COLUMBIA | SIPA
Center for Development Economics and Policy

CDEP-CGEG WORKING PAPER SERIES

CDEP-CGEG WP No. 25

The Performance of Elected Officials: Evidence from State Supreme Courts

Elliott Ash and W. Bentley MacLeod

March 2016

COLUMBIA | SIPA
Center on Global Economic Governance

The Performance of Elected Officials:
Evidence from State Supreme Courts

Elliott Ash and W. Bentley MacLeod*

March 17, 2016

Abstract

This paper provides evidence on the effect of electoral institutions on the performance of public officials. Using panel data on state supreme courts between 1947 and 1994, we measure the effects of changes in judicial electoral processes on judge work quality – as measured by citations by later judges. Judges selected by non-partisan elections write higher-quality opinions than judges selected by partisan elections. Judges selected by technocratic merit commissions write higher-quality opinions than either partisan-elected judges or non-partisan-elected judges. Election-year politics reduces judicial performance in both partisan and non-partisan election systems. Giving stronger tenure to non-partisan-selected judges improves performance, while giving stronger tenure to partisan-selected judges has no effect. These results are consistent with the view that technocratic merit commissions have better information about the quality of candidates than voters, and that political bias can reduce the quality of elected officials.

1 Introduction

The goal of this paper is to contribute to our understanding of the labor market for elected officials. As Epstein et al. (2013) observe for federal judges, the decision-making powers of public officials can have large impacts upon our lives, yet their pecuniary rewards are by design only weakly related to their performance. In consequence, a variety of concerns – including career rewards (Ferejohn, 1986; Alesina and Tabellini, 2007; Dewatripont et al., 1999), professionalism (Wilensky, 1964), and prosociality (Benabou and Tirole, 2006) – can be decisive in determining the behavior of public officials. These motivations are in large part intrinsic to the individual, meaning that the performance of public individuals is determined in part by the type of person who is selected to serve in the public interest.

In this paper we exploit the fact that the method used to select and reappoint judges to state appellate courts varies over time and across states. In contrast to the U.S. Supreme Court, where justices have lifetime tenure, most U.S. states use one of three types of regular review: partisan elections, in which judges are explicitly affiliated with a political party on the ballot; non-partisan systems, where there is a vote, but party affiliation is not listed; and finally, a merit system in which judges are nominated by a commission of experts – senior attorneys and retired judges – and confirmed by the governor.

There is a lively debate regarding which is the superior system. The fact that states have experimented with different systems illustrates that it is not clear which system is optimal. There is a body of research that shows that the political affiliation of a judge at the margin affects the decisions that they make (Huber and Gordon, 2004; Lim, 2013; Canes-Wrone et al., 2014). Yet regardless of party affiliation, judges are tasked with interpreting and applying the law as written. In a common law system where judges follow their predecessors, the quality of a decision can have a large impact on the evolution of legal rules. The goal of this paper is to assess how variations in the appointment system affects the quality of judicial decisions.

To address these questions we have created a large panel dataset consisting of 400,000 opinions written by more than 1500 judges for all fifty states for the years 1947 through 1994. With this data we are able to construct a large number of diagnostic performance measures, for example the number of decisions written, the length of decisions, and how often those decisions are cited by later judges. To make the results intuitive, we construct five performance indexes from the individual performance variables. These include Total Output, Effort Per Case, Discretionary Opinions, Case Quality, and Total Impact. These indexes provide useful summaries of how judges change their behavior in response to changes

in electoral procedures.

This data is uniquely suited to study the impact of appointment systems upon the performance of public officials. First, the job of judging has varied little over this time period, meaning that we can more credibly assign variations in outcomes to variations in the individual performance by judges. Second, there is great deal of experimentation over time by states in the selection and retention processes.

This variation allows us to carry out a set of natural experiments to study how the quality and performance of judges responds to electoral reforms. First, following the approach in Ash and MacLeod (2015), we can explore within-judge changes of the electoral cycle. Specifically, we compare the performance of a judge in a year in which he is up for election with years in which he is not up for election. We find that in contested systems, election-year politics takes away time from work. In an election year performance is reduced, in both partisan and non-partisan elections. In uncontested systems where judges do not have a challenger, there is no decrease in performance during election years.

In addition to the electoral cycle, we study the within-judge effect of reforming the retention process. Moving from partisan to non-partisan elections reduces performance, while moving from non-partisan to uncontested elections increases performance. However, there is no effect on performance moving from partisan to uncontested elections. We demonstrate that partisan-selected judges do not change their electoral behavior – even after the reform, they reduce performance during election years.

A more challenging question is to measure the selection effects of the different electoral processes. We do this by comparing the performance of judges on the same court, making decisions in the same year, but selected under different systems. We add a number of controls, and carry out some robustness checks, and find that compared to judges selected by voters, there is consistent evidence that judges selected by a merit commission are better at their jobs.

In a recent paper, Choi et al. (2010) explore a similar set of issues using data from 1998, 1999 and 2000. First, they find that the correlation between appointment systems and measures of judge effort are quite unstable and sensitive to the control variables that they use. In their most highly controlled specification, the partisan judges are estimated to work harder (write more decisions) than judges selected under other systems. The results on quality are more stable, but tend to be close to zero, with a judge selected under a partisan system having a slightly negative effect upon quality.

These results are interesting for two reasons. First, they illustrate how the cross-section

can give a very different picture from estimates that are able to more tightly control for judge characteristics. Second, as a practical matter, state legislators would not have access to estimates such as ours, but would have to base their choice of appointment system upon observations of their current system and how it compares to other states. From our perspective, if the cross-sectional estimates are unstable, this implies that the choice of appointment system is more likely to be random, and hence our identification strategy is more likely yield a measure of the causal effect of the change in appointment system.

To help with the interpretation of our results, we introduce a simple model of the appointment system based on Condorcet's (1785) observation that elections are a way of aggregating information. Specifically, we suppose that the representative voter gets a noisy signal of judge ability. In such a model a merit plan can be viewed as a system in which the representative voter (governor) receives a higher-quality signal of performance, and accordingly the expected ability of the selected judge is higher than under a system that relies upon the public's impression of a judge.

Partisan elections can be distinguished from non-partisan elections by supposing that the representative voter prefers a judge from her preferred political party. This is modeled by adding a bias b in favor of the voter's party. As the bias increases, the expected ability falls and eventually approaches the expected ability that a one-candidate election would produce.

We find that this simple model is broadly consistent with the evidence on state supreme court judges. This evidence is more broadly consistent with the early rational-choice approaches of Downs (1957) and Ferejohn (1986), in which voters use their information to make the best decisions they can, conditional upon their policy preferences. But more information is not always better; more information on candidate quality can improve performance (see Pande, 2011), but more information on political affiliation can reduce performance.

The rest of the paper is organized as follows. Section 2 provides an institutional background on state supreme court selection and retention. Section 3 introduces a model of the selection and incentive effects of judicial elections. Section 4 discusses the issue of measuring judge performance. Sections 5, 6, and 7 report the results, respectively. Section 8 provides a concluding discussion.

2   Background

This section provides relevant background for the theoretical and empirical analysis. First, Subsection 2.1 describes the electoral institutions that provide our treatment variation. Sub-

Table 1: Judicial Selection and Retention Systems

State (Years)	Selection	Retention	State (Years)	Selection	Retention
Alaska	Merit	Uncontested	Mississippi	Partisan	Partisan
Alabama	Partisan	Partisan	Montana	Non-Partisan	Non-Partisan
Arkansas	Partisan	Partisan	North Carolina	Partisan	Partisan
Arizona (-1974)	Non-Partisan	Non-Partisan	North Dakota	Non-Partisan	Non-Partisan
Arizona (1975-)	Merit	Uncontested	Nebraska (-1962)	Partisan	Partisan
Colorado (-1966)	Partisan	Partisan	Nebraska (1963-)	Merit	Uncontested
Colorado (1967-)	Merit	Uncontested	New Mexico (-1988)	Partisan	Partisan
Florida (-1971)	Partisan	Partisan	New Mexico (1989-)	Partisan	Uncontested
Florida (1972-1976)	Non-Partisan	Non-Partisan	Nevada	Non-Partisan	Non-Partisan
Florida (1977-)	Merit	Uncontested	New York (1978-)	Partisan	Partisan
Georgia (-1984)	Partisan	Partisan	Ohio	Partisan	Non-Partisan
Georgia (1985-)	Non-Partisan	Non-Partisan	Oklahoma (-1967)	Partisan	Partisan
Iowa (-1962)	Partisan	Partisan	Oklahoma (1968-)	Merit	Uncontested
Iowa (1963-)	Merit	Uncontested	Oregon	Non-Partisan	Non-Partisan
Idaho	Non-Partisan	Non-Partisan	Pennsylvania (1969-)	Partisan	Uncontested
Illinois (-1964)	Partisan	Partisan	South Dakota (-1980)	Non-Partisan	Non-Partisan
Illinois (1965-)	Partisan	Uncontested	South Dakota (1981-)	Merit	Uncontested
Indiana (-1970)	Partisan	Partisan	Tennessee (-1971)	Partisan	Partisan
Indiana (1971-)	Merit	Uncontested	Tennessee (1972-1977)	Merit	Uncontested
Kansas (-1958)	Partisan	Partisan	Tennessee (1978-)	Partisan	Partisan
Kansas (1959-)	Merit	Uncontested	Texas	Partisan	Partisan
Kentucky (-1975)	Partisan	Partisan	Utah (-1951)	Partisan	Partisan
Kentucky (1976-)	Non-Partisan	Non-Partisan	Utah (1952-1985)	Non-Partisan	Non-Partisan
Louisiana	Partisan	Partisan	Utah (1986-)	Merit	Uncontested
Maryland (-1976)	Non-Partisan	Non-Partisan	Washington	Non-Partisan	Non-Partisan
Maryland (1977-)	Merit	Uncontested	Wisconsin	Non-Partisan	Non-Partisan
Michigan	Partisan	Non-Partisan	West Virginia	Partisan	Partisan
Minnesota	Non-Partisan	Non-Partisan	Wyoming (-1972)	Non-Partisan	Non-Partisan
Missouri	Merit	Uncontested	Wyoming (1973-)	Merit	Uncontested

Notes. This table lists the elections systems for state supreme court judges observed in our data. Election-system reforms indicated by cell borders.

section 2.2 provides an overview of our data sources. Subsection 2.3 describes some related literature.

2.1 Institutions

Our institutional setting is the set of state supreme courts, also known as state courts of last resort. As described in greater detail in Ash and MacLeod (2015), these courts serve as the state judiciary's analogue to the U.S. Supreme Court, where judges review state court cases rather than federal court cases. In each case, a judge writes an opinion explaining the decision. The job of a supreme court judge does not change much over the course of the career, and it does not vary across states.

While the work tasks are the same, the rules for selecting and retaining appellate judges vary across states and over time. These rules are listed in Table 1, with rule changes indicated by cell borders. These changes are used in our empirical section to identify the incentive and selection effects of changing electoral systems.

We study three major regimes for selecting and retaining appellate judges. There is a large literature in political science and political economy examining how these systems affect voter behavior and the politics of judicial decision-making (e.g. Shepherd, 2009; Canes-Wrone et al., 2010; Lim and Snyder, 2015). There is also a separate legal scholarship discussing the implications of these systems for legal rulemaking (e.g. Pozen, 2010). For a discussion of the political motivations behind reforms to these regimes see Hanssen (2004).

The first system, partisan elections, is used for both selection of new judges and retention of incumbent judges. For these elections, judges are members of a political party, Republican or Democrat. They must win a primary election for their party before running in a general election, where their political affiliation is labeled on the ballot.[1] Incumbent judges rarely face a credible challenge in the primary, but in the general election they usually face a challenger from the opposition political party.

Second, non-partisan elections are also used for both selection and retention. In this system there are competitive elections, but there are no primaries and party affiliations are not on the ballot. There are generally two candidates, an incumbent and a challenger, but the incumbent is not identified as such.

The third major system is merit selection with uncontested retention elections, also known as the Missouri Plan. In this system, judges are nominated by a commission of experts – senior attorneys and retired judges – and confirmed by the governor. Incumbent judges face an up-or-down retention vote with no challenger. This system is designed to be more meritocratic, and to impose weaker political incentives, than electoral selection. In a fourth hybrid system, judges are initially selected through partisan elections but thereafter face uncontested retention elections.

These institutions provide the variation in selection and incentives that we study in the empirical analysis. In the next section we formally analyze the key differences between these procedures.

---

[1]Ohio and Michigan state judicial elections are difficult to classify within the partisan/non-partisan dichotomy because they have partisan primaries and nomination processes, but the political party is not on the ballot in general elections. Following Nelson et al. (2013), we classify these states as partisan elections. However, coding them as non-partisan, or leaving them out of the analysis, does not change our results.

2.2 Data Overview

The dataset used for the empirical analysis is an extension of that used in Ash and MacLeod (2015). It merges information on judge biographies, state-level court institutions, and published judicial opinions. These data allow panel estimates on the effects of court institutions on judge performance.

We have biographical data on almost all the judges working at state supreme courts between 1900 and today. Table 2 reports summary statistics on the characteristics of judges working in one of the three selection systems discussed in Section 2.1. For many of the variables, the systems are comparable. Relative to the partisan judges, the non-partisan and merit judges are more likely to be female. Merit judges are the most likely to have judicial experience, while partisan judges are the most likely to have political experience. Non-partisan and merit judges have longer career lengths. Merit judges are the least likely to lose re-election.

Our performance measures were constructed from published state supreme court opinions for the years 1947 through 1994, obtained (along with some annotated metadata) from bloomberglaw.com. The full sample includes 1,025,461 cases. Because we are interested in studying the behavior of individual judges, we drop opinions that do not have a named author (per curiam decisions). We also drop cases that are less than seven sentences in length – these are summary orders such as cert denials. The restricted sample includes 387,905 majority opinions (plus attached discretionary opinions), about 25 cases per judge per year on average.

2.3 Literature

As previously mentioned, Choi et al. (2010) find in the cross section that elected judges write more opinions but merit-selected judges write more highly cited opinions. Other work in this vein includes Hall and Bonneau (2006), who find that a judge's qualifications – experience, salary, and other observable characteristics – increase the chance of being reelected.

Lim and Snyder (2015) provide especially useful evidence in our context. They find that bar association evaluations of judge candidate quality have a large effect on voting and electoral success in non-partisan elections, reflecting that voters care about judge quality. In partisan elections, however, the bar association evaluations have no effect on voter choices – the information on quality is crowded out by the information on political affiliation. In uncontested elections, the bar association evaluation correlates with voting but does not

Table 2: Summary Statistics on Judge Characteristics by Selection System

	Partisan Elections		Non-Partisan Elections		Merit Selection	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Background						
Start Age	53.6969	8.8354	52.8235	8.3454	52.1143	7.8792
Female	0.0305	0.1721	0.0663	0.2491	0.0616	0.2410
Top School	0.0973	0.2966	0.1040	0.3058	0.1081	0.3114
Previous Experience						
Private Practice	0.6862	0.4644	0.8141	0.3897	0.6645	0.4737
Judiciary	0.6082	0.4886	0.5630	0.4969	0.6818	0.4673
Politics	0.2818	0.4503	0.2269	0.4197	0.1118	0.3162
Academia	0.0879	0.2834	0.1076	0.3105	0.1060	0.3088
Partisan Affiliation						
Republican	0.5517	0.4675				
Democrat	0.4483	0.4412				
Career Length	11.8401	8.6620	13.6512	9.6248	13.2120	7.0797
How Ended						
Retired	0.3293	0.4703	0.3896	0.4883	0.3825	0.4871
Resigned	0.1111	0.3145	0.2098	0.4077	0.0553	0.2291
Died in Office	0.1070	0.3094	0.1144	0.3188	0.0553	0.2291
Lost Election	0.0650	0.2468	0.0409	0.1983	0.0138	0.1170
Impeached	0.0054	0.0735	0.0000	0.0000	0.0046	0.0679
Judges	738		367		217	

Notes. Biographical information by judge election system. Observation is a judge. *Female* is a dummy for being female. *Start Age* is judge age upon joining the court. *Female* is a dummy for being female. *Top School* means the judge attended law school at Yale, Harvard, Columbia, Stanford, or Chicago. The *Previous Experience* items equal one if the judge has previous experience in the respective area. *Republican* is a dummy for being Republican, *Democrat* for being Democrat. *Career Length* is number of years working on the court, conditional on having left the court before 2014. The *How Ended* items equal one if the judge's state supreme court judgeship ended for this reason.

affect electoral success because virtually all incumbents are retained.

