Judicial Elections and the Merit Plan: Adverse Selection and Moral Hazard Issues in State Appellate Courts

Kenneth Coriale

November 10, 2019

Abstract

I examine the effect of the merit plan for selecting and retaining judges on judicial decision-making. In the merit plan, a nominating commission nominates three candidates, one of whom is chosen by the Governor to serve as a judge; judges are retained through unopposed retention elections. Using panel data on state appellate courts from 1952 to 1995, I identify and examine moral hazard and adverse selection effects associated with transitions between selection and retention methods. I contribute to the literature by investigating the relationship between judicial selection and retention methods and judicial decisions, as well as by introducing methodology that allows me to simultaneously estimate separate moral hazard and adverse selection effects. I find that judges appointed through the merit plan are 4% more likely to support civil suit plaintiffs, while the institution of retention elections has no statistically significant effect on civil suit decisions. In criminal cases, by contrast, I find mixed evidence of an effect of merit selection on support for criminal prosecutors, but do find suggestive evidence that judges appointed before the merit plan are 10% less likely to support criminal prosecutors after retention elections are introduced. All four of these findings conform to theoretical predictions based on the incentives faced by judges.

I. Introduction

The judicial branch forms a vital part of the American system of government. The role of the judicial branch is to "interpret the meaning of laws, apply laws to individual cases, and decide if laws violate the Constitution" (USA.gov 2019). In interpreting and applying the laws, a key characteristic of the judiciary is that it does so impartially. In the interest of keeping judges impartial, states seek to balance two elements - accountability, to ensure the judges do not indulge their own biases, and independence, to ensure the system does not induce bias in the judges. States are free to choose and change their own judicial selection and retention systems, and different states use different methods that are designed to balance these elements of accountability and independence.

The goal of this paper is to examine the effect of these different selection and retention methods on judicial decisions. To do so, I use state-level variation in judicial selection and retention methods over time, within a difference-in-difference framework, in order to analyze how a state's choice of judicial selection and retention methods affects how a judge votes in a given case. In so doing, I am also able to estimate separate adverse selection and moral hazard effects based on changes in, respectively, judicial selection methods and judicial retention methods. This paper thus informs both the choice of methods of judicial selection and retention by each state and the broader literature on adverse selection and moral hazard for public officials.

In the United States, the federal government selects judges through presidential appointment and senate confirmation, who then retain their positions for life.[1] However, this system is quite rare among state governments. At the state level, the debate over judicial selection methods has largely focused on three methods: partisan elections, nonpartisan elections, and nomination by a committee on the basis of merit. Similarly, the debate over judicial retention methods has focused on partisan elections, nonpartisan elections, and unopposed elections (which are often called retention elections). These debates have been quite fervent on both sides. Many judges, in particular, have consistently decried judicial elections as making judges into politicians - Sandra Day O'Connor, for example, has spent much of her post-Supreme Court work advocating for the Merit Plan, a system where judges are initially nominated by a committee based on merit and retained through unopposed elections.[2] Several academics have, with similar fervor, opposed the Merit Plan, principally

[1]Federal judges in the United States retain their office during "good behavior," which in practice means they can only lose their position in the event of criminal malfeasance.

[2]Together with the Institute for the Advancement of the American Legal System, O'Connor has gone so

on the grounds that judicial elections are no worse for independence than the Merit Plan and substantially better for accountability (Bonneau and Hall 2009).

A significant literature exists on the relationship between electoral pressure and judicial rulings, where electoral pressure is the pressure a judge feels to prepare themselves for facing reelection and varies across different retention methods and the amount of time until the next election. Several papers show that judges respond to electoral pressure by expropriating out-of-state litigants in tort cases (Tabarrok and Helland 1999; Helland and Tabarrok 2002), by pandering on hot-button issues (Canes-Wrone, Clark, and Park 2012), and by issuing harsher criminal sentences near elections (Lim 2013). Several papers also find that these electoral pressures are highest in nonpartisan elections, though even appointed judges may respond to electoral cycles for behavioral reasons (Lim, Snyder, and Stromberg 2015; Berdejo and Chen 2013).

A related literature has sought to investigate the relationship between judicial selection methods and the quality of judges on the bench. Using various potential measures of judge quality, such as number of opinions written, words per opinion, citations by other judges, and frequency of having rulings overturned by higher courts, these papers find candidate quality is important for nonpartisan elections, but crowded out by party affiliation in partisan elections (Lim and Snyder 2015) and that elected judges write many more opinions than appointed judges, while receiving only slightly fewer citations (Choi, Gulati, and Posner 2010).

While all of this previous work has approached these issues from a cross-sectional perspective, it is easy to think that a state's choice of a method of judicial selection is endogenous to several elements of a state's economic and legal environment. For example, we might think that cities with large concentrations of businesses and start-ups (e.g. New York City, San Francisco) have large numbers of high-quality lawyers. The large number of high-quality lawyers may lead to more competition for spots on the state high court, resulting in higher quality judges and may also make the state bar association more powerful, leading to the state being more likely to adopt the Merit Plan. A cross-sectional approach, then, may lead us to incorrectly conclude that there is a causal relationship between the Merit Plan and judicial quality.

Responding to this criticism of the cross-sectional approach, Ash and MacLeod (2019) introduce the usage of variation in judicial selection and retention methods over time to

far as to create a comprehensive model for the institution of the Merit Plan (see O'Connor 2014).

form a causal identification strategy. This usage of variation over time clearly marks the paper as the closest to my own. Ash and MacLeod apply this variation over time to the previous literature on judicial quality. They examine separately the effects of being up for re-election, being appointed under the Merit Plan, or being subject to retention elections on various potential measures of judge quality. Ash and MacLeod find that, according to their measures, selection of judges under the Merit Plan or under nonpartisan elections leads to higher quality than those selected through partisan elections. For retention, they find that those retained in retention elections are higher quality than those retained in non-partisan elections.

My paper's main contribution to the literature here is three-fold. The first contribution is to investigate the relationship between judicial selection and retention methods and judicial rulings. This is an important synthesis between the two fields of literature mentioned. Current research on electoral pressure and judicial rulings is unable to examine the effect of the selection and retention methods themselves as the choice of the methods likely correlates with time-invariant state-level characteristics. Thus, my usage of variation over time allows this field to expand to address the effects of judicial selection and retention methods directly. As for research on the relationship between judicial selection and retention methods and judge quality, my concern is that it is both theoretically uncertain what makes a judge "high quality" and empirically unclear how to determine which judges are "high quality." The theoretical uncertainty comes from uncertainty over the proper balance between independence and accountability (i.e. how much weight judges should put on the beliefs of the broader public, rather than their own beliefs), while the empirical uncertainty comes from difficulty comparing observable measures to whatever has been determined to be true judicial quality. Measures such as number of opinions written and number of citations by other judges may or may not correlate with actually being a high quality judge - judges may actually sacrifice quality (i.e. making the correct decision) to write more opinions, and citations may have more to do with being a part of the "old boys network" or with the citation culture within a state than with making quality decisions. Given this uncertainty over the meaning and measurement of judge quality, I choose instead to ask more basic questions - Do different judicial selection and retention methods lead to judges making different decisions? If so, how do the adverse selection and moral hazard effects of these methods change the decisions? In answering these basic questions, I hope to provide some evidence for future researchers to build upon in comparing different judicial selection and retention methods.

