressional impeachment and conviction.

Treatment. Treatment occurs during enumeration periods shortly before regime changing elections, and shortly after inaugurations that follow them. For each judge, treatment consists of changing the party of the incumbent president from different from, to the same as, the political party
of the president who first appointed them to the federal judiciary. Characteristics of judges are assumed to not change meaningfully from the start of the pre-election enumeration period to the end of the post-inauguration period. These characteristics include judges’ perceptions of their own health, personal finances, job satisfaction, desire to retire, and similar.

**Outcomes.** For any of the 11 regime-changing elections considered here, three outcomes are possible: Judges can take no trigger action; they can take a trigger action in the pre-election period; or they can take a trigger action in the post-inauguration period.

**Effect Measures.** PDH predicts that, if treated judges terminate full-time service about the time of a regime changing election, they are more likely to do so post-inauguration than pre-election. Thus, for any particular regime changing election, the treatment effect is the difference between the number of treated judges who terminate full-time service in the post-inauguration and the number of treated judges who terminate full-time service in the pre-election period. Growth of the federal judiciary from 1920 to 2018 would affect these numbers, so results are also expressed as proportions, odds and odds ratios, per common statistical practice (Agresti 1990). Counts and proportions can be recovered from n’s, odds and odds ratios.

**An Example.** To illustrate and clarify the analysis plan, Figure 1 provides a schematic diagram of the analysis design, the hypotheses it tests, and treatment effect measures for a single election-inauguration (2008-2009; won by the Democratic candidate) and enumeration periods of 270 days before election and 270 days after inauguration. Symbols and terms are defined in Table 1.

Row labels on the left side of Figure 1 distinguish untreated (Republican) appointees in the bottom row from treated (Democratic) appointees above them. Across the top, column labels distinguish pre-election periods on the left from post-inauguration periods to the right. Judges in Group A were first appointed by Democratic presidents. After the election, those same Democratic appointee
judges appear in “Group B.” Hypothesis 1 asserts that the number of trigger actions by judges in Group B after the election ($\sum^{d}Y_{aRD}$) exceeds the number of trigger actions by those very same judges before the election ($\sum^{d}Y_{bRD}$) when they constitute Group A. Without loss of information, the numbers of triggers in Group A and Group B can be divided by the number of Democratic appointee judges $N_{d}$ to obtain proportions, and the hypothesis becomes $H_{A}$: $(\sum^{d}Y_{aRD}/N_{d}) - (\sum^{d}Y_{bRD}/N_{d}) > 0$. Re-scaling proportions to odds and comparing them by division instead of subtraction yields the odds ratio, $\text{Diff} = \frac{O_{aRD}}{O_{bRD}}$, and the hypothesis becomes $H_{A}$: $\text{Diff} > 1$, where super- and sub-scripts retain their meaning as earlier defined, $O$ replaces $Y$ to indicate the odds of a trigger action rather than a count of trigger actions, and $\text{Diff}$ is defined as written here.

We also compute Difference-in-Differences (hereafter, DiD) which is the ratio of $\text{Diff}$ for judges appointed by presidents of the same party as the winner of the most recent presidential election to the same ratio for judges appointed by presidents of the same party as the loser of the most recent presidential election. DiD controls for the possibility that some unrecognized agent has appeared in the form of a secular trend or a random shock to increase trigger actions after inauguration by all judges, regardless of the party of the president who first appointed them to the federal bench.

Further, we consider a measure we call Directional Diff in Diff (hereafter, DDD), which compares $\text{Diff}$ to the end-of-term odds ratio measure of the effect of the pre-election president’s political party on terminations by judges first appointed by presidents of that party. DDD is useful in addressing the secondary hypothesis that political influence on trigger action timing declines as the presidential term of office approaches expiration.

For the 2008 election and 2009 inauguration shown in Figure 1, there are 755 judges appointed by Republican presidents and 499 appointed by Democrats. Thirteen Republican appointees and 12 Democratic appointees take trigger actions in the 270 days preceding the 2008 presidential election. In the 270 days following the 2009 inauguration, 15 Republican appointees
and 26 Democratic appointees take trigger actions. Odds and odds ratios are computed with the usual continuity correction of 0.5 (Agresti 1990:68), yielding the following results:

1. The odds ratio, Diff, equals 2.18, indicating that,

   
   as the political influence hypothesis predicts, the odds that Democratic appointees take a trigger action in the post-inauguration period are more than twice the odds that they do so in the pre-election period.

2. The value of DID, the ratio of Diff for Democratic appointees to the same odds ratio for Republican-appointed judges in the same period, is 1.90, indicating that,

   
   even if a secular trend or aberrant influence increased post-inauguration departures from full-time judging, the increased odds ratio for Democratic appointees predicted by the political influence hypothesis remains almost twice (1.90 times) the size of the odds ratio for Republican appointees.