The literature on elections and judge quality is part of a much larger literature examining the effects of judge elections on the content of judicial rulings. For example, a range of papers have shown that judges impose harsher criminal sentences in response to stronger electoral pressure (Huber and Gordon, 2004; Gordon and Huber, 2007; Lim, 2013; Berdejo and Yuchtman, 2013; Iaryczower et al., 2013; Park, 2014). More generally, previous papers have demonstrated that the politics of selection matter for the ideology of the selected judges (Landes and Posner, 2009; Epstein et al., 2013), and that incumbent judges respond to changes in the political preferences of the body responsible for retaining them (Shepherd, 2009; Canes-Wrone et al., 2010).

More broadly, our results add to the emerging empirical literature in political economy on how to design the institutions for selecting and rewarding public officials. These papers include Besley and Case (1995), Besley and Coate (2003), List and Sturm (2006), Besley et al. (2010), and Ash et al. (2015). Our focus on the information available to voters is relevant to the literature on transparency, which includes Snyder and Stromberg (2010), Ferraz and Finan (2011), and Pande (2011).

Deserving special mention are the models in Alesina and Tabellini (2007, 2008), analyzing the differences in incentives for elected politicians versus tenured bureaucrats. One can view the variation in the way state judges are (re-)appointed as a natural test bed for these ideas. Judges selected and retained by partisan or non-partisan elections can be treated as "politicians," while judges selected by merit commissions and given strong tenure can be treated as "bureaucrats." Our evidence that merit-selected judges produce more highly cited decisions is consistent with the hypothesis that in the case of appellate court judges, individuals selected to be good "bureaucrats" perform as well as or better than elected politicians.

3 Model

In this section we introduce a model based upon Condorcet's (1785) jury theorem that views voting as an information revelation problem.[2] The model provides a simple framework that is sufficiently rich to make clear predictions for the cases we consider.[3] It is assumed that each voter has a noisy measure of judge quality that is used to make their decisions. In

---

[2] see Young (1988) for a discussion

[3] See Ashworth and de Mesquita (2008) and Ashworth et al. (2015) for more sophisticated versions of this class of models.

addition they care about the political views of judges, which is modeled as a bias in favor of judges from their preferred party.

More precisely, suppose that there is an opening for a judge from which there are two candidates, A and B. One of these could be an incumbent, but we abstract from this and suppose that each judge j has a quality level q_j drawn from a normal distribution: $q_j \sim N(0,1), j \in \{A, B\}$. It is assumed that these draws are uncorrelated, though different jurisdictions may have different distributions. The socially desirable outcome is to choose the most able judge, though a judge's political views may bias this decision.

The remaining subsections analyze how differences in information on judge candidates may influence the expected quality q_j of the judge selected, as well as the judge's performance once he is in office. Subsection 3.1 introduces a merit selection baseline where the better judge is always selected. Subsection 3.2 considers the consequences of electoral selection, where voters do not have perfect information, and may be biased by politics. Subsection 3.3 looks at the effects on an incumbent judge of electoral campaign demands.

3.1 Merit Selection and Governor Appointment

The salient feature of merit selection is that there is a committee that looks carefully at each potential candidate. We model this by supposing that q_j is observable to members of the commission. The merit commission is assumed to be able to communicate its finding clearly to the governor, who in turn will select the more able candidate. Thus, the expected quality of a judge under an appointment system is the first order statistic:

$$\bar{q}^M = \mathbb{E}\left\{\max\left\{q_A, q_B\right\}\right\} = \frac{1}{\sqrt{\pi}} > 0.$$

If the expected ability of a randomly chosen candidate is 0, then selecting the better one from a pool of only two judges results in positive expected quality. Increasing the size of the pool would simply increase the expected quality of the appointed judge; it is the same logic as Condorcet's (1785) jury theorem.

We can compare this to an appointment system where political bias enters. As a matter of convention we suppose that the governor (and later the representative voter) prefers Judge A. We can model this as a bias b and suppose that Judge A is chosen if and only if:

$$q_A + b \geq q_B. \tag{3.1}$$

Let $I(q_A, q_B, b) = 1$ if (3.1) and zero otherwise. Let

$$\bar{q}^G(b) = \mathbb{E}\{q_A I(q_A, q_B, b) + (1 - I(q_A, q_B, b)) q_B\}. \tag{3.2}$$

In the appendix we show:

Proposition 1. *The average quality of judges chosen under an unbiased merit panel is higher than that under governor appointment with bias:* $\bar{q}^M = \bar{q}^G(0) > \bar{q}^G(b), b \neq 0$. *The difference in quality rises as the level of political bias increases:* $\bar{q}^G(b)$ *is strictly decreasing in b.*

This rather intuitive result illustrates the cost associated with bias. In the absence of any bias the best candidate is chosen. However, preference for one or the other candidate can lead to the less able individual being chosen in some cases.

3.2 Selection of Judges by Election

Next we consider the effect on quality of selecting judges by election. This is modeled by supposing that the quality of information held by the electorate is lower than that of the merit panel. Suppose that the representative voter gets a signal of judge $j's$ quality:

$$s_j = q_j + \epsilon_j$$

where ϵ_j is normally distributed with mean zero and variance σ_j^2. The precision is defined by $\rho_j = 1/\sigma_j^2$. The representative voter observes the two signals and then assesses the relative quality of the judges.

We distinguish partisan and non-partisan electoral systems by introducing bias b. As a matter of convention suppose that judge A comes from the same party as the representative voter, where b represents the voter's utility weight on partisan affiliation. In a non-partisan system $b = 0$, while a partisan system is characterized by $b > 0$.

After observing s_j, the voter's posterior distribution on q_j is normal with mean

$$\mathbb{E}\{q_j|s_j\} = \pi_j s_j$$

and precision $1 + \rho_j$, where $\pi_j = \frac{\rho_j}{1+\rho_j}$ is the weight assigned to s_j. The representative voter selects Judge A if and only if

$$\pi_A s_A + b \geq \pi_B s_B.$$

As the bias in favor of a judge from the same party increases, the probability that Judge A is selected increases. This can be understood as reducing the competitiveness of the election. The expected quality of a judge selected under an electoral system with bias b is defined by:

$$\bar{q}^E\left(b\right) = \mathbb{E}\left\{q_A I\left(\pi_A s_A, \pi_B s_B, b\right) + q_B\left(1 - I\left(\pi_A s_A, \pi_B s_B, b\right)\right)\right\}. \tag{3.3}$$

In the appendix we show:

Proposition 2. *When voters do not perfectly observe judge quality, the average quality of elected judges is lower than that of merit-selected judges:*

$$q^M \geq \bar{q}^G\left(b\right) > \bar{q}^E\left(b\right).$$

Average judge quality falls with the strength of political bias, and therefore quality with partisan elections is lower than that with non-partisan elections: $\bar{q}^E\left(b\right)$ falls with b.

As in the previous case, bias reduces the effectiveness of the electoral system.

3.3   Campaign Incentives

We now build upon the previous framework to analyze the incentives for a judge seeking re-election. The most direct way to introduce campaign effort is to suppose that effort enhances the quality of the signal observed by voters.

We formalize this idea as follows. We suppose that the individuals have a normal level of effort for their work, given by $\bar{y}_A$ and $\bar{y}_B$ for the incumbent A and the challenger B, respectively. In an election year the individuals divert effort to election-year politics. While B is a challenger and is not sitting on the court, for simplicity we assume he faces the same decision problem as the incumbent A. This approximates the situation where B is a judge on another court – a federal court for example, or the state's intermediate appellate court.

Thus in an election year it is assumed that the individuals supply y_A and y_B to their jobs, resulting in election year effort:

$$e_A = \bar{y}_A - y_A \geq 0,$$

$$e_B = \bar{y}_B - y_B \geq 0.$$

The consequence is that the representative voter chooses judge A over judge B if and only if

$$\pi_A(s_A + e_A) + b \geq \pi_B(s_B + e_B).$$

The probability of A winning is:

$$p_A(e_A, e_B | q_A, q_B) = \mathbb{E}\left\{ I\left(\pi_A(s_A + e_A), \pi_B(s_B + e_B), b\right) | q_A, q_B \right\}.$$

Correspondingly, define $p_B(e_A, e_B | q_A, q_B) = 1 - p_A(e_A, e_B | q_A, q_B)$.

We suppose that candidate j has preferences:

$$U_j = Bp_j(e_A, e_B | q_A, q_B) - C(e_j),$$

where B is the intrinsic value from winning the election and $C_j(e_j) = C_j(\bar{y}_j - y_j)$ is the utility cost of campaign effort. The campaigning cost $C(e)$ is assumed to be twice differentiable in e and satisfies $C_j(0) = C_j'(0) = 0, C_j'' > 0$. This guarantees an interior solution.

Let us suppose that A is a sitting judge, while B is a potential challenger. In our data we can observe the output of judges, and hence both $\bar{y}_A$, the output before an election year, and y_A, the output in an election year, are observable. Consider first an uncontested elections (the "Missouri Plan"), in which judges do not face a challenger. This can be understood in the model notation as $e_B = 0$; the challenger sets zero campaign effort. The incumbent judge A sets e_A accordingly.[4]

Next, we consider the equilibrium when there is an active challenger (details in the appendix). If we suppose that $\rho_A = \rho_B$, the problem is symmetric and we have $e_A = e_B$. The first-order conditions for effort in this case are given by:

$$C_j'(e_j) = \sqrt{\frac{\rho}{2}} \phi\left(\sqrt{\frac{\rho}{2}}(q_A - q_B) + b\left(\frac{1+\rho}{\sqrt{2\rho}}\right) \right). \tag{3.4}$$

where $\phi(\cdot)$ is the standard normal pdf. Since $\phi(x)$ achieves its maximum value at $x = 0$, we see that effort is highest when:

$$(q_A - q_B) + b\left(\frac{1+\rho}{\rho}\right) = 0. \tag{3.5}$$

---

[4]The only caveat is for judges who feel they may not get re-elected for whatever reason (for example, bad press from a high-profile case). Thus, there may be some judges who do exert effort, in which case e_A may be positive. There is never any reason to observe a negative effort level.

These observations can be summarized as follows:

Proposition 3. *When voters have the same quality of information regarding candidates, the candidates choose the same level of campaign effort. Moreover, the amount of effort is highest in the most competitive races - when (3.5) is small. In particular, campaign effort decreases with the bias b. This means that campaigns reduce judging effort more under non-partisan elections than under partisan elections.*

In the appendix we prove that an equilibrium to the campaign effort game exists and that the effort of Judge A is greater than candidate B if and only if the electorate has a better measure of Judge A's quality.

This proposition has the following implications in our data. First, uncontested elections are the least competitive and have the weakest electoral incentives. Among the electoral systems, they should have a smaller effect on judging effort than partisan elections or non-partisan elections. Second, if non-partisan elections have less bias, then they are more competitive than partisan elections. Therefore non-partisan elections should have a larger negative effect on judging effort than partisan elections.

4 Measuring Judge Performance

In this section we discuss the problem of measuring judicial performance. We have a large number of performance variables that could each be used to assess judge performance, making interpretation of the results difficult. Looking at the separate treatment effects on all of these outcomes would present a multiple-comparisons problem.[5] We resolve this issue by aggregating the variables into a set of five performance indexes designed to summarize the effects of the treatments on the work components of judging.

The set of performance variables, along with justifications of how they were divided into indexes, is discussed in Subsection 4.1. The formal definitions of the indexes are described in 4.2.

4.1 Performance Variables

The set of performance variables are listed by index in Table 3.[6] The table also reports the mean and standard deviation, where the data are constructed at the judge-year level. The

---

[5]We report the effects of treatments on these individual measures in Appendix B.2.

[6]See Ash and MacLeod (2015) for a detailed discussion of these variables.

Table 3: Summary Statistics on Judge-Year Performance Variables

Outcome Variable	Mean	Std. Dev.	ML Factor Scores
Case Output Index			
Majority Opinions Written	25.25	16.46	0.0564
Total Words in Majority Opinions	55235.35	33630.15	0.64594
Total Sentences in Majority Opinions	2849.436	1937.005	0.21498
Previous Cases Cited in Majority Opinions	510.24	387.92	0.09826
Effort Per Case Index			
Words Per Majority Opinion	2453.38	1348.8	0.68332
Sentences Per Majority Opinion	129.8098	98.83284	0.22265
Previous Cases Cited Per Majority Opinion	22.62	16.92	0.11854
Discretionary Opinions Index			
Discretionary Opinions Written	6.15	9.1	0.20884
Total Words in Discretionary Opinions	8034.01	15360.4	0.35855
Previous Cases Cited in Discretionary Opinions	86.17	179.71	0.45563
Case Quality Index			
Positive Cites Per Opinion	13.03	12.86	0.27311
Distinguishing Cites Per Opinion	2.14	2.74	0.07903
Discuss Cites Per Opinion	2.96	2.75	0.29684
Quoted Cites Per Opinion	3.3	4.22	0.35495
Out-of-State Cites Per Opinion	1.81	2.45	0.06997
Total Impact Index			
Total Positive Cites	291.31	275.64	0.2971
Total Distinguishing Cites	44.92	55.54	0.04631
Total Discuss Cites	65.35	54.15	0.35981
Total Quoted Cites	70.12	72.73	0.26187
Total Out-of-State Cites	43.1	79.49	0.06529

Notes. Observation is a judge-year, N=16,084. These statistics are constructed from each judge's yearly output of cases. "Per Opinion" measures are divided by the number of majority opinions written that year. See variable definitions in the accompanying text.

right-most column (ML Factor Scores) will be discussed further in Subsection 4.2. Before indexes are constructed, all the metrics are transformed using the inverse hyperbolic sine.[7]

The first set of variables constitute a Case Output Index. At the state supreme court level, if judges accept more cases for review they are taking on more work. An additional measure of total work is the total number of words written, and total number of sentences written, in majority opinions. Similarly, the total amount of caselaw research performed – as measured by the number of previous cases cited a in a judge's opinions – is included.

The second index is Effort Per Case. This includes two basic opinion length measures – the average number of words, and average number of sentences, per majority opinion. We also have a measure of the amount of research a judge engages in – the Previous Cases Cited measure gives the number of previous authorities cited in her opinions. We include both the number of sentences and the number of words partly because it is unclear a priori which is a better measure of language output. It also solves the problem that three measures are needed to construct ML Factors, as described in Subsection 4.2.

The third index, Discretionary Opinions, includes variables related to effort on discretionary opinions. Whether to write a discretionary opinion—a concurrence or a dissent—is up to the judge's discretion and involves willingly taking on more work. Further, the number of words and number of previous cases cited in those opinions are components of the time spent on discretionary opinions. In previous versions of the paper, discretionary opinions were merged in with majority opinions, but the covariance matrix for performance suggests that they are separate factors in judicial decision making.

Fourth, we look at the Case Quality Index. To measure the quality of decision-making, we use the number of citations to a judge's opinions by other judges. In our data, Bloomberg Law staff attorneys have categorized citations as positive, distinguishing, or negative. A positive cite is a clear signal that a decision is found useful by a future judge. A distinguishing cite means that part of the ruling is useful, but needs to be clarified – so this is perhaps a weaker signal of opinion quality. In the set of positive citations, we also use information about whether a case is discussed by the future court (rather than cited without comment) and whether it is directly quoted by the citing court. These measures can be understood as more direct signals that the citing court finds the opinion useful. The Out-of-State Cites measure includes positive cites from out-of-state courts; as noted by Choi et al. (2010) among others, this is perhaps the best measure because the cited case serves as persuasive rather

---

[7]Defined as $\sinh^{-1}(x) = \log(x + \sqrt{1 + x^2})$, and used instead of the log transformation to allow for zeros in the data (Burbidge et al., 1988). Our results are robust to using levels or logs of the dependent variable. The adjusted R^2 is usually higher in the IHS or log specification than in levels.

than binding precedent. Note that, while these citations provide a good signal of expert evaluation, they may or may not reflect voter evaluation or what decision is best for social welfare.

Fifth and finally, the Total Impact Index is a combined quality and quantity measure. It gives the total number of positive, distinguishing, discussion, quoted-in, and out-of-state cites to a judge's work in a year. This serves to complement Case Output as a measure of quantity, and Case Quality as a measure of quality.

In the results sections, the main text reports the effects of our treatments on the performance indexes. In the appendix we report the effects on the individual variables listed, as well as a larger set of performance variables not listed here. We break out the effects on concurrences and dissents, for example, and show negative cites and number of cases overruled. See Appendix B.2 for details.

