UC San Diego
UC San Diego Previously Published Works

Title

Does Judicial Selection Affect Judicial Performance? Evidence from a Natural Experiment

Permalink

https://escholarship.org/uc/item/8231p477

Journal

The Forum, 18(1)

ISSN

1540-8884

Authors

Vakilifathi, Mona
Kousser, Thad

Publication Date

2020-04-01

DOI

10.1515/for-2020-1002

Peer reviewed

Does Judicial Selection Affect Judicial Performance?
Evidence from a Natural Experiment

Mona Vakilifathi, New York University
Thad Kousser, UC San Diego

Abstract. Do judges selected by merit review commissions perform better than elected judges or those directly appointed by elected officials? This is a central question in both the academic study of state judicial institutions and the policy discourse about how to reform them. To address it, we take advantage of the variation in the means of the selection for trial court judges within Arizona, a state comprised of appointed, elected, and merit-selected trial court judges. This unique context allows us to use an objective measure of judicial performance – the reversal rate of trial court cases appealed to Arizona's state appellate courts – to evaluate judges by their means of selection. We gather an original dataset on 2,919 cases heard by 176 judges, estimating multivariate models that control for characteristics of cases and of judges. Overall, we find that elected judges have a lower reversal rate than merit-selected judges. Our findings question the conventional wisdom in the state courts literature in favor of merit selection and against judicial elections, and encourage further work on the effects of judges' means of selection beyond state supreme courts to include state appellate and trial courts.

Should judges be appointed, elected, or nominated by a merit review commission? The state courts literature remains divided on the most appropriate means of selection for judges. Proponents for gubernatorial appointment claim that the appointment process insulates nominees and judges from political influence (National Center for State Courts 2002; American Bar Association 2003). Advocates for judicial elections argue that the practice of elections ensures accountability between judges and their constituents (Hall 1992, 2001, 2007; Bonneau and Hall 2009; Caldarone et al. 2009). Merit selection supporters claim that a blended means of selection where candidates are nominated by an independent commission, chosen by the governor, and subject to retention elections ensure the best features of the appointment and election processes (Harrison et al. 2006; O'Connor and Jones 2008). Scholars from these different disciplines thus advocate for distinct and conflicting goals, making it difficult to compare these three means of selection and complicating this long-lasting debate (Cann 2007; Hall 2011; Gill 2013).

Building on recent efforts to use an objective measure of judicial performance, we take advantage of a natural experiment in Arizona to compare trial court judges' reversal rates by their means of selection within a state (Cass 1995; Cross and Lindquist 2009; Posner 2000; Choi et al. 2012; Hartzell 2014; Epstein et al. 2013; Sen 2015; Owens et al. 2015). All three systems – traditional elections, unfettered gubernatorial nominations, and merit review commissions – operate at once in this state.

The law the put this system in place has a celebrated political lineage. Then a little-known Arizona state senator, future Supreme Court Justice Sandra Day O'Connor was a major proponent of the constitutional amendment that created Arizona's merit review commissions in 1974. She and her allies argued that "the election system was not doing a good enough job at ensuring the selection of competent and qualified judges." (O'Connor

and Jones 2008, 17). O'Connor soon became one of the first judges to be selected to the Court of Appeals through the new process. Decades later, she co-authored a law review article concluding that "Merit selection may not be a perfect solution, but it is clearly better than the pure election system" (O'Connor and Jones 2008, 24).

Is there clear evidence that merit review, pure elections, or unfettered appointments produces more competent and qualified trial court judges? Arizona's judicial system gives us the unique opportunity to evaluate all three systems on the same scale, by looking at the rates at which trial court decisions are reversed on appeal by a state appellate court. We can thus measure the performance of judges along one objective and comparable metric to compare the means of selection within a state.

Compared to merit-selected judges' decisions, we find that elected judges issue decisions that are less likely to be reversed, in whole or in part, by the state's courts of appeals. Our results support past works claiming that elected judges perform better because these judges are motivated to perform on the basis of likely information shortcuts to voters, such as perceived productivity, rather than opinion quality (Choi et al. 2008). Elected judges likely aim for lower reversal rates to signal judicial quality to their constituents, especially to those involved litigants who may be unsatisfied with the judges' rulings.

This paper begins with an overview of the legal and political science literature regarding judges' means of selection, laying out three theoretical approaches and the empirical hypotheses that follow from each. Then, we describe Arizona's natural experiment and how it allows us to test these theories. We present results on the descriptive characteristics of judges by their means of selection and our findings using our measure of judges' reversal rates. To rule out potential alternative explanations of the differences that our analysis reveals, we then conduct a series of robustness checks. We conclude by

discussing the implications of our findings and future directions for using this measure and this natural experiment.

I. Judicial Selection and Objective Measures of Judicial Performance

A central debate in the state courts literature is whether a judge's means of selection - elections, appointment, or merit selection – affects her judicial performance. Proponents for each side of the debate largely disagree on the appropriate means of selection because of their different definitions of ideal judicial performance (Cann 2007; Hall 2011; Gill 2013). Proponents of appointment and merit selection, mostly legal scholars, value a judge's independence from public opinion and partisanship. Defenders of judicial elections, mostly political scientists, value accountability provided by elections as well as the congruence of a judge's rulings with public opinion. Proponents of merit selection value the blend of independence and accountability provided by a commission's nomination, gubernatorial appointment from one of the commission's nominees, and subsequent retention elections for the selected judge. In this section, we briefly review the rationale behind the arguments on behalf of each of the three systems, as well as the evidence that prior work has uncovered.

The findings from this literature substantiate each side of the debate by demonstrating unelected judges, which includes both appointed and merit-selected judges, as more independent from political pressure and elected judges as more accountable to public opinion. Hanssen (1999) and Iaryczower et al. (2013) find a positive relationship between appointed judges and levels of uncertainty in litigation cases, which indicates appointed judges as more independent than elected judges. However, Shepherd (2009a) finds judges to rule in accordance with their governor or legislature's preferences, where judges facing

upcoming reappointment rule in support of the current majority political party's policy preferences.

In support for judicial elections, Hall (1992, 2001, 2007), Bonneau and Hall (2009), Caldarone et al. (2009) demonstrate that judicial elections, especially partisan elections, to increase accountability between judges and their voters. Tabarrok and Helland (1999), Gordon and Huber (2007), Shepherd (2009b), and Canes-Wrone et al. (2012) argue that judges facing upcoming retention elections, especially partisan retention elections, rule in accordance to the public opinion of their prospective voters. Caldarone et al. (2009), Franklin (2002), and Canes-Wrone et al. (2014) finds judges selected by nonpartisan elections to be more attentive to public opinion than judges selected by partisan elections. Brace and Boyea (2008) and Besley and Payne (2013) finds the practice of judicial elections to influence the composition of those state supreme courts as well as those judges' decisions to uphold death sentences. Huber and Gordon (2004) claim judges facing reelection impose longer sentences on criminal cases as their reelection approaches. Ware (1999), Walternburg and Lopeman (2000), McCall (2003), McCall and McCall (2007), Cann (2007, 2012), and Kang and Shepherd (2011) demonstrate the impact of campaign contributions on elected judges' decisionmaking. However, Cann (2002) and Bonneau and Hall (2009) argue against such an effect.