3. Finally, we obtain a value of 2.51 for DDD, indicating that

   the boost in odds of trigger actions by Democratic appointees during the first 270 days of this regime-changing Democratic presidency is about two and one half times as large as the disparity between Republican appointee odds of trigger action during the last 270 days before the election, when the president was Republican, and Republican appointee odds of trigger action during the 270 days after the inauguration. This result for DDD is consistent with the hypothesis that political influence effects decline as the end of the presidential term in office approaches.

Identification of effect measures in these analyses is explicated formally by Hahn et al (2001); see also Imbens and Lemieux (2009: 217-19); Lee and Lemieux (2010); Cattaneo and Escanciano (2017); and Cattaneo et al. (2017). Informally, identification is apparent from several design features of this research. First, there is no self-selection for treatment: assignment to control and treatment groups is determined by the outcome of a presidential election, and therefore beyond control by any individual judge. Second, temporal ordering and close conjunction of treatment and outcome are assured by strictly-defined periods in which the outcome is measured and the treatment
is either entirely present or completely absent. If unspecified individual characteristics of subjects affect outcomes, their effects are cancelled by division in calculation of Diff. And, third, effects are measured by comparisons of treated individuals to themselves when not treated, thereby permitting an assumption that unobserved characteristics of treated and untreated subjects do not differ. Formally, this last comparison is stratification on retirement/resignation/accession to senior status (retirement): everyone in the analysis is leaving full-time judging during an interval that straddles an election and inauguration. The estimand of interest compares the is the ratio of the probability of actual departure during the term of the outgoing president. As described famously by Frangakis and Rubin (2002), this stratification on retirement renders retirement invariant in the analyses and therefore without effect on the estimand (DIFF), obviating any need to specify an instrument for retirement. For a comparison to instrumental variables estimation, see

*Replication and Stratification.* We apply the regression discontinuity design method just explicated to federal judicial trigger actions immediately before and after each of the 11 regime-changing presidential elections between 1920 to 2016, using data from 1919 through 2018, when we began preparation of data for these and other analyses. Because 6 of those regime-changing elections were won by Republicans, and 5 were won by Democrats, the replication also stratifies the analysis by the party of the presidential election winner.

*Significance tests.* We perform separate, disjoint tests of PDH, one for each regime changing presidential election from 1920 to 2018. Absent any PDH effect, and other things equal, probabilities of retirement before and after the election would be equal so that $d_{Y_{aRD}} = d_{Y_{bRD}}$. Following Agresti (1990: 352), the null hypothesis of no presidential party effect on voluntary terminations is

$$H_0: \text{difference} = d_{Y_{aRD}} - d_{Y_{bRD}} = 0,$$

and $d_{Y_{aRD}} - d_{Y_{bRD}} > 0$ is distributed as Bernoulli (binomial) trials with $p=0.5$ and $n=11$. The probability of 8 or more successes is 0.113, which is the test significance level. For 9, 10 or 11
successes, significance levels are .033, .006 and .0005 respectively. In 6 analyses of Republican
appointees, probabilities of 5 or more, or 4 or more successes are .109 and .344 respectively. For
n=5 analyses of Democratic appointees, the probability of 4 or more successes is .188, and the
probability of 3 or more is .500. These tests do not address compound null hypotheses.

Data

Primary data examined here were produced by extensive checks, corrections and re-codes of
data downloaded from the Federal Judicial Center (n.d.a.) on April 28, 2018. Most corrections are
based on consistency checking and comparison with records and on-line biographies from the the
Federal Judicial Center (n.d.b.), resulting in a file of 86,316 judge-year records for all 3,516
individuals who were nominated by presidents to Article III judicial positions, confirmed by the
Senate and commissioned in office, from 1789 to April 2018.

11 Results for Diff in 9-Month Observation Periods. Table 2 reports values of Diff in
column (3) for analyses in which pre-election and post-inauguration periods are both 270 days, for
all regime-changing elections from 1920 through 2016.

Per Column (3), Diff exceeds one in 10 of 11 analyses, and is consistent with the first hypothesis at
a significance level of 0.0059. Consistent with PDH, the mean of Diff is 3.23: on average, the odds
of trigger action in the post-inaugural period is 3.23 times the odds of a trigger action in the
immediately preceding pre-election period.

Figure 2 plots Diff for 270-day enumeration periods, from 1919 to 2018, with a line fitted by
Cleveland’s (1979) “robust locally weighted regression” method. The main finding, by inspection
of the solid line in Figure 3, as from Column (3) of Table 2, is that temporal variation in Diff
reflects atypically large values at the elections of 1952 and 1960, and is consistent with PDH.
**DiD Results for 9-month observation periods.** Consistent with PDH, Column (5) of Table 2 shows the mean of DiD as 4.19. So, on average, Diff is 4.19 times as large for judges first appointed by presidents of the same party as the newly-inaugurated president (concordant party judges) as for those first appointed by presidents of the other party (discordant party judges). Consistent with PDH, DiD exceeds one in 10 of 11 analyses, for a binomial test significance level of .00059.

