Measuring Merit in Rhode Island's Natural Experiment in Judicial Selection

Mirya R. Holman

Florida Atlantic University
Measuring Merit in Rhode Island’s Natural Experiment in Judicial Selection

Mirya R. Holman, Ph.D.*

INTRODUCTION

Rhode Island’s decision to change the method by which the State selects its Supreme Court justices has provided the world of judicial behavior with an exciting natural experiment: a way to examine whether changing the selection method for judges does, in fact, matter. Proponents of a particular kind of selection method – whether it be partisan election, merit selection, appointment, or non-partisan election – argue that their preferred selection method will produce a court with “better” judges.¹ In the case of Rhode Island, a series of scandals involving Supreme Court justices resulted in a move to alter the selection process from one involving the Grand Committee to a less partisan process by which the justices would be selected based on merit.² Those

* Mirya R. Holman is an Assistant Professor of Political Science at Florida Atlantic University. The author would like to extend her thanks to Hala Furst and Andrew Redman for research assistance, to Michael Yelnosky for the invitation to present and for providing essential pieces of information, and to Mitu Gulati, Stephen Choi, and Eric Posner for use of their data and methods.


advocating for the change of selection method argued that the change would result in better justices. But has the change actually resulted in these changes? To answer this question, I will present a series of measures of judicial quality that were developed by Stephen Choi, Mitu Gulati, and Eric Posner. Specifically, I will present information on the production, influence, expertise, verbosity, and independence for Rhode Island Supreme Court justices selected prior to the selection change (those on the Court from 1989 to 1991), a group of justices split between those selected prior to the change and those selected post-change (1998 to 2000), and a collection of justices selected entirely by the merit selection method (2005 to 2007).

I. Measures

The dataset used in this paper is based on several interesting pieces of information about Supreme Court justices on the Rhode Island Supreme Court. The paper looks at three specific sets of justices on the Supreme Court: first, a group of justices (specifically Fay, Kelleher, Weisberger, Murray, Shea) who served on the Court from 1989 to 1991, which was prior to the selection method change. The remainder of the paper will refer to these justices as the "pre-merit" or the "early group" of justices. The second period examined is from 1998 to 2000 and includes Justices Bourcier, Flanders, Goldberg, Lederberg, and Weisberger. In the


3. See Wayne Miller, Right Now! Urges Reform in Choosing Judges; Cronyism, the Group Says, Must Forever be Removed from the Process, PROVIDENCE J., Aug. 27, 1993, at A8. As Barton P. Jenks articulates, "Although the new judicial selection procedure is flawed, it unquestionably is an improvement over its predecessor." Jenks, supra note 2, at 65.

4. The data on this period comes, with permission, from: Stephen J. Choi et al., Judicial Evaluations and Information Forcing: Ranking the State
paper, this Court is referred to as the “hybrid Court” or the “mid Court” (as it contains justices selected for the court from both before and after the merit change). The final period of analysis is from 2005 to 2007 and includes Justices Williams, Goldberg, Flaherty, Robinson, and Suttell. This period will be referred to as the “post-merit” or “later period” for the remainder of the paper.

For each of these justices, data was collected on multiple measures, including: (1) the number of published opinions; (2) the numbers of citations from outside the state (that is, non-precedent driven citations); (3) the number of citations from outside Rhode Island’s federal district; (4) the length of opinions; (5) the subject matter of the opinions; and (6) the amount of open disagreements (dissenting opinions) between justices with the same political background. These measures are based on a significant body of research. There are alternate measures that have been proposed by those that question the usefulness and validity of these measures. However, for this research, I will leave the debate to other scholars. While these measures are certainly imperfect, they should produce equal levels of bias and variance, regardless of the year examined. Essentially, I assume that any inadequacies of these measures are not the result of the year of measurement, so looking year to year will allow me to engage in a reasonable comparison.