My second contribution to the literature is in my methodology. Using variation over time within a state, I am able to identify the adverse selection and moral hazard effects of changes in judicial selection and retention methods separately. This is possible because I am able to estimate these effects using two different comparisons - the adverse selection effect is based on the comparison between judges on the court together who were appointed under different judicial selection methods, while the moral hazard effect is based on the comparison of the same judge before and after a change in retention method. Because my identification of these effects comes from different comparisons, I am able to estimate both the adverse selection and moral hazard effects simultaneously. This is important because, if I instead were to estimate those effects separately, I would have difficulty identifying the true effects of each underlying policy, as changes in selection and retention policies are often implemented simultaneously. Even if I instead estimated the effects separately and split the sample along the identification axis I have described here (e.g. if I estimated the moral hazard effect on the sample of only judges present before and after a change in retention method), the result would still be inferior to the method I lay out in this paper, as my method uses the full sample to better estimate the state and time fixed effects in the estimation, and thus to better estimate the adverse selection and moral hazard effects of interest.

My third contribution to the literature is the development of a large new dataset. I scrape and parse the set of appellate court decisions in the states and years indicated and develop a never-before-used set of data on judge votes in appellate cases, case characteristics, and case outcomes. This data is likely to prove useful in answering a number of additional questions on judicial decision-making. I describe this data further in Section III.

II. Institutional Background

The individual states of the United States have consistently struggled to settle upon a system for selecting and retaining judges that ensures the impartiality of the judiciary. At the founding of the United States, all thirteen original states had judges that were appointed by either the legislature or the governor. The federal government also adopted this method of appointment, which continues to this day. During the nineteenth century, many states felt the need to increase judicial independence from the influence of the legislative and executive branches, as well as increase accountability to the public, so they transitioned to partisan judicial elections. However, with the growth of strong political machines in the late nineteenth century, some states worried that judges had become dependent on political parties (and party bosses). During the Progressive era of the early 1900's, several states moved to non-partisan judicial elections as a means of increasing judicial independence. Nevertheless, many states still felt that the need to campaign for election limited the independence of the judiciary, as well as creating the "unseemly" impression that judges were politicians (American Judicature Society 2011).

In response to these concerns, what is commonly known as the Merit (or Missouri) Plan was created. This plan was first implemented in Missouri, in 1940, and has since spread to many other states. The Merit Plan seeks to ensure judicial independence with a panel that nominates judges based on merit. The (typically three) nominees are then sent to the governor, who must select one for approval. The panel composition is generally a mixture of citizens appointed by the governor or legislature (usually required to be non-lawyers) and lawyers appointed by the state bar association. In Missouri, often considered the model state, the panel for appellate courts consists of three lawyers who are elected by members of the Missouri Bar Association, three non-lawyers selected by the governor, and a justice of the Missouri Supreme Court elected by the members of that court (American Judicature Society 2012). Exact panel composition varies among states using the Merit Plan, but Missouri's panel constitutes a reasonable representation of the composition of an average panel.

Under the Merit Plan, after a judge is appointed to office, the judge is subject to periodic retention elections to ensure some degree of accountability. A retention election differs from a normal election in that there is no opposition candidate. Instead, voters are asked a yes or no question - should the judge be retained in office? If the voters choose yes, then the judge will serve another term. If not, then the nominating panel will prepare a new slate of nominees for the vacancy (American Judicature Society 2011).

It should be clear that the Merit Plan may affect the incentives for judges and potential judges in ways that differ from the incentives under a system of judicial elections. Of particular interest is the fact that these systems have implications for both the selection effect for judges and the moral hazard behavior of judges in office. With regard to moral hazard, while judges under an electoral system must consider the effect of their case vote on their future reelection prospects, judges under the Merit Plan face virtually no electoral effects of their case voting as judicial candidates in retention elections are virtually always retained.[3] With regard to the selection effect, the type of judges who are successful at winning an initial election, and even the type who are willing to apply for a judgeship requiring election, may differ from those selected through the process of merit appointment. Given these differing incentives in different selection and retention systems, I use state-level variation in judicial selection and retention methods over time to identify both the adverse selection and moral hazard effects of a change in judicial selection and retention systems.

While I use this variation to test a broad range of potential effects, the changing incentives do yield two theoretical predictions for my outcomes. The first is that the larger influence wielded by the members of the bar association under the Merit Plan (with the bar association choosing a large portion of the nominating commission, rather than constituting only a very small portion of the broader electorate) may lead to Merit Plan judges who are more pro-plaintiff in civil cases. Furthermore, since this effect comes from the change in the selection process, we should expect to see this effect through the adverse selection channel.[4]The second theoretical prediction is that, as judges may attempt to curry favor with voters by being tough on crime under an electoral system (see Lim 2013 for a further discussion and more evidence for this theoretical prediction), the institution of the Merit Plan may result in judges making fewer decisions in favor of the prosecution in criminal cases. Furthermore, this effect should be seen in the moral hazard channel, since the effect is coming from the change in retention method.

III. Data

For my analysis, I use variation over time in judicial selection and retention methods within a state. I observe twelve different combinations of selection method and retention method over the time period studied (1952-1995). In practice, with one exception,[5] states move

[3]Studies estimate that greater than ninety-eight percent of those who seek retention are retained (see Owens et al. [2015] for a discussion of the current research).

[4]We may expect the bar association to favor pro-plaintiff judges as larger civil awards mean larger fees for plaintiff attorneys, who often receive a percentage of the judgement. Defense attorneys are not harmed by these larger civil awards, as they receive a payment that is not dependent on the result, and in fact should benefit as we would expect the increase in civil awards to increase demand for defense services.

[5]Tennessee changed from the Merit Plan to partisan elections in 1974, three years after having replaced partisan elections with the Merit Plan. Tennessee later moved the high court back to the Merit Plan in 1994. Throughout this process, Tennessee has experienced continuing controversy over the Merit Plan and a string of lawsuits claiming the system violates the state constitution (in contrast to most other states, Tennessee implemented the Merit Plan as a law, rather than a constitutional amendment). This uncertainty continues even today, as there is still debate in the state over the selection method moving forward and the legislative and executive branches are currently in disagreement on the meaning of current law and on whom

monotonically in the direction described in Section II (from the federal system to partisan elections, then nonpartisan elections, then the Merit Plan). For my purposes, I focus on states that move from partisan or nonpartisan elections to the Merit Plan. In this paper, I define a state as using the Merit Plan if it fulfills two criteria: judicial appointment is through a merit commission that nominates candidates to the governor, and judicial retention is through retention elections.[6] Data on state judicial selection methods and transitions used in this paper come from the American Judicature Society.