4.2   Performance Indexes

We implemented two methods for constructing performance indexes. These include the Z-score Index and the Maximum Likelihood Factor. Both of these indexes are used in the regressions reported in the results sections.

First, the Z-score Index refers to the standard aggregation method used in O'Brien (1984), Kling et al. (2007), and Deming (2009). For this index, each of the performance variables is residualized on a state and year fixed effect. Then these residuals are standardized by dividing by the standard deviation. The index is constructed from the average of these standardized variables for each judge-year observation.

The second index, Maximum Likelihood Factor, uses factor analysis and is based on Rao (1955) and Akaike (1987). Defining this measure provides an intuitive way of understanding judge quality and its impact on observed output.

Let $k \in \{1, 2, 3, 4, 5\}$ index the set of factors underlying judge performance. In our case those include case quality, effort per case, etc. Let $i \in \{1, 2, ..., m_k\}$ index the observed measures of factor k, for example the four variables representing Case Output in Table 3. Let z_{ijt}^k represent the observed level of performance measure i in factor k of judge j in year t, after being residualized and standardized as done for the Z-score Index. Therefore each performance measure has zero mean and variance 1. This is natural, as we do not have an absolute scale for judge performance and are interested in changes rather than levels.

In our model we suppose that rather than choosing the individual measure z_{ijt}^k, the judge

17

chooses the factor, y_{jt}^k. The factor y_{jt}^k is related to measure z_{ijt}^k by:

$$z_{ijt}^k = \alpha_i^k \times y_{jt}^k + \epsilon_{ijt}^k.$$

Given that z_{ijt}^k is standardized for all i and k, factor analysis begins by supposing that y_{jt}^k is normally distributed with mean zero and unit variance over the whole population.

Now we apply the results from Rao (1955). If the number of indexes is greater than or equal to three ($m_k \geq 3$), then we can estimate α_i^k, the loading for measure z_{ijt}^k on factor y_{jt}^k, as follows. Given that the variances are all normalized to be 1, then $\alpha_i^k \in (-1, 1)$. However, if our interpretation is correct, then each component is positively correlated with z_{ijt}, hence we should find $\alpha_i^k \geq 0$. This is indeed the case, providing further evidence in support of our interpretation.

In general, factor analysis allows for several unobserved factors per group of observed measures. But in our case the natural focus is a single factor model, where all of the measures in a group are driven by a single factor. We found that a full set of 20 performance measures is well-explained by a five-factor model, using the standard information criterion. This is consistent with our interpretation of one factor per group, and allows a more natural interpretation for each factor.

Now that α_i is estimated, notice that

$$x_{ijt}^k \equiv \frac{z_{ijt}^k}{\alpha_i^k} = y_{jt}^k + \frac{\epsilon_{ijt}^k}{\alpha_i^k} \tag{4.1}$$

is an unbiased estimate of the factor y_{jt}^k. We can compute the empirical covariance of $\vec{x}_{jt}^k = \left\{ x_{1jt}^k, x_{2jt}^k, ..., x_{m_kjt}^k \right\}$ for the full sample, denoted by Σ^x. By construction the diagonal elements will be $1/\alpha_{ik}^2 > 1$. Next let $\vec{y}_{jt}^k$ be a vector of scalars, all equaling y_{jt}^k, with length m_k (the number of performance measures in group k). The covariance of $\vec{y}_{jt}^k$ is JJ^T, where J^T is a vector of ones with length m_k and JJ^T is an $m_k \times m_k$ matrix of ones. Finally, let

$$\Sigma = \Sigma^x - JJ^T$$

be the covariance matrix for the vector of error terms $\frac{\epsilon_{ijt}^k}{\alpha_i^k}$.

Next we form the predicted factor $\vec{y}_{jt}^k$. Since x_{ijt}^k is an unbiased estimate of the factor

Table 4: Summary Correlations on Performance Indexes

Population Correlations

	Case Output	Case Effort	Discretionaries	Case Quality	Total Impact
Z-score Indexes					
Case Output	1				
Case Effort	0.2525	1			
Discretionaries	0.3574	0.1498	1		
Case Quality	0.1595	0.5635	0.0815	1	
Total Impact	0.8078	0.2154	0.2828	0.5664	1
Maximum Likelihood Factors					
Case Output	1				
Case Effort	0.3402	1			
Discretionaries	0.3406	0.1224	1		
Case Quality	0.2044	0.5448	0.0893	1	
Total Impact	0.8027	0.187	0.2786	0.5535	1

Mean Within-Judge Correlations

	Case Output	Case Effort	Discretionaries	Case Quality	Total Impact
Z-score Indexes					
Case Output	1				
Case Effort	0.2684	1			
Discretionaries	0.3397	0.1423	1		
Case Quality	0.1906	0.4652	0.0694	1	
Total Impact	0.7818	0.2194	0.2689	0.6088	1
Maximum Likelihood Factors					
Case Output	1				
Case Effort	0.3424	1			
Discretionaries	0.3196	0.1091	1		
Case Quality	0.2088	0.4390	0.0706	1	
Total Impact	0.7877	0.1944	0.2675	0.5805	1

y_{jt}^k, it follows from O'Brien (1984, p.1082) that the best estimate of the factor is:

$$\bar{y}_{jt}^k = \frac{J^T \Sigma^{-1} \vec{x}_{jt}^k}{\left(J^T \Sigma^{-1} J\right)}. \tag{4.2}$$

This prediction has variance $\left(J^T \Sigma^{-1} J\right)^{-1}$.

We repeat this process for each of the five groups of measures to produce the five ML Factor indexes used in the regression analysis. The estimated weight α_i^k for each measure is given in the right-most column of Table 3. By construction, the mean of $\hat{y}_{jt}^k$ is zero. The correlation matrices for each factor y_{jt}^k with all the other factors are given in Table

4. The table reports the population correlations as well as the means of the within-judge correlations. Judges who work hard overall, will also work hard per case, and thus we observe correlation between factors.

5 Effect of Being Up For Election

This section examines how judges change their behavior over time in response to the election cycle. Ash and MacLeod (2015) show that contested elections reduce performance. We add to that analysis by distinguishing between partisan and non-partisan elections. In theory, if judges wish to be re-elected then they should put effort into election year politics, as implied by the theory. This in turn leads to a reduction in output on the court. The question is whether the way a judge is elected affects this effort?

5.1 Empirical Strategy

The empirical strategy for examining the effects of electoral demands on judicial behavior is to exploit the staggered election cycle for identification of stronger electoral incentives. The election schedule is arbitrarily assigned by history, so it is reasonable to assume that the schedule is uncorrelated with other institutional or socioeconomic factors that might affect individual judge performance. For this analysis we used data provide by Kritzer (2011).

The electoral cycle is represented in our regressions as a vector of dummy variables E_{ist}, which equals one for years that a judge is up for election. There is a different element of the vector for partisan, non-partisan, and uncontested retention elections. The dummy variable is coded as a one regardless of whether the judge actually ran for election – this is needed to avoid endogeneity problems from the judge's choice whether to actually run.

One possible source of bias in this analysis comes from time-invariant characteristics of individual judges. Some judges may have higher or lower performance than others on average due to unobservable characteristics, and they may be up for election more often or less often for any number of reasons. To deal with this possibility, we include a full set of judge-specific fixed effects. Therefore any estimated election coefficients are relative to a judge's personal average.

A second major source of bias comes from the time-varying changes in the court work environment which may be correlated with the electoral schedule. For example, there may be campaigning demands during election years on all judges – not just those up for election – if they are asked to assist fellow members of their political party. To deal with this possibility,

we include a full set of state-year fixed effects. Therefore any estimated election coefficients are also relative to the court average in each year. This means they effectively compare judges sitting on the same court, working at the same time, but who are in different stages of the electoral cycle.

Formally, we estimate

$$y_{ist} = \text{JUDGE}_i + \text{STATE}_s \times \text{YEAR}_t + E'_{ist}\rho + \epsilon_{ist} \tag{5.1}$$

where JUDGE_i is a judge fixed effect, $\text{STATE}_s \times \text{TIME}_t$ is a state-year fixed effect for each s and year t, and E_{ist} includes the election-year treatments. Standard errors are clustered by state.

5.2   Results

The coefficient estimates from Equation (5.1) are reported in Table 5. Each row is from a separate regression, where the three columns report the estimate for partisan, non-partisan, and uncontested elections, respectively. These coefficients can be interpreted as the standard-deviation change in a judge's performance when he is up for election relative to his own average and the state-year average of his colleagues.

Columns 1 and 2 show the effects of being up for elections under contested elections. Under partisan elections, there is a decrease in performance across the board. Under non-partisan elections, there is a decrease in discretionary opinions and on total impact.

Column 3 gives the effect of elections for uncontested systems. The effects are the opposite. Instead of a negative effect, there is a positive change in total output, effort per case, discretionary opinions, and total impact.

5.3   Discussion

The fact that the point estimates in the partisan and non-partisan elections are negative and have the same order of magnitude, while the estimated effects for uncontested elections is, if anything positive, is consistent with the expectation that election year politics take time. The results are generally supportive of the idea that judges reduce judging effort during election years to spend more time on campaigning. The least competitive election system, uncontested elections, has the smallest effect on behavior as we would expect given Proposition 3. There may actually be some positive effects, which might be consistent with

Table 5: Effect of Being Up For Election

Outcome	Partisan Election Year (1)	Non-Partisan Election Year (2)	Uncontested Election Year (3)
Case Output			
Z-index	-0.106*	-0.154	0.119+
	(0.0473)	(0.0959)	(0.0609)
ML Factor	-0.113*	-0.164	0.126+
	(0.0493)	(0.100)	(0.0643)
Effort Per Case			
Z-index	-0.0543+	-0.0253	0.0576*
	(0.0281)	(0.0281)	(0.0290)
ML Factor	-0.0558+	-0.0361	0.0582+
	(0.0286)	(0.0298)	(0.0318)
Discretionary Opinions			
Z-index	-0.0640+	-0.0625**	0.0703+
	(0.0340)	(0.0231)	(0.0387)
ML Factor	-0.0627+	-0.0633**	0.0699+
	(0.0369)	(0.0233)	(0.0406)
Case Quality			
Z-index	-0.0643+	-0.0415	0.0252
	(0.0366)	(0.0381)	(0.0418)
ML Factor	-0.0701+	-0.0532	0.0155
	(0.0383)	(0.0383)	(0.0408)
Total Impact			
Z-index	-0.109*	-0.165*	0.0840+
	(0.0462)	(0.0821)	(0.0479)
ML Factor	-0.112*	-0.182*	0.0856+
	(0.0487)	(0.0882)	(0.0492)
Treated States	23	17	19
Treated Judges	437	270	277
Election Events	810	517	451

N= 16,084 judge-years. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01. Each row is from a separate regression for the stated outcome variable. Treatment variable is a dummy equaling one for years judge is facing reelection. Regressions include a state-year fixed effect and judge fixed effect, estimated using Stat's reg2hdfe module.

a desire to do a better job in an election year, though this would only be speculation.[8]

In contrast, there are negative effects on judging effort in the contested systems, where judges spend time campaigning during election years. The results for the non-partisan elections are consistent with the idea from Lim and Snyder (2015) that they are competitive and require campaign work. What is possibly more surprising is the large negative effect in the case of partisan elections. This is surprising because we would expect that given voters tend to follow party lines, then one would not expect these elections to be as competitive as the non-partisan elections. It would be interesting to know what role elections play in campaign financing, and if there are spillover effects between judicial campaigns and other campaigns that are occurring at the same time.

6 Effect of the Selection Process on Judge Quality

In this section we investigate how changes to the procedure to select judges affects the quality of chosen judges. This analysis is motivated by Proposition 2. Selection mechanisms that use better information about candidates or have less bias should, all else equal, select better candidates on average.

A priori, there is no reason to suppose that a judge chosen by the Missouri Plan faces less bias than in, say, a non-partisan election. However, the intent of using a merit commission is to create a pool of better qualified judges. Similarly, political parties have an incentive to choose qualified judges that are consistent with the party's views. Hence, it is an empirical question whether or not the judges chosen by the Missouri Plan or by a partisan election system are of higher or lower quality that those selected under a non-partisan system. What the theory illustrates is that the presence of bias reduces quality, while more precise signals increase quality.

6.1 Empirical Strategy

This subsection describes the empirical strategy for measuring the effects on judge quality of different judge selection systems. The source of identification used is the set of reforms to the judicial selection systems, depicted in Table 1. Three states changed from partisan selection

---

[8]Note that since the mid-1990s, third-party funding for negative advertising in Missouri Plan elections has increased significantly. Our results may not extend to more recent years (our panel ends in 1994). This is an important area for future research.

to non-partisan selection: Georgia, Kentucky, and Utah.[9] Six states moved from partisan selection to merit selection: Colorado, Iowa, Indiana, Kansas, Nebraska, and Oklahoma.[10] Three states moved from non-partisan selection to merit selection: Arizona, Maryland, and South Dakota.[11] The goal of the empirics is to compare the performance of judges selected before these reforms to the performance of judges selected after these reforms.

We control for time-varying state-specific factors by including a full set of state-year (interacted) fixed effects. This specification effectively compares the performance of judges sitting on the same court at the same time, but selected under different regimes. We carry out some robustness check to ensure that timing issues, such as the age of the judge, do not explain our results. We do this by including a full set of dummies for years of judge experience. This means that any estimates are made relative to other judges of the same experience level.

Second, the regressions include a full set of dummies for the judge's starting year. This set of controls complements the years of experience, with the goal of controlling for cohort-specific effects on performance. For example, judges beginning in the 1970s may be systematically better than judges beginning in the 1980s, due to changes in the economy. These indicators control for national variation in the market for judges as a function of time.

Third, the treatment indicators are active only for years where there are at least two judges selected from each system working on the court during that year. This is done to make a clean comparison that is not biased by outlier pre-reform judges who remain on the bench long after the other pre-reform judges. Appendix Table B8 reports the results when all years are included – they are similar.

The estimating equation for performance variable y_{ist} for judge i in state s at year t is

$$y_{ist} = \text{STATE}_s \times \text{YEAR}_t + X'_{ist}\beta + S'_{ist}\rho + \epsilon_{ist} \tag{6.1}$$

where $\text{STATE}_s \times \text{YEAR}_t$ includes the state-year fixed effects, X_{ist} includes the indicators for

---

[9]Florida also moved from partisan to non-partisan, but it is not included in this section because it changed to merit selection five years later.

[10]Tennessee moved to merit selection in 1972, but moved back to partisan selection in 1978. It is not included in this analysis.

[11]Florida also moved from non-partisan to merit, but it is not included in this section because it had changed from partisan to non-partisan elections five years prior. Utah also moved from non-partisan to merit, but our data set does not extend long enough to get observations with two merit-selected judges. Wyoming also moved from non-partisan to merit, but there were not any years where there were more than two judges selected from each system. Wyoming is therefore included in Table B8, where there are similar results.

years of judge experience and judge's starting year, and S_{ist} includes the treatment indicators equaling one for judges selected under the post-reform system. Standard errors are clustered by state.

Given the inclusion of the fixed effects, the coefficients ρ procure the average difference in performance between judges selected under the new system and judges selected under the old system, controlling for other time-varying state-level factors, for years of experience, and for cohort effects.[12]

6.2 Results

Table 6 reports the estimates from Equation (6.1). Column 1 compares non-partisan-selected judges to partisan-selected colleagues. Column 2 compares merit-selected judges to partisan-selected judges. Column 3 compares merit-selected judges to non-partisan-selected judges.

The results can be summarized as follows. Non-partisan-selected judges have lower effort-per-case but higher total impact than partisan-selected judges on average. Merit-selected judges have higher effort-per-case and higher case quality than partisan-selected judges. Merit-selected judges have higher discretionary effort and case quality than non-partisan-selected judges.

6.3 Discussion

First, non-partisan judges have higher total impact than partisan judges. Lim and Snyder (2015) find that party affiliation drives voter behavior, and hence our result is consistent with bias, where having a public political affiliation results in worse candidates. The results on merit selection suggest that merit commissions select better judges than elections. This is consistent with the model's notion that merit commissions have more information about judge quality than voters.

It is worth pointing out that Choi et al. (2010), using 3 years of data and identifying the effect from a cross section of judges, find a much larger effect of election system upon output (measured by number of opinions).[13] The difference in results illustrates the effect

---

[12]Note that in the electoral selection systems, the judges may be initially appointed by the governor to fill a vacant seat, rather than being initially selected through a competitive electoral process. We still code the appointed judges as being selected under the electoral system – since the predecessor's choice whether to step down is endogenous to the system.