1. Publication Rates

Those looking to measure judicial quality have often turned to the published opinion as the base for such measures; the published judicial opinion is, after all, the “heart of the common law system.” Previous research by Choi, Gulati and Posner has argued that publishing an opinion is more work than issuing an unpublished disposition. In addition, by publishing an opinion, the authoring judge risks bringing greater external scrutiny, as their attitudes and reasoning are more widely available. A judge who publishes an opinion also is engaging in an act that benefits the system, as a published opinion represents an attempt by the judge to share his/her reasoning with the parties and with others who seek to understand the resolution of the dispute. Further, the publication of an opinion gives it greater precedential weight. As such, there are two possible competing hypotheses that might be generated by the data examined here. First, an observer might contend that the merit selection process should bring a higher quality of justice to the bench than was previously selected through the Grand Committee, which was notoriously partisan. If this is so, the expectation should be that the justices selected under the merit system would be selected for meritorious reasons, not partisan grounds. Thus, these judges may be “better” at their


9. See Choi et al., Professionals or Politicians?, supra note 5.


11. See Garland, R.I. Changes the Method of Selecting its Justices, supra note 2.
jobs and more productive.

Hypothesis 1a: Justices Selected by Merit, Instead of by the Grand Committee, will Publish more Opinions than Those Selected by the Grand Committee.

Another possibility is that those justices selected for the Supreme Court via the merit system may be more aware of the public perception of their actions.\(^1\) This may be especially true in Rhode Island, where a series of scandals were the impetus behind the change in selection methods.\(^2\) As such, the justices on the post-merit court may be less likely to publish opinions, as a published opinion is more widely available (including to their peers, the media, and the general public), thus subjecting their decisions to scrutiny. Thus, it might be reasonable to expect that post-merit courts will publish fewer opinions.

Hypothesis 1b: Justices Selected by Merit will Publish less Opinions than Those Selected by the Grand Committee.

2. Citations

Looking next for a rough measure of influence, researchers have commonly used citations by outside authorities.\(^3\) There are three sets of influence as measured by citations presented here: all citations to majority opinions published by the state high court in question from (1) other state courts; (2) federal courts (other than the home federal circuit); and (3) secondary sources, such as law reviews. Citation counts can be used to measure the quality of

---


analysis, nimbleness in writing, and creativity. If the extant scholarship is accurate, I expect that the justices selected through a non-partisan mechanism will be cited more often than those justices from the early period. As compared to a court where justices can be selected because of personal or political reasons, a justice selected for reasons of merit should be producing opinions of greater quality of analysis. Thus, I expect that justices on the hybrid or post-merit Courts will have higher levels of citation than the pre-merit Court.

Hypothesis 2: Justices Selected by Merit Selection will be Cited to More Often – Regardless of the Measure of Citations – than Justices Selected by the Grand Committee.

3. Length of Opinions

There are diverse opinions about what the page length of a judicial opinion means. Generally, scholars who have examined the length of opinions argue that a long opinion is indicative of the amount of effort by the judge to justify her opinion in the case; legitimacy is derived from explaining the reasoning behind the opinion fully; i.e. “[h]e is right because he has satisfied us that his judgment rests on good grounds.” However, others have suggested that a shorter opinion is more concise, and, as such, is harder to write than a long opinion. Regardless of who is correct, both sides – and many scholars – believe that page length

15. See Choi et al., Are Judges Overpaid?, supra note 5; Choi et al., Judicial Evaluations, supra note 4; Choi et al., Professionals or Politicians?, supra note 5.
17. See Richard A. Posner, Judicial Behavior and Performance: An Economic Approach, 32 FLA. ST. U. L. REV. 1259, 1276-77 (2005) (suggesting that creative judges are likely to be cited more often because they are more often reversed).
MEASURING MERIT

2010

is an important measure of judicial effort.\textsuperscript{20} The disagreements in expectations produce the possibility for conflicting hypotheses.\textsuperscript{21}

Hypothesis 3a: Justices Selected by the Merit System will Publish Longer Opinions.