**DDD results for 9-month observation periods.** Directional Diff in Diff (DDD) compares beginning-of-presidential term PDH effects to end-of-term PDH effects. The mean of DDD in column (7) of Table 2 is 3.30, indicating that the impact of party concordance is more than three times as large at the start of a president’s term as at the end. DDD exceeds unity in 8 of 11 election-inauguration sequences, with a significance level of 0.113.

---Insert Table 3 about here---

**Party Differences.** Rows 2 and 3 of Table 3 compare values of Diff, DiD and DDD for all 11 regime-changing presidential election-inauguration sequences from 1920 to 2018, separately for the 6 elections won by Republicans, and the 5 elections won by Democrats. At every observation period length, Diff is larger on average when Republicans win than when Democrats win. Indeed, for 14 of these 15 comparisons of row (2) to row (3) of Table 3, the average values of Diff, DiD and DDD obtained under Republican presidents exceeds the average value obtained under Democrats. These results are consistent with the claim that exit timing of Republican appointees is more influenced by the political party of the newly-elected president than exit timing of Democratic appointees. We know of no previously-published hypotheses of party differences in PDH effects, so we only note them, and wait for future research to properly test for and explain their existence.

**Enumeration Period Length Effects.** Results presented so far pertain to 270-day observation periods (about 9 months) before regime-changing elections and after regime-changing inaugurations. Row 1 of Table 3 summarizes results for periods of 180, 270, 365, 547 and 730 days – about 6, 9, 12, 18, and 24 months – for Diff, DiD, and DDD.\(^\text{xi}\) As observation periods lengthen,
Table 3 shows that average values of Diff and DDD decline strictly monotonically. DiD declines similarly, although its value in 1-year observation periods is larger than for the 9-month periods. These patterns are consistent with the assertion that judges who wish to leave full-time service honor principles of enduring ideology or party reciprocity, but only up to a point. That point seems to be based on the time they must linger in full time jobs they wish to leave.

**Amalgamated Results.** Table 4 re-tabulates voluntary terminations in 270-day enumeration periods, by concordance of the party of the presidential election winner with the party of the appointing president, for all 11 regime-changing election-inauguration periods from 1920 to 2017.

Although not a proper test of PDH, Table 4 is consistent with it: 225 judges appointed by presidents of the same party as the recently elected president resigned or took senior status in these enumeration periods, triggering new presidential appointments. Of these, 36.0 percent did so in the pre-inauguration period, and, consistent with PDH, 1.8 times as many (64.0 percent) did so in the post-election interval – a difference of 28.0 percent. For judges appointed by presidents of the election-losing party, the corresponding difference is -7.6 percent, and the difference between these differences is 35.6 percent, which is all consistent with PDH.

**Discussion and Conclusion.** This paper considers the politicized departure hypothesis, a venerable but still controversial assertion that as Article III judges approach the ends of their careers, they tend to adjust the timing of their departures so that the rights to name their replacements are given to presidents of the same political party as the president who first appointed them to the federal bench. Previous research on politicized departure is abundant, but questions remain, we think, because, first, previous research gives little attention to judges of courts below the Circuit Courts of Appeal, and, second, because judicial ethics and custom discourage judges from providing information about their health, family circumstances, job attitudes, work
satisfaction, and similar things that have been shown to affect voluntary job termination and retirement in the general population.

To escape the problems of unmeasured and unknown variables, and to expand coverage to all Article III judges, we apply sharp regression discontinuity methods, with and without the “difference in differences” estimator to the entire Article III judiciary. To apply SRD, we seek situations in which the political party of the sitting U.S. president changes abruptly over a span of time that is too short for retirement-related characteristics of judges to change much, if at all. We observe that such situations occur repeatedly, shortly before regime changing presidential elections and shortly after the presidential inaugurations that follow them. Our application of regression discontinuity methods to the politicized departure hypothesis appears to be novel, but neither regression discontinuity research nor potential outcomes methods are new (see Thistlethwaite and Campbell, 1960; Holland 1986; Haavelmo 1943, 1944). As we compare periods just before regime changing elections to periods of equal length immediately after those elections, we find, consistent the politicized departure hypothesis, that Article III judges are more likely to retire when their party’s candidate wins the election and sits in the White House, than in the pre-election period, when the president is of the other party.

SRD, like other potential outcomes research designs, gains much of its power by a strategy that is characteristic of scientific experiments, and very uncharacteristic of survey research: it focuses on times and conditions in which treatment effects are apparent – even if those circumstances are atypical – and ignores other circumstances altogether. When regime changing presidential elections occur, the politicized departure hypothesis predicts more retirements in the post-inauguration period than in the pre-election period before it, for judges who were first appointed by presidents of the same party as the recently elected president. We report that difference as Diff, as well as a difference-in-differences (DiD) estimator, and related quantities. This SRD pseudo experiment is replicated 11 times between 1920 and 2018. For pre-election and
post-inauguration observation periods of 270 days, we find values of Diff and DiD that are consistent with PDH in 10 of these 11 replications. Treating these 11 analyses as binomial trials leads to rejection of the null hypothesis of no PDH effects. Less formally, results lend credence to the PDH.