[13]Rather then estimate the causal effect by comparing the judges on the same court, they run a single model with a large set of controls. The results are reported in Table 6; in Model 1, the three coefficients for partisan, non-partisan, and merit selection are: 1.219** (4.930), 0.738** (3.180), 0.651* (2.240) (standard errors in

Table 6: Effect of Judicial Selection System on Judge Quality

Outcome	Non-Partisan Judges Relative to Partisan Judges	Merit-Selected Judges Relative to Partisan Judges	Merit-Selected Judges Relative to Non-Partisan Judges
	(1)	(2)	(3)
Case Output			
Z-index	-0.0041	-0.084	-0.0261
	(0.0914)	(0.0750)	(0.169)
ML Factor	-0.0428	-0.0734	0.00604
	(0.105)	(0.0743)	(0.192)
Effort Per Case			
Z-index	-0.280*	0.289*	0.305
	(0.108)	(0.141)	(0.213)
ML Factor	-0.317*	0.270+	0.32
	(0.125)	(0.155)	(0.209)
Discretionary Opinions			
Z-index	0.341	-0.00129	0.428+
	(0.249)	(0.150)	(0.246)
ML Factor	0.357	0.0262	0.434
	(0.224)	(0.159)	(0.261)
Case Quality			
Z-index	0.046	0.194*	0.405+
	(0.0919)	(0.0767)	(0.206)
ML Factor	0.0221	0.194*	0.422*
	(0.106)	(0.0829)	(0.191)
Total Impact			
Z-index	0.141+	-0.0349	0.0774
	(0.0798)	(0.0802)	(0.196)
ML Factor	0.115	-0.0614	0.0619
	(0.0844)	(0.0799)	(0.191)
Treated States	3	6	3
Treated State-Years	24	86	24
Treated Judges	14	54	16

N= 16,084 judge-years..Estimate of the average difference between judges selected under a new system, relative to to judges selected under the old system, limited to years in which there are at least two judges on the court selected from each system. Regressions include a state-year fixed effect, a full set of dummies for years of experience, and a full set of dummies for starting years. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01.

of research design upon the estimated effects. Given all our controls, one might argue that our results are a lower bound on the effect of the selection system.

It is worth highlighting the fact that when jurisdictions choose a particular election system they must rely upon rather crude information regarding the causal effect of a reform. In particular, the Choi et al. (2010) show that comparing the experience of two jurisdictions with different systems can easily lead to larger perceived effects of the selection system.

Choi et al. (2010) suggest that out-of-state citations provide the best measure of quality. They find that merit commissions have zero effect. Since they run a single regression they are only measuring the correlation between a merit plan and cites. Our identification strategy attempts to measure the effect of changing from a partisan to merit commission. Here we find large significant effects on out-of-state citations (see Appendix B.2), consistent with the hypothesis that merit commissions have access to better information regarding judicial performance.

7 Variation in Response to Incentives

This section examines differences in the response to incentive changes based on how the judge was selected.

7.1 Effect of Judge Retention Process

This subsection reports the results on how changing the system for judge retention affects the performance of sitting judges. Subsection 3.3 discusses the model mechanism for the effects of retention elections on incumbent judge behavior. More competitive elections result in more campaigning, which will reduce effort spent on judging. We examine this issue using judge fixed effects and institutional reforms to the retention system.

7.1.1 Empirical Strategy

Identification comes from discrete changes in the rules for retaining state supreme court judges. The timing of these reforms is illustrated in Appendix Figure 1. Four states changed from partisan retention elections to non-partisan retention elections: Florida, Georgia, Kentucky, and Utah. Eight states moved from partisan retention to uncontested retention elec-

---

parentheses). Given that we have a much longer time period, and more tightly controlled comparisons, the fact that these coefficients are so large and so significant suggests that the results are driven by variation across states rather than the variation in electoral system.

tions: Colorado, Illinois, Iowa, Indiana, Kansas, Nebraska, New Mexico, and Oklahoma. Six states moved from non-partisan retention to uncontested retention: Arizona, Florida, Maryland, South Dakota, Utah, and Wyoming.[14]

The regression framework is a standard differences-in-differences approach based on Bertrand et al. (2004). To control for time-invariant judge characteristics that may be correlated with the retention system in various states, we include judge fixed effects. To control for national trends in performance, we include year fixed effects. To control for pre-existing state trends in performance that may be confounded with the reforms, we include state-specific linear trends.

As in Ash and MacLeod (2015), we measure effects in a ten-year window around the reforms. The regressions include an indicator equaling one for the baseline time window of ten years before and ten years after a change to the retention system. The treatment variable is a dummy for the ten years after the change. Thus, with the inclusion of the judge fixed effects, the estimates can be interpreted as the average difference in within-judge performance for the ten years after the policy change relative to the ten years before the policy change. In a handful of states, we shrank the time window if the reform occurred close to the beginning or end of the sample.[15]

Formally, we estimate

$$y_{ist} = \text{YEAR}_t + \text{JUDGE}_i + \text{STATE}_s \times t + \bar{R}'_{st}\bar{\rho} + R'_{st}\rho + \epsilon_{ist} \qquad (7.1)$$

where YEAR_t is a fixed effect for year t, JUDGE_i is a judge fixed effect, and $\text{STATE}_s \times t$ is a state-level linear time trend for state s. The term $\bar{R}_{st}$ is a vector of indicators equaling one for the baseline time windows of ten years before and ten years after each of the retention reforms. R_{st} is a vector of treatment indicators for the ten years after each rule change. Standard errors are clustered by state.

With the inclusion of the judge fixed effects, the estimates for the elements of ρ can be interpreted as the average difference in within-judge performance for the ten years after the policy change relative to the ten years before the policy change. Notice that these results

---

[14]For some of these treatments, there were other types of judicial reforms occurring around the same time. See Appendix B.1 for more details and robustness checks.

[15]These reforms are mostly enacted by voters through ballot referendums administered in November and officially going into effect the subsequent January. In these cases the dummy variable would turn on in the year following the vote. In cases where the policy is effective in the first half of the year, it is coded as turning on in that year. Note that Florida changes from partisan to non-partisan and then to uncontested elections. In the table regressions it is coded using the years depicted in the figure. Our results do not change substantially if Florida is left out of the analysis.

apply to different types of judges. For example, moving from a partisan to non-partisan system measures the effect of the change upon a judge selected under a partisan system.

7.1.2 Results

Table 6 reports our estimates for $\vec{\rho}$ from Equation 6.2. Each row is from a separate regression, with the first column giving the partisan-to-nonpartisan effect, the second column giving the partisan-to-uncontested effect, the the third column giving the nonpartisan-to-uncontested effect. Each regression includes a year fixed effect, judge fixed effect, and state trend.

Column 1 gives the incentive effect on sitting judges of moving from a partisan system to a non-partisan system. There is a negative effect on case output. Column 2 has the effect of moving from a partisan system to an uncontested system. Here we see no effects on performance. In Column 3, we see that a move from non-partisan to uncontested elections is associated with an increase in discretionary effort and on case quality.

7.1.3 Discussion

Begin with the move from partisan to non-partisan. There is a decrease in case output. Without placing too much emphasis on these estimates (and noting the small sample of states for this reform), this is consistent with the point in the model that nonpartisan elections are more competitive than partisan elections, and therefore impose greater electoral constraints on a judge's time.

What about the effect of moving from partisan to uncontested? There are no effects. This could mean two things. This would be consistent with the idea from the model and Lim and Snyder (2015) that partisan systems impose weak electoral incentives, so moving to an uncontested system wouldn't change incentives very much. This could also mean that partisan systems select for judges that don't care about quality judging, so reducing electoral incentives does not result in increased judging effort.

We see a positive effect on work quality when moving from non-partisan retention to uncontested retention. There is a statistically significant increase in discretionary effort and case quality. In response to the weaker electoral incentives, the non-partisan judges improve performance. This suggests that the strong electoral demands were taking time away from judging.

Table 7: Effect of Changing the Retention System on Incumbent Judge Performance

Outcome	Partisan Retention to Non-Partisan Retention	Partisan Retention to Uncontested Retention	Non-Partisan Retention to Uncontested Retention
	(1)	(2)	(3)
Case Output			
Z-index	-0.183*	0.0445	-0.0257
	(0.0862)	(0.0567)	(0.103)
ML Factor	-0.167+	0.0355	-0.0324
	(0.0849)	(0.0568)	(0.100)
Effort Per Case			
Z-index	0.0748	-0.0931	0.225
	(0.196)	(0.118)	(0.166)
ML Factor	0.0969	-0.0953	0.181
	(0.216)	(0.108)	(0.182)
Discretionary Opinions			
Z-index	-0.00425	-0.0132	0.216+
	(0.0652)	(0.136)	(0.115)
ML Factor	0.0169	-0.018	0.226+
	(0.0644)	(0.143)	(0.127)
Case Quality			
Z-index	0.0352	-0.0716	0.283*
	(0.133)	(0.104)	(0.111)
ML Factor	0.039	-0.0722	0.262*
	(0.166)	(0.120)	(0.115)
Total Impact			
Z-index	-0.142	0.0313	0.0595
	(0.0918)	(0.0813)	(0.122)
ML Factor	-0.154	0.0498	0.016
	(0.101)	(0.0929)	(0.140)
Treated States	4	8	6
Treated Judges	25	65	35

N= 16,084 judge-years. Estimate of the average treatment effect of changing the judge retention system on incumbent judges at the time of the reform. Regressions include a judge fixed effect, year fixed effect, and state trends. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01.

7.2 Relative Election-Year Effect on Judges Selected by Different Processes

Finally we look at whether judges selected under different systems respond differently to electoral demands. As in Section 6, we focus on the states that changed their procedures for selecting judges. We then look at the effect of the electoral cycle separately for judges selected under different systems.

7.2.1 Empirical Strategy

Again we have the vector of electoral dummies E_{ist} to equal one for judges that are up for election in year t. In addition, we have the vector of dummies S_i for the process under which a judge is selected. We estimate

$$y_{ist} = \text{JUDGE}_i + \text{STATE}_s \times \text{YEAR}_t + E'_{ist}\rho + S_i E'_{ist}\eta + \epsilon_{ist} \tag{7.2}$$

where we have included a judge and state-year fixed effect as in Section __. While S_i isn't identified at the judge level, the interactions with the electoral cycle are identified within and across judges.

   We are interested in the following estimates. First, ρ will give the baseline electoral-cycle effects for non-partisan judges (in the states that went from partisan to non-partisan elections). It will also give the baseline electoral-cycle effect for merit judges (in the states that moved from elections to merit selection). The vector of coefficients η will include the effect of non-partisan elections on partisan-selected judges relative to their non-partisan-selected counter-parts. It will also include the effect of uncontested elections on partisan-selected judges relative to their merit-selected counter-parts. Finally, it will include the effect of uncontested elections on non-partisan-selected judges relative to their merit-selected counter-parts.

7.2.2 Results

Table 8 reports the relative effects of the electoral cycle by the process a judge is selected.

   Column 1a gives the baseline effect of non-partisan elections on non-partisan selected judges. Column 1a has negative effects on discretionary opinions and total impact, similar to the baseline election-cycle results. Column 1b shows the relative effect of non-partisan elections on partisan judges. These are zeros.

Table 8: Relative Election-Year Effect On Judges Selected by Different Processes

Outcome	Effect of Non-Partisan Elections on Non-Partisan-Selected Judges (1a)	Relative Effect of Non-Partisan Elections on Partisan-Selected Judges (1b)	Effect of Uncontested Elections on Merit-Selected Judges (2a)	Relative Effect of Uncontested Elections on Partisan-Selected Judges (2b)	Relative Effect of Uncontested Elections on Non-Partisan-Selected Judges (2c)
Case Output					
Z-index	-0.157	0.136	0.158*	-0.198*	0.0663
	(0.0982)	(0.178)	(0.0642)	(0.0984)	(0.150)
ML Factor	-0.168	0.131	0.170*	-0.217*	0.0469
	(0.103)	(0.188)	(0.0678)	(0.0994)	(0.162)
Effort Per Case					
Z-index	-0.025	-0.0133	0.0836*	-0.128*	0.0336
	(0.0272)	(0.204)	(0.0358)	(0.0551)	(0.111)
ML Factor	-0.0356	-0.0178	0.0893*	-0.139*	-0.0118
	(0.0291)	(0.189)	(0.0377)	(0.0559)	(0.123)
Discretionary Opinions					
Z-index	-0.0616*	-0.0363	0.0694	0.00347	0.00336
	(0.0239)	(0.103)	(0.0501)	(0.0637)	(0.117)
ML Factor	-0.0622*	-0.0429	0.0737	-0.0117	-0.0197
	(0.0241)	(0.0801)	(0.0513)	(0.0687)	(0.128)
Case Quality					
Z-index	-0.0486	0.279	0.0626	-0.145	-0.0964
	(0.0368)	(0.277)	(0.0384)	(0.128)	(0.135)
ML Factor	-0.0593	0.239	0.0566	-0.159	-0.111
	(0.0370)	(0.329)	(0.0404)	(0.134)	(0.119)
Total Impact					
Z-index	-0.173*	0.302	0.129*	-0.198*	-0.0234
	(0.0837)	(0.234)	(0.0509)	(0.0783)	(0.0987)
ML Factor	-0.189*	0.261	0.132*	-0.206**	-0.0226
	(0.0902)	(0.255)	(0.0538)	(0.0770)	(0.104)
Treated States	2	2	11	8	4
Treated Judges	7	4	119	51	10
Election Events	8	5	201	90	16

N= 16,084 judge-years. Standard errors clustered by state in parentheses. + $p < .1$, * $p < .05$, ** $p < .01$. Each row is from a separate regression for the stated outcome variable. The estimated coefficient is a dummy equaling one for years judge is facing reelection, interacted with a dummy for if the judge is selected under the new selection system. Regressions include a state-year fixed effect, judge fixed effect, and the baseline coefficient for the election-year

Column 2a gives the baseline effect of uncontested elections on merit-selected judges. As with the baseline results, there are actually positive effects estimated for the election-cycle effect in an uncontested system. When one only looks at the merit-selected judges, the effect is stronger. Column 2b gives the relative effect of uncontested elections on partisan-selected judges. There are significant negative effects. The coefficients are larger in absolute value than the coefficients from Column 2a, meaning that uncontested elections actually have a negative effect on performance for partisan-selected judges. Finally Column 2c gives the relative effect of uncontested elections for non-partisan-selected judges, relative to merit judges. There aren't any significant differences here.

7.2.3 Discussion

The interesting effect in this section is that after changing to uncontested retention elections, partisan-selected judges still demonstrate the same election-year behavior as they did under partisan elections. This perhaps explains why there was no within-judge effect of moving from partisan to uncontested elections – these judges are responding the same way to the electoral cycle as they had been doing before the reform. This adds further evidence of a difference in preferences between partisan and non-partisan systems. Partisan judges prefer to reduce performance in election years, even when those are uncontested elections where most everyone is retained. One possible interpretation is that because they have more partisan preferences, they feel a desire to be involved in campaigning for other non-judge candidates when they are up for election.

8 Conclusion

The goal of this paper has been to evaluate the effect of election processes on the quality of individuals and the effort they put into their jobs. To address this question we exploit the fact that the work of judging has remained relatively stable over time, which allows us to build performance measures based on a database of written state appellate court decisions.

We exploit the fact that U.S. states have experimented with different methods to appoint judges. This allows us to measure the causal effect of a change in the system upon performance. We can also evaluate the selection effect by comparing the performance of judges selected by different systems, but serving at the same time. Our results are summarized in Table 9.

Table 9: Summary of Results

	Partisan Judges	Non-Partisan Judges	Merit Judges
Selection Process Effects			
Relative to Partisan Judges		↑	↑
Relative to Non-Partisan Judges			↑
Electoral Cycle Effects			
Partisan Election Year	↓		
Non-Partisan Election Year	↓	↓	
Uncontested Election Year	↓	~	↑
Retention Reform Effects			
Move to Non-Partisan	↓		
Move to Uncontested	~	↑	

Summary of results. The left-most column indicates the treatment, and the other column headers indicate the sample of judges upon which the effect is being measured. Arrows indicate a positive or negative effect on judge performance. A tilde (~) indicates no effect. See text for details.

The evidence suggests that non-partisan elections select better judges than partisan elections. This is consistent with a selection model where a stronger signal on party affiliation crowds out information on candidate quality, so candidates are lower quality on average. The evidence also suggests that an expert, merit-based selection process selects better judges than an election system. This is consistent with a selection model where better-informed experts can choose more high quality officials than voters on average. In the realm of selecting public officials, more information is not always better for the quality of the person chosen.