My data set consists of seven treated states, which contain a total of eleven courts. All courts included for analysis are state intermediate appellate courts or state high courts, as Merit Plan usage for trial court judges is rare and case data for the transitional period for trial court judges is not currently available. It is also important to note that, except for 2 counties in Arizona, trial court judges do not transition to the merit plan at the same time as appellate courts. These seven states are chosen as they are the only U.S. states that receive the treatment during the study period. The other forty-three states either have never used the Merit Plan as defined or switched to the Merit Plan before 1960.

My data set also contains five control states. These states are chosen by a two-step process. The first step is to select states in the same U.S. Census regions as treatment states (these are the Mountain, West North Central, East North Central, and West South Central census regions). The second step is to select states from those Census regions that are completely untreated during the period studied, i.e. those who do not change their selection and retention methods and use either partisan elections, nonpartisan elections, or the Merit Plan during the treatment period.[7]

Given my seven treatment states and five control states, I now divide the states into groups that I will refer to as "types." This division into types is a quasi-matching process at the state level. Type one consists of those treatment states carrying out the common transition from partisan elections to Merit Plan, as well as those control states that use

has the power to select judges. I exclude Tennessee from my analysis for two reasons, one practical and one theoretical. The practical reason is that Tennessee only used the Merit Plan from 1971 to 1974, during which only one judge joined the court (in 1973). Three years and one judge provide insufficient variation for any meaningful analysis. The theoretical reason is that, given the frequent changes, challenges, and uncertainty, it is unclear how Tennessee judges may have viewed the permanence of the method of judicial retention.

[6]It is not uncommon for states to use a hybrid system that mixes features of the federal system, partisan or nonpartisan elections, and the Merit Plan. Hybrid states are not included in my sample.

[7]Since no treatment state has switched to the Merit Plan since 1980, states that switch systems after the treatment period are allowed as controls - the first such state is New Mexico, which switched to a hybrid system in 1988, and thus New Mexico cases are only included up to 1987).

partisan elections (or, in the case of Missouri, transitioned to the Merit Plan prior to the period in question and thus always use the Merit Plan). Type two consists of treatment states moving from nonpartisan elections to the Merit Plan (South Dakota) and control states that use nonpartisan elections (North Dakota). Type three consists of treatment and control states that feature an interesting judicial feature: bifurcated high courts. In both Texas and Oklahoma, one court is the high court for civil cases, while another court is the high court for criminal cases. During the study period, Oklahoma transitions from partisan elections to the Merit Plan, while Texas continues to use partisan elections.

For each of the twelve states in the data set, I collect two sets of judicial opinions for all sample years from the Lexis Nexis archives. For treatment states, I include a twenty-five year span - ten years before the transition and fifteen years after the transition.[8] For control states, I collect a span of years that covers the entire span of years covered by any treatment state of the same "type". As mentioned, I collect two sets of opinions for each state. The first set is tort or contract cases, which are defined as all cases that contain at least one of the words "tort" and "contract."[9]I will henceforth refer to this first set of opinions as civil cases. The second set of opinions are those from criminal cases.

After data collection from Lexis Nexis, I use Python to parse the decisions to form useable data elements. I do this by taking advantage of predictable formulations that recite key elements of the case. However, I am largely unable to collect data from the narrative sections of the opinion, as such sections are written in no predictable fashion. I am able to collect data on which judges participated in the case; which judges wrote the opinion of the court, the concurrences, and the dissents (including those who joined the concurrences and the dissents); the date the case was decided; and whether the court ruled in favor of the original plaintiff or original defendant. I also attempt to collect the type (or types) of crime involved in a case for criminal cases. I use crime categories from the Federal Bureau of Investigation's Uniform Crime Reporting Program, which are homicide, rape, robbery, assault, burglary, and larceny.[10]

[8]There is one exception to this, which results in a twenty-three year span: Data for South Dakota are collected for only eight years before the transition, as the election system had been modified in 1972 (eight years before the transition to the Merit Plan). Data for North Dakota as a control are similarly truncated.

[9]Sampling has shown this to be a strong proxy for cases where one side is seeking an award of money from the other. This definition allows us to exclude two types of cases which are not informative to the questions at hand: Suits against and between government officials involving the enforcement of rights (where it is difficult to determine, in any systematic way, what group would gain an advantage from a ruling for or against the plaintiff) and administrative proceedings before the court (e.g. discipline of lawyers).

[10]I do not use two categories from the UCR, Arson and Motor Vehicle Theft. Arson is excluded because

Note that data on judgement award amounts and criminal sentence lengths are not available. This is because all cases in my data set are appellate decisions and such cases rarely directly concern the award amount or sentence length. Nevertheless, the case data I have can be interpreted as ruling on award amount and sentence length. To do so, we need only realize that a trial court's decision resolves only some of the uncertainty in a case. Even after the trial court's decision, there remains the potential for an appeals court to overturn the verdict or mandate a re-examination of certain evidence (this is why we often witness lawsuit settlement agreements that occur after a case has gone to trial, as the trial court verdict has altered the strength of the parties' respective positions but has not resolved all uncertainty). Thus we may view the appellate court's decision as resolving some of the uncertainty in the favor of one of the parties to the case. For example, if a tort plaintiff wins an appeal and is thus able to present more evidence favorable to the plaintiff to the jury, it is reasonable to view this as increasing the expected value of the award resulting from the case, even though the appellate decision did not directly involve the award amount.

IV. Methodology

I now turn to my empirical analysis of the effects of different judicial selection and retention regimes. My source of identification is the variation over time in judicial selection and retention methods within a state. This variation will allow me to use difference-in-difference estimation to identify the effect of the switch to the Merit Plan on judicial decision-making. My estimation strategy will also allow me to separately identify the effect of the switch in selection method, which I will interpret as an adverse selection effect, and the effect of the switch in retention method, which I will interpret as a moral hazard effect.

For both my civil cases and my criminal cases, I use identical regression frameworks. Estimation is done at the judge-case level (so the vote of each judge in a case constitutes a separate observation). I control for state-specific factors by introducing a set of state fixed effects. I also include a set of type-year fixed effects to control for time-varying type-specific factors. I use type-year fixed effects to allow different types to follow different time trends. I am unable to use state-year fixed effects to allow state specific time trends as within state-year variation is limited. [11]

it was added as a category in 1979, after the beginning of the study period. Motor Vehicle Theft is not included because it is not reliably reported in the court decisions.

[11]Within a state-year, there may be as few as five judges, all hearing the same cases, for whom there is

My basic estimation equation here is

$$(1)\ Outcome_{jcspt} = \alpha + \beta_1 AfterRetention_{ts} + \beta_2 Selection_j + \lambda_s + \gamma_{pt} + \varepsilon_{jcspt}$$

where Outcome in tort cases is ProPlaintiff, an indicator equal to one if the judge votes for the original plaintiff in the case, and Outcome in criminal cases is ProProsecution, an indicator equal to one if the judge votes for the prosecution ("the state") in the case. λ is a set of state fixed effects, γ is a set of type*year fixed effects, AfterRetention is an indicator variable equal to one if the case was decided after the state had moved to a system of retention elections, and Selection is an indicator variable equal to one if the judge was appointed via the Merit Plan. β_1 is thus the estimate for the effect of switching to retention elections (i.e. the moral hazard effect) and β_2 is the estimate for the effect of switching to merit selection (i.e. the adverse selection effect). The t subscript is time in years, the s subscript is state, the c subscript is case, the j subscript is judge, and the p subscript is type (as discussed in section III).