The clarity of SRD is valuable, but not costless. In particular, in the 98 years from 1920 to 2018, there have been 25 elections, of which less than half (11) are regime-changing and suitable for the regression discontinuity method that we apply. Similarly, data and method used here do not allow much partitioning of judges into subsets based on organizational, demographic or political characteristics, so little can be said about, for example, differences or nondifferences between SCOTUS justices, judges of the Circuit Courts of Appeals, and judges of district courts. Potential outcomes analyses specific to the SCOTUS and the Circuit Courts of Appeals would require methods more suited to small n’s than those we apply here. It appears that some invention would be needed to create those methods.

Although we did not hypothesize party differences before undertaking this research, we observe stronger average gross PDH effects for Republican appointees than for Democratic appointees. These effects and differences are gross, rather than adjusted, insofar as results for Republican and Democratic appointees are measured at different times, and therefore, perhaps under different conditions. Like any results not hypothesized in advance of their detection, these differences are harder to distinguish from statistical noise than if they were predicted a priori. To wit, one could as easily conjure a post hoc expectation of this finding as its opposite, or a finding of no difference at all. For that reason, examination of party differences might require a different method or different data than we use here. For example, it may be worthwhile for future research to consider the hypothesis that Republican presidents are more likely than Democrats to appoint party stalwarts, such as those who have run for public office as party candidates. Or one might hypothesize that party differences in this judicial behavior are the result of party differences in
grooming and systematic persuasion after judges take office. To wit, Teles (2008: chapter 1) offers a model of judicial influence in which presidential nomination is a mere first step in a diffuse, ongoing, career-long and fully institutionalized pattern of effort by ideologues and commercial interests to influence the perceptions and decisions of federal judges.

Finally, it seems important to revisit the fundamental question that motivates tests of PDH: Are Article III judges influenced by politics while in office? The politicized beginnings of Article III judicial careers are apparent from the nominations of these judges by the politicians who serve as elected presidents, and their confirmations by politicians who serve as elected Senators. But PDH suggests that judges themselves tend to behave politically at the final moment of their full-time courtroom careers, without discernible incentives, financial or otherwise, long after their confirmation hearings. If it is apparent that judicial careers are politically vetted at their start, and if sharp regression discontinuity analysis of objective data indicates that judges tend to act politically at career end, then we think there is reason to believe that politics has been an active influence on many of these judges in the interim.
Figure 1 – Simplified Nonparametric Regression Discontinuity Design for Analyses of Judicial Trigger Actions, when Republican President is incumbent before election and Democrat is inaugurated after election

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Time Period Before Election and After Inauguration (2008 Election, 2009 Inauguration, 270 day Enumeration Period)</th>
<th>Tests of Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-Election Year (270 days before Election)</td>
<td>Election-to-Inauguration Interval</td>
</tr>
<tr>
<td>Treated Judges</td>
<td><strong>Group A</strong></td>
<td><strong>Group B</strong></td>
</tr>
<tr>
<td><strong>Democratic Appointees</strong> (i.e. Nominated by Democratic Presidents)</td>
<td>( \sum_{d}^{Y_{bRD}} = ) Trigger Actions by Democratic appointees in 270 days before election, when president before election is Republican and election winner is Democrat</td>
<td>( \sum_{d}^{Y_{aRD}} = ) Trigger Actions by Democratic appointees in 270 days after inauguration, when president before election is Republican and election winner is Democrat</td>
</tr>
<tr>
<td>Not Treated Judges</td>
<td><strong>Group C</strong></td>
<td><strong>Group D</strong></td>
</tr>
<tr>
<td><strong>Republican Appointees</strong> (i.e. Nominated by Republican Presidents)</td>
<td>Not Treated (Hypothesis 2)</td>
<td>Treated (Hypothesis 3)</td>
</tr>
</tbody>
</table>

Notes:
1. Outcomes are “Trigger Actions:” retirements, resignations, and senior service accessions that trigger new appointments.
2. Excluded are a) any who die before taking a trigger action before the end of the 270 days post inauguration period; b) any who are removed from office for “bad behavior,” or whose departure occurs after the election but before inauguration of the new president.
3. This research design illustrates the case when party of president prior to election is Republican, and the Democratic candidate wins election.
4. If the incumbent before the election is Democratic and the election winner is Republican, then the top row of this table would pertain to Republican appointees, rather than Democratic appointees shown, and the bottom row would pertain to Democratic appointees, rather than Republican appointees as shown.
Figure 2 – Diff, DiD and DDD, by Election Year, for 270 Day Enumeration Periods, with Values Smoothed by Cleveland’s Robust Locally Weighted Regression (Bandwidth=0.9)
Table 1 – Symbols and Definitions