For incumbent judges, we find that stronger electoral incentives reduce performance in election years, and that contested elections reduce performance more than uncontested elections. This is consistent with a simple model in which campaign effort takes time away from judging. Moving from partisan to non-partisan elections reduces performance for incumbent partisan-selected judges, which is consistent with the idea that partisan elections are less competitive because voters are biased by political affiliations. Moving from non-partisan to uncontested elections increases performance, consistent with the notion that non-partisan contested elections are more demanding of a judge's time than uncontested elections.

There is no within-judge effect of moving from partisan to uncontested elections. We show that this occurs because the partisan-selected judges do not change their electoral behavior – they continue to reduce performance during election years after the reform, even though uncontested elections are not competitive. Merit-selected judges actually increase some measures of performance during election years. These results highlight that these electoral systems select for different types of individuals. These differences in abilities and preferences result in measurable differences in their legal output.

A great deal of work remains. Even though we have a long panel, and arguably good identification, the effects are often relatively small or barely significant. This fact may not be all that surprising. If a single system had strong, consistent results, then we would have expected the market to have moved in that direction quickly, consistent with Posner's (1987) view that legal institutions move in the direction of efficient exchange.

Yet, the fact that we do find a pattern of effects that are consistent with our simple model helps explain why there is experimentation. The results are consistent with the hypothesis that merit commissions select better judges, followed by non-partisan judges, and finally partisan judges. Yet, judging is not a purely technical activity. There is a large literature in political science showing that the political views of judges color their decisions, which may explain why many jurisdictions prefer to allow the democratic process to be informed by

the political views of judges.[16] In this paper we provide some evidence on the performance consequences of these choices, which will hopefully inform decision-making on this important institutional question.

---

[16]See Epstein et al. (2013) for a discussion of federal judges, and copious citations to this large literature.

References

Akaike, H. (1987). Factor analysis and aic. *Psychometrika 52*(3), 317–332.

Alesina, A. and G. Tabellini (2007). Bureaucrats or politicians? part i: A single policy task. *The American Economic Review 97*(1), 169–179.

Alesina, A. and G. Tabellini (2008). Bureaucrats or politicians? part ii: Multiple policy tasks. *Journal of Public Economics 92*(3-4), 426–447.

Ash, E. and W. B. MacLeod (2015, forthcoming). Intrinsic motivation in public service: Theory and evidence from state supreme courts. *Journal of Law and Economics*.

Ash, E., M. Morelli, and R. Van Weelden (2015). Elections and divisiveness: Theory and evidence. Working Paper 21422, National Bureau of Economic Research.

Ashworth, S. and E. B. de Mesquita (2008). Electoral selection, strategic challenger entry, and the incumbency advantage. *Journal of Politics 70*(4), 1006–1025.

Ashworth, S., E. B. de Mesquita, and A. Friedenberg (2015, May). Accountability and information in elections. mimeo, University of Chicago.

Benabou, R. and J. Tirole (2006). Incentives and prosocial behavior. *American Economic Review 96*(5), 1652–1678.

Berdejo, C. and N. Yuchtman (2013). Crime, punishment, and politics: an analysis of political cycles in criminal sentencing. *Review of Economics and Statistics 95*(3), 741–756.

Bertrand, M., E. Duflo, and S. Mullainathan (2004, Feb). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Besley, T. and A. Case (1995). Does electoral accountability affect economic-policy choices - evidence from gubernatorial term limits. *Quarterly Journal of Economics 110*(3), 769–798.

Besley, T. and S. Coate (2003). Elected versus appointed regulators: Theory and evidence. *Journal of the European Economic Association 1*(5), 1176–1206.

Besley, T., T. Persson, and D. M. Sturm (2010). Political competition, policy and growth: Theory and evidence from the us. *The Review of Economic Studies 77*(4), pp. 1329–1352.

Burbidge, J. B., L. Magee, and A. L. Robb (1988). Alternative transformations to handle extreme values of the dependent variable. *Journal of the American Statistical Association 83*(401), 123–127.

Canes-Wrone, B., T. S. Clark, and J. P. Kelly (2014, 2). Judicial selection and death penalty decisions. *American Political Science Review 108*, 23–39.

Canes-Wrone, B., T. S. Clark, and J.-K. Park (2010). Judicial independence and retention elections. *Journal of Law, Economics & Organization 28*(2), 211.

Choi, S. J., G. M. Gulati, and E. A. Posner (2010). Professionals or politicians: The uncertain empirical case for an elected rather than appointed judiciary. *Journal of Law, Economics, and Organization 26*(2), 290.

Condorcet, M. d. (1785). Essay on the application of analysis to the probability of majority decisions.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. *American Economic Journal: Applied Economics*, 111–134.

Dewatripont, M., I. Jewitt, and J. Tirole (1999). The economics of career concerns, part ii: Application to missions and accountability of government agencies. *Review of Economic Studies 66*(1), pp.199–217.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy 65*(2), pp. 135–150.

Epstein, L., W. M. Landes, and R. A. Posner (2013). *The Behavior of Federal Judges*. Harvard University Press.

Ferejohn, J. (1986). Incumbent performance and electoral control. *Public Choice 50*(1-3), 5–25.

Ferraz, C. and F. Finan (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review 101*(4), 1274–1311.

Gordon, S. and G. Huber (2007). The effect of electoral competitiveness on incumbent behavior. *Quarterly Journal of Political Science 2*(2), 107–138.

Hall, M. and C. Bonneau (2006, January). Does quality matter? challengers in state supreme court elections. *American Journal of Political Science 50*(1), 20–33.

Hanssen, F. (2004). Is there a politically optimal level of judicial independence? *The American Economic Review 94*(3), 712–729.

Huber, G. A. and S. C. Gordon (2004). Accountability and coercion: Is justice blind when it runs for office? *American Journal of Political Science 48*(2), 247–263. Times Cited: 79 Huber, GA Gordon, SC Huber, Gregory/A-5950-2012 Huber, Gregory/0000-0001-6804-8148 79.

Iaryczower, M., G. Lewis, and M. Shum (2013). To elect or to appoint? bias, information, and responsiveness of bureaucrats and politicians. *Journal of Public Economics 97*, 230–244.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Kritzer, H. M. (2011). Competitiveness in state supreme court elections, 1946–2009. *Journal*

of Empirical Legal Studies 8(2), 237–259.

Landes, W. M. and R. A. Posner (2009). Rational judical behavior: A statistical study. *The Journal of Legal Analysis 1*, 775–831.

Lim, C. and J. M. Snyder (2015, April). Is more information always better?party cues and candidate quality in u.s. judicial elections. *Journal of Public Economics*.

Lim, C. H. S. (2013). Preferences and incentives of appointed and elected public officials: Evidence from state trial court judges. *American Economics Review*.

List, J. A. and D. M. Sturm (2006). Elections matter: Theory and evidence from environmental policy. *Quarterly Journal of Economics 121*(4), 1249–1281. Times Cited: 50 50.

Nelson, M. J., R. P. Caufield, and A. D. Martin (2013). Oh, mi: A note on empirical examinations of judicial elections. *State Politics & Policy Quarterly*, 1532440013503838.

O'Brien, P. C. (1984). Procedures for comparing samples with multiple endpoints. *Biometrics*, 1079–1087.

Pande, R. (2011). Can informed voters enforce better governance? experiments in low-income democracies. *Annual Review of Economics 3*(1), 215–237.

Park, K. H. (2014, April). *Judicial Elections and Discrimination in Criminal Sentencing*. Ph. D. thesis, Harris School, University of Chicago, Chicago, IL.

Posner, R. A. (1987). The law and economics movement. *The American Economic Review 77*(2), pp.1–13.

Pozen, D. (2010). Judicial elections as popular constitutionalism. *Columbia Law Review 110*, 2047–2134.

Rao, C. R. (1955). Estimation and tests of significance in factor analysis. *Psychometrika 20*(2), 93–111.

Shepherd, J. M. (2009). The influence of retention politics on judges' voting. *The Journal of Legal Studies 38*(1), 169–206.

Snyder, J. M. and D. Stromberg (2010). Press coverage and political accountability. *Journal of Political Economy 118*(2), 355–408.

Wilensky, H. L. (1964). The professionalization of everyone? *American Journal of Sociology 70*(2), pp.137–158.

Young, H. P. (1988). Condorcet's theory of voting. *The American Political Science Review 82*(4), pp. 1231–1244.

A Model Appendix

This appendix enumerates the proofs for the major theoretical results from Section 3. Subsection A.1 formalizes the effects of bias and noise on the quality of selected judges. A.2 formalizes the role of bias and noise in campaign incentives for judge effort.

A.1 Effect of bias and noise on judge quality

Let ϕ, Φ respectively denote the standard normal's probability density and cumulative distribution functions. The expected quality of judges selected by the governor, expression 3.2, can be written as:

$$
\begin{aligned}
\bar{q}^{G}(b) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(q_A + (q_B - q_A) I\left(q_A, q_B, b\right)\right) \phi\left(q_B\right) \phi\left(q_A\right) dq_B dq_A. \\
&= \int_{-\infty}^{\infty} \left(q_A + \int_{q_A+b}^{\infty} (q_B - q_A) \phi\left(q_B\right) dq_B\right) \phi\left(q_A\right) dq_A
\end{aligned}
\tag{A.1}
$$

Clearly $\bar{q}^{M} = \bar{q}^{G}(0)$. Notice that:

$$
\begin{aligned}
\frac{d\bar{q}^{G}(b)}{db} &= \int_{-\infty}^{\infty} \left(-b\phi\left(q_A + b\right)\right) \phi\left(q_A\right) dq_A \\
&= -b \int_{-\infty}^{\infty} \phi\left(q_A + b\right) \phi\left(q_A\right) dq_A \\
&= -\frac{1}{\sqrt{\pi}} b \exp\left(-\frac{b^2}{4}\right) < 0.
\end{aligned}
\tag{A.2}
$$

This shows that a small amount of bias has a small negative effect on quality, that gets larger with b. This proves Proposition 1.

Next we consider the expected quality with elections. In this case the expected payoff is over q_A and q_B, with selection determined by the signals:

$$
\bar{q}^{E}(b) = \int \int \left(q_A + (q_A - q_B) \Pr\left[s_B > s_A + b | q_A, q_B\right]\right) \phi\left(q_A\right) \phi\left(q_B\right) dq_A dq_B.
$$

Notice that $(q_A - q_B) I\left(q_A, q_B, b\right) > (q_A - q_B) \Pr\left[s_B > s_A + b | q_A, q_B\right]$ and hence we have immediately that $\bar{q}^{G}(b) > \bar{q}^{E}(b)$. Also since

$$
\frac{d \Pr\left[s_B > s_A + b | q_A, q_B\right]}{db} < 0
$$

for all q_A, q_B, we have that expected ability of judges falls with b. This implies Proposition

2.

A.2 Effect of campaign incentives on effort

We can write the signals observed by the voters as:

$$
\begin{aligned}
s_j &= m_j + r_j \epsilon_j, \\
&= \pi_j (x_j + e_j) + \pi_j \sigma_j \epsilon_j,
\end{aligned}
$$

where ϵ_j follows a standard normal distribution. Let us compute:

$$
\Pr \left[m_A + r_A \epsilon_A + b \geq m_B + r_B \epsilon_B \right].
$$

The inequality can be rewritten as:

$$
m_A + b + e_A - e_B - m_B \geq r_B \epsilon_B - r_A \epsilon_A = \sqrt{r_B^2 + r_A^2} \epsilon,
$$

where ϵ is a standard normal distribution. Hence, we have:

$$
\Pr \left[m_A + b + e_A - e_B + r_A \epsilon_A \geq m_B + r_B \epsilon_B \right] = \Phi \left(\frac{m_A + b - m_B}{\sqrt{r_B^2 + r_A^2}} \right)
$$

where $F(\cdot)$ is the standard normal cdf. In our case we have

$$
m_j = \frac{\rho_j}{1 + \rho_j} (x_j + e_j)
$$

and

$$
\begin{aligned}
r_j &= \frac{\rho_j}{1 + \rho_j} \sigma_j \\
&= \frac{\sqrt{\rho_j}}{1 + \rho_j}.
\end{aligned}
$$

Taking the effort of the other judge as given, the first order condition for a judge defines an optimal effort choice:

$$C'_j(e_j) = B\frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi\left(\frac{m_A + b - m_B}{\sqrt{r_B^2 + r_A^2}}\right). \tag{A.3}$$

$$= B\frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi\left(\frac{\frac{\rho_A}{1+\rho_A}(q_A + e_A) + b - \frac{\rho_B}{1+\rho_B}(q_B + e_B)}{\sqrt{r_B^2 + r_A^2}}\right). \tag{A.4}$$

Observe that if $\pi_A = \pi_B$, both judges choose the same level of effort, and this has *no effect* on the probability of winning – it is a negative sum game.

Assumption Effort costs are *strongly convex given $\rho_j, i \in \{A, B\}$* if for every $x \in \Re$ the solution to the following equation is unique:

$$C'_j(e) = B\frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi\left(\frac{\pi_j e}{\sqrt{r_B^2 + r_A^2}} + x\right), i \in \{A, B\}.$$

Such functions exist because $\phi > 0$ and ϕ', ϕ'' are bounded, and $C_j(0) = C'_j(0) = 0, C''_j > 0$. More generally, given any function $C(e)$ satisfying $C(0) = C'(0) = 0, C''' > 0$, and precisions ρ_j for $j \in A, B$, one can choose $\gamma_j > 0$ sufficiently large that this condition holds for $C_j(e) = \gamma_j C(e)$.

Proposition 4. *If effort costs are strongly convex given $\rho_j, i \in \{A, B\}$ then there exists a Nash equilibrium in campaign effort. Moreover Judge A chooses more effort than Judge B ($e_A > e_B$) if and only if the quality of information regarding Judge A is higher ($\pi_A > \pi_B$).*

Proof. Notice that the maximum effort possible for judge j is:

$$C'_j\left(e_j^{max}\right) = \frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi(0).$$

Let $m = \max\{\pi_A e_A^{max}, \pi_B e_B^{max}\}$ and define the function:

$$h : [-m, m] \rightarrow [-m, m]$$

by:

$$h(x) = \frac{\rho_A}{1 + \rho_A}e_A(x) - \frac{\rho_B}{1 + \rho_B}e_B(x)$$

where:

$$C_j'\left(e_j(x)\right) = B\frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi\left(\frac{\frac{\rho_A}{1+\rho_A}q_A + b - \frac{\rho_B}{1+\rho_B}q_B + x}{\sqrt{r_B^2 + r_A^2}}\right).$$

Strong convexity ensures that $e_j(x)$ is a uniquely defined continuous function of x that maximizes the payoff of judge j given the effort of the other judge. Hence $h(x)$ is continuous, and by Brower's fixed point theorem we have the existence of x^* such that $h(x^*) = x^*$, which is in turn by construction a Nash equilibrium, where:

$$
\begin{aligned}
C_j'\left(e_j^*\right) &= B\frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi\left(\frac{\frac{\rho_A}{1+\rho_A}q_A + b - \frac{\rho_B}{1+\rho_B}q_B + x^*}{\sqrt{r_B^2 + r_A^2}}\right), \\
&= B\frac{\pi_j}{\sqrt{r_B^2 + r_A^2}}\phi\left(\frac{\frac{\rho_A}{1+\rho_A}(q_A + e_A^*) + b - \frac{\rho_B}{1+\rho_B}(q_B + e_B^*)}{\sqrt{r_B^2 + r_A^2}}\right).
\end{aligned}
$$

.   □

B   Empirical Appendix

This appendix includes some further notes on the data and the institutional reforms, as well as further regression specifications.

B.1   Notes on Institutional Reforms

This section provides some notes on the institutional reforms. The key point is that there were often coterminous reforms, such as the introduction of an intermediate appellate court. To deal with this we ran all the regressions while leaving one state out. None of the results were substantially changed in these checks. Note that these coterminous reforms only threaten identification in the analysis of retention-process reforms. When we look at the electoral cycle and when we look at selection effects, we are holding court-specific incentives constant.

Colorado instituted an intermediate appellate court in 1971, four years after the election reform. Changing Colorado to a four year window does not change the results. Florida moved from partisan to non-partisan elections in 1972, then moved from non-partisan to merit-uncontested in 1977. Florida is not included in the selection process regressions. In the retention-process regressions we treat these as separate reforms with five-year effect windows. Removing Florida from the regressions does not change the results.