I also introduce a second way to estimate my parameters of interest. I introduce a set of individual judge fixed effects and estimate the equation

$$(2)\ Outcome_{jcspt} = \beta_1 AfterRetention_{ts} + \delta_j + \gamma_{pt} + \varepsilon_{jcspt}$$

where the judge fixed effects, δ, have replaced the prior Selection indicator (and the state fixed effects, which they subsume). β_1 is still the effect of switching to retention elections, but I must perform a second step to recover the estimate for the effect of switching to merit selection. To recover this estimate, I record the judge fixed effect estimates for each judge in the sample. I then calculate the difference between average judge fixed effects of those appointed under the Merit Plan and average judge fixed effects for those entering office under the electoral system, as shown in equation 3.1, where judges 1 to n entered the court under the electoral system and judges n+1 to m entered the court through the Merit Plan.

$$(3.1)\ \beta_2 = \frac{\sum_{n+1}^{m} \delta_j}{m-n} - \frac{\sum_{1}^{n} \delta_j}{n}$$

I also introduce as second way of estimating β_2, as shown in equation 3.2, where c_j is the number of case votes cast by that judge.

$$(3.2)\ \beta_2 = \frac{\sum_{n+1}^{m} c_j\delta_j}{c(m-n)} - \frac{\sum_{1}^{n} c_j\delta_j}{cn}$$

no variation in retention method, and for whom there may be no variation in selection method.

In equation 3.2, rather than the effect of merit appointment being the difference of simple averages of judge fixed effects (which implicitly weights each judge equally), I now instead compute the effect of merit appointment as the difference between the weighted average judge fixed effects of those appointed under the Merit Plan and the weighted average judge fixed effects for those entering office under the electoral system, where the weights are the number of cases of that type (tort or criminal) decided by each judge. To the extent they differ, estimates from equation 3.2 should likely be preferred to estimates from equation 3.1. From an estimation perspective, equation 3.2 should better estimate the effects because it puts more weight on judge fixed effects that are better estimated (i.e. those computed from more judge-case observations). From a validity perspective, equation 3.2 also better estimates the real effect of the policy change (since judges who decide more cases inherently affect more cases).[12]
The method used in equation 2 provides a second way to estimate the effects of interest. Furthermore, I am imposing different constraints on the estimation, as I no longer impose a single fixed effect for each state and a single coefficient for Selection. Instead, I only impose that each judge has a single fixed effect. By imposing fewer constraints, my estimation using equation 2 may be more efficient that the estimation from equation 1.

In case of worries about variation in case characteristics that is not adequately controlled for by type*year fixed effects, I introduce case fixed effects, as shown in equations 4 and 5, with η as a set of case fixed effects.

$$(4)\ Outcome_{jcst} = \alpha + \beta_2 Selection_j + \eta_c + \varepsilon_{jcst}$$

$$(5)\ Outcome_{jcst} = \delta_j + \eta_c + \varepsilon_{jcst}$$

With case fixed effects, I am unable to estimate the effect of retention elections, as all the judges on the same case face the same retention method, but I am able to estimate the effect of merit appointment while completely controlling for the cases heard, as I am only comparing, on the same case, judges entering office through an electoral system to judges entering office through merit selection. The ability to control for case characteristics, even unobserved ones, suggests that I am able to purge almost all potential bias from the results with case fixed effects.

In order to estimate the effects on case outcomes (i.e. how the court as a whole actually rules), I also run an estimation at the case level

[12] Note also that equation 1 implicitly weights the estimate $\beta_2$by number of cases

$$(6)\ Outcome_{cspt} = \alpha + \beta_1 AfterRetention_{ts} + \beta_2 Selection_c + \lambda_s + \gamma_{pt} + \varepsilon_{cspt}$$

where Outcome is now an indicator equal to one if the court rules for the original plaintiff / prosecutor and Selection is a continuous variable between 0 and 1 representing the percentage of the court selected through merit selection.

Finally, the likely correlation of error terms between decisions in the same state (particularly considering the judges will often be the same) leads to a desire to cluster standard errors. In my specifications, I cluster all standard errors at the state level to address this issue. In section 5.3, I alternatively use the Wild Cluster Bootstrap as a robustness check.

V. Results

5.1 Civil Cases

I begin with the results for my set of tort and contract cases. Table 1 shows my results for equations 1 and 2. Odd numbered columns use type*year fixed effects, while even numbered columns instead use year fixed effects for robustness. Columns (1) and (2) estimate equation 1, columns (3) and (4) estimate equation 2 using equation 3.1, and columns (5) and (6) estimate equation 2 using equation 3.2. In table 2, column (1) estimates equations 4, column (2) estimates equation 5 using equation 3.1, and column (3) estimates equation 5 using equation 3.2.

As for the interpretation of these estimates, the effect of instituting retention elections on civil cases is uniformly insignificant in all specifications. The point estimates are all negative and should be interpreted as percentage point changes in the probability of a judge-case vote being pro-plaintiff. For example, the estimate for table 1, column (1) is that the institution of retention elections decreases the probability of a judge voting for the plaintiff in a case by 1.08 percentage points (but is not statistically significant). The standard errors for all columns suggest that I can rule out a decrease of more than 4 to 5 percentage points.

In interpreting the estimate of the effect of merit selection on civil cases, it is helpful to look column by column. In table 1, we see a positive estimate, significant at the 5% level, when estimating equation 1 (columns (1) and (2)). When estimating equation 2, the estimate is no longer significant, though its magnitude does grow as we move from weighting by judge (columns (3) and (4)) to weighting by case (columns (5) and (6)). Finally, when I use my best controlled estimation with case fixed effects and weighting by case (table 2, column (3)), I once again find positive results, this time significant at the 1% level. My estimates suggest that appointing a judge through merit appointment increases his or her probability of voting in favor of a civil plaintiff by between 1.4 and 2.2 percentage points. From a baseline probability of voting for the civil plaintiff of 46%, this represents an increase of 3 to 4.8 percent.

Table 2, columns (4) to (6) and table 3 show various robustness checks. In table 2, columns (4) to (6), I repeat columns (1), (3), and (5) from table 1 for a restricted sample. This restricted sample consists of judge who hear at least 25 civil cases (thus throwing out judges who are very poorly estimated and who may be systematically different from other judges). These results are very similar to those in table 1, except that the effect of merit appointment, as estimated with equation 2 using equation 3.1, becomes more positive. [13] This seems to suggest that those judges who hear very few cases are more favorable towards the plaintiff than other judges (note that the overwhelming majority of those judges with few cases are judges entering through the electoral system).

In table 4, I show results for the estimation of equation 6. I estimate that instituting retention elections decreases the probability of a plaintiff winning the case by 4.6 to 5 percentage points. This estimate is significant at the 5% or 10% level, depending on the specification. I also find that converting the court from no Merit-appointed judges to all Merit-appointed judges increases the probability of a plaintiff victory by 7.2 to 7.9 percentage points, which is significant at the 5% level. Note that there is no conflict with the judge-case level estimates, as the relationship between the probability of changing a judge-case vote and the probability of changing an overall case decision depends on whether the judges who change their votes are more or less likely to be pivotal. Here, my estimates show that the judges changing their votes in civil cases are more likely to be pivotal.