While this literature has generally focused on the differences between appointed and elected judges, the "Missouri Plan" of merit selection has been lauded by legal scholars and advocacy groups (Harrison et al. 2006; O'Connor and Jones 2008). Proponents claim that merit review commissions incorporate the best of both worlds regarding independence and accountability – judges are nominated by a nonpartisan commission and face retention elections to ensure accountability to the voters. However, there seems to be little empirical

support for this hybrid means of selection from legal scholars and political scientists (Reddick 2002; Gill 2013). Canon (1972), Flango and Ducat (1979), Glick and Emmert (1987), and Hurwitz and Lanier (2003, 2008) find no support for merit-selected judges to be any different in their characteristics than appointed or elected judges. Furthermore, prior findings suggest there are no differences between the decisionmaking (Watson and Downing 1969; Atkins and Glick 1974; O'Callaghan 1991; Romero et al. 2002) or the judicial quality (Watson and Downing 1969) of merit-selected and elected judges. Canes-Wrone et al. (2014) finds merit-selected judges to be less influenced by public opinion but only during the era prior to the defeat of California's state Supreme Court justice Rose Bird in 1986 – an event that has indicates the beginning of the recent new-style judicial campaigns (Canes-Wrone et al. 2014, 37). Hall (2001) and Brace and Boyea (2008) show that merit-selected judges subject to retention elections are influenced by public opinion in their decisionmaking. However, Aspin (2007) and Hall (2007) show that retention elections have generally been uncompetitive, which weakens the promise of accountability for merit-selected judges.

Recent works acknowledge the tradeoffs for each means of selection but advocate for a change in framing for future research by evaluating each alternative with a consistent measure of judicial performance. Once the ideal standards are agreed upon, it will be easier to determine the most appropriate means of selection (Cann 2007; Choi et al. 2008; Hall 2011; Gill 2013). There is relatively little discussion as to how judicial performance can be measured to compare different means of selection in a similar and objective manner. To date, the most commonly-used measure of objective judicial performance is reversal rates, the rate a judge's trial decisions are overturned in part or in its entirety (Cass 1995; Cross and Lindquist 2009; Posner 2000; Choi et al. 2012; Hartzell 2014; Epstein et al. 2013; Sen 2015; Owens et al. 2015). Reversal rates indicate a judge's level of agreement among her colleagues

(Cass 1995). Reversal rates are an important indicator of performance for judges who seek reelection to their current position or promotion to a higher court (Cass 1995; Owens et al. 2015). Given the prominence of this measure in the legal and political science literature, we use reversal rates to compare judicial means of selection by judicial performance.

Certainly, other works use alternative measures of judicial performance to evaluate judicial means of selection. Cann (2007) surveyed state supreme, appellate, and trial judges regarding the performance of the state court system in their own state and found states where most judges are elected have a lower evaluation of judicial quality than judges selected by appointment or merit review commission. Choi et al. (2008, 2009) build off past rankings of judicial influence, which were largely based on citations of a state supreme courts' opinions (Mott 1936; Friedman et al. 1981; Caldeira 1983; Comparato 2002; Dear and Jessen 2007), to create an index of judicial quality comprised of effort/productivity (the number of opinions a judge publishes in a year), skill/opinion quality (the number of times an out-of-state court cites an opinion), and independence (the probability a judge votes with the opposite party affiliation). Choi et al. (2008) find that appointed judges have higher skill/quality than elected judges, but elected judges have high effort/productivity than appointed judges. Iaryczower et al. (2013) measure the information quality a judge receives on a case and likelihood for a judge to vote to overturn an appeal to predict a judge's rate of voting correctly on a case. Iaryczower et al. (2013) find small but significant differences between incorrect decision rates of state court judges, where elected judges more likely to make an incorrect ruling than an appointed judge.

In response to the demand from the state courts literature to use consistent measures of judicial quality across judicial means of selection, we use the judges' reversal rates as our measure of judicial performance. In the next section, we apply the arguments for each means

of judicial selection to determine whether merit-selected judges are overturned on appeal more or less often than appointed or elected judges. We attempt to lay out each argument in an evenhanded way, developing alternative hypotheses that we then test in the remainder of the paper.

II. Hypotheses

Proponents of the appointment method argue that judges have more information, less bias, and more uncertainty or independence in their decisionmaking (Hanssen 1999; Iaryczower et al. 2013). Appointed judges are also expected to produce higher quality opinions than merit-selected and elected judges, in order to improve their reputation among their colleagues (Choi et al. 2008). Proponents also claim that appointed judges are superior to merit-selected and elected judges because appointed judges are generally more insulated from public opinion. Using reversal rates as our measure of judicial quality, these proponents would expect appointed judges to have lower reversal rates than merit-selected and elected judges, because appointed judges are more independent and are more incentivized to produce high-quality opinions.

> *H_1: Judges selected by gubernatorial appointment will have lower reversal rates than elected judges or judges nominated by a merit review commission.*

On the other hand, defenders of judicial elections argue that elected judges are more accountable to public opinion, more consistent in their rulings, and more productive by issuing more opinions than appointed or merit-selected judges. (Hall 1992, 2001, 2007, 2011; Hanssen 1999; Choi et al. 2008; Bonneau and Hall 2009; Caldarone et al. 2009; Iaryczower et al. 2013) These defenders would expect elected judges to have lower reversal rates since reversal rates could serve as a signal of judicial quality to potential voters. Like the number of

opinions per judge (as used in Choi et al. 2008), a minimal reversal rate is an easy signal an elected judge can use to demonstrate her performance to her constituents – especially to those voters least satisfied with the judge's unpopular rulings.

> H_2: Elected judges will have lower reversal rates than judges selected by gubernatorial appointment or judges nominated by a merit review commission.

Proponents of merit selection argue that the selection and accountability of these judges make them superior to appointed and elected judges. The nomination of judges by a nonpartisan commission comprised of attorneys capable of selecting nominees for the governor to select from ultimately ensures the appointment of a high-quality judge. The means of appointment ensures the advantages an appointed judge generally has over an elected judge. The use of retention elections for these judges also ensures the accountability of these judges to public approval. These proponents expect merit-selected judges to have the lowest reversal ratings since the judges are nominated by practicing attorneys, which forces the governor to choose one of these candidates rather than any judge that aligns with the governors' ideology.

> H_3: Judges nominated by a merit review commission will have lower reversal rates than judges selected by gubernatorial appointment or elected judges.

III. Arizona's Natural Experiment

Arizona's unique system of judicial selection – in which superior court judges are chosen in different ways, depending on a county's population or the timing of a position becoming open – provides a fortuitous research design for testing the impact of selection mechanisms. It was first proposed by the state bar in 1959 and slowly gained momentum, but bills to enact it failed twice in the early 1970s in the state legislature. Undeterred,

reformers such as Sandra Day O'Connor were able to enact the new system through an initiative that successfully amended the state constitution in 1974 to put merit selection in place for trial court judges in Arizona's largest counties, while allowing small counties to continue using a mixture of elections and gubernatorial appointments (O'Connor and Jones 2008). This natural experiment is especially strong for discerning whether elections or unfettered gubernatorial appointments lead to stronger judicial performance. In this section, we describe each of the three selection mechanisms and how they fit into our research design.

Election. The route to a judgeship in the Grand Canyon State depends, first, on where a lawyer lives. In any county with a population below 250,000 in the last U.S. Census, there are two possible avenues to the bench. When a seat becomes open, a member of the state bar may run for election to a four-year judicial term. Elected judges are subject to reelection in subsequent terms. These contests, which were the norm across Arizona from statehood in 1912 through 1974, progress from a partisan primary to a non-partisan general election (National Center for State Courts 2015). That is, in Arizona's semi-closed primary, judicial candidates run for their party's nomination, with an electorate consisting of party members and any independent voters who choose that party's ballot during the primary (The Center for Voting and Democracy 2012). The victors then move to the general election, in which they run without a party label. As in all three selection methods, to serve in the Superior Court, a lawyer must be at least thirty years old, a member of the bar, a resident of the state for at least five years, and "of good moral character" (Arizona Judicial Branch 2015b).