At the same time that Illinois changed from partisan retention to uncontested retention (November 1962), the state also increased judge term lengths from nine years to ten years. However, the term-lengths change went into effect in January 1963, two years before the election reform went into effect.

At the same time it moved from partisan to merit-uncontested, Indiana increased term lengths from six years to ten years.

Kentucky instituted an intermediate appellate court at the same time that it moved from partisan to non-partisan elections.

The Maryland governor began selecting new appointees by merit commission beginning in 1971. When it moved from non-partisan retention to uncontested retention, the term length was reduced from 15 years to 10 years.

Oklahoma instituted an intermediate appellate court at the same time it moved from partisan to merit-uncontested.

In 1973, South Dakota increased its term length from six years to eight years, eight years before the non-partisan to merit-uncontested reform.

Tennessee moved from partisan to merit-uncontested in 1972, then moved back to partisan elections in 1975. It is not included in the analysis.

Utah instituted an intermediate appellate court in 1988, two years after the reform from non-partisan to merit-uncontested.

B.2   Additional Regression Results

This appendix reports additional empirical results.

Table B1 reports summary statistics for the additional set of outcome variables. We have dissents and concurrences reported separately. The outcome used in the main tables, discretionary opinions written, includes concurrences, dissents, and opinions that are concurring in part and dissenting in part.

We also report in the appendix an additional set of citations measures. These measures were excluded from the main text for brevity and because they are relatively rare, as can be seen in the summary statistics. Negative citations are coded as negative by Bloomberg staff attorneys. Federal circuit cites includes citations from federal circuit courts. "Multiple-use" cites means that an opinion is cites multiple times by a later court, which is similar to the discussion cites measure used in the text. A case can be overruled by the state supreme court at a later date, or it can be overruled by the U.S. Supreme Court on appeal. Finally, a case can be superseded by statute – which means the state legislature passes a law to reverse

Table B1: Summary Statistics on Judge-Year Performance Variables (Additional Outcomes)

Outcome Variable	Mean	Std. Dev.	Min	Max
Dissents Written	3.755	5.694	0	129
Concurrences Written	1.816	3.448	0	60
Negative Cites Per Opinion	0.575	0.895	0	20.625
Federal Circuit Cites Per Opinion	0.649	3.624	0	153.11
Multiple-Use Cites Per Opinion	1.288	2.024	0	96
Proportion of Cases Overruled	0.052	0.194	0	5.52
Proportion of Cases Superseded by Statute	0.047	0.106	0	2.5
Total Negative Cites	11.629	15.932	0	246
Total Federal Circuit Cites	16.509	106.101	0	3503
Total Multiple-Use Cites	26.141	31.348	0	800
Total Cases Overruled	1.063	3.174	0	116
Total Cases Superseded by Statute	0.947	1.868	0	28

Notes. Observation is a judge-year, N=16,084. These statistics are constructed from each judge's yearly output of cases. "Per Opinion" measures are divided by the number of majority opinions written that year. See variable definitions in the accompanying text.

a court ruling.

Tables B7, B4, B11, and B16 report regression results from the same equations estimated in the corresponding main tables but with the additional outcome measures. These results generally line up with those in the main tables. Note that there are often strong effects on negative cites, but that may be due to the initially low average for this measure (see Table B1).

Tables B2 and B3 show the effects of the electoral cycle on individual performance variables. As shown in Column 1, partisan elections are associated with reduced performance/output. First, there is a decrease in the number of majority opinions and discretionary opinions written. The point estimate indicates about an 8% decrease in the number of words written, though the estimate is quite noisy. Average length of each opinion decreases as well. Opinion quality is going down slightly (a decrease in positive cites), which means that total cites are going down significantly for all of the measures.

Column 2 shows the effect for non-partisan elections. The number of discretionary opinions and total words written decrease by 7% and 11% respectively, but as in the case of a partisan election the estimate is quite noisy. There aren't significant effects on average opinion quality, except discussion cites. But the combined effect of slightly fewer opinions and slightly lower quality results in statistically significant decreases in most of the total

Table B2: Effect of Being Up For Election (Output and Effort)

Outcome	Partisan Election Year (1)	Non-Partisan Election Year (2)	Uncontested Election Year (3)
Majority Opinions Written	-0.0583*	-0.101	0.0654
	(0.0296)	(0.0658)	(0.0425)
Words in Majority Opinions	-0.0786*	-0.119	0.0896+
	(0.0350)	(0.0739)	(0.0470)
Cases Cited in Majority Opinions	-0.0816*	-0.108	0.100*
	(0.0396)	(0.0771)	(0.0453)
Discretionary Opinions Written	-0.0736**	-0.0699*	0.061
	(0.0285)	(0.0354)	(0.0383)
Words in Discretionary Opinions	-0.208+	-0.154	0.362**
	(0.123)	(0.128)	(0.126)
Cases Cited in Discretionary Opinions	-0.1	-0.131**	0.0711
	(0.0854)	(0.0404)	(0.0873)
Words Per Majority Opinion	-0.0196+	-0.0166	0.0218+
	(0.0108)	(0.0130)	(0.0128)
Cases Cited Per Majority Opinion	-0.0211	-0.00571	0.0331*
	(0.0168)	(0.0204)	(0.0163)
Words Per Discretionary Opinion	-0.142	-0.0881	0.333**
	(0.117)	(0.138)	(0.103)
Cases Cited Per Discretionary Opinion	-0.0377	-0.0499	0.0531
	(0.0678)	(0.0608)	(0.0667)
Treated States	23	17	19
Treated Judges	437	270	277
Election Events	810	517	451

N= 16,084 judge-years. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01. Each row is from a separate regression for the stated outcome variable. Treatment variable is a dummy equaling one for years judge is facing reelection. Regressions include a state-year fixed effect and judge fixed effect, estimated using Stat's reg2hdfe module.

Table B3: Effect of Being Up For Election (Quality and Impact)

Outcome	Partisan Election Year (1)	Non-Partisan Election Year (2)	Uncontested Election Year (3)
Positive Cites Per Opinion	-0.0390*	-0.0222	-0.00167
	(0.0183)	(0.0231)	(0.0186)
Distinguishing Cites Per Opinion	-0.038	-0.0297	0.0252
	(0.0246)	(0.0275)	(0.0354)
Discuss Cites Per Opinion	-0.026	-0.0273+	0.0109
	(0.0186)	(0.0166)	(0.0153)
Quoted Cites Per Opinion	-0.0273+	-0.0229	0.00125
	(0.0151)	(0.0150)	(0.0204)
Out-of-State Cites Per Opinion	-0.0213	0.00414	0.0231
	(0.0177)	(0.0254)	(0.0232)
Total Positive Cites	-0.106*	-0.158+	0.0661
	(0.0416)	(0.0882)	(0.0451)
Total Distinguishing Cites	-0.124*	-0.188*	0.0839
	(0.0552)	(0.0907)	(0.0692)
Total Discuss Cites	-0.0888*	-0.160*	0.0795+
	(0.0444)	(0.0715)	(0.0427)
Total Quoted Cites	-0.0930*	-0.161*	0.0685
	(0.0424)	(0.0725)	(0.0416)
Total Out-of-State Cites	-0.0970*	-0.104	0.0929*
	(0.0448)	(0.0782)	(0.0431)
Treated States	23	17	19
Treated Judges	437	270	277
Election Events	810	517	451

N= 16,084 judge-years. Standard errors clustered by state in parentheses. + $p < .1$, * $p < .05$, ** $p < .01$. Each row is from a separate regression for the stated outcome variable. Treatment variable is a dummy equaling one for years judge is facing reelection. Regressions include a state-year fixed effect and judge fixed effect, estimated using Stat's reg2hdfe module.

Table B4: Effect of Being Up For Election (Additional Outcomes)

Outcome	Partisan Election Year (1)	Non-Partisan Election Year (2)	Uncontested Election Year (3)
Number of Concurrences Written	-0.0282	-0.0227	-0.00666
	(0.0320)	(0.0294)	(0.0348)
Number of Dissents Written	-0.0563**	-0.0621*	0.0616*
	(0.0201)	(0.0254)	(0.0270)
Negative Cites Per Opinion	-0.0244*	-0.0229+	0.018
	(0.0124)	(0.0120)	(0.0188)
Federal Circuit Cites Per Opinion	-0.0154*	-0.0042	-0.00141
	(0.00679)	(0.00532)	(0.00971)
Multiple-Use Cites Per Opinion	-0.0132	-0.0199+	0.00565
	(0.0105)	(0.0114)	(0.0175)
Proportion of Cases Overruled	-0.000786	0.00634	0.00553
	(0.00566)	(0.00889)	(0.00973)
Proportion of Cases Superseded by Statute	-0.00555	-0.000988	0.00865
	(0.00355)	(0.00352)	(0.00636)
Total Negative Cites	-0.115*	-0.169**	0.1
	(0.0505)	(0.0654)	(0.0680)
Total Federal Circuit Cites	-0.101**	-0.0664	0.0262
	(0.0305)	(0.0491)	(0.0468)
Total Multiple-Use Cites	-0.0819*	-0.142**	0.0586
	(0.0413)	(0.0512)	(0.0489)
Cases Overruled	-0.0332	-0.0485	0.0504
	(0.0329)	(0.0385)	(0.0583)
Cases Superseded by Statute	-0.0454*	-0.0251	0.0760+
	(0.0224)	(0.0252)	(0.0413)

cites measures. Though the coefficients are imprecisely estimated, the point estimates are of the same order of magnitude as in the case of a partisan election.

Column 3 shows the effect of uncontested elections. There aren't any negative effects from the electoral cycle in this system. There are actually some positive effects, with an increase in total words written, majority opinion length, length of table of cases, and some of the cites measures.

In Table B4, note that in partisan and non-partisan election years, there is a decrease in the number of dissents, but no effect on concurrences. In uncontested elections, meanwhile, there is a large positive election-year effect on dissents.

Tables B5 and B6 report the estimates from Equation 6.1 for the individual variables. Column 1 estimates the average difference in performance between non-partisan judges and partisan judges. Relative to the partisan judges, the non-partisan judges write shorter opinions, but they are higher quality. Opinions written by non-partisan judges have more positive

Table B5: Effect of Judicial Selection System on Judge Quality (Output and Effort)

Outcome	Non-Partisan Judges Relative to Partisan Judges	Merit-Selected Judges Relative to Partisan Judges	Merit-Selected Judges Relative to Non-Partisan Judges
	(1)	(2)	(3)
Majority Opinions Written	0.0844	-0.148+	-0.111
	(0.0546)	(0.0804)	(0.115)
Words in Majority Opinions	-0.0525	-0.0566	0.00523
	(0.0750)	(0.0569)	(0.147)
Cases Cited in Majority Opinions	-0.0897	0.0186	0.0108
	(0.124)	(0.0581)	(0.126)
Discretionary Opinions Written	0.368	-0.13	0.453*
	(0.373)	(0.147)	(0.208)
Words in Discretionary Opinions	0.579	0.159	1.235
	(0.766)	(0.456)	(0.817)
Cases Cited in Discretionary Opinions	0.891**	0.152	0.837
	(0.265)	(0.341)	(0.536)
Words Per Majority Opinion	-0.135*	0.0918	0.116
	(0.0507)	(0.0595)	(0.0722)
Cases Cited Per Majority Opinion	-0.167	0.166**	0.121
	(0.117)	(0.0571)	(0.0856)
Words Per Discretionary Opinion	0.206	0.32	0.863
	(0.445)	(0.372)	(0.676)
Cases Cited Per Discretionary Opinion	0.509**	0.296	0.447
	(0.106)	(0.245)	(0.389)
Treated States	3	6	3
Treated State-Years	24	86	24
Treated Judges	14	54	16

N= 16,084 judge-years..Estimate of the average difference between judges selected under a new system, relative to to judges selected under the old system, limited to years in which there are at least two judges on the court selected from each system. Regressions include a state-year fixed effect, a full set of dummies for years of experience, and a full set of dummies for starting years. Standard errors clustered by state in parentheses. + $p < .1$, * $p < .05$, ** $p < .01$.

Table B6: Effect of Judicial Selection System on Judge Quality (Quality and Impact)

Outcome	Non-Partisan Judges Relative to Partisan Judges	Merit-Selected Judges Relative to Partisan Judges	Merit-Selected Judges Relative to Non-Partisan Judges
	(1)	(2)	(3)
Positive Cites Per Opinion	0.0692+	0.0690+	0.114
	(0.0382)	(0.0391)	(0.0765)
Distinguishing Cites Per Opinion	0.108+	0.0992+	0.251+
	(0.0582)	(0.0518)	(0.131)
Discuss Cites Per Opinion	0.0355	0.0702*	0.154*
	(0.0652)	(0.0290)	(0.0593)
Quoted Cites Per Opinion	-0.0751	0.0942*	0.239**
	(0.0562)	(0.0442)	(0.0891)
Out-of-State Cites Per Opinion	-0.0217	0.121**	0.194
	(0.0674)	(0.0341)	(0.136)
Total Positive Cites	0.148**	-0.0777	-0.00391
	(0.0553)	(0.0776)	(0.156)
Total Distinguishing Cites	0.418**	-0.0102	0.13
	(0.131)	(0.106)	(0.217)
Total Discuss Cites	0.124	-0.0708	0.0372
	(0.0966)	(0.0648)	(0.152)
Total Quoted Cites	-0.0242	-0.0255	0.123
	(0.0785)	(0.0669)	(0.176)
Total Out-of-State Cites	0.0324	0.0308	0.0847
	(0.145)	(0.0793)	(0.232)
Treated States	3	6	3
Treated State-Years	24	86	24
Treated Judges	14	54	16

N= 16,084 judge-years..Estimate of the average difference between judges selected under a new system, relative to to judges selected under the old system, limited to years in which there are at least two judges on the court selected from each system. Regressions include a state-year fixed effect, a full set of dummies for years of experience, and a full set of dummies for starting years. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01.

cites and more distinguishing than opinions written by partisan judges in the same court and same year. The effect is larger and more significant when we consider the total cites.

Column 2 estimates the performance measure differential for merit-selected judges relative to partisan-selected judges. While merit judges write fewer opinions than the partisan elected judges, they are higher-quality on a range of measures. The merit judges have more caselaw research, as seen in the Length of Table of Cases. They also have more citations on all of our metrics – positive, distinguishing, discussed-in, quoted-in, and out-of-state.

In Column 3 we look at the difference between merit-selected judges and non-partisan-selected judges. First, the merit-selected judges write more discretionary opinions than the non-partisan-selected judges. In terms of opinion quality, merit-selected judges write higher-quality opinions than non-partisan-selected judges for most of the citation measures. For distinguishing cites, discussion cites, and quoted cites, the estimates are statistically significant.

Table B8 reports an alternative specification for the selection-process results. Recall that for the Table 6 estimates, we only included years where there were at least two judges from each system working together on the court. Table B8 includes all years, so it includes years where there is only one judge from one of the systems – so the effect is identified off that judge's difference from the rest of the court. The estimated coefficients are different in this table, but the results all go in the same direction. Non-partisan judges are better than their partisan colleagues. Merit judges are better than their election colleagues.

Tables B9 and B10 report retention-system reform effects on individual performance variables. Column 1 gives the incentive effect on sitting judges of moving from a partisan system to a non-partisan system. We see small negative coefficients for number of majority opinions, total words written, and total discussion cites. These are only marginally significant, however.

Column 2 has the effect of moving from a partisan system to an uncontested system. Here we see no effects on performance.

Finally we look at Column 3. In contrast with the other reforms, moving from non-partisan to uncontested elections is associated with an increase in performance on a range of measures. While the number of majority opinions doesn't change, the number of discretionary opinions does increase. Caselaw research also increases. There are large positive effects on the quality of opinions written, as reflected in positive cites, distinguishing cites, discuss cites, and quoted cites. The total cites measures are more noisy and less significant, but still positive.

