Judicial retention elections have been part of Arizona’s governmental system for more than 30 years. Enacted in 1974 as an element of the state’s judicial merit selection plan, retention elections are held for all judges in Maricopa and Pima counties and for appellate and Supreme Court judges statewide. Since their debut in 1976, 16 sets of retention elections have occurred in Arizona, and a total of 735 candidates have run for retention.

What, if anything, can we learn from these elections? The question is far from academic.

For one thing, the issue of judicial selection has been a perpetual source of controversy in Arizona, dating back even before the state’s admission to the union and continuing to this day. In addi-
tion, retention elections are generally not well understood. Surprisingly little research has been done on the topic, which has hampered an informed debate on the relative merits of different judicial selection methods. In my doctoral dissertation and subsequent research, I have sought to address the latter problem, analyzing in detail the results of judicial retention elections in many different jurisdictions. In this article are the results of my analysis of Arizona’s retention elections. For the most part, I find that their outcomes are comparable to those of other retention jurisdictions—but not entirely so. Moreover, in examining data on the judges who have run for retention, I have uncovered limited but undeniable evidence that Arizona’s switch to merit selection has resulted in a measurable improvement in the quality of the state’s judiciary.

Retention Approval Rates

The voter approval rates for Arizona retention candidates exhibit three general characteristics.

First, defeats of judges have been rare occurrences, and even near-defeats have been rare. Only two judges have failed to attain the 50 percent requirement for retention, and only seven others have fallen as low as 60 percent. This is in line with the results of other populous retention-election states. In Missouri, for instance, only two judges have been defeated since retention elections were instituted in 1948. And in Illinois, which in 1972 raised its approval requirement from 50 percent to 60 percent, more than 98 percent of judges have been retained, even under the more stringent requirement.

Second, the approval rates for most or all of the judges who...
appear on the same ballot cluster tightly together. This, too, is par for the course among the retention-election states. It occurs because most voters cast blocks of all-yes or all-no votes rather than singling out any particular candidates (though there are exceptions; see the next section).

Third, in 1990 there was a sudden and substantial drop in approval rates statewide. Figures 1 and 2 illustrate the extent of the drop for the various categories of Arizona candidates. This drop occurred not only in Arizona but in nearly every retention jurisdiction in the nation. Its cause was an influential national anti-incumbent campaign, a grassroots campaign that was the precursor for Ross Perot’s 1992 presidential candidacy. The campaign was directed only against the U.S. Congress, but in my dissertation research I discovered that it had a huge spillover effect, causing the defeats of state and local incumbents in unprecedented numbers all over the country. And it also affected approval rates for retention judges. This is not surprising, because retention candidates are specifically identified on the ballot as incumbents.

It is significant that although approval rates in every other retention state rebounded quickly from the 1990 drop, Arizona’s rates fell even lower in 1992. And while a rebound did occur in Arizona two years later, in Maricopa County the rates have to this day remained substantially below their pre-1990 levels. Polling data that would conclusively explain these unique results does not exist; however, I believe there are two obvious suspects.

One is the 1992 Perot candidacy, which reinforced the government-is-broken message of the 1990 campaign and which was more successful in Arizona than in all but a handful of the other U.S. states. As for
Maricopa County, there may well be a hardening of the attitudes of a portion of the local electorate toward the criminal justice system in the wake of the 1992 election of a new Maricopa County sheriff, Joe Arpaio’s high-profile, controversial tenure and the contemporaneous reduction in Maricopa’s retention approval rates seem unlikely to be coincidental.

Studies on all types of judicial elections have found that voters are not well informed about judicial candidates. The events of 1990 and thereafter suggest that many voters in retention elections rely on general attitudes toward politics or government as a substitute for specific information about judges.

Voter Targeting of Individual Judges

One focus of my research has been on instances in which a judge is singled out for negative votes by a significant proportion of the electorate. What causes voters to target a specific retention candidate?

In Arizona, as in the other states I have studied, the answer is clear. Information about a judge’s performance has only a minimal impact on its own—it is never sufficient to defeat a judge. Rather, it is negative publicity, whether due to performance alone or otherwise, and usually occurring just before the election, that in every case explains the most substantial voter reactions. In Arizona, this is the case for all nine of the instances of defeats or near-defeats. In two other instances, negative publicity was counteracted by visible campaigns of support for retention candidates, significantly reducing the “hit” that the candidates would have taken at the polls.

As for performance information, evaluations of the job performance of retention candidates have been available to Arizona voters throughout the state’s retention history. From 1976 to 1992, the state’s major bar associations conducted lawyer surveys on the performance of retention candidates. Since 1994, the evaluation function has been performed by the state’s Judicial Performance Review Commission. But despite the reliable provision of detailed information from these sources, candidates who have received poor evaluations without any accompanying negative publicity have seen only a slight voter reaction. This has been the case under both systems of evaluation.