Elections in the smaller counties thus resemble the general model of judicial elections followed in many states, with all of the potential vices and virtues. These campaigns

can raise money, though judicial candidates are banned from personally soliciting campaign contributions in Arizona just as they are in 30 of the 39 states which feature judicial elections (Liptak 2015). Arizona reformers have worried that having to run for election brings "the potential for unduly influencing judicial decisions," (Lee 1973, 63) or that it might discourage good lawyers from running if they "view campaigning as undignified" (Lee 1973, 62). Yet the elections also generally feature high participation rates, as they take place during November, and allow a wide range of potential judges to run.

Unfettered Appointment. In all of these ways, elections stand in stark contrast to the other route to a judgeship in counties with fewer than 250,000 residents. When a judgeship becomes vacant due to its holder stepping down, reaching the mandatory retirement age of 70, or passing away, Arizona's governors have the unfettered ability to appoint their replacement. The interim judge then serves until the next election, in which she runs as an incumbent. These sorts of appointments are quite common in Arizona, with one state legislator claiming that 75 percent of judges were "appointed at the whim of the Governor" (quoted in O'Connor and Jones 2008, 18-19) and reformers worrying that "judges who were appointed before they were elected were not subject to any official screening process" (O'Connor and Jones 2008, 18). A governor's office may internally screen candidates, sometimes in consultation with local judicial leaders, but there are not constitutional or statutory obligations to work through any particular process. In the words of one critic of this system, speaking at the time that Arizona was considering the introduction of its own Missouri Plan, "We now have a king maker who appoints judges. [The Governor] appoints them without the advice of anyone" (quoted in Roll 1990, 854).

This creates a clear comparison between elections and unfettered appointment systems, both taking place in the same geographic areas. They operate simultaneously and

draw from identical legal pools, with "treatment" of an interim appointment determined by the exogenous forces that lead to a judicial vacancy. In this way, Arizona's natural experiment allows for a strong causal inference. Any difference we observe between the characteristics and performance of elected judges versus those chosen through unfettered appointments is likely due to the selection mechanism itself.

Merit Selection. Arizona's third route to the Superior Court is the one that must be taken by any lawyer who lives in a county of more than 250,000 residents. This has included Maricopa County, which contains Phoenix, as well as Pima County, which contains Tucson, ever since merit selection was first introduced in 1974. After crossing the population threshold in the 2010 Census, Pinal County (which lies between Maricopa and Pima) switched from elections and appointments into merit selection.[1] The system was first proposed by the Arizona State Bar Association in 1959, kept alive by the Citizens Association on Arizona Courts, pushed in unsuccessful legislation in the early 1970s, and finally instituted through an initiative amending the state constitution in 1974 (Harrison et al. 2006; O'Connor and Jones 2008). It is based on the Missouri Plan, adapted to Arizona and slightly altered during the process of legislative bargaining. The system was amended into its current form by Proposition 109 in 1992, which responded to criticism that the system lacked public accountability by requiring the consideration of diversity in appointments and creating a unique system of judicial performance review for judges selected under this system (O'Connor and Jones 2008, 245-6).

Separate merit review commissions nominate the state's Supreme Court justices, the

---

[1] Over the next decade, as more merit-selected judges take the bench in Pinal County, its transition over the population threshold will serve as an exceptionally strong "regression discontinuity" research design. At the time of our data collection for this project, unfortunately, only a single judge had been selected through the new system in Pinal, preventing us from analyzing it in this manner.

state appellate court judges, and all superior court judges in each of the three most populous counties. These commissions are made up of five lawyers and ten members of the general public, who review judicial applicants in great depth before forwarding at least three nominees to the governor, no more than 60% of whom may come from the same party (Arizona State Bar 2013). The governor then picks from among this greatly constrained set of choices – which is, in practice, almost always no more than two nominees if the governor wants to choose a judge from her own party – to make the appointment. After finishing their terms, judges then face retention elections. Yet this electoral pressure is far different from the pressure put on those who run against other candidates for an open seat; of the 735 merit-selected judges who appeared on Arizona ballots from 1976 through 2006, only two failed to receive the 50 percent support needed to retain their seats (Klumpp 2008, 12).

This was, of course, precisely the idea of the reform. It was a solution to the perceived ills of judicial elections, as described by Harrison et al. (2006, 241): "The reality of judicial politics forced most observers to acknowledge that typical voters were unaware of the candidates, the issues, or even the existence of contested judicial races." The new process could also draw on a broader group of potential judges. According to the Arizona state bar, "Merit selection produces a better pool of applicants because the most qualified attorneys are often unwilling to risk their practice on an expensive political campaign that would be required if they needed to run for election" (Arizona State Bar 2013).

According to Vanessa Haney, the program manager for the Judicial Nominating Commissions, the sixty-day process of screening and selection nominees is indeed thorough and transparent. It begins with commission members reviewing applications and letters of recommendation, hearing public testimony, and then deciding which candidates to interview. Both the paid staff of the commissions and the volunteer commissioners then perform due

diligence on the interviewees during the screening process. The interviews take place at a second open meeting in which members of the public are again invited to testify.[2] Former Arizona State Bar President Joe Kanefield, who has often advised those who seek nomination, describes the process as analogous to other political efforts. "You have to wage a lobbying campaign when you apply for one of these positions. You have to educate, and the commissioners want to hear from the public, people who have appeared in court. So you want 10-15 people who can speak intelligently about you in a professional way."[3]

Overall, Arizona provides a unique context to compare three different means of judicial selection within a state. However, in order to make a systematic comparison of these three systems, we must compare the behavior of trial court judges in Arizona's two metropolitan areas to the actions of judges in the outlying counties. That highlights a potential weakness of our research design. It is possible the merit review commissions in the Phoenix and Tucson areas may have a stronger pool of lawyers to draw from, because these cities host the state's major law firms and most prestigious law schools. In both the characteristics of judges and in their eventual performance, Maricopa and Pima counties may surpass the smaller counties because of the supply of strong lawyers, not because of the "treatment" of the merit selection system. If their judges performance better, we should be wary of attributing all of the advantages that we might see in these courts to merit selection alone. Yet the weakness of this design will only confound our causal inferences only if the merit-selected judges do better. If we find that elected or appointed judges in the small counties outperform the merit-selected judges, despite the fact that they are drawn from a weaker pool of lawyers, we can be confident that this effect is primarily due to the means of

---

[2] Interview with Vanessa Haney December 17, 2013, Phoenix, Arizona.
[3] Interview with Joe Kanefield December 17, 2013, Phoenix, Arizona.

selection.

IV. Data on Judicial Characteristics and Performance

In order to make a comparison of performance across different systems of selection, we collected an original dataset on 176 Arizona Superior Court judges and on the outcomes of 2,919 appeals of their decisions. The primary basis on which we compare judges is the reversal rate in these appeals. We will also examine the types of judges selected through all three systems, and then use these characteristics as control variables to isolate the effects of judicial selection mechanisms in our models. These models include variables measured at the levels of judges and at the case level. In this section, we first introduce our key dependent variable of reversal rates before describing our explanatory and control variables at the case level and then at the judge level.