Hypothesis 3b: Justices Selected by the Merit System will Publish Shorter Opinions.

4. Subject Areas

An additional area of judicial behavior that is quantifiable is an evaluation of the most common subject areas in which judges choose to publish. Choi, Gulati, and Posner developed fifteen categories of the most common subject areas after conducting a broad survey of the cases that were heard by state Supreme Court justices.\textsuperscript{22} Theoretically, there is reason to believe that a justice selected by the merit system might be less likely to specialize in a specific kind of case and would be more broadly trained than those justices selected by the Grand Committee.

Hypothesis 4: Courts Comprised of Justices Selected by Merit Selection will Publish on a more Diverse Group of Cases than Courts with Justices Selected by the Grand Committee.

5. Disagreement

The fourth measure looks at the willingness of a judge to disagree with co-partisans, either by dissenting from their

\textsuperscript{20} See Goutal, supra note 18; Wells, supra note 19; Leach, supra note 18.

\textsuperscript{21} Alternately, some suggest that the length of opinion holds little meaning in the process of judging judges; some cases call for more elaboration, while other opinions can be expresses succinctly. As such, there is the possibility that the length of opinion will have no relationship at all with the change in selection mechanism.

\textsuperscript{22} The subjects include administration, campaign finance, church and state, criminal, environmental, business, first amendment, government, courts, labor, other, private law, rights, property, tax, and torts. See Choi et al., Are Judges Overpaid?, supra note 5; Choi et al., Professionals or Politicians?, supra note 5; Choi et al., Judicial Evaluations, supra note 4, at 1323.
opinions or writing majority opinions that induce dissents. In calculating this measure, I look at dissenting opinions, which are open and public statements of disagreements. First, the number of disagreements by a judge against co-partisans divided by the total number of disagreements by the justice is examined, which provides a rough idea as to how likely a judge is to agree or disagree with people of her same political party. However, if a justice is the only Republican or Democrat on a specific court, she will necessarily oppose opposite party judges (due to the lack of any co-partisans). While partisanship is far from a completely reliable measure, previous scholars have found that justices that deviate from their partisan roots are vastly the exception, not the rule. In order to account for this possible bias, the total number of majority opinions by other justices of the same party is divided by the total (overall) number of majority opinions written by the court.

Independence is then defined as the percent of disagreements with co-partisans (i.e. how often the justice dissents against members of their own party) over the percent of co-partisan majority opinions (i.e. how many of the majority opinions are written by co-partisans). The scale is calibrated so that a more negative score corresponds to a justice who writes opposing opinions against opposite-party justices more frequently than the background pool of majority opinions authored by opposite-party justices. On the other hand, a more positive score corresponds to an authoring justice who writes opposing opinions less frequently against opposite-party justices compared with the background pool of opinions (and thus more frequently against co-partisans). Generally, Choi, Gulati and Posner have viewed a more positive

23. See Choi et al., Are Judges Overpaid?, supra note 5, at 50; Choi et al., Professionals or Politicians?, supra note 5; Choi et al., Judicial Evaluations, supra note 4, at 1323.


25. See Choi et al., Are Judges Overpaid?, supra note 5, at 50; Choi et al., Professionals or Politicians?, supra note 5; Choi et al., Judicial Evaluations, supra note 4, at 1323.
score as indicative of a more independent justice. Other scholarship has suggested that disagreement among judges is a negative— a sign of disagreeability or cantankerousness.

In the context of the Rhode Island Supreme Court selection method change, one might expect that a justice selected based on merit and not based on a heavily partisan selection process would be less likely to be swayed by partisan alignments on the Court.

Hypothesis 5: Those Justices Selected Through the Merit Process will be more Independent than Those Selected by the Grand Committee.

Overall, these five hypotheses and their subparts suggest that the landscape of measuring judicial quality is wide, diverse, and often difficult to quantify. Here, I attempt to use existing measures and specific hypotheses to examine the effect of selection change on judicial behavior. The presentation of findings will proceed in the order of the hypotheses.