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( Y_i )</td>
<td>The outcome for the ( i )th subject. ( Y_i = 1 ) if the subject takes a trigger action, and ( Y_i = 0 ) else</td>
</tr>
<tr>
<td>( dY_{bRD} )</td>
<td>Count of outcomes for subjects in a specified subgroup of subjects. Pre-superscript ( d ) or ( r ) indicates party (Democratic or Republican, respectively) of president who nominated judge to the federal bench. First post-subscript indicates if ( Y ) is measured before ( (b) ) or after ( (a) ) election. Post-subscripts ( R ) and ( D ) indicate parties of presidents before election and after election, in that order: RD indicates Republican president before election and Democratic president after election. DR indicates Democratic president before election, followed by Republican. Regression discontinuity occurs only when a presidential election changes the party of the incumbent president, so RR and DD would not occur in analyzed data.</td>
</tr>
<tr>
<td>( \tau )</td>
<td>The estimand for testing the first hypothesis is the average treatment effect at the point of discontinuity: ( \tau = E[Y_i(1) - Y_i(0)</td>
</tr>
</tbody>
</table>
### Table 2 – Analyses of Trigger Actions 270 days before 11 Regime-changing Presidential Elections and 270 days after Subsequent Inaugurations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1920</td>
<td>Republican</td>
<td>3.04 (+)</td>
<td>9.28 (+)</td>
<td>5.00 (+)</td>
<td>+</td>
<td>69</td>
<td>60</td>
</tr>
<tr>
<td>1932</td>
<td>Democratic</td>
<td>3.07 (+)</td>
<td>1.83 (+)</td>
<td>5.15 (+)</td>
<td>+</td>
<td>174</td>
<td>41</td>
</tr>
<tr>
<td>1952</td>
<td>Republican</td>
<td>9.53 (+)</td>
<td>13.87 (+)</td>
<td>6.54 (+)</td>
<td>+</td>
<td>72</td>
<td>246</td>
</tr>
<tr>
<td>1960</td>
<td>Democratic</td>
<td>5.71 (+)</td>
<td>3.40 (+)</td>
<td>9.57 (+)</td>
<td>+</td>
<td>160</td>
<td>205</td>
</tr>
<tr>
<td>1968</td>
<td>Republican</td>
<td>2.20 (+)</td>
<td>3.20 (+)</td>
<td>1.52 (+)</td>
<td>+</td>
<td>152</td>
<td>346</td>
</tr>
<tr>
<td>1976</td>
<td>Democratic</td>
<td>2.46 (+)</td>
<td>6.47 (+)</td>
<td>0.93 (+)</td>
<td>+</td>
<td>330</td>
<td>304</td>
</tr>
<tr>
<td>1980</td>
<td>Republican</td>
<td>3.37 (+)</td>
<td>2.41 (+)</td>
<td>4.71 (+)</td>
<td>+</td>
<td>331</td>
<td>314</td>
</tr>
<tr>
<td>1992</td>
<td>Democratic</td>
<td>0.47 (+)</td>
<td>0.54 (+)</td>
<td>0.42 (+)</td>
<td>+</td>
<td>620</td>
<td>371</td>
</tr>
<tr>
<td>2000</td>
<td>Republican</td>
<td>1.63 (+)</td>
<td>1.19 (+)</td>
<td>2.23 (+)</td>
<td>+</td>
<td>626</td>
<td>539</td>
</tr>
<tr>
<td>2008</td>
<td>Democratic</td>
<td>2.18 (+)</td>
<td>1.90 (+)</td>
<td>2.51 (+)</td>
<td>+</td>
<td>755</td>
<td>499</td>
</tr>
<tr>
<td>2016</td>
<td>Republican</td>
<td>1.82 (+)</td>
<td>2.00 (+)</td>
<td>1.66 (+)</td>
<td>+</td>
<td>667</td>
<td>676</td>
</tr>
<tr>
<td><strong>Mean</strong></td>
<td></td>
<td><strong>3.23</strong></td>
<td><strong>4.19</strong></td>
<td><strong>3.30</strong></td>
<td><strong>359.6</strong></td>
<td><strong>327.4</strong></td>
<td><strong>8</strong></td>
</tr>
</tbody>
</table>

**Source:** Computed by author from data downloaded from the Federal Judicial Center (n.d.a.) on April 28, 2018, and subsequently corrected by author.

**Notes:**
- In columns (4), (6) and (8), “+“ indicates that the odds ratio is greater than 1; a blank indicates that the relevant odds ratio does not exceed 1.
- In Columns (3) (5) and (7), “Winner’s Odds” are the odds of a trigger action by judges first appointed by a president of the same party as the presidential election winner.
Table 3 – Results of 55 Election- and Inauguration-Specific Analyses of Trigger Actions, by Parties of Appointing President and Election Winner

<table>
<thead>
<tr>
<th>Election/Inauguration Years and Aggregation Method</th>
<th>Diff Odds Ratio by Length of Pre-Election and Post-Inauguration Periods: Mean of Diff and (Number of Analyses for Which Diff &gt; 1)</th>
<th>Diff in Diff (DiD) Odds Ratio by Length of Pre-Election and Post-Inauguration Periods: Mean of DiD and (Number of Analyses for Which DiD &gt; 1)</th>
<th>Directional Diff in Diff (DDD) Odds Ratio by Length of Pre-Election and Post-Inauguration Periods: Mean of DDD and (Number of Analyses for Which DDD &gt; 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>180 days</td>
<td>270 days</td>
<td>365 days</td>
</tr>
<tr>
<td>(2) 6 Presidential Regime-changing Elections/Inaugurations won by Republicans 1920-2016</td>
<td>5.88</td>
<td>3.60</td>
<td>3.52</td>
</tr>
<tr>
<td>(3) 5 Presidential Regime-changing Elections/Inaugurations won by Democrats 1920-2016</td>
<td>2.16</td>
<td>2.78</td>
<td>2.64</td>
</tr>
</tbody>
</table>