5.2 Criminal Cases

I now move on to the results for my set of criminal cases. My tables here are similar in format to those shown for civil cases, with one modification. I run each criminal regression both with and without an additional set of controls - case-level crime type indicators. As mentioned previously, I collect data on the crime or crimes concerned in each case. I then categorize the crime or crimes involved using the crime categories from the Federal Bureau

[13] Note that this is exactly where we would expect to see an effect, as this is the judge-weighted estimation while the other two are case weighted.

of Investigation's Uniform Crime Reporting Program, which are homicide, rape, robbery, assault, burglary, and larceny.[14] I then run each regression with a vector of these six crime type indicators and without the crime type indicator vector, except for the case fixed effects model, where the case fixed effects subsume the crime type indicators.

Tables 5 and 6 show my results for equations 1 and 2. Table 5 uses type*year fixed effects, while table 6 uses year fixed effects. The regressions show in the tables are otherwise the same. Within each table, even numbered columns use the vector of crime type indicators, while odd numbered columns do not. Just as in the civil cases, columns (1) and (2) estimate equation 1, columns (3) and (4) estimate equation 2 using equation 3.1, and columns (5) and (6) estimate equation 2 using equation 3.2.

In table 7, I follow the same pattern as in civil table 2. Table 7's first three columns estimate the effects using case fixed effects, while the other 3 columns estimate the effects on a restricted sample of only those judges who hear at least 25 criminal cases.

In table 7, column (1) estimates equations 4, column (2) estimates equation 5 using equation 3.1, and column (3) estimates equation 5 using equation 3.2. In table 7, columns (4) to (6), I repeat columns (1), (3), and (5) from table 5 for a restricted sample. This restricted sample consists of judge who hear at least 25 criminal cases (thus throwing out judges who are very poorly estimated and who may be systematically different from other judges).

The estimates in tables 5 and 6 for the effect of retention elections on criminal cases are insignificant when using equation 1, though they are negative and fairly large (at 2.6 to 3.9 percentage points). When I instead use the potentially better estimated judge fixed effects, I find a statistically significant and even larger negative effect of 8 percentage points. In my robustness tests, when I restrict the sample and when I implement the Wild Cluster Bootstrap, I see nearly identical results as well. This decline of 8 percentage points, from a baseline of 82% pro-prosecution decisions, represents a large decline in decisions in favor of the prosecution. That being said, it does support the theoretical prediction from Lim 2013 that judges under the Merit Plan no longer need to cater to voters by being tough on crime.

As for the effect of merit appointment on criminal cases judge-votes, I find very mixed evidence. When I estimate the effect using equation 1, I see a small negative effect. When

[14]Note that not all cases involve crimes that fit into these categories. For crimes that do not fit into a category (e.g. "interfering with the lawful operation of a hunting party"), I simply set the value for each crime type indicator to zero.

I estimate the effect with equation 2 using equation 3.1, the results are positive and fairly large. Finally, when I estimate the effect with equation 2 using equation 3.2, the results are large, positive, and significant at the 5% level. This pattern holds in the restricted sample as well, but when I use case fixed effects (likely the best method for estimating the effect of merit appointment), the result is negative and insignificant. The Wild Cluster Bootstrap results once again look very similar, supporting my estimation strategy despite the relatively small number of states.

In table 9, I show results for the estimation of equation 6. I estimate that instituting retention elections decreases the probability of the prosecution winning the case by 1.4 to 2.7 percentage points. I also find that converting the court from no Merit-appointed judges to all Merit-appointed judges increases the probability of a prosecution victory by 3.9 to 5 percentage points. None of these case-level results are statistically significant and, as noted in the civil section, there is no conflict with the judge-case level estimates as the relationship between the probability of changing a judge-case vote and the probability of changing an overall case decision depends on whether the pivotal-ness of the voter changing judges.

I observe that the introduction of crime type indicators has no effect on the estimates in any of the specifications. Note, however, that this does not mean that judges disregard crime type, but rather the degree to which they are harsher or more lenient based on the crime committed is constant across the changes here (i.e. constant regardless of retention or selection method).

One final point in the discussion of the criminal results is that many of the estimates are not statistically significant, despite being larger than significant civil case results. This is clearly driven by large increases in the standard errors. Why are the standard errors growing by such a large amount? Part of the reason is the decreased sample size (judges in the sample simply hear more civil cases than criminal cases). However, the driving factor is that there is a large divergence in behavior between judges in these criminal cases. This is particularly true for retention elections, where some judges experience a large pro-defense shift when retention elections are introduced, while others see no shift and a few even have a pro-prosecution shift. To give an example of three Nebraska judges, Judge Brower saw his pro-prosecution percentage drop from 87.5% to 77.5% with the introduction of retention elections, while Judge Boslaugh's percentage stayed nearly constant (88.9% to 89.6%). At the same time, Judge Messmore saw his pro-prosecution percentage increase from 63.3% to 82.1%.

That we see these large divergences in pro-prosecution percentage in criminal cases much more than with pro-plaintiff percentage in tort cases is not surprising for two reasons. The first is that it seems reasonable to believe that both judges and the public have stronger views on criminal law than they do on civil law. To give one example, interest by both lawyers and the public in criminal cases before the Supreme Court is far higher (rivaled only by Constitutional issues not included in my dataset) than interest in civil cases before the Supreme Court. If the public pays more attention to criminal cases, judges facing reelection may be more careful to decide those cases in a pro-public (likely pro-prosecution) way. Similarly, if judges have stronger views on criminal cases, then in the absence of public pressure (since retention elections empirically often mean lifetime tenure), the judge is more likely to depart from their belief on what the public wants and instead make the decision based on their own belief. The second reason we might expect these large divergences comes from Lim (2013), where she lays out a theoretical explanation for why we would expect judges selected under an electoral system would have more divergent personal preferences in criminal cases, as there is little selection on views on criminal cases (since the increased public pressure in criminal cases imposes greater orthodoxy once on the bench).

5.3 Robustness

In considering the robustness of my results, two primary concerns may arise. The first concern is that, when I cluster standard errors at the state level, the relatively small number of states in my sample (12) may bring into question the asymptotic behavior of my estimator. This concern is illustrated by Cameron, Gelbach, and Miller (2008), showing that asymptotic tests can over-reject the null hypothesis when they have fewer than 30 clusters. Therefore, I also run the Wild Cluster Bootstrap, Cameron, Gelbach, and Miller's proposed solution, as a robustness test. The Wild Cluster Bootstrap is an asymptotic refinement which offers improved asymptotic behavior for small numbers of clusters - Monte Carlo simulations have shown the bootstrap to reduce rejection rates for small numbers of clusters to the standard 5% rate (i.e. 5% of randomly generated estimates are statistically significant at the 5% level).