II. FINDINGS

1. Production

Looking first at the number of published cases for each justice, the justices in the pre-merit period published at a higher rate than those in the post-merit period. Specifically, the earlier group of justices published an average of 28.7 cases per year per justice, while the middle group of justices published 18.2 cases per year, and the later group of justices published an average of 21.7 cases per year per justice (See Figure 1). The difference between the pre-merit court and either the hybrid (Pr(T > t) = 0.0231) or the post-merit (Pr(T > t) = 0.0004) courts is statistically significant. As Figure 1 also shows, overall, there is little variation between the justices on each Court.

26. Choi et al., Are Judges Overpaid?, supra note 5, at 50; Choi et al., Professionals or Politicians?, supra note 5; Choi et al., Judicial Evaluations, supra note 4, at 1323.
This suggests that there is preliminary evidence that Hypothesis 1b is correct; justices selected by the merit system seem to be publishing at a lower rate than those selected by the Grand Committee. One interesting finding is that Weisberger, the one pre-merit Justice on the hybrid court, had a much lower publication rate in 1998-2000 than he did in 1989 to 1991, suggesting that either: (1) his production fell as he approached retirement; or (2) there were court-level changes in the expectation of publication or performance post-merit.

2. Citations

The next substantial measure that is used to measure the quality of judicial performance is the rate at which judge's
opinions are cited by other courts. Presented here are three different citation count measures: (1) citations from outside Rhode Island's federal district; (2) citations from outside Rhode Island; and (3) secondary citations.

Looking first at the average rate of citation by courts outside Rhode Island's district, we see that earlier justices were cited at a higher rate than justices serving in 2005-2007 (See Figure 2). Keep in mind that citation counts for both sets of justices were only gathered for the two years immediately following the time span. For the cases from 1989 to 1991, citations through January 1, 1993 were collected; for the later cases, citations through January 1, 2009 were collected. Therefore, there should not be any bias towards the earlier cases, as the term of the citations is equal. However, as Figure 2 shows, opinions written by the earlier Court are cited more often than opinions written by the later Court. This difference is not statistically significant ($Pr(T > t) = 0.0823$). There are also no significant differences in the citation rates between justices on each Court. This data demonstrates that Hypothesis 2 is incorrect — those justices chosen by the merit system are not more likely to be cited by courts outside of Rhode Island's district.

---

27. See Choi et al., *Are Judges Overpaid?*, supra note 5, at 50; Choi et al., *Judicial Evaluations*, supra note 4, at 272; Choi et al., *Professionals or Politicians?*, supra note 5, at 3; Landes et al., *supra* note 14, at 272.
Looking at a larger pool of citations — those citations from courts outside of Rhode Island — similar patterns emerge. As Figure 3 shows, the average out of state citation rate for the earlier Court is slightly less than three times that of the later Court. The difference in average out of state citation rates between the pre-merit and post-merit Courts is statistically significant ($Pr(T > t) = 0.0000$). The differences between the widely varying rate of citation among the justices on the 1989-1991 Court is just outside the standard cut-off for significance ($Prob > F = 0.1095$). This is interesting given the consistency of citation rates among the justices on the 2005-2007 court. Generally, the data on out of district citations demonstrates that Hypothesis 2 is incorrect, thus forcing me to accept the null that merit selection does not lead to more influential justices. Further evidence for the null is that the sole Justice (Weisberger) remaining from the pre-merit Court on the hybrid Court had the highest average rate of out of state citations.
While citation counts by other courts are a good measure of the view that other judges have about a justice or court’s ability to make law, looking at citations by law reviews serves to identify the perception of the Court among academics and those that study the law.\footnote{See Landes et al., \textit{supra} note 13 (analyzing judicial opinions’ citations to other courts).} Looking at the rate of secondary citations, there are very few differences between the earlier and later Courts. In fact, as Figure 4 shows, the early and late Courts have an almost identical overall average rate of secondary source citations, a difference that is unsurprisingly insignificant.
The hybrid Court has a slightly higher rate of secondary citations, but the difference is, again, insignificant. Justices Lederberg and Goldberg have, by far, the highest average number of secondary citations. These results, again, demonstrate the inaccuracy of Hypothesis 2; there is no evidence in the data that justices selected by the merit system are cited more by law reviews than justices selected by the Grand Committee.