Note: 547 and 730 Day Post-Inauguration Enumeration Periods for 2016 Election are Truncated to 536 days
Table 4 – Consolidated Counts of Trigger Actions, 1920-2018, by Concordance of Party of Appointing President and Party of Election Winner, 270 days before and after 11 regime changing elections

<table>
<thead>
<tr>
<th>Party of Election Winner and Party of Appointing President</th>
<th>When Trigger Action is Taken</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-election</td>
<td>Post-inauguration</td>
</tr>
<tr>
<td>(1) Judges Appointed by president of same party as election victor</td>
<td>%</td>
<td>36.0%</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>81</td>
</tr>
<tr>
<td>(2) Judges Appointed by president of same party as election loser</td>
<td>%</td>
<td>53.8%</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>78</td>
</tr>
<tr>
<td>Total</td>
<td>n</td>
<td>159</td>
</tr>
</tbody>
</table>

Notes: The number of judges varies from 1920 to 2017. See text for explanation of use of counts in this table rather than odds, probabilities and proportions.
REFERENCES


*An Act to Establish a Uniform Time for Holding Elections for Electors of President and Vice President in all the States of the Union*, U.S. Statutes at Large, 28th Congress, 2nd Sess., p. 721 (Jan. 23, 1845).


Martin, John Levi. 2015. “What is ideology?” *Sociologia, Problemas e Práticas* [Online], 77 | 2015, Online since 06 April 2015, connection on 19 April 2019. URL: http://journals.openedition.org/spp/1782


Wasserman, Larry 2003 *All of Statistics*. Heidelberg: Springer.


doi.org/10.1207/S15324834BASP2103_8.

Wright, Erik Olin 1997 *Classes*. London [u.a.]: Verso.


Appendix
Factor Analysis of Ideology and Political Measures for 31 Supreme Court justices, 1960-2018

I. Analysis Results

Table A1 -- Pearson and Point Biserial Correlations Among Ideology Scores and Party of Appointing President

<table>
<thead>
<tr>
<th>Measure</th>
<th>Bailey Score Mean</th>
<th>Bailey Score S.D.</th>
<th>JCS Score Mean</th>
<th>JCS Score S.D.</th>
<th>MQ Score Mean</th>
<th>MQ Score S.D.</th>
<th>11 Democratic Appointees</th>
<th>20 Republican Appointees</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bailey Score</td>
<td>1</td>
<td>.979</td>
<td>.963</td>
<td>-.784</td>
<td>.758</td>
<td>.337</td>
<td>.850</td>
<td></td>
</tr>
<tr>
<td>JCS Score</td>
<td>.979</td>
<td>1</td>
<td>.970</td>
<td>-.254</td>
<td>.347</td>
<td>.245</td>
<td>.412</td>
<td></td>
</tr>
<tr>
<td>MQ Score</td>
<td>.963</td>
<td>.970</td>
<td>.970</td>
<td>-1.350</td>
<td>1.550</td>
<td>.639</td>
<td>.1644</td>
<td></td>
</tr>
<tr>
<td>Republican Appointee*</td>
<td>.656</td>
<td>.611</td>
<td>.636</td>
<td>0</td>
<td>n/a</td>
<td>1</td>
<td>n/a</td>
<td></td>
</tr>
</tbody>
</table>

Note:
* Point biserial correlations in this row.

Table A2 – Principal Components Factor Analysis (two factors retained)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Eigenvalue</th>
<th>Difference</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor1</td>
<td>3.324</td>
<td>3.268</td>
<td>0.994</td>
</tr>
<tr>
<td>Factor2</td>
<td>0.056</td>
<td>0.017</td>
<td>0.017</td>
</tr>
</tbody>
</table>

Table A3 -- Factor loadings

<table>
<thead>
<tr>
<th>Variable</th>
<th>Loading on Factor 1</th>
<th>Loading on Factor 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bailey2020</td>
<td>0.988</td>
<td>0.015</td>
</tr>
<tr>
<td>JCS2020</td>
<td>0.983</td>
<td>-0.123</td>
</tr>
<tr>
<td>MQ2019</td>
<td>0.976</td>
<td>-0.026</td>
</tr>
<tr>
<td>Republican Appointee</td>
<td>0.654</td>
<td>0.200</td>
</tr>
</tbody>
</table>
Table A4 – Lifetime Average Ideology Scores and Party of First Appointing President of Supreme Court Justices, 1960-2018