51

Table B7: Effect of Judicial Selection System on Judge Quality (Additional Outcomes)

Outcome	Non-Partisan Judges Relative to Partisan Judges (1)	Merit-Selected Judges Relative to Partisan Judges (2)	Merit-Selected Judges Relative to Non-Partisan Judges (3)
Number of Concurrences Written	0.22	-0.065	0.269+
	(0.324)	(0.0602)	(0.141)
Number of Dissents Written	0.301+	-0.106	0.211
	(0.162)	(0.132)	(0.247)
Negative Cites Per Opinion	0.0455**	0.0395**	0.133**
	(0.0159)	(0.0146)	(0.0452)
Federal Circuit Cites Per Opinion	0.00357	0.0249+	0.0677**
	(0.0124)	(0.0141)	(0.0107)
Multiple-Use Cites Per Opinion	-0.0439	0.0484*	0.202**
	(0.0554)	(0.0197)	(0.0671)
Proportion of Cases Overruled	0.0171	0.00511	-0.00256
	(0.0144)	(0.00806)	(0.00931)
Proportion of Cases Superseded by Statute	0.00881	0.00505	0.0365**
	(0.00934)	(0.00498)	(0.0133)
Total Negative Cites	0.246**	0.0213	0.141
	(0.0373)	(0.0540)	(0.110)
Total Federal Circuit Cites	0.0509	0.00275	0.224+
	(0.0875)	(0.0588)	(0.124)
Total Multiple-Use Cites	0.000095	0.0148	0.162
	(0.137)	(0.0404)	(0.181)
Cases Overruled	0.00862	-0.0263	-0.00949
	(0.0542)	(0.0493)	(0.0616)
Cases Superseded by Statute	0.0473	0.0241	0.141
	(0.0324)	(0.0412)	(0.0895)

Table B8: Effect of Judicial Selection System on Judge Quality (All Years)

Outcome	Non-Partisan Judges Relative to Partisan Judges (1)	Merit-Selected Judges Relative to Partisan Judges (2)	Merit-Selected Judges Relative to Non-Partisan Judges (3)
Majority Opinions Written	0.149**	-0.0925	-0.0976
	(0.0451)	(0.0894)	(0.0783)
Discretionary Opinions Written	0.204	-0.173	0.119
	(0.297)	(0.119)	(0.195)
Total Words Written	0.0873	-0.015	-0.0829
	(0.0786)	(0.0646)	(0.141)
Length of Majority Opinion	-0.0706+	0.0899+	0.05
	(0.0390)	(0.0484)	(0.102)
Length of Table of Cases	-0.143	0.124*	0.0534
	(0.0925)	(0.0575)	(0.0903)
Positive Cites Per Opinion			
	-0.01	0.0362	0.125*
Distinguishing Cites Per Opinion	(0.0570)	(0.0379)	(0.0602)
	-0.00849	0.0668	0.123
Discuss Cites Per Opinion	(0.0596)	(0.0466)	(0.0762)
	0.022	0.0389	0.132*
Quoted Cites Per Opinion	(0.0286)	(0.0274)	(0.0586)
	-0.044	0.0478	0.171*
Out-of-State Cites Per Opinion	(0.0517)	(0.0368)	(0.0775)
	0.00122	0.0697**	0.139+
	(0.0373)	(0.0246)	(0.0806)
Total Positive Cites			
	0.134	-0.0689	0.0256
Total Distinguishing Cites	(0.0998)	(0.0800)	(0.140)
	0.168	0.0345	0.0392
Total Discuss Cites	(0.144)	(0.0967)	(0.179)
	0.176*	-0.0454	0.0427
Total Quoted Cites	(0.0825)	(0.0761)	(0.131)
	0.0725	-0.0232	0.0906
Total Out-of-State Cites	(0.116)	(0.0731)	(0.146)
	0.148+	0.034	0.0888
	(0.0781)	(0.0709)	(0.143)

Table B9: Effect of Retention Process (Output and Effort)

Outcome	Partisan Retention to Non-Partisan Retention	Partisan Retention to Uncontested Retention	Non-Partisan Retention to Uncontested Retention
	(1)	(2)	(3)
Majority Opinions Written	-0.153+	0.0602	-0.0899
	(0.0852)	(0.0590)	(0.100)
Words in Majority Opinions	-0.115+	0.0274	-0.0363
	(0.0587)	(0.0418)	(0.0713)
Cases Cited in Majority Opinions	-0.155+	0.033	0.0694
	(0.0780)	(0.0579)	(0.0786)
Discretionary Opinions Written	-0.101	0.0174	0.190*
	(0.0738)	(0.120)	(0.0906)
Words in Discretionary Opinions	0.0974	-0.149	0.743
	(0.299)	(0.479)	(0.459)
Cases Cited in Discretionary Opinions	0.116	-0.0221	0.43
	(0.124)	(0.278)	(0.268)
Words Per Majority Opinion	0.0371	-0.0317	0.0516
	(0.0832)	(0.0376)	(0.0682)
Cases Cited Per Majority Opinion	-0.00619	-0.0272	0.155*
	(0.0781)	(0.0766)	(0.0654)
Words Per Discretionary Opinion	0.219	-0.186	0.597
	(0.279)	(0.409)	(0.461)
Cases Cited Per Discretionary Opinion	0.235*	-0.0457	0.28
	(0.0933)	(0.202)	(0.238)
Treated States	4	8	6
Treated Judges	25	65	35

N= 16,084 judge-years.Estimate of the average treatment effect of changing the judge retention system on incumbent judges at the time of the reform. Regressions include a judge fixed effect, year fixed effect, and state trends. Standard errors clustered by state in parentheses. + $p < .1$, * $p < .05$, ** $p < .01$.

Table B10: Effect of Retention Process (Quality and Impact)

Outcome	Partisan Retention to Non-Partisan Retention	Partisan Retention to Uncontested Retention	Non-Partisan Retention to Uncontested Retention
	(1)	(2)	(3)
Positive Cites Per Opinion	0.0492	-0.0173	0.111*
	(0.0841)	(0.0716)	(0.0551)
Distinguishing Cites Per Opinion	0.0328	-0.0745	0.304**
	(0.0505)	(0.0559)	(0.0949)
Discuss Cites Per Opinion	0.0144	-0.0315	0.0915+
	(0.0718)	(0.0500)	(0.0478)
Quoted Cites Per Opinion	-0.00402	-0.0355	0.102*
	(0.0786)	(0.0474)	(0.0481)
Out-of-State Cites Per Opinion	-0.00613	-0.013	0.0768
	(0.0461)	(0.0357)	(0.0680)
Total Positive Cites	-0.108	0.0574	0.0143
	(0.0966)	(0.0867)	(0.126)
Total Distinguishing Cites	-0.0996	-0.0713	0.308*
	(0.0692)	(0.0833)	(0.134)
Total Discuss Cites	-0.140+	0.0497	-0.00665
	(0.0791)	(0.0876)	(0.125)
Total Quoted Cites	-0.149	0.0288	0.00288
	(0.0955)	(0.0719)	(0.120)
Total Out-of-State Cites	-0.157	0.0629	0.00356
	(0.105)	(0.0706)	(0.121)
Treated States	4	8	6
Treated Judges	25	65	35

N= 16,084 judge-years. Estimate of the average treatment effect of changing the judge retention system on incumbent judges at the time of the reform. Regressions include a judge fixed effect, year fixed effect, and state trends. Standard errors clustered by state in parentheses. $+ p < .1$, $* p < .05$, $** p < .01$.

Table B11: Effect of Retention Process (Additional Outcomes)

Outcome	Partisan Retention to Non-Partisan Retention	Partisan Retention to Uncontested Retention	Non-Partisan Retention to Uncontested Retention
	(1)	(2)	(3)
Number of Concurrences Written	-0.191**	0.00801	0.0557
	(0.0624)	(0.0567)	(0.106)
Number of Dissents Written	-0.0375	0.034	0.0386
	(0.0731)	(0.0772)	(0.0617)
Negative Cites Per Opinion	0.0206	-0.0189	0.134**
	(0.0137)	(0.0300)	(0.0474)
Federal Circuit Cites Per Opinion	-0.0406**	-0.00426	0.00738
	(0.00522)	(0.0113)	(0.0221)
Multiple-Use Cites Per Opinion	-0.0108	-0.0187	0.0694
	(0.0294)	(0.0235)	(0.0501)
Proportion of Cases Overruled	0.0238*	0.00099	-0.00551
	(0.0118)	(0.00827)	(0.00503)
Prop. Cases Superseded by Statute	0.00336	-0.00738*	0.0274**
	(0.00398)	(0.00365)	(0.00710)
Total Negative Cites	-0.00581	0.0377	0.307*
	(0.0654)	(0.0624)	(0.124)
Total Federal Circuit Cites	-0.168**	0.0341	-0.0255
	(0.0465)	(0.0575)	(0.101)
Total Multiple-Use Cites	-0.146+	0.0569	0.0127
	(0.0852)	(0.0709)	(0.103)
Cases Overruled	0.0992	0.0377	-0.136*
	(0.0616)	(0.0348)	(0.0546)
Cases Superseded by Statute	0.0288	-0.0141	0.216**
	(0.0395)	(0.0255)	(0.0715)

In Table B11, note that the effect of the non-partisan to uncontested reform has interesting effects. Under the nonpartisan-to-uncontested reform, judges are overruled less often by later courts, but they are overruled more often by the legislature. There are also higher negative cites per opinion. This may be a sign of greater judicial independence.

Tables B12 and B13 report the election-selection interaction effects on individual performance variables. Column 1a gives the baseline effect of non-partisan elections on non-partisan selected judges. Column 1a has negative effects, which are similar to the estimates for non-partisan elections in Table 5. Column 1b shows the relative effect of non-partisan elections on partisan judges. These are mostly zeros, with a likely spurious positive effect on out-of-state cites.

Column 2a gives the baseline effect of uncontested elections on merit-selected judges. This column is similar to column 3 from Table 5, which gave the average effect of uncontested elections. As with that table, there are actually positive effects estimated for the election-cycle effect in an uncontested system. When one only looks at the merit-selected judges, the effect is stronger.

Column 2b gives the relative effect of uncontested elections on partisan-selected judges. There are significant negative effects. The coefficients are larger in absolute value than the coefficients from Column 2a, meaning that uncontested elections have a negative effect on performance for partisan-selected judges. This means that partisan judges respond in the opposite direction due to elections compared to merit-selected judges.

Finally Column 2c gives the relative effect of uncontested elections for non-partisan-selected judges, relative to merit judges. There aren't any significant differences here.

Notice that the point estimates on out of state citations for partisan judges are very large. When facing a competitive non-partisan election there is a 36% increase in citations, but an 18% *decrease* in an uncontested election. The pattern is consistent with the hypothesis that partisan judges are more sensitive to incentives. This provides some direct evidence that the characteristics of the judges vary by the selection procedure.

Table B14 reports additional outcomes for the interacted study of incentives and selection. The estimates are similar to those in Table 8. Partisan-selected judges respond to uncontested elections with a reduction in negative cites, circuit cites, and multiple-use cites.

B.3 Effect of Retention Process in Election Years

In this section we look at the retention reforms and the electoral cycle together. We look at the interacted effect of a retention process reform in years that a judge is up for election, to

Table B12: Relative Election-Year Effect on Judges Selected by Different Processes (Output and Effort)

Outcome	Effect of Non-Partisan Elections on Non-Partisan-Selected Judges	Relative Effect of Non-Partisan Elections on Partisan-Selected Judges	Effect of Uncontested Elections on Merit-Selected Judges	Relative Effect of Uncontested Elections on Partisan-Selected Judges	Relative Effect of Uncontested Elections on Non-Partisan-Selected Judges
	(1a)	(1b)	(2a)	(2b)	(2c)
Majority Opinions Written	-0.104	0.0996	0.0853+	-0.0994	0.0322
	(0.0673)	(0.129)	(0.0459)	(0.0732)	(0.0954)
Words in Majority Opinions	-0.122	0.103	0.123*	-0.157*	0.0194
	(0.0757)	(0.137)	(0.0495)	(0.0701)	(0.122)
Cases Cited in Majority Ops	-0.112	0.164	0.131**	-0.166*	0.096
	(0.0788)	(0.149)	(0.0475)	(0.0739)	(0.134)
Disc. Opinions Written	-0.0696+	-0.0107	0.0398	0.0684	0.106
	(0.0364)	(0.242)	(0.0520)	(0.0568)	(0.126)
Words in Discretionary Ops	-0.151	-0.135	0.372*	-0.0276	-0.0784
	(0.131)	(0.172)	(0.162)	(0.198)	(0.374)
Cases Cited in Discr. Ops	-0.128**	-0.107	0.1	-0.0966	-0.14
	(0.0420)	(0.130)	(0.102)	(0.178)	(0.352)
Words Per Majority Opinion	-0.0167	0.00379	0.0347*	-0.0544*	-0.0177
	(0.0130)	(0.0616)	(0.0151)	(0.0215)	(0.0489)
Cases Cited Per Maj. Op	-0.00733	0.0638	0.0429+	-0.0607+	0.0592
	(0.0201)	(0.104)	(0.0228)	(0.0347)	(0.0567)
Words Per Discy Opinion	-0.0849	-0.127	0.367***	-0.105	-0.188
	(0.142)	(0.243)	(0.126)	(0.159)	(0.352)
Cases Cited Per Disc Op	-0.0474	-0.101	0.102	-0.165	-0.228
	(0.0623)	(0.205)	(0.0724)	(0.123)	(0.306)
Treated States	2	2	11	8	4
Treated Judges	7	4	119	51	10
Election Events	8	5	201	90	16

N= 16,084 judge-years. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01. Each row is from a separate regression for the stated outcome variable. The estimated coefficient is a dummy equaling one for years judge is facing reelection, interacted with a dummy for if the judge is selected under the new selection system. Regressions include a state-year fixed effect, judge fixed effect, and the baseline

Table B13: Relative Election-Year Effect on Judges Selected by Different Processes (Quality and Impact)

Outcome	Effect of Non-Partisan Elections on Non-Partisan-Selected Judges	Relative Effect of Non-Partisan Elections on Partisan-Selected Judges	Effect of Uncontested Elections on Merit-Selected Judges	Relative Effect of Uncontested Elections on Partisan-Selected Judges	Relative Effect of Uncontested Elections on Non-Partisan-Selected Judges
	(1a)	(1b)	(2a)	(2b)	(2c)
Positive Cites Per Opinion	-0.0254	0.127	0.0214	-0.098	-0.0289
	(0.0232)	(0.120)	(0.0172)	(0.0813)	(0.0483)
Disting. Cites Per Opinion	-0.0336	0.153	0.0416	-0.0899	0.0549
	(0.0277)	(0.140)	(0.0317)	(0.0870)	(0.119)
Discuss Cites Per Opinion	-0.0285+	0.0463	0.0216	-0.037	-0.0439
	(0.0158)	(0.138)	(0.0166)	(0.0411)	(0.0393)
Quoted Cites Per Opinion	-0.0255+	0.101	0.022	-0.0759	-0.0701
	(0.0144)	(0.179)	(0.0212)	(0.0606)	(0.0633)
Out-of-State Cites Per Op.	-0.00167	0.229*	0.0402+	-0.0465	-0.12
	(0.0244)	(0.116)	(0.0232)	(0.0530)	(0.101)
Total Positive Cites	-0.164+	0.258	0.110*	-0.202*	0.00656
	(0.0903)	(0.200)	(0.0519)	(0.0814)	(0.0951)
Total Distinguishing Cites	-0.197*	0.364	0.132*	-0.240+	0.0794
	(0.0925)	(0.294)	(0.0639)	(0.127)	(0.159)
Total Discuss Cites	-0.164*	0.159	0.113*	-0.147*	-0.0152
	(0.0732)	(0.211)	(0.0456)	(0.0637)	(0.0986)
Total Quoted Cites	-0.166*	0.215	0.113***	-0.192**	-0.0455
	(0.0742)	(0.284)	(0.0439)	(0.0669)	(0.0888)
Total Out-of-State Cites	-0.115	0.434*	0.132***	-0.147*	-0.12
	(0.0784)	(0.171)	(0.0478)	(0.0722)	(0.0944)
Treated States	2	2	11	8	4
Treated Judges	7	4	119	51	10
Election Events	8	5	201	90	16

N= 16,084 judge-years. Standard errors clustered by state in parentheses. + p < .1, * p < .05, ** p < .01. Each row is from a separate regression for the stated outcome variable. The estimated coefficient is a dummy equaling one for years judge is facing reelection, interacted with a dummy for if the judge is selected under the new selection system. Regressions include a state-year fixed effect, judge fixed effect, and the baseline

Table B14: Relative Election-Year Effect on Judges Selected by Different Processes (Additional Outcomes)

Outcome	Effect of Non-Partisan Elections on Non-Partisan-Selected Judges	Relative Effect of Non-Partisan Elections on Partisan-Selected Judges	Effect of Uncontested Elections on Merit-Selected Judges	Relative Effect of Uncontested Elections on Partisan-Selected Judges	Relative Effect of Uncontested Elections on Non-Partisan-Selected Judges
	(1a)	(1b)	(2a)	(2b)	(2c)
Number of Concurrences Written	-0.0196	-0.127*	-0.0243	0.0542	0.132
	(0.0301)	(0.0578)	(0.0455)	(0.0657)	(0.134)
Number of Dissents Written	-0.0659*	0.155	0.0635+	-0.014	0.0122
	(0.0262)	(0.144)	(0.0336)	(0.0667)	(0.146)
Negative Cites Per Opinion	-0.0243*	0.0573	0.0254	-0.0438	0.0117
	(0.0118)	(0.0846)	(0.0213)	(0.0490)	(0.0423)
Federal Circuit Cites Per Opinion	-0.0058	0.0654	0.0116	-0.0611**	-0.0302
	(0.00591)	(0.0707)	(0.0111)	(0.0231)	(0.0497)
Multiple-Use Cites Per Opinion	-0.0232*	0.136*	0.0107	-0.0303	0.00815
	(0.0105)	(0.0669)	(0.0201)	(0.0355)	(0.0554)
Proportion of Cases Overruled	0.00626	0.00308	0.00515	-0.0001	0.00694
	(0.00911)	(0.0164)	(0.0120)	(0.0112)	(0.0129)
Proportion of Cases Superseded by Statute	-0.000715	-0.0111	0.0109	-0.0134	0.00431
	(0.00357)	(0.0151)	(0.00779)	(0.0135)	(0.0200)
Total Negative Cites	-0.176**	0.277	0.157*	-0.276*	-0.106
	(0.0669)	(0.231)	(0.0754)	(0.133)	(0.163)
Total Federal Circuit Cites	-0.0733	0.281	0.0836	-0.279**	-0.106
	(0.0509)	(0.307)	(0.0543)	(0.0565)	(0.117)
Total Multiple-Use Cites	-0.153**	0.444**	0.0950+	-0.201**	0.011
	(0.0512)	(0.127)	(0.0504)	(0.0559)	(0.123)
Cases Overruled	-0.0493	0.0286	0.0496	-0.0431	0.15
	(0.0393)	(0.208)	(0.0701)	(0.0590)	(0.0986)
Cases Superseded by Statute	-0.0194	-0.232+	0.0962*	-0.0899	-0.0634
	(0.0257)	(0.141)	(0.0460)	(0.114)	(0.180)

see whether the observed effect is due to changes in campaigning behavior.