Reversal Rates. A judge's reversal rate is based on the instances a higher court disputed with the judge's decisions and indicates the level of criticism of a judge's decisionmaking from her colleagues (Cass 1995). In the broader literature on judicial performance, the rate at which a judge's decisions are reversed upon appeal (or its converse, the "affirmance rate") is a common metric. It has been used as a measure of judicial performance to study state supreme courts' means of selection (Owens et al. 2015), federal district court judges (Cass 1995; Cross and Lindquist 2009; Choi et al. 2012; Epstein et al. 2013; Sen 2015), to evaluate entire federal appeals circuits (Posner 2000), to predict a lawyer's chances of succeeding on appeal (Hartzell 2014), and debated during Senate confirmation hearings (Dinan 2009; Schlesinger 2009). Although reversal rates may be an imperfect measure of judicial performance, a judge's reversal rate is an information shortcut voters, commissions, and legislatures regularly use as a measure of judicial performance to

determine appointments to higher courts. A judge seeking reelection to their current position or promotion to a higher court is motivated to satisfy this metric of judicial performance by minimizing their reversal rate.

In Arizona, Arizona Superior Court judges' decisions are subject to appeal in one of two state appellate courts, covering the northern and southern portions of the state. All of the Arizona Appeals Court judges are chosen through merit selection, meaning that Superior Court judges all have their cases reviewed by the same pool of judges, effectively evaluating them on a common scale. Using the public websites of both appellate districts,[4] we worked with a team of five research assistants at our home university to track the outcome of each appeal that was decided in 2012, 2013, and the first half of 2014.[5] We gathered the outcome of both published and memorandum Superior Court decisions, in order to get the most comprehensive measure possible of each judge's performance. There are many potential outcomes of appeals, ranging from clear affirmance to partial reversal to clear reversal. Though there is no completely standardized method of classifying them present in the literature (Hartzell 2014), we following the approach of other scholars in placing outcomes such as "Affirmed" or "Appeal Dismissed" into our *affirmed* category, outcomes like

---

[4] The website for Division I, based in Phoenix, can be found at
http://azcourts.gov/coa1/searchdecisions.aspx, and the website of Division II, based in Tucson, can be found at http://www.appeals2.az.gov/ODSPlus/caseInfo.cfm
[5] Because we could not find a measure of the total number of decisions made per judge, the reversal rates that we use are the rate at which judges are reversed, conditional on a case being appealed. As such, they overstate the portion of rulings that are reversed, but likely do so by the same amount for judges from all three types of selection mechanisms. If there is any difference across counties in the rate at which decisions are appealed, it is likely the more populous counties that see more challenges with lower chances of succeeding (with their larger density of lawyers, and more affluent clients). If so, then merit-selected judges might appear to perform better on this measure than judges in smaller counties. The other possible drawback of using reversal rates is that, because all appellate judges are chosen through merit selection, the Superior Court judges chosen through merit selection see their cases reviewed by a more similar set of judges, who may share the same jurisprudential approach. Again, this might lead merit-selected trial court judges to score more highly on our metric than their underlying performance would indicate. And again, this would be a reason to be skeptical of findings that the merit-selected judges outperform judges picked in either of the other two selection methods, but more confidence in our conclusions if, on our performance metric, Arizona's "Missouri Plan" judges do worse.

"Affirmed in Part/Reversed in Part," "Affirmed in Part/Vacated in Part," or "Affirmed as Modified" into our *partially reversed* category, and "Reversed and Remanded" or "Vacated and Remanded" into our *reversed* category.

Type of Case. When we gathered information on the outcome of each decision, we also noted whether the case was *civil*, *criminal*, *juvenile*, or another type of case. In our sample, 54% of appealed cases were criminal, 28% were civil, and 9% were juvenile, and our models compare each of these to the *other* category in order to control for possible differential rates of success across the types of cases, which could vary across types of judges.

Year of Appeals Court Decision. We also note and control for the year in which an appeals court decision was issued, to guard against the threats to inference for any chronological trends that might be correlated with the mixture of cases across types of judges.

Type of Selection Mechanism. The first judge-level variable that we collected is our key independent variable, the method through which a Superior Court judge first came into office. We located this through a variety of sources. In Maricopa and Pima counties, in which merit selection is the exclusive means of coming to office, this simply required finding a judge's county and confirming this through rosters of *merit selection* appointments. In the less populous counties, we used county court websites, biographies of judges, election results, newspaper articles, and personal contact with county clerks in order to code a judge as one selected through *unfettered appointment* or an *election*. Our dataset of judges contains 118 judges chosen through merit selection, 33 judges picked directly by a governor, and 25 elected judges.

Ranking of Law School. In order to gauge the prestige of the law school from which each judge graduated, we used the rankings contained in the *US News and World*

Report's 2014 law school rankings. Our sample of judges included 32 who attended Arizona State University's Sandra Day O'Connor College of Law, which was ranked 31st that year, and 59 from the University of Arizona's James E. Rogers College of Law, ranked 40th. Law schools that did not appear in the rankings were given an equal 150th ranking.

Gender of Judge. We also coded whether a judge was male or female, drawing upon online biographical sources when gender was not obvious from a judge's first name. Gathering this variable enables us to measure the gender diversity of judges selected through each system, as well as letting us hold gender constant in our analysis of reversal rates. Overall, 69% of the judges in our dataset are male.

Years on the Bench and **Years on the Arizona Bar**. In order to measure judicial experience, we drew upon our biographical sources[6] to note how long a judge had served and how long each had been a member of the state bar. Our most experienced judge first came to office in 1985, with the average judge serving ten years on the bench by 2014. Our member with the longest membership in the bar first joined in 1969, with judges passing the bar an average of 29 years ago.

V. Findings

Drawing on our original dataset, we can see how each of Arizona's three selection mechanisms shapes the state's bench. We begin with the preliminary question of what types of judges are chosen, looking at the characteristics of judges selected through each mechanism. We then turn to our primary dependent variable, reversal rates. After conducting a straightforward bivariate comparison of reversal rates under each system, we

---

[6] In particular, we referred to the Arizona Bar Association's "Find a Lawyer" website (http://www.azbar.org/FindaLawyer) to find such information.

estimate multivariate models that control for both the characteristics of cases and the attributes of judges to isolate the impact of selection mechanisms on judicial performance. Both types of analyses yield a consistent and surprising verdict: elected judges in Arizona see fewer of their decisions reversed than judges chosen through merit selection.

We begin, in Figure 1 by looking at the types of judges chosen through each system. It is important to remember that merit selection, which operated exclusively in Maricopa and Pima counties during the time of our study, possibly draws on a different pool of lawyers than the unfettered appointments and elections in Arizona's smaller counties. With Phoenix and Tucson lawyers to choose among, merit review commissioners could be expected to pick judges with greater experience or stronger bona fides. Or, if these positions are highly sought after by many strong applicants, the judges who eventually emerge from each system might all look alike. Our interviews with experts involved in judicial nominations in Arizona support the latter claim. Joe Kanefield, the Arizona Bar president who formerly oversaw the screening of judicial appointments as the chief counsel to Gov. Jan Brewer, acknowledged that the initial pool of applicants when a Superior Court position came up in one of the smaller counties was a bit more "raw." But after his office screened applicants, a process which included taking public feedback on applicants, the quality of judges in the outlying counties resembled the judges selected in the Maricopa and Pima: "Once you got to the top 3-4 in those counties, they were similar to the folks sent to you by the commissions."[7]

Yet the data in Figure 1 shows that some strong differences emerge in the education backgrounds of judges. The Maricopa and Pima County judges chosen through merit selection graduated from law schools that had an average ranking of 45th in the US News ratings. On average, merit-selected judges attend more prestigious law schools than

---

[7] Interview with Joe Kanefield, December 17, 2013, Phoenix, Arizona.

appointed and elected judges from smaller counties. In fact, the difference between merit-selected judges and ones chosen directly by the governor is statistically significant at above the 95% confidence level. Within the small counties – where the natural experiment is strongest, because the pool of lawyers is held constant – we also see a statistically significant difference. Appointed judges went to law schools with an average ranking of 81st, while elected judges attended schools averaging 57th. Elected judges, this figure shows, went to significantly more prestigious law schools. This may be surprising to the legal scholars who argue that elections discourage qualified lawyers from seeking a judgeship. Based on law school prestige, judges elected from small counties look more like merit-selected judges from Arizona's major cities than they resemble the appointed judges in their areas.