3. Page Length

One remarkable difference between the judicial performance in the pre-merit selection time and the hybrid and post merit selection period is the length of opinions by the justices of the
Rhode Island Supreme Court. Specifically, there are dramatic differences between the average length of opinions for the pre-merit and post-merit periods. As Figure 5 shows, the average length of opinions in the pre-merit period was 4.1 pages, with Justices Murray, Weisberger and Fay at the high end, publishing an average opinion length of 4.3 pages, and Justice Kelleher at the low end, publishing an average length of 3.8 pages. Looking at the hybrid Court, there is a small increase in the average length of opinions, with Justice Flanders (7 pages per opinion) and Justice Boucier (6.4 pages per opinion) at the upper end of opinion length. Finally, turning to the post-merit Court, Figure 5 demonstrates a dramatic increase in the average page length, rising to an overall average of 9.8 pages per opinion. On this Court, Justice Williams is on the lower end, with 8.6 pages per opinion (or twice the longest average of the pre-change court), and Justice Suttell is the most expansive, with 10.6 pages per opinion. While the difference between the earlier term and the later term is statistically significant (Pr(T < t) = 0.0000), as is the difference between the middle term and the later term, the difference between the early and middle term is not statistically significant.
There are no statistically significant differences between the justices on each Court. A similar pattern emerges when we look at the total number of pages written by each justice. As with the average number of pages, the later Court has written dramatically more pages of opinions than the earlier Court. Specifically, as displayed in Figure 6, the earlier Court has an average of 357.6 pages per justice, a high of 391 pages, and a low of 299 pages.
The 1998-2000 Court has an average of 327.4 pages, with a high of 486 (by Justice Flanders) and a low of 263. The 2005-2007 Court has an average of 637.2 pages per judge, a low of 571, and a high of 669. As with the average length, the difference in total pages between the early or middle Courts and the later Court is statistically significant (Pr(T < t) = 0.0000), but there are no statistically significant differences between the early and middle Courts, or between each Court’s justices’ total pages.
4. Subject Area

Turning next to the subject areas that justices publish in most often, Figure 7 shows that the overall patterns of subject areas have changed very little from 1989 to 2007. Overall, the rate of criminal cases as the subject matter of published opinions is on the rise in the Rhode Island Supreme Court, while cases relating to labor disputes are in decline.

**Fig. 7: Subject Area Breakdown**

- **Criminal**: 40%
- **Private Law**: 16%
- **Property**: 12%
- **Other**: 8%
- **Labor**: 13%
- **Torts**: 15%
- **Tax**: 5%
- **Rights**: 1%
- **Government**: 3%
- **Courts**: 2%
- **Administration**: 7%

Data labels:
- 2005 to 2007
- 1998 to 2000
- 1989 to 1991
Looking closely at the top subject areas, these subject areas account for 75% of cases in the earlier term, 84% of the middle term, and 87% of cases in the later term. It is also worth noting that the top subject areas of published cases change very little from 1989 to 2007. For example, criminal is the most common subject area for both sets of years, although there are far more criminal cases published in the later term (40% of total cases) compared to the middle term (34% of cases) and the earlier term (27% of cases). Furthermore, as Figures 8, 9, and 10 demonstrate, the top subject areas do not change. The 1989-1991 court also published heavily in private law (16% of total cases), property (12% of total cases), labor (11% of total cases) and tort (9% of total cases).
Looking at the middle term, the same five subject areas appear in the top five subject areas, including criminal, labor, private law, property, and torts. Compared to the early term, the middle term published more on torts (19%) and criminal (34%), and less on labor (9%) and private law (11%).