This subsection describes the empirical strategy for looking at the election-year effects of the retention process reforms. The regression approach combines the approach from Subsection 6.1 on the electoral cycle with the approach from Subsection 6.2 on the retention process reforms.

As before, we have the vector of election dummies E_{ist} that equal one when judge i from state s is up for election at year t, with a separate set of dummies for each retention system. We have the vector of treatment indicators for the retention treatments, R_{st}, which go into effect relative to the 10-year treatment window $\bar{R}_{st}$ as described in Subsection 6.2.1. As in Subsection 6.2, our regressions include year fixed effects, judge fixed effects, and state-specific time trends.

The regressions include the full set of interactions $E_{ist}R'_{st}$. Specifically, we estimate

$$y_{ist} = \text{TIME}_t + \text{JUDGE}_i + \text{STATE}_s \times t + \bar{R}'_{st}\bar{\rho} + E_{ist}R'_{st}\rho + \epsilon_{ist} \tag{B.1}$$

where again we cluster standard errors by state. The components of ρ include the effects of the reform in non-election years ($E = 0$) as well as the effects in election years ($E = 1$). Because the interactions are included, this is relative to the election-year average before the reform.

An additional specification is reported in Appendix Table B18. In that specification, the election-year effect is measured relative to a baseline for all years after the reform (rather than looking at non-election years and election years separately).

The results on the effect of the retention process reform on election and non-election years are reported in Table 7. The "a" columns report the effect in non-election years. The "b" columns report the effect in election years. We report the results for partisan to non-partisan, and non-partisan to uncontested. The results from partisan to uncontested (they are all zeros) are in Appendix Table B17.

Columns 1a and 1b look at the election/non-election effects for the partisan to non-partisan reform. These results bolster what was found in Subsection 6.2. We see significant negative effects when we look at election years specifically. There is a decrease in majority opinions written, total positive cites, total discuss cites, and total quote cites.

Columns 2a and 2b show the effect of moving from nonpartisan-to-uncontested reform. The baseline effect is comparable to the estimate from Table 6, with clear improvements in opinion quality. Moreover, as seen in Column 2b, this effect is even stronger in election years.

Table B15: Effect of Retention Process in Election Years

Outcome	Partisan Retention to Non-Partisan Retention		Non-Partisan Retention to Uncontested Retention	
	Non-Election Years (1a)	Election Years (1b)	Non-Election Years (2a)	Election Years (2b)
Majority Opinions Written	-0.137	-0.252**	-0.0972	-0.0771
	(0.0839)	(0.0882)	(0.0968)	(0.0863)
Discretionary Opinions Written	-0.0905	-0.0898	0.136+	0.183*
	(0.0639)	(0.0644)	(0.0762)	(0.0754)
Total Words Written	-0.0965+	-0.236	-0.022	0.0609
	(0.0536)	(0.142)	(0.0694)	(0.0687)
Length of Majority Opinion	0.0395	0.0187	0.047	0.0757
	(0.0795)	(0.116)	(0.0678)	(0.0805)
Length of Table of Cases	-0.0165	0.0127	0.135*	0.201**
	(0.0613)	(0.144)	(0.0604)	(0.0736)
Positive Cites Per Opinion	0.0377	0.0596	0.0919+	0.177*
	(0.0739)	(0.0754)	(0.0461)	(0.0672)
Distinguishing Cites Per Opinion	0.0149	0.0851	0.221**	0.291**
	(0.0347)	(0.0893)	(0.0707)	(0.0740)
Discuss Cites Per Opinion	0.00857	0.0184	0.0678+	0.119**
	(0.0530)	(0.0686)	(0.0382)	(0.0419)
Quoted Cites Per Opinion	-0.00617	-0.000161	0.0712+	0.155**
	(0.0600)	(0.0768)	(0.0392)	(0.0543)
Out-of-State Cites Per Opinion	-0.0147	0.0548	0.0513	0.0875
	(0.0361)	(0.0511)	(0.0503)	(0.0572)
Total Positive Cites	-0.0914	-0.183*	-0.00721	0.125
	(0.101)	(0.0858)	(0.127)	(0.103)
Total Distinguishing Cites	-0.0882	-0.144	0.274*	0.403**
	(0.0624)	(0.182)	(0.129)	(0.119)
Total Discuss Cites	-0.118	-0.252**	-0.0272	0.0699
	(0.0812)	(0.0870)	(0.124)	(0.0972)
Total Quoted Cites	-0.125	-0.290*	-0.0251	0.12
	(0.104)	(0.124)	(0.116)	(0.0995)
Total Out-of-State Cites	-0.139	-0.127	-0.0156	0.0527
	(0.106)	(0.114)	(0.111)	(0.114)
Treated States	4		6	
Treated Judges	25		35	

N= 16,084 judge-years..Each row is a separate regression. The "a" columns give the baseline effect of changing the judge retention system on incumbent judges at the time of the reform, while the "b" columns give the additional effect during judge election years. Regressions include a judge fixed effect, year fixed effect, and state trends. Standard errors clustered by state in parentheses. $+ p < .1, * p < .05, ** p < .01$.

Table B16: Effect of Retention Process in Election Years (Additional Outcomes)

| Outcome | Partisan Retention to Non-Partisan Retention | | Non-Partisan Retention to Uncontested Retention | |
	Non-Election Years (1a)	Election Years (1b)	Non-Election Years (2a)	Election Years (2b)
Number of Concurrences Written	-0.184**	-0.25	0.0619	0.0219
	(0.0541)	(0.174)	(0.0998)	(0.151)
Number of Dissents Written	-0.0427	-0.00385	0.0388	0.0925+
	(0.0781)	(0.0564)	(0.0622)	(0.0544)
Negative Cites Per Opinion	0.0244+	0.00355	0.133**	0.151**
	(0.0130)	(0.0389)	(0.0490)	(0.0403)
Federal Circuit Cites Per Opinion	-0.0405**	-0.0372	0.0106	-0.0234
	(0.00662)	(0.0254)	(0.0209)	(0.0324)
Multiple-Use Cites Per Opinion	-0.016	0.0306	0.0645	0.0909+
	(0.0291)	(0.0458)	(0.0502)	(0.0527)
Proportion of Cases Overruled	0.0191*	0.0594	-0.00372	-0.0134
	(0.00837)	(0.0535)	(0.00589)	(0.00986)
Prop. Cases Superseded by Statute	0.00463	-0.00642	0.0257**	0.0387**
	(0.00449)	(0.00577)	(0.00776)	(0.00802)
Total Negative Cites	0.0127	-0.147	0.282*	0.338*
	(0.0658)	(0.131)	(0.134)	(0.135)
Total Federal Circuit Cites	-0.151**	-0.309*	-0.0176	-0.177
	(0.0490)	(0.121)	(0.0971)	(0.158)
Total Multiple-Use Cites	-0.129	-0.295*	-0.0215	0.0576
	(0.0852)	(0.117)	(0.109)	(0.110)
Cases Overruled	0.0976+	0.0904	-0.147*	-0.123
	(0.0568)	(0.141)	(0.0620)	(0.0841)
Cases Superseded by Statute	0.0459	-0.125**	0.197**	0.248+
	(0.0472)	(0.0431)	(0.0732)	(0.129)

The non-partisan-to-uncontested results show that there is both a durable effect across the whole term, as well as an especially large effect from relieving electoral campaigning demands. Overall, these results substantiate that the effects of these reforms are due in part to the weakening of electoral demands.

Table B17 shows the estimates for the partisan-to-uncontested reform by election year, which were left out of Table 7. These are almost all zeros – there is no within-judge electoral effect of this reform.

Table B18 extends the analysis from Table 7 but looks at the *relative* effect of election years to a baseline after the reform – rather than looking at the effects on non-election years and election years separately. In this specification, we lose statistical significance in the partisan-to-nonpartisan reform In the partisan to uncontested reform, we see a couple of more positive effects of the reform in election years. In the nonpartisan-to-uncontested meanwhile, the coefficients are positive in both columns, meaning that there is a statistically significant additional positive effect during election years.

The partisan-to-non-partisan reform shows the stronger electoral demands from non-partisan elections. Relative to the case before the reform (partisan elections), the election years in non-partisan elections are more demanding and cause a larger reduction in performance. This is consistent with those elections being more competitive. Particularly noteworthy is the 25% decline in writing majority opinions in election years. This is consistent with anecdotal evidence that in election years other judges on the court help reduce the load on judges up for re-election. It seems that this pro-social behavior is more evident on non-partisan benches.

Table B17: Effect of Partisan-to-Uncontested Retention Reform in Election Years

Outcome	Non-Election Years (2a)	Election Years (2b)	Outcome	Non-Election Years (2a)	Election Years (2b)
Majority Opinions Written	0.0531 (0.0611)	0.117+ (0.0651)	Number of Concurrences Written	0.00274 (0.0556)	0.0428 (0.0691)
Discretionary Opinions Written	-0.00694 (0.0902)	0.168 (0.117)	Number of Dissents Written	0.0158 (0.0723)	0.133 (0.127)
Total Words Written	0.0062 (0.0441)	0.0531 (0.0728)	Negative Cites Per Opinion	-0.019 (0.0303)	-0.0258 (0.0364)
Length of Majority Opinion	-0.0263 (0.0385)	-0.061 (0.0385)	Federal Circuit Cites Per Opinion	-0.00115 (0.0113)	-0.0177 (0.0172)
Length of Table of Cases	-0.0166 (0.0711)	-0.0668 (0.0783)	Multiple-Use Cites Per Opinion	-0.0144 (0.0245)	-0.0365 (0.0326)
Positive Cites Per Opinion	-0.00175 (0.0601)	-0.102 (0.0889)	Proportion of Cases Overruled	0.00146 (0.00867)	-0.000169 (0.00919)
Distinguishing Cites Per Opinion	-0.0525 (0.0412)	-0.0956 (0.0690)	Proportion of Cases Superseded by Statute	-0.00836* (0.00331)	-0.00244 (0.00714)
Discuss Cites Per Opinion	-0.0155 (0.0384)	-0.0605 (0.0403)	Total Negative Cites	0.0432 (0.0637)	0.0485 (0.151)
Quoted Cites Per Opinion	-0.0175 (0.0355)	-0.0688 (0.0487)	Total Federal Circuit Cites	0.0453 (0.0638)	0.0112 (0.0829)
Out-of-State Cites Per Opinion	-0.00572 (0.0273)	-0.0228 (0.0463)	Total Multiple-Use Cites	0.0737 (0.0800)	0.0236 (0.131)
Total Positive Cites	0.0732 (0.0957)	-0.0116 (0.165)	Cases Overruled	0.0361 (0.0343)	0.0765 (0.0588)
Total Distinguishing Cites	-0.048 (0.0809)	-0.0783 (0.177)	Cases Superseded by Statute	-0.0211 (0.0320)	0.0548 (0.0753)
Total Discuss Cites	0.0565 (0.0906)	0.0327 (0.107)			
Total Quoted Cites	0.0405 (0.0734)	0.00651 (0.126)			
Total Out-of-State Cites	0.0621 (0.0756)	0.0966 (0.0978)			

Table B18: Relative Effect of Retention Process in Election Years

Outcome	Partisan Retention to Non-Partisan Retention		Partisan Retention to Uncontested Retention		Non-Partisan Retention to Uncontested Retention	
	Baseline Effect (1a)	Election Years (1b)	Baseline Effect (2a)	Election Years (2b)	Baseline Effect (3a)	Election Years (3b)
Majority Opinions Written	-0.104 (0.0676)	-0.0747 (0.166)	-0.0175 (0.101)	0.0957* (0.0420)	-0.117 (0.126)	-0.0132 (0.118)
Discretionary Opinions Written	-0.0461 (0.0750)	-0.0715 (0.132)	-0.0321 (0.0860)	0.188* (0.0723)	0.104 (0.0849)	0.0329 (0.0921)
Total Words Written	-0.0764+ (0.0442)	-0.0483 (0.304)	-0.0689 (0.0879)	0.0668 (0.0460)	-0.0436 (0.0975)	0.0504 (0.126)
Length of Majority Opinion	0.0226 (0.0657)	0.0274 (0.135)	-0.0242 (0.0298)	-0.0412 (0.0272)	0.0426 (0.0704)	0.0318 (0.0434)
Length of Table of Cases	-0.0338 (0.0510)	0.159 (0.188)	-0.0318 (0.0541)	-0.0142 (0.0462)	0.148* (0.0683)	0.0722+ (0.0389)
Positive Cites Per Opinion	0.0328 (0.0688)	0.0609 (0.0713)	-0.0195 (0.0470)	-0.145+ (0.0752)	0.110* (0.0512)	0.0720+ (0.0395)
Distinguishing Cites Per Opinion	-0.00419 (0.0310)	0.119 (0.123)	-0.0557 (0.0358)	-0.0392 (0.0556)	0.245** (0.0782)	0.105** (0.0279)
Discuss Cites Per Opinion	0.00112 (0.0485)	0.02 (0.0618)	-0.0241 (0.0292)	-0.0195 (0.0240)	0.0764+ (0.0432)	0.0936* (0.0420)
Quoted Cites Per Opinion	-0.0186 (0.0504)	0.0406 (0.0694)	-0.0252 (0.0274)	-0.0659 (0.0397)	0.0851+ (0.0477)	0.118** (0.0370)
Out-of-State Cites Per Opinion	-0.0137 (0.0353)	0.0754 (0.0814)	-0.0189 (0.0223)	0.024 (0.0405)	0.0573 (0.0533)	0.032 (0.0275)
Total Positive Cites	-0.0677 (0.100)	0.019 (0.250)	-0.0526 (0.126)	-0.105 (0.156)	0.00315 (0.151)	0.0707 (0.134)
Total Distinguishing Cites	-0.091 (0.0727)	0.0472 (0.437)	-0.146+ (0.0815)	0.0413 (0.172)	0.315* (0.124)	0.165 (0.146)
Total Discuss Cites	-0.0955 (0.0898)	-0.0646 (0.249)	-0.0502 (0.104)	0.0502 (0.0628)	-0.0272 (0.146)	0.11 (0.140)
Total Quoted Cites	-0.112 (0.0956)	-0.0486 (0.282)	-0.0652 (0.100)	-0.0259 (0.107)	-0.0148 (0.136)	0.135 (0.143)
Total Out-of-State Cites	-0.122 (0.104)	0.176 (0.291)	-0.0555 (0.125)	0.192** (0.0682)	-0.0132 (0.133)	0.0043 (0.127)