Figure 1 also reports on gender diversity through each of the three systems. The two methods of appointment lead to benches with memberships that were 70% male. Women do better through elections, with only 60% of these seats occupied by men, but the difference between this and the other percentages falls short of statistical significance. Looking at legal experience, Figure 1 shows that the three systems also produce judges who are similar in terms of their legal experience at the time they are selected. By subtracting a judge's years on the bench from his or her total years as a member of the Arizona Bar (our two measures that we analyze separately in our regressions), we see that judges came to office in all three systems with approximately two decades of legal experience. None of the differences are statistically significant.

Overall, our look at judges' characteristics shows similarity in the gender balance and legal experience, but some difference in law school prestige. Merit selection, which draws from perhaps a stronger pool of lawyers in the Phoenix and Tucson area, leads to the selection of judges who come from more prestigious law schools than judges appointed

directly by the governor in the small counties. We cannot be certain of whether this is due to the pool of lawyers or to the selection mechanism. But in these smaller counties, in which the legal pool is held constant, elections surprisingly lead to the selection of judges who graduate from significantly more prestigious law schools than the unfettered appointment process. This is an initial piece of evidence in favor of the idea that elections can serve as a rational screening process for judges.

When we turn to judicial performance, again we see evidence in favor of the idea that elections can create incentives for judges to rule consistently with their state's legal standards. Table 1 reports the rates at which Superior Court judges see their decisions affirmed, reversed in part, or fully reversed at the state appellate court. First, it is important to note that the vast majority of cases withstand appeal, indicating that all three systems are producing high-quality judges. For merit-selected judges, 80.7% of their decisions were upheld on appeal, 12.5% reversed in part, and only 6.8% fully reversed. Yet judges chosen through the other methods perform even better on this metric. The judges appointed directly by governors had 88.3% of their cases affirmed, 5.6% reversed in part, and 6.2% reversed. Elected judges performed similarly with 87.3% of cases affirmed, and had the fewest cases fully reversed (only 5.4%). Although the judges chosen through merit selection do quite well, Superior Court judges chosen through Arizona's other two systems do even better. Part of what makes this especially surprising is that all of the Appeals Court judges gained that position through merit selection, and even though this might give them more of a common jurisprudential approach to the trial court judges chosen through merit selection, they are more likely to reverse the work of merit-selected Superior Court judges.

Yet these apparent patterns may be an artifact of the different types of cases heard by the different types of judges, or a function of each judge's background. In the multivariate

models reported in Table 2, we add these two groups of variables to our basic model in order to hold them constant, isolating the impact of selection mechanisms. This allows us to asks the hypothetical question, "How likely is it that judges selected through different mechanisms are reversed on appeal, if they are all ruling on a civil (or, perhaps, criminal) case in the same year and each went to a similar law school and then entered the bench with the same level of experience?"

In Table 2, we report the results of ordered logit models, with a dependent variable that progresses from the category of "affirmed" to "reversed in part" to "reversed," growing larger when a more of a decision is overturned. Negative coefficients, then, indicate that a variable makes a judge less likely to be reversed. Standard errors are clustered at the level of the judge, because if we underpredict the chances that one of a judge's decisions is reversed, we are likely to make the same error for all of that judge's decisions.

The result of all of our models, which build from a concise approach to one that controls for all of our independent variables, is that elected judges are clearly less likely than merit-selected judges to see their decisions reversed on appeal. This finding is significant at the 99% confidence level in a two-tailed test in all models. To get a sense of the scale of the effect, we calculated first differences using the Clarify statistical package (King, Tomz, and Wittenberg 2000), setting all variables constant at their median levels. We do this with our full multivariate model, containing both case-level and judge-level characteristics. If two judges with identical experience levels who graduated from similar law schools were deciding on the same type of case, the probability that the elected judge's decision would be upheld on appeal would be 4.6 percentage points higher than the decision of a merit-selected judge. The merit-selected judge's decision would be 3.0 percentage points more likely to be partially reversed and 1.6 percentage points more likely to be fully reversed. These are incremental

effects, to be sure, but every case matters greatly to those involved, and over the thousands of appeals court decisions made in Arizona, the differences begin to add up.

Although our simplest model also suggests that the unfettered appointment system leads to lower reversal rates than merit selection, this finding drops below conventional levels of statistical significance when we add control variables to the model. This indicates that much of the performance advantage for these judges that was apparent in Table 1 was in fact due to the types of cases that these judges decided. The coefficients on case types also show that these indeed matter for the outcome of appeals. Civil and, to a lesser extent, criminal cases are more likely to be reversed on appeal than juvenile cases and the "other" types of issues that trial courts address. Cases decided in 2013 and 2014 lead to more reversals than those in 2012. The prestige of judge's law school and his or her years of experience did not lead to significantly different outcomes. The fundamental finding of our analysis, then, is that Arizona's elected judges are less likely to be reversed than the state's merit-selected judges, all else equal.

VI. Robustness Analysis

While we have controlled for the characteristics of judges and of cases in these models, it is of course possible that we have not held constant other relevant attributes of the legal and social climates in which these judges operate. Merit-selected judges serve exclusively in Maricopa and Pima's Superior Courts during the period of our study (excepting Pinal County's one merit-selected judge) because these two large urban areas must use that procedure under state law. The strength of this natural experiment is the exogenous constraint imposed by state law: neither a county's leaders nor its lawyers have the chance to choose its selection procedure. Yet the weakness is that there are other potentially relevant

differences between these two counties and the rest of the state. In this section, we consider

three alternative explanations of the differences that we observe between elected and merit-

selected judges, introduce new control variables that capture these explanations, and ask

whether our findings are robust to their inclusion in multivariate models.[8]

County Case Workload. It could be that judges in Maricopa and Pima Counties

face far more cases or hold more trials than judges in Arizona's outlying counties, and that

these workload differences account for the variation in reversal rates. To determine whether

differences in reversal rates may be a function of the varying workloads that judges face

across Arizona's counties, we collected data on the average number of filings in each

county's Superior Courts, as well as the number of cases that went to trial. We took both

figures from 2014, gathering them from the Arizona Supreme Court Administrative Office

of the Court's annual report on Superior Court case activity by county (Arizona Judicial

Branch 2015a). We also used our dataset of appellate outcomes to calculate the total

number of appeals that each judge faced during the time period of our study, in order to

hold constant the volume of appeals.