To be fair, the current system of distributing performance information through the state’s Voter Information Pamphlet may yet prove to be influential with more voters. A full analysis is not yet possible because the number of judges who have received anything less than a unanimous or near-unanimous endorsement from the JPR Commission has simply been too small. Nonetheless, the unequivocal record from Arizona and from other populous retention states offers a strong message to groups that evaluate retention candidates: Even the most negative of performance evaluations is by itself of little electoral consequence.

Improvement in Judicial Performance

As part of my overall analysis, I gathered and reviewed historical data from the performance evaluations of Arizona retention candidates. My original purpose in doing so was to compare evaluation results to election results, as discussed above. But in taking a closer look at the evaluation results, I found that they have an important story to tell on their own.

Regrettably, the original source materials from the 1976–1992 bar surveys have been destroyed, and I was unable to locate a complete set of scores from any other source. Fortunately, however, back issues of Arizona Attorney provided me with sets of summary scores from the 1976, 1978 and 1990 elections. I also acquired the JPR Commission’s vote results for the five elections between 1998 and 2006. These eight elections allow for at least a partial analysis, based on the percentages of attorneys or commission members who favored the retention of each candidate.

Table 1 summarizes the scoring data (see p. 18). The table shows that between the 1976–1978 period and 1990, mean and median ratings for judges increased noticeably, and the proportion of judges receiving particularly low scores declined considerably. The same also occurred between 1990 and the 1998–2006 period. Statistical testing of the full sets of scores found that in both instances the score increases are formally statistically significant.

What makes these increases so important? In the 1970s, nearly all of the judges who were being evaluated had reached the bench through partisan elections. As time passed, though, judges chosen through merit selection became a larger and larger proportion of the judiciary, and today they comprise 100 percent of the retention candidates. In other words, the transition from a partisan-elected to a merit-appointed judiciary and the improvement in evaluation scores have occurred simultaneously. This is doubtless not a coincidence.

Remarkably, the increases have occurred even as the judicial system has steadily expanded in size. The growing numbers of retention candidates throughout the 1976–2006 period mean that there have been higher proportions of newer, less experienced judges than if the system were more stable size-wise. It could reasonably be expected that new judges with little or no experience...
would receive relatively lower scores, reducing the overall averages. But this has not happened.

Two further comments should be made regarding the figures in Table 1.

First, it has been suggested by some critics that the JPR Commission has been too lenient in its voting; if true, that would contribute to the high scores seen in the 1998–2006 period. However, even if this accusation has merit, it would not explain the score improvements that predate the commission’s establishment.

Second, the scores that I analyzed are based simply on the question of whether a judge performs at a baseline level worthy of retention. The scores do not indicate whether judges are performing at higher-than-minimal levels. While more research is needed on the question, Table 1 suggests indirectly that there are more judges worthy of “A” and “B” grades today than in the past, and that there are fewer “D” and “F” judges.

Conclusion

I should make clear that the goal of my research is not to advocate for or against any particular type of judicial selection system. And in fact, some of what I have presented here will be well received by advocates of partisan judicial elections. In particular, the near-100 percent success rate of retention candidates certainly does not refute the argument that retention elections by themselves are inadequate to hold substandard judges accountable to the public. Similarly, considering that the whole point of retention elections is to protect incumbent judges from facing challengers and to retain or remove them based on performance, the rates of performance-based voting are extremely low.

Nevertheless, the bigger picture here suggests that the sets of candidates who have sought retention have become more competent and more retention-worthy under the state’s merit selection plan. Furthermore, anecdotal evidence suggests that the rigorous judicial performance evaluation programs that were adopted in 1992 have discouraged questionable judges from even seeking retention in the first place. To the extent that this is the case, then a retention election process without removals is actually a desirable consequence. Whether or not a substandard judge occasionally seeking and winning retention is an acceptable tradeoff for these improvements is an issue that Arizonans will have to weigh carefully.

Debate in the states over judicial selection systems is invariably long on rhetoric and argument and short on factual support. Continued research of the type presented here can play an invaluable role in evaluating the claims made by proponents and opponents of different systems. Ideally, such research will help Arizona address the subject with less conflict in its second century than in its first.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Candidates</th>
<th>Mean Score</th>
<th>Median Score</th>
<th>Candidates Scoring Below 75</th>
<th>Candidates Scoring Below 50</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>20</td>
<td>82.3</td>
<td>87</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>1978</td>
<td>36</td>
<td>83.5</td>
<td>87</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>1990</td>
<td>49</td>
<td>89.3</td>
<td>93</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>1998</td>
<td>50</td>
<td>97.1</td>
<td>100</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2000</td>
<td>56</td>
<td>98.3</td>
<td>100</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2002</td>
<td>60</td>
<td>98.6</td>
<td>100</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>73</td>
<td>97.5</td>
<td>100</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>2006</td>
<td>64</td>
<td>99.6</td>
<td>100</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note: 1976-1990 scores are percentages of responding attorneys favoring retention; 1998-2008 scores are calculated from JPR Commission votes.

4. Id.
5. Klumpp, supra note 2.
9. A full set of the 1976–1992 survey results would allow for such an analysis. If any reader happens to be in possession of those results, please contact the author.