I implement the Wild Cluster Bootstrap and show the estimates in table 3 for civil cases and table 8 for criminal cases. The columns in table 3 correspond respectively to: table 1, column (1); table 1, column (2); table 2, column (1); and table 2, column (4). There is little difference between the original standard errors and the Wild Cluster Bootstrap standard errors, which suggests that despite the small number of states in my sample, my estimators are functioning well. The columns in table 8 correspond respectively to: table 5, column (1); table 5, column (2); table 6, column (1); table 6, column (2); table 7, column (1); and table 7, column (4). Once again, the Wild Cluster Bootstrap standard errors are very similar to the original standard errors, thus showing my estimators are functioning well.

The second concern is that other inputs in the appellate judge case-decisionmaking process change at the same time as selection and retention methods. I'll break this concern into two questions. The first is: Do other institutional changes occur at the same time as selection and retention changes? For example, if lower courts are also switching their selection and retention methods, case outcomes that reach the appellate courts could be different than under the counterfactual, calling my identification into question. As far as I can observe, the answer to this question is no. Constitutional amendments changing appellate selection and retention methods do not include other changes to laws or precedents that would affect judicial decisions. Lower courts are also not facing changes in selection and retention methods at the same time as appellate courts.[15]

The second question is: Are there behavioral responses from other parties to the selection and retention changes? For example, if the behavior of lower-court judges or attorneys changes at the discontinuity, leading in turn to changes in the case mixture that appellate courts are seeing, it could lead to concerns about the strength of my identification. I am limited in my ability to measure case mixture as my scraped data mostly lacks case characteristics, but I am able to test an important measure of behavior - percentage of appeals from each party. If lower courts are changing their behavior and ruling more often in favor of one party, or if attorneys are changing their appeal decisions based on changing expectations regarding appellate court decisions, these changes should show up in this measure. Graphical results of this analysis are shown in figures 3 and 4. In both cases, I normalize the first year after the transition date for all transitioning states to be year zero. In figure 4, I show the results for criminal cases. The y-axis is the percentage of all appeals that are made by the prosecution. The y-axis values move smoothly through the transition, suggesting no behavioral responses from other parties to the change.

In figure 3, I show the results for my civil cases. The y-axis is the percentage of all appeals that are made by the original plaintiff. At the discontinuity, there is no statistically

[15]The only exception is two counties in Arizona which change methods at the same time. I am not able to precisely identify in my data the county from which a case is appealed, so I cannot run my estimation excluding just these two counties. I do run my estimation without Arizona as a robustness, and do not observe any significant differences in results.

significant change, which could be interpreted as showing no behavioral responses from other parties. That being said, there is a jump of about 2.5 percentage points, so perhaps the change is not statistically significant due to the size of the standard errors (since I'm looking at averages of all appeals, the number of observations for this analysis is small and thus the analysis is relatively low-powered). If the reader wants to take this jump as meaningful, we can think of what this jump means. After the transition, original plaintiffs have become a little more likely to appeal, an effect which grows a little more as we move away from the transition as well. What could cause this result? The best explanation, keeping in mind my finding that the appellate judges become more pro-plaintiff after the transition, is that plaintiff's attorneys are becoming more likely to appeal because they believe the appellate court may be more friendly to plaintiffs.[16] This actually biases my results towards zero (since it suggests that the appellate court after the transition is receiving more marginal appeals from plaintiffs), suggesting that, if you believe the not statistically significant jump in this graph is nonetheless meaningful, my results actually underestimate the effect of the Merit Plan on civil cases.

VII. Conclusion

In this paper, I present evidence on the causal relationship between judicial selection and retention systems and judicial decisions. Using variation over time in state judicial selection systems, I am able to use a difference-in-difference model to identify separate effects of retention elections and merit appointment on civil and criminal case decisions. Through various specifications making different assumptions, I show that the institution of retention elections has no effect on civil cases, while the institution of merit appointment results in a small but significant increase in voting in favor of the civil plaintiff. This aligns with the theoretical prediction that the bar association can use its increased power to favor civil plaintiffs and thus benefit members of the bar.

In examining results in criminal cases, I find suggestive evidence that the institution of retention elections leads to a large decrease in voting in favor of the prosecution, which aligns well with my theoretical prediction that contested elections encourage judges to pander by

[16]The alternative interpretation of this jump would be that lower-court judges become less pro-plaintiff after the transition, though it is hard to think of an explanation for why this would be. Since there are no institutional changes occuring for lower-court judges, this would suggest that these judges are either more determined to disagree with precedents from higher courts or want to increase their probability of being overturned, which would be at odds with the literature suggesting that judges should not want to be overturned (for reasons of either pride or just wanting to do less work).

being "tough on crime." Furthermore, I observe suggestive evidence that the institution of retention elections leads to a greater dispersion in judicial views on criminal cases, which nicely matches theoretical predictions of Lim (2013). I find mixed results of the effect of merit appointment on criminal judicial decisions.

My contribution to the literature here is three-fold, with the first part being my methodology which allows the simultaneous estimation of separate moral hazard and adverse selection effects of judicial selection and retention methods. The second part is my development of a new and extensive dataset, while the third, and largest, contribution is the opening of the study of the causal relationship between judicial selection systems and judicial rulings. I find empirical results that show that there is an important causal relationship here that is worthy of further study. The largest open question is a theoretical one, that of determining the desirability of an electoral system vis-a-vis the Merit Plan based on the differences in decisions made by judges under the two systems (e.g. is it a good or bad thing that criminal decisions become more pro-defendant and exhibit higher dispersion under the Merit Plan?). Empirical open questions also remain, in examining how judicial decision-making on various topics differs by judicial selection system. Nevertheless, this paper provides evidence that these differences do exist and that it is vital to consider these differences as states consider which judicial selection methods to utilize.

References

American Judicature Society (2011). "Judicial Selection in the States." National Center for State Courts. www.judicialselection.us. Accessed 5/25/2019.

American Judicature Society (2012). "Judicial Selection in the States: Missouri." National Center for State Courts. http://www.judicialselection.us/judicial_selection/index.cfm?state=MO. Accessed 9/3/2019.

Ash, Elliot and W. Bentley MacLeod (2015). "Intrinsic Motivation in Public Service: Theory and Evidence from State Supreme Courts." Journal of Law and Economics. 58(4): 863-913.

Ash, Elliot and W. Bentley MacLeod (May 2019). "Selection and Incentive Effects of Elections: Evidence from State Supreme Courts." NBER Working Paper.

Berdejo, Carlos and Daniel Chen (2017). "Electoral Cycles among U.S. Courts of Appeals Judges." Journal of Law and Economics. 60(3): 479-496.

Bonneau, Chris W. and Melinda Gann Hall (2009). In Defense of Judicial Elections. Routledge: New York.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008). "Bootstrap-Based Improvements for Inference with Clustered Errors." Review of Economics and Statistics. 90(3): 414-427.

Canes-Wrone, Brandice, Tom S. Clark, and Jee-Kwang Park (2012). "Judicial Independence and Retention Elections." Journal of Law, Economics, and Organization. 28(2): 211-234.

Choi, Stephen J., G. Mitu Gulati, and Eric 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): 290-336.