County Demographics. The demographics in this diverse state vary dramatically

across counties – while 80.7% of the residents of Maricopa County are white, similar to the

79.3% of residents in Pima County, the populations of Navajo County (49.1% white) and

Apache County (23.8% white) are majority minority. Coconino and La Paz Counties also

have significant non-white populations. We have no clear intuition about how the diversity

---

[8] One factor that we cannot hold constant, of course, is a county's population, because a county's population
determines whether or not the county will use merit selection. We see no theoretical reason why population
size by itself should directly influence reversal rates. Instead, factors that often correlate with population – such
as a county's case workload, the demographic characteristics of it residents, and the state of its economy –
could affect reversal rates, and it is the impact of these factors that we directly address in our robustness
analysis.

of the social setting in which a judge operates should impact reversal rates; a plausible case could be made for effects going in either direction. Regardless, because this is a potentially important difference to take into account, we gather the white population percentage in each county from the ACS Demographic and Housing Estimates, from the U.S. Census Bureau's 2009-2013 5-Year American Community Survey.

County Economics. The economic resources of a county's residents could determine their ability to appeal a case, as well as the strength of the representation that they can afford when the case is initially heard. This could impact how Superior Court judges rule, how many of their cases are appealed, and the strength of those appeals. Across Arizona's counties, economic conditions vary significantly. Average unemployment levels in 2009-2013 were 8.8% in Maricopa and 9.9% in Pima, but 12.7% in Yuma, 13.8% in Mohave, 18.7% in Navajo, and 19.4% in Apache county, according to the Employment Status data in the U.S. Census Bureau's 2009-2013 5-Year American Community Survey.

Table 3 presents the results of our three multivariate models investigating the robustness of our main finding. Each column reports the coefficients of a multivariate model that builds on our main model, holding constant the characteristics of judges and of cases, and then adding the new explanatory factors to rule out three possible alternative explanations. Each model shows that the difference in reversal rates between elected and merit-selected judges remains strong and statistically significant. Demographics is the only factor that appears to matter, with judges less likely to be reversed in the more homogenously white counties. But the impact of elections, compared to merit selection, is robust to the inclusion of each new set of controls.

VII. Conclusions

The main finding provided by each of these regressions, and indeed throughout our empirical analysis, is one that challenges the conventional wisdom of many legal scholars who have trumpeted the advantages of appointment, especially merit selection, in producing better-qualified judges who outperform elected judges on the bench. Observers of Arizona's judicial system, some of whom have been participants themselves in the state judiciary, have also strongly recommended merit selection over election: "We further conclude that judges are best able to perform their constitutionally prescribed role in a hybrid merit based system like Arizona now has, featuring both appointment and retention election… In practice, the Arizona experience shows us that merit selection can work to promote both accountability and independence, and to select highly qualified, competent judges who will uphold the judicial promise to be fair and impartial arbiters of the law" (O'Connor and Jones 2008, 16, 24). "While not a perfect system, merit selection is functioning commendably and has significant advantages over a system of traditional partisan or non-partisan elections" (Harrison 2006, 262). By contrast, we find that on the metric of reversal rates, elected judges outperform those selected through merit selection. This finding is strongly statistically significant and robust across our bivariate comparisons and multivariate models. This is an original finding in a long literature which has found no differences between the decisionmaking (Watson and Downing 1969; Atkins and Glick 1974; O'Callaghan 1991; Romero et al. 2002) or the judicial quality (Watson and Downing 1969) of merit selection and elected judges.

Our new measure of reversal rates, of course, is not the only way to measure judicial quality. There are many other legitimate means of evaluating judicial performance, independence, and responsiveness. A full analysis of judicial selection systems could compare the types of sentences (Huber and Gordon 2004; Brace and Boyea 2008; Besley

and Payne 2013), the congruence with public opinion (Tabarrok and Helland 1999; Gordon and Huber 2007; Shepherd 2009b; Canes-Wrone et al. 2012), or the level of uncertainty (Hanssen 1999; Iaryczower et al. 2013) in the decisions handed out by judges chosen through different means. Arizona's trial courts would be a perfect context in which to do so. It is beyond the scope of this paper to conduct all of these analyses, which necessarily limits the scope of our conclusions. We also focus only on the performance of trial court judges, not on the court of last resort judges that have been the focus of so much of the literature. But we believe that trial courts – the bodies that deliver justice for the vast majority of people who come into contact with the legal system – and the state appellate courts deserve the close attention of scholars in future work.
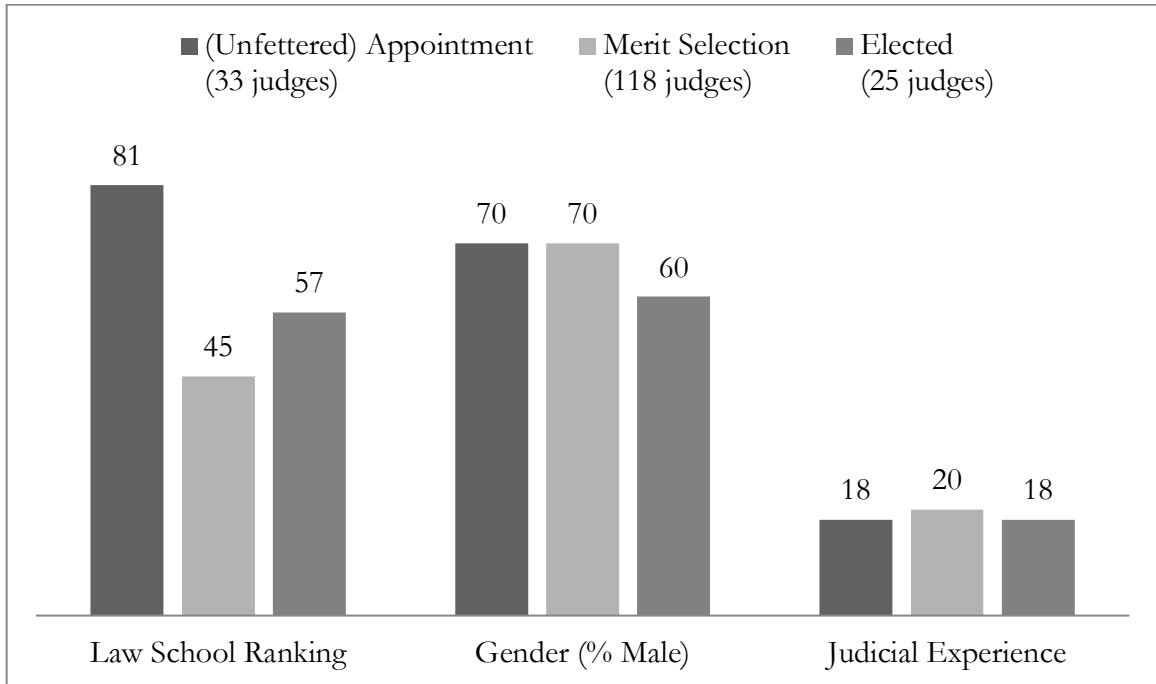
Figure 1. Selection Mechanisms and Judges' Personal Characteristics

Table 1. Selection Mechanisms and Reversal Rates, bivariate analysis

	Affirmed	Reversed in Part	Reversed
Merit Selection (2,190 cases)	80.7%	12.5%	6.8%
Unfettered Appointment (358 cases)	88.3%	5.6%	6.2%
Election (371 cases)	87.3%	7.3%	5.4%