In the later term, criminal law (40%), torts (15%), property (14%), private law (12%) and labor (6%) are also the most common subject areas, although their order of frequency changes. These results suggest that while there may be some small changes in the subject matter of cases published, the justices on the Rhode Island Supreme Court hear and write about a stable group of subject areas.
5. Independence

As we turn next to an evaluation of the average rates of independence for the three Courts, it is important to note that Rhode Island has ranked very high on the average rate of independence by previous studies\(^\text{29}\) with an average rate of independence of 0.19 in 1998 to 2000 (See Figure 11). Earlier, I hypothesized that the Courts where justices are selected by the merit system will be more independent than those that are selected by the previous system. However, as Figure 12 shows, there is little evidence that the Court grew more independent with

29. See Choi et al., Professionals or Politicians?, supra note 5.
the change in measures. In fact, the hybrid Court has the highest level of independence of all the courts examined, followed by the pre-merit Court, with the post-merit Court falling at the tail end (See Figure 11). Despite the lower level of independence for the 2005 to 2007 Court, it is important to put these rates of independence in context; Choi, Gulati and Posner's 2009 research found that the average rate of independence for all the state high courts combined was less than -0.03. As such, the lowest rate of independence (0.12) is still much higher than the overall average rate of independence (See Figure 11).

![Fig. 11: Average Rate of Independence](image)

30. See Choi et al., *Judicial Evaluations*, supra note 4, at 1325.
31. See id.
Examining the data within the context of Hypothesis 5, there is little evidence in support of the theory that merit selection results in higher levels of independence.

CONCLUSION

In conclusion, this attempt to examine the effects of the change in selection method through a standard empirical matrix finds that the change to a traditional merit selection system has had very few effects on traditional measures of judicial quality. In Rhode Island, there has been a decline in the rate of opinion publication and in the rate of citation to opinions written by the state Supreme Court justices (See Figures 1, 2, and 3). While opinions have gotten longer (resulting both an increase in the average length of published opinions and the total number of pages written; see Figures 5 and 6), there has not been a significant effect on either the subject areas that the justices choose to publish in, or the rate of independence (See Figures 7, 8, 9, 10, and 11).

Overall, these results can be interpreted in two ways. First, they might suggest that the change in selection process has not dramatically improved the performance of the Rhode Island Supreme Court in areas where we have established measures. It may be that, when collegiality, fairness, and efficiency are examined, the post-change courts are demonstrably better. However, there is little evidence that the method of selection change has positively affected the production, influence, or independence of the court. Second, it may be that these results are simply evidence that the human-nature focused task of judging is difficult to measure. Looking at citations from an outside court is an excellent example: the high average citations from outside courts in the earlier court was largely the product of a single justice (Weisberger) who is very well regarded and cited to frequently by judges in other districts (See Figure 3). The independence scores can be similarly examined; the process by which a judge's party is identified is far from foolproof. If the process of determining a justice's party is error-ridden, this introduces a level of error into the measurement itself. In addition, in a state like Rhode Island, where Democrat and Republican are often imprecise labels that do not correspond to ideology, using a measure based on partisan may be a fool's errand.
The general lessons from the data presented here are this: from examining the most common subject areas, it does not appear that the substance of cases appearing before the Court has changed much over the last twenty years. The data presented on citation counts and publication rates suggest that the ranking a Court enjoys on these measures of quality can be highly influenced by the actions of a single or a few justices. Finally, a very high independence score can speak to several findings; first, that justices in Rhode Island may be very independent. Second, it is possible that the measure is innately flawed, as it relies on a piece of data – partisanship – that is either flawed in collection (as a consistent measure of partisanship is difficult to find) or useless in a context like Rhode Island where partisanship is uncompetitive and does not necessarily correspond to ideology. Third, measures like independence may be best interpreted in the context of a large number of states. All together, the data presented in this article offers many clues but few answers as to the effect of the change in selection method on the quality of judging in Rhode Island's court of last resort.