<table>
<thead>
<tr>
<th>Last Name</th>
<th>Bailey Score</th>
<th>JCS Score</th>
<th>MQ Score</th>
<th>Republican Appointee</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alito</td>
<td>1.13</td>
<td>0.58</td>
<td>1.82</td>
<td>1</td>
</tr>
<tr>
<td>Black</td>
<td>-1.61</td>
<td>-0.42</td>
<td>-1.76</td>
<td>0</td>
</tr>
<tr>
<td>Blackmun</td>
<td>-0.07</td>
<td>0.07</td>
<td>-0.03</td>
<td>1</td>
</tr>
<tr>
<td>Brennan</td>
<td>-1.13</td>
<td>-0.44</td>
<td>-1.78</td>
<td>1</td>
</tr>
<tr>
<td>Breyer</td>
<td>-0.65</td>
<td>-0.33</td>
<td>-1.23</td>
<td>0</td>
</tr>
<tr>
<td>Burger</td>
<td>0.94</td>
<td>0.59</td>
<td>1.89</td>
<td>1</td>
</tr>
<tr>
<td>Clark</td>
<td>0.23</td>
<td>0.26</td>
<td>0.46</td>
<td>0</td>
</tr>
<tr>
<td>Douglas</td>
<td>-1.90</td>
<td>-0.71</td>
<td>-4.72</td>
<td>0</td>
</tr>
<tr>
<td>Fortas</td>
<td>-1.09</td>
<td>-0.37</td>
<td>-1.33</td>
<td>0</td>
</tr>
<tr>
<td>Frankfurter</td>
<td>0.34</td>
<td>0.26</td>
<td>0.52</td>
<td>0</td>
</tr>
<tr>
<td>Ginsburg</td>
<td>-0.93</td>
<td>-0.43</td>
<td>-1.73</td>
<td>0</td>
</tr>
<tr>
<td>Goldberg</td>
<td>-0.97</td>
<td>-0.29</td>
<td>-1.08</td>
<td>0</td>
</tr>
<tr>
<td>Goresuch</td>
<td>1.02</td>
<td>0.42</td>
<td>0.98</td>
<td>1</td>
</tr>
<tr>
<td>Harlan</td>
<td>0.77</td>
<td>0.53</td>
<td>1.62</td>
<td>1</td>
</tr>
<tr>
<td>Kagan</td>
<td>-0.72</td>
<td>-0.43</td>
<td>-1.58</td>
<td>0</td>
</tr>
<tr>
<td>Kavanaugh</td>
<td>0.77</td>
<td>0.30</td>
<td>0.54</td>
<td>1</td>
</tr>
<tr>
<td>Kennedy</td>
<td>0.42</td>
<td>0.33</td>
<td>0.68</td>
<td>1</td>
</tr>
<tr>
<td>Marshall</td>
<td>-1.47</td>
<td>-0.58</td>
<td>-2.83</td>
<td>0</td>
</tr>
<tr>
<td>O'Connor</td>
<td>0.56</td>
<td>0.41</td>
<td>1.01</td>
<td>1</td>
</tr>
<tr>
<td>Powell</td>
<td>0.49</td>
<td>0.42</td>
<td>0.97</td>
<td>1</td>
</tr>
<tr>
<td>Rehnquist</td>
<td>1.37</td>
<td>0.68</td>
<td>2.97</td>
<td>1</td>
</tr>
<tr>
<td>Roberts</td>
<td>0.65</td>
<td>0.39</td>
<td>0.93</td>
<td>1</td>
</tr>
<tr>
<td>Scalia</td>
<td>1.18</td>
<td>0.66</td>
<td>2.51</td>
<td>1</td>
</tr>
<tr>
<td>Sotomayor</td>
<td>-1.17</td>
<td>-0.61</td>
<td>-2.68</td>
<td>1</td>
</tr>
<tr>
<td>Souter</td>
<td>-0.52</td>
<td>-0.17</td>
<td>-0.77</td>
<td>1</td>
</tr>
<tr>
<td>Stevens</td>
<td>-1.02</td>
<td>-0.39</td>
<td>-1.81</td>
<td>1</td>
</tr>
<tr>
<td>Stewart</td>
<td>0.15</td>
<td>0.25</td>
<td>0.40</td>
<td>1</td>
</tr>
<tr>
<td>Thomas</td>
<td>1.45</td>
<td>0.75</td>
<td>3.60</td>
<td>1</td>
</tr>
<tr>
<td>Warren</td>
<td>-0.85</td>
<td>-0.34</td>
<td>-1.26</td>
<td>1</td>
</tr>
<tr>
<td>White</td>
<td>0.16</td>
<td>0.25</td>
<td>0.44</td>
<td>0</td>
</tr>
<tr>
<td>Whittaker</td>
<td>0.59</td>
<td>0.46</td>
<td>1.17</td>
<td>1</td>
</tr>
</tbody>
</table>

Data Sources:


Biographical information, dates of service, and party of first appointing president: Data downloaded Sept 20, 2021 [https://www.supremecourt.gov/about/members_text.aspx](https://www.supremecourt.gov/about/members_text.aspx)
NOTES

i “There are currently four Article III courts: The Supreme Court of the United States, the U.S. courts of appeals, the U.S. district courts and the U.S. Court of International Trade. Congress has abolished, combined or reorganized several other Article III courts over time.” Notable exclusions are Article I courts, including military courts martial tribunals and administrative law courts. Judges of Article I courts lack life tenure and compensation protections of Article III judges. (https://www.fjc.gov/history/courts/courts-brief-overview)

ii In using the word, “ideology,” we follow prior, related judicial and political research (e.g. Pinello 1999, Bonica et al. 2019; Segal and Cover 1989; Wetstein et al. 2009; Epstein et al. 2007). This usage of “ideology” neither requires nor prohibits a fully-elaborated, institutionalized political ideology as described in sociological treatments, e.g. Wright (1997), Althusser (2014), and Bell (1960).

iii Ideal point measures differ from commonplace dominance scale measures, e.g. Likert scale measures, that range monotonically from low to high scale values (e.g. from “not at all liberal” to “very liberal”) (see Tay and Ng 2018). In contrast, ideal point measures recognize that respondents may prefer moderation (e.g. slightly liberal or slightly conservative) to extremes (e.g. staunchly conservative or extremely liberal), just as they might react with equal distaste to overcooked and undercooked food. Ideal point and dominance scale measures of the same phenomenon can, but need not, be negatively correlated for certain ranges of scale values.

iv Related analyses include Chabot’s (2019) JCS-based examination of Supreme Court departures, Zigerell’s (2013) use of Bailey’s ideological scores, Bonica and Sen (2017), Segal et al. (2011), Farnsworth (2007) and Spruk and Kovac (2019). Some effort has been made to apply
MQ methods to judges beyond the Supreme Court (Spruk and Kovac 2019), but judicial ideology measures are generally unavailable for judges of federal district courts, who comprise the bulk of Article III judges (but see Bonica et al 2017).

* A reader of this paper insightfully suggest that party effects might differ noticeably from ideology effects (and therefore become visible) when political parties move suddenly and significantly to the right or left. We think that those effects also might be revealed if individual judges change their ideologies after appointment. But parties do not appear to shift their ideologies often, and ideological shifts by individual lower court judges are as yet unmeasured. Analysis of changes in political party ideologies and attitudes of individual judges would be useful, if they could be done well, but they would also be a diversion from present concerns with PDH. More generally, it appears that small ideological differences between individuals might have measurable effects on their chances of presidential nominations to the judiciary, because presidents have many individuals to choose among. However, when judges consider retirement and resignation, they can, at most, choose only between a Republican or a Democratic president, and making minor adjustments in party ideology seems unlikely to have much effect on the party identification of judges.

*vi In passing, this paper neither accepts nor rejects any claim of causal relationships between party and ideology.

*vii By law since 1845, elections occur on the Tuesday between November 2 and November 8, inclusive (*An Act to Establish a Uniform Time for Holding Elections 1845*). In 1921, 1929 and 1933 inaugurations occurred on March 4, following elections by about four months (e.g. 122 days in 1922). After 1933 (starting in 1953) inaugurations occur between 73
and 78 days later, on January 20, per the 20th Amendment to the U.S. Constitution, passed in 1933.

A thoughtful reader asks if the method used here is transformed from a sharp to a soft regression discontinuity model when the outcomes of some future elections are prognosticated in advance. We respond that “sharpness” here is produced by the transfer of power to nominate judges, which occurs definitively as the newly elected president completes recitation of the oath of office. Well-known examples of unsuccessful presidential aspirants who were widely predicted to win include Thomas E. Dewey, Albert Gore, and Hilary Clinton. More generally, the popular presidential election vote is often close, and therefore seemingly difficult to predict. To wit, in the 11 regime changing elections considered here, 5 were won by a candidate who received less than 50 percent of the popular vote, 7 were won by a candidate who received less than 51 percent of the popular vote, and only Harding, Roosevelt and Eisenhower took more than 55 percent of the popular vote (source: https://www.archives.gov/electoral-college/results accessed 10/24/2021).

Two judges were excluded from analyses reported here because they served less than 730 days before terminating service: Lindley Beckworth resigned in 1968 after 546 days in office, and George Mitchell resigned in 1980 after 224 days in office. Both were first appointed by Democratic presidents. Mitchell resigned his judicial post for an appointment to fill a vacant seat in the U.S. Senate, where he served until 1994. More generally, from 1790 through 2016, 42 judges terminated judicial service with less than 730 days of service. Of these, 14 terminated due to abolition of the court on which they served, one retired and 27 resigned. After 1919, one judge retired and 11 resigned after less than 730 days of service.
The Supreme Court decision of *Bush v. Gore* in December, 2000 might be an exception to this claim, except that no Supreme Court justices left office just before or after this case came to the Supreme Court.

Because data processing for this research began in 2018, 536 days after the 2017 inauguration, the 547 and 730 day post-inauguration enumeration periods for the 2016 election are truncated to 536 days, and results for them are not comparable to results for the same enumeration periods in earlier elections.