Table 2. Models of Selection Mechanisms and Reversal Rate

	Selection Mechanisms Only	Case-Level Variables	Case- and Judge-Level Variables
Judge was Elected	-0.49**	-0.45**	-0.54**
	(0.18)	(0.17)	(0.19)
Judge Appointed by Gov.	-0.56**	-0.49*	-0.43
	(0.25)	(0.27)	(0.29)
Ranking of Law School			-0.0003
			(0.002)
Male Superior Court Judge			-0.07
			(0.14)
Years on the Bench			-0.003
			(0.02)
Years on the Arizona Bar			-0.01
			(0.01)
Criminal Case		1.21**	1.30**
		(0.40)	(0.42)
Civil Case		2.17**	2.24**
		(0.40)	(0.42)
Juvenile Case		0.57	0.62
		(0.44)	(0.45)
Decided in 2013		0.34**	0.36**
		(0.15)	(0.16)
Decided in 2014		0.60**	0.61**
		(0.16)	(0.16)
Observations	2,919	2,919	2,819

Notes: Results from ordered logit models, with robust standard errors clustered on individual judges in parentheses. In the last model, 100 cases lost due to missing information on judges' backgrounds. ** $p<0.05$, * $p<0.10$

Table 3: Robustness Analysis

	Adding County Workloads	Adding County Demographics	Adding County Economics
Judge was Elected	-0.61**	-0.54**	-0.54**
	(0.24)	(0.19)	(0.22)
Judge Appointed by Gov.	-0.56*	-0.45	-0.43
	(0.34)	(0.29)	(0.30)
Ranking of Law School	-0.0002	0.00006	0.00009
	(0.002)	(0.002)	(0.002)
Male Superior Court Judge	-0.05	-0.07	-0.08
	(0.14)	(0.13)	(0.14)
Years on the Bench	-0.001	-0.003	-0.003
	(0.02)	(0.02)	(0.02)
Years on the Arizona Bar	-0.01	-0.013	-0.012
	(0.02)	(0.02)	(0.02)
Criminal Case	1.30**	1.49**	1.49**
	(0.42)	(0.41)	(0.41)
Civil Case	2.23**	2.44**	2.43**
	90.44)	(0.41)	(0.41)
Juvenile Case	0.62	0.81*	0.81*
	(0.47)	(0.44)	(0.44)
Decided in 2013	0.36**	0.37**	0.38**
	(0.16)	(0.16)	(0.16)
Decided in 2014	0.621**	0.62**	0.63**
	(0.16)	(0.16)	(0.16)
Number of Appeals	-0.0002		
	(0.004)		
Case Filings per Judge	0.00004		
	(0.0002)		
Trials per Judge	-0.02		
	(0.02)		
Percent White		-0.02**	
		(0.008)	
Unemployment Rate			0.003
			(0.006)
Observations	2,819	2,813	2,813

Notes: Results from ordered logit models, with robust standard errors clustered on individual judges in parentheses. Each model has fewer cases than our main models because of missing information on judges' backgrounds. ** p<0.05, * p<0.10

References

American Bar Association Commission on the 21st Century Judiciary. 2003. "Justice in Jeopardy." Chicago: American Bar Association.

Arizona Judicial Branch. 2015a. "2014 Case Activity by County." 2015. Accessed at https://www.azcourts.gov/statistics/Annual-Data-Reports/2014-Data-Report/2014-Case-Activity-by-County in June 2015.

Arizona Judicial Branch. 2015b. "Superior Courts." Accessed at http://www.azcourts.gov/AZCourts/SuperiorCourt.aspx in April 2015.

Arizona State Bar. "Judicial Merit Selection in Arizona." 2013. Accessed at http://www.azbar.org/aboutus/leadership/boardofgovernors/importantissues/fairandimpartialcourts/meritselection in December 2013.

Aspin, Larry. 2007. "Trends in Judicial Retention Elections, 1964-2007." *Judicature* 90(5): 208-213.

Atkins, Burton M. and Henry R. Glick. 1974. "Formal Judicial Recruitment and State Supreme Court Decisions." *American Politics Quarterly* 2(4): 427-449.

Besley, Timothy and Abigail Payne. 2013. "Implementation of Anti-Discrimination Policy: Does Judicial Selection Matter?" *American Law and Economics Review* 15(1): 212-251.

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

Brace, Paul and Brent D. Boyea. 2008. "State Public Opinion, the Death Penalty, and the Practice of Electing Judges." *American Journal of Political Science.* 52(2): 360-372.

Caldarone, Richard P., Brandice Canes-Wrone, and Tom S. Clark. 2009. "Partisan Labels and Democratic Accountability: An Analysis of State Supreme Court Abortion Decisions" *Journal of Politics* 71(2): 560-573.

Caldeira, Gregory A.. 1983. "On the Reputation of State Supreme Courts." *Political Behavior* 5(1): 83-108.

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.

Canes-Wrone, Brandice, Tom S. Clark, and Jason P. Kelly. 2014. "Judicial Selection and Death Penalty Decisions." *American Political Science Review* 108(1): 23-39.

Cann, Damon. 2002. "Campaign Contributions and Judicial Behavior." *American Review of Politics* 23(3): 261-274.

Cann, Damon. 2007. "Beyond Accountability and Independence: Judicial Selection and State Court Performance." *Judicature* 90(5): 226-232.

Cann, Damon M., Chris W. Bonneau, and Brent D. Boyea. "Campaign Contributions and Judicial Decisions in Partisan and Nonpartisan Elections," in *New Directions in Judicial Politics*, ed. Kevin T. McGuire, 38-52. New York: Routledge, 2012.

Canon, Bradley C.. "The Impact of Formal Selection Processes on the Characteristics of Judges – Reconsidered." *Law and Society Review* 6(4): 579-594.

Cass, Ronald A. 1995. "Judging: Norms and Incentives of Retrospective Decision-Making." *Boston University Law Review* 75:941–96.

The Center for Voting and Democracy. "Congressional and Presidential Primaries: Open, Closed, Semi-Closed, and 'Top-Two.'" 2012. Accessed at http://www.fairvote.org/research-and-analysis/presidential-elections/congressional-and-presidential-primaries-open-closed-semi-closed-and-top-two/ in April 2015.

Choi, Stephen J., G. Mitu Gulati, and Eric A. Posner. "Professionals or Politicians: The Uncertain Empirical Case for an Elected Rather than Appointed Judiciary." *Journal of Law, Economics, and Organization* 26(2): 290-336.

Choi, Stephen J., Mitu Gulati, and Eric A. Posner. "Judicial Evaluations and Information Forcing: Ranking State High Courts and their Judges." *Duke Law Journal* 58(7): 1313-1381.

Choi, Stephen J.; Gulati, Mitu; Posner, Eric A. 2012. "What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals" *Journal of Law, Economics & Organization*: 28(3):518.

Comparato, Scott A.. "On the Reputation of State Supreme Courts Revisited." Paper presented at the annual meeting for the Midwest Political Science Association, Chicago, Illinois, April 2002.

Cross, Frank B., Lindquist, Stefanie A. 2009. "Judging the Judges." *Duke Law Journal* 58:1383–437.

Dear, Jake and Edward W. Jessen. 2007. ""Followed Rates" and Leading State Cases, 1940-2005." 41: 683-711, 1665-1670.

Dinan, Stephen. 2009. "Sotomayor Reversed 60% by High Court." *Washington Times*, May 27, 2009.

Epstein, Lee, William M. Landes, and Richard A. Posner. 2013. *The Behavior of Federal Judges: A Theoretical and Empirical Study of Rational Choice,* Cambridge, MA: Harvard University Press.