Helland, Eric and Alexander Tabarrok (2002). "The Effect of Electoral Institutions on Tort Awards." American Law and Economics Review. 4(2): 341-370.

Lim, Claire S. H. (2013). "Preferences and Incentives of Appointed and Elected Public Officials: Evidence from State Trial Court Judges." American Economic Review. 103(4): 1360-1397.

Lim, Claire S.H. and James M. Snyder, Jr. (2015). "Is More Information Always Better? Party Cues and Candidate Quality in U.S. Judicial Elections." Journal of Public Economics. 128: 107-123.

Lim, Claire S.H., James M. Snyder, Jr., and David Stromberg (2015). "The Judge, the

Politician, and the Press: Newspaper Coverage and Criminal Sentencing across Electoral Systems." American Economic Journal: Applied Economics. 7(4): 103-135.

Maskin, Eric and Jean Tirole (2004). "The Politician and the Judge: Accountability in Government." The American Economic Review. 94(4): 1034-1054.

O'Connor, Sandra Day (June 2014). "The O'Connor Judicial Selection Plan." Institute for the Advancement of the American Legal System.

Owens, Ryan J., Alexander Tahk, Patrick C. Wohlfarth, and Amanda C. Bryan (2015). "Nominating Commissions, Judicial Retention, and Forward-Looking Behavior on State Supreme Courts: An Empirical Examination of Selection and Retention Methods." State Politics & Policy Quarterly. 15(2): 211-238.

Tabarrok, Alexander and Eric Helland (1999). "Court Politics: The Political Economy of Tort Awards." Journal of Law & Economics. 42(1): 157-188.

USA.gov. "Branches of Government." www.usa.gov/branches-of-government. Accessed 5/1/2019.

Figure 1: Treated States

State	Transition Year	High Court	Appeals Court	Judgeships
Arizona	1974	Yes	Yes	27
Indiana	1970, Effective 1972	Yes	Yes	20
Iowa	1962	Yes	None	7
Nebraska	1962	Yes	None	7
Oklahoma	1967	Both Civil and Criminal	None	14
	1987	N/A (Already Merit)	Yes (Civil Only)	12
South Dakota	1980	Yes	None	5
Wyoming	1972	Yes	None	5

Figure 2: Control States

States	Selection Method	High Court	Appeals Court	Judgeships
Arkansas	Partisan	Yes	Yes	19
Missouri	Merit	Yes	Yes	39
North Dakota	Nonpartisan	Yes	None	5
New Mexico	Partisan	Yes	Yes	15
Texas	Partisan	Both Civil and Criminal	Yes	98

Figure 3: Civil Cases - Percentage of Appeals by Original Plaintiff through the Transition

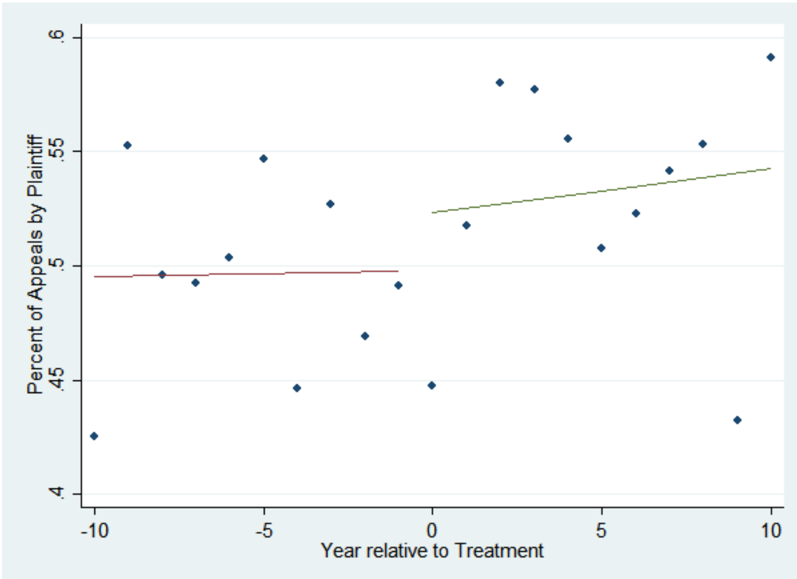

Figure 4: Criminal Cases - Percentage of Appeals by Prosecution through the Transition

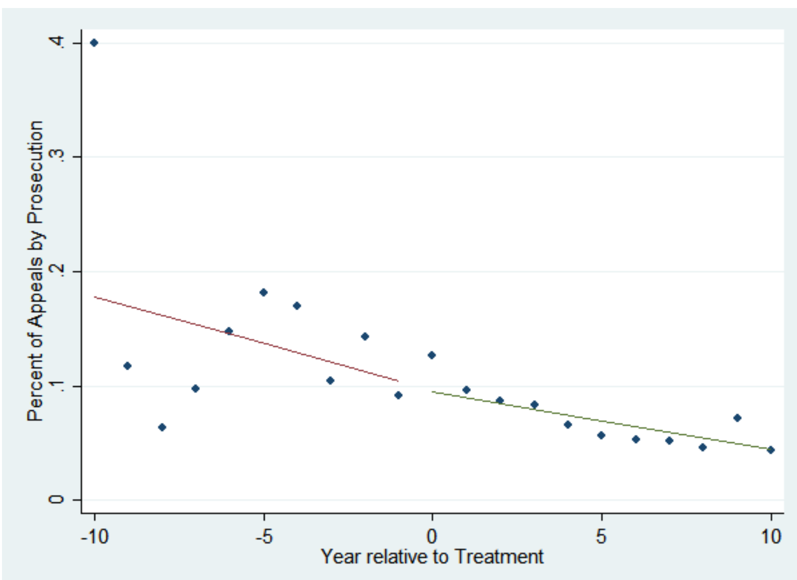

Table 1: Standard Results for Civil Cases

	(1)	(2)	(3)	(4)	(5)	(6)
Retention Elections	-0.0108	-0.0134	-0.0326	-0.0295	-0.0326	-0.0295
	(0.0185)	(0.0226)	(0.0219)	(0.0245)	(0.0219)	(0.0245)
Merit Appointment	0.0223**	0.0218**				
	(0.0099)	(0.0104)				
Merit Appointment			-0.0034	0.0015	0.0014	0.0097
			(0.0095)	(0.0107)	(0.0072)	(0.0076)
Intercept	0.430***	0.429***				
	(0.0406)	(0.0412)				
N	43225	43225	43225	43225	43225	43225
TypeYearFE	X		X		X	
YearFE		X		X		X
JudgeFE			X	X	X	X
JudgeWeighted					X	X