Flango, Victor Eugene and Craig R. Ducat. 1979. "What Difference Does Method of Judicial Selection Make? Selection Procedures in State Courts of Last Resort." *Justice System Journal* 5(1): 25-44.

Franklin, Charles H."Behavioral Factors Affecting Judicial Independence," in *Judicial Independence at the Crossroads: An Interdisciplinary Approach*, eds. Stephen B. Burbank and Barry Friedman, 148-159. Sage Publications and the American Academy of Political and Social Science, 2002.

Friedman, Lawrence M., Robert A. Kagan, Bliss Cartwright, and Stanton Wheeler. 1981. "State Supreme Courts: A Century of Style and Citation." *Stanford Law Review* 33(5): 773-818.

Gill, Rebecca D. Gill. 2013. "Beyond High Hopes and Unmet Expectations: Judicial Reforms in the States." *Judicature* 96(6): 278-295.

Glick, Henry R. and Craig F. Emmert. 1987. "Selection Systems and Judicial Characteristics: The Recruitment of State Supreme Court Judges." *Judicature* 70(4): 228-235.

Gordon, Sanford C. and Gregory A. Huber. 2007. "The Effect of Electoral Competitiveness on Incumbent Behavior." *Quarterly Journal of Political Science* 2(2): 107-138.

Hall, Melinda Gann. 1992. "Electoral Politics and Strategic Voting in State Supreme Courts." *Journal of Politics* 54(2): 427-446.

Hall, Melinda Gann. 2001. "State Supreme Courts in American Democracy: Probing the Myths of Judicial Reform." *American Political Science Review* 95(2): 315-330.

Hall, Melinda Gann. 2007. "Voting in State Supreme Court Elections: Competition and Context as Democratic Incentives." *Journal of Politics* 69(4): 1147-1159.

Hall, Melinda Gann. "On the Cataclysm of Judicial Election and Other Popular Antidemocratic Myths," in *What's Law Got to Do with It? What Judges Do, Why They Do It, and What's at Stake*, ed. Gardner Geyh, 223-247. Palo Alto, Stanford University Press, 2011.

Hanssen, F. Andrew. 1999. "The Effect of Judicial Institutions on Uncertainty and the Rate of Litigation: The Election versus Appointment of State Judges." *Journal of Legal Studies* 28(1): 205-232.

Harrison, Mark I., Sara S. Greene, Keith Swisher, and Meghan H. Grabel. 2006. "On the Validity and Vitality of Arizona's Judicial Merit Selection System: Past, Present, and Future." *Fordham Urban Law Journal* 34(1): 239-263.

Hartzell, Jerry. 2014. "Probability of Success on Appeal: Reversal Rates for the Fourth Circuit and the North Carolina Court of Appeals." *Trial Briefs*, April 30, 2014:30-34.

Holmstrom, Bengt and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24-52.

Huber, Gregory A. and Sanford C. Gordon. 2004. "Accountability and Coercion: Is Justice Blind When It Runs for Office?" *American Journal of Political Science* 48(2): 247-263.

Hurwitz, Mark S. and Drew Noble Lanier. 2003. "Explaining Judicial Diversity: The Differential Ability of Women and Minorities to Attain Seats on State Supreme and Appellate Courts." *State Politics and Policy Quarterly* 3(4): 329-352.

Hurwitz, Mark S. and Drew Noble Lanier. 2008. "Diversity in State and Federal Appellate Courts: Change and Continuity Across 20 Years." *Justice System Journal* 29(1): 47-70.

Iaryczower, Matias, Garrett Lewis, and Matthew Shum. 2013. "To Elect or to Appoint? Bias, Information, and Responsiveness of Bureaucrats and Politicians." *Journal of Public Economics* 97(1): 230-244.

Kang, Michael S. Joanna M. Shepherd. 2011. "The Partisan Price of Justice: An Empirical Analysis of Campaign Contributions and Judicial Decisions." *New York University Law Review* 86: 69-130.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44(2): 347−61.

Klumpp, Albert J. "Arizona Judicial Retention: Three Decades of Elections and Candidates." 2008. *Arizona Attorney.* November 2008:12-18.

Lee, Stephen E. 1973. "Judicial Selection and Tenure in Arizona." *Law and the Social Order* 4 (1): 51-80.

Liptak, Adam. "Supreme Court to Review Bans on Solicitations in Judge Races." *The New York Times,* January 18, 2015.

McCall, Madhavi. 2003. "The Politics of Judicial Elections: The Influence of Campaign Contributions on the Voting Patterns of Texas Supreme Court Justices, 1994-1997." *Politics and Policy* 31(2): 314-343.

McCall, Madhavi M. and Michael A. McCall. 2007. "Campaign Contributions, Judicial Decisions, and the Texas Supreme Court: Assess the Appearance of Impropriety." *Judicature* 90(5): 214-225.

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

Mott, Rodney L. 1936. "Judicial Influence" *American Political Science Review* 30(2): 295-315.

National Center for State Courts. 2002. "Call to Action: Statement of the National Summit on Improving Judicial Selection." Williamsburg, Virginia: National Center for State Courts.

National Center for State Courts. "Methods of Judicial Selection: Arizona." 2015. Accessed at http://www.judicialselection.com/judicial_selection/methods/selection_of_judges.cfm?state=AZ in April 2015.

O'Callaghan, Jerome. 1991. "Another Test for the Merit Plan." *Justice System Journal* 15(1): 477-485.

O'Connor, Sandra Day and RonNell Andersen Jones. 2008. "Reflections on Arizona's Judicial Selection Process" *Arizona Law Review* 50:15-24.

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" *State Politics & Policy Quarterly* 15(2): 211-238.

Posner, Richard A. 2000. "Is the Ninth Circuit Too Large? A Statistical Study of Judicial Quality" *The Journal of Legal Studies* 29(2):711-719.

Reddick, Malia. 2002. "Merit Selection: A Review of the Social Scientific Literature." *Dickinson Law Review* 106(4): 729-745.

Roll, John M. 1990. "Merit Selection: The Arizona Experience." *Arizona State Law Journal* 22(3): 837-894.

Romero, Francine Sanders, David W. Romero, and Victoria Ford. "The Influence of Selection Method on Racial Discrimination Cases: A Longitudinal State Supreme Court Analysis." Paper presented at the annual meeting for the Midwest Political Science Association, Chicago, Illinois, April 2001.

Schlesinger, Robert. 2009. "The Myth of Sotomayor's 60 Percent Supreme Court Reversal Rate." *US News and World Report*, May 28, 2009.

Sen, Maya. 2015. "Is Justice Really Blind? Race and Reversal in US Courts." *Journal of Legal Studies.* 44(1): 187-229.

Shepherd Joanna M. 2009(a). "Are Appointed Judges Strategic Too?" *Duke Law Journal* 58(7): 1589-1626.

Shepherd, Joanna M. 2009(b). "The Influence of Retention Politics on Judges' Voting." *Journal of Legal Studies* 38(1): 169-203.

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

Waltenburg, Eric N. and Charles S. Lopeman. 2000. "Tort Decisions and Campaign Dollars." *Southeastern Political Review* 28(2): 241-263.

Ware, Stephen J. 1999. "Money, Politics, and Judicial Decisions: A Case Study of Arbitration Law in Alabama." *Journal of Law and Politics* 15: 645-686.

Watson, Richard A., and R.G. Downing. *The Politics of the Bench and the Bar: Judicial Selection Under the Missouri Nonpartisan Court Plan.* New York: John Wiley and Sons, 1969.