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the plaintiff. Regressions 1-2 estimate the effect of merit appointment through a dummy variable. Regressions 3-6 estimate the effect of merit appointment as the difference in judge fixed effects between judges appointed under an electoral system and judges appointed under the merit system. Regressions 5 and 6 weight judge fixed effects by number of cases.All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Civil Cases: Case Fixed Effects and Restricted Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Retention Elections				-0.0114	-0.0327	-0.0327
				(0.0174)	(0.0218)	(0.0218)
Merit Appointment	-0.0131			0.0224**		
	(0.0175)			(0.0100)		
Merit Appointment		0.0021	0.0144***		0.0119	0.0019
		(0.0104)	(0.0029)		(0.0095)	(0.0072)
Intercept	0.503***	0.496***	0.496***	0.434***		
	(0.0070)	(0.0028)	(0.0033)	(0.0400)		
N	43225	43225	43225	43075	43075	43075
JudgeFE		X	X		X	X
JudgeWeighted			X			X
CaseFE	X	X	X			
RestrictedSample				X	X	X

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the plaintiff. Regressions 1 and 4 estimate the effect of merit appointment through a dummy variable. Regressions 2, 3, 5, and 6 estimate the effect of merit appointment as the difference in judge fixed effects between judges appointed under an electoral system and judges appointed under the merit system. Regressions 3 and 6 weight judge fixed effects by number of cases. All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: Civil Cases: Wild Cluster Bootstrap

	(1)	(2)	(3)	(4)
Retention Elections	-0.0108	-0.0134		-0.0114
	(0.0184)	(0.0223)		(0.0168)
Merit Appointment	0.0223**	0.0218**	-0.0131	0.0224**
	(0.0099)	(0.0102)	(0.0176)	(0.0097)
Intercept	0.430***	0.429***	0.503***	0.434***
	(0.0401)	(0.0405)	(0.0083)	(0.0381)
N	43225	43225	43225	43075
TypeYearFE	X			X
YearFE		X		
CaseFE			X	
RestrictedSample				X

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the plaintiff. All regressions estimate the effect of merit appointment through a dummy variable.

All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Civil Cases: Case-Level Results

	(1)	(2)
Retention Elections	-0.0498**	-0.0458*
	(0.0219)	(0.0277)
Percent Merit Appointment	0.0792**	0.0725**
	(0.0285)	(0.0345)
Intercept	0.421***	0.420***
	(0.0498)	(0.0505)
N	7856	7856
TypeYearFE	X	
YearFE		X

Standard errors in parentheses

Dependent variable in all regressions is a court decision in favor of the plaintiff.

All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Standard Results for Criminal Cases, Type-Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Retention Elections	-0.0277	-0.0264	-0.0795**	-0.0789**	-0.0795**	-0.0789**
	(0.0331)	(0.0321)	(0.0408)	(0.0399)	(0.0408)	(0.0399)
Merit Appointment	-0.0217	-0.0215				
	(0.0232)	(0.0238)				
Merit Appointment			0.0501	0.0496	0.0629**	0.0607**
			(0.0424)	(0.0496)	(0.0264)	(0.0248)
Intercept	0.878***	0.857***				
	(0.0101)	(0.0171)				
N	30695	30695	30695	30695	30695	30695
CrimeControls		X		X		X
JudgeFE			X	X	X	X
JudgeWeighted					X	X

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the prosecution. Regressions 1 and 2 estimate the effect of merit appointment through a dummy variable. Regressions 3-6 estimate the effect of merit appointment as the difference in judge fixed effects between judges appointed under an electoral system and judges appointed under the merit system. Regressions 5 and 6 weight judge fixed effects by number of cases. All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 6: Standard Results for Criminal Cases, Year Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)
Retention Elections	-0.0388	-0.0368	-0.0831**	-0.0824**	-0.0831**	-0.0824**
	(0.0278)	(0.0265)	(0.0392)	(0.0387)	(0.0392)	(0.0387)
Merit Appointment	-0.0196	-0.0196				
	(0.0237)	(0.0243)				
Merit Appointment			0.0556	0.0552	0.0601**	0.0563**
			(0.0415)	(0.0382)	(0.0278)	(0.0267)
Intercept	0.878***	0.858***				
	(0.00970)	(0.0168)				
N	30695	30695	30695	30695	30695	30695
CrimeControls		X		X		X
JudgeFE			X	X	X	X
JudgeWeighted					X	X

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the prosecution. Regressions 1 and 2 estimate the effect of merit appointment through a dummy variable. Regressions 3-6 estimate the effect of merit appointment as the difference in judge fixed effects between judges appointed under an electoral system and judges appointed under the merit system. Regressions 5 and 6 weight judge fixed effects by number of cases. All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Criminal Cases: Case Fixed Effects and Restricted Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Retention Elections				-0.0305	-0.0806**	-0.0806**
				(0.0316)	(0.0400)	(0.0400)
Merit Appointment	-0.0227			-0.0220		
	(0.0197)			(0.0233)		
Merit Appointment		-0.0389	-0.0475		0.0521	0.0628**
		(0.0327)	(0.0335)		(0.0357)	(0.0242)
Intercept	0.771***	0.779***	0.791***	0.848***		
	(0.0080)	(0.0093)	(0.0092)	(0.0119)		
N	30695	30695	30695	30395	30395	30395
JudgeFE		X	X		X	X
JudgeWeighted			X			X
CaseFE	X	X	X			
RestrictedSample				X	X	X

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the prosecution. Regressions 1 and 4 estimate the effect of merit appointment through a dummy variable. Regressions 2-3 and 5-6 estimate the effect of merit appointment as the difference in judge fixed effects between judges appointed under an electoral system and judges appointed under the merit system. Regressions 3 and 6 weight judge fixed effects by number of cases. All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Criminal Cases: Wild Cluster Bootstrap

	(1)	(2)	(3)	(4)	(5)	(6)
Retention Elections	-0.0277	-0.0264	-0.0460	-0.0467		-0.0305
	(0.0331))	(0.0322)	(0.0330)	(0.0564)		(0.0265)
Merit Appointment	-0.0217	-0.0215	-0.0064	-0.0065	-0.0227	-0.0220
	(0.0232)	(0.0231)	(0.0046)	(0.0081)	(0.0236)	(0.0233)
Intercept	0.507***	0.515***	0.658***	0.654***	0.771***	0.848***
	(0.0123)	(0.0157)	(0.0137)	(0.0139)	(0.0091)	(0.0123)
N	30695	30695	30695	30695	30695	30395
CrimeControls		X		X		X
TypeYearFE	X	X				
YearFE			X	X		

Standard errors in parentheses

Dependent variable in all regressions is a vote in favor of the prosecution. All regressions estimate the effect of merit appointment through a dummy variable. All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 9: Criminal Cases: Case-Level Results

	(1)	(2)	(3)	(4)
Retention Elections	-0.0142	-0.0138	-0.0274	-0.0258
	(0.0249)	(0.0259)	(0.0223)	(0.0221)
Percent Merit Appointment	0.0390	0.0420	0.0478	0.0495
	(0.0636)	(0.0697)	(0.0596)	(0.0640)
Intercept	0.855^{***}	0.831^{***}	0.856^{***}	0.832^{***}
	(0.0267)	(0.0364)	(0.0255)	(0.0355)
N	5571	5571	5571	5571
TypeYearFE	X	X		
YearFE			X	X
CrimeControls		X		X

Standard errors in parentheses

Dependent variable in all regressions is a court decision in favor of the prosecution.

All regressions estimate the effect of merit appointment through a dummy variable.

All regressions are clustered at the state level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$