Judicial Merit Selection: Beliefs about Fairness and the Undermining of Gender Diversity on the Bench

Nancy B. Arrington[*]

Working Draft: January 2020

Abstract

Much attention is paid to how mechanisms for selecting political officials shape which types of officials are selected to positions of power, but sometimes selection procedures do not produce the desired outcomes. This article proposes an explanation for why the enthusiasm for judicial merit selection procedures for increasing gender diversity in the judiciary has not been realized. I argue that observers' beliefs that merit selection procedures are more "fair" than unilateral selection procedures leads observers to be less critical of suboptimal outcomes such as gender homogeneity. Using evidence from two survey experiments, I find that respondents perceive of merit selection procedures as more fair than gubernatorial selection procedures. In turn, respondents are *less critical* of all-male courts when judges are selected through a merit selection procedure relative to gubernatorial selection procedures. These findings contribute to our understanding of the ways in which selection institutions shape prospects for diversity.

Introduction

Both scholars and practitioners express expectations that judicial selection institutions can shape the quality and characteristics of judges selected to the bench, and a growing literature addresses how different selection institutions shape prospects for gender or racial diversity in the judiciary (Reddick, Nelson and Caufield 2010; Hurwitz and Lanier 2003; Bratton and Spill 2002; Alozie 1990). Merit selection – in particular – has garnered substantial attention for the ways in which it may (or may not) promote diversity in the judiciary. Typically under

merit selection procedures, a merit commission of several individuals accepts applications and seeks candidates for a vacancy. The commission then generates a short list of candidates from which the executive chooses a judge to fill the vacancy.[1]

Merit selection was famously employed by President Carter for the selection of lower-level federal courts.[2] His use of merit selection coincided with a successful attempt to diversify the federal judiciary.[3] Berkson, Carbon and Neff (1979) write, "When President Jimmy carter established the U.S. Circuit Judge Nominating Commission on February 14, 1977, he gave it a two-fold mission: to select circuit judges on the basis of professional merit [...] and to correct for past discrimination by affirmatively seeking women and members of minority groups for the bench" (p. 105). Of the 262 judges Carter appointed to district and circuit courts, 15% were women and 15% were Black, Hispanic, or Latino (Solberg and Bratton 2005; Walker and Barrow 1985), a record-breaking level of diversity at the time (Babcock 1980).

Proponents of merit selection argue that members of the nominating commission will focus on the qualifications of judges rather than on political or personal criteria, which should ensure the non-biased selection of the most qualified candidates. Because women have been less politically active, connected, and powerful than men, de-emphasizing those traits in the selection process should allow women to compete equally for the prestigious posts (Martin 1981). Ruth Bader Ginsburg, describing Carter's diversification efforts at an address at the Annual Conference of the National Association of Women Judges in 1995 said,

> "the key to Carter's success in recruiting women was not to ignore merit, but to
> rely upon it. The 'old boys' network could not be trusted to identify all the able

---

[1]There is, of course, variation across merit selection systems and the ways in which systems are classified Goelzhauser (2018*a*).

[2]Popularized by President Carter, the idea of a two-stage, commission-assisted selection procedure had been around long before Carter's presidency. See Berkson (1980, p 6-7) for a brief history of commission-assisted selection processes.

[3]See Slotnick (1982) for a discussion of affirmative action for judgeships under the Carter administration.

candidates, so Carter encouraged Senators to establish nominating committees that would evaluate a broad range of candidates for district court judgeships... With more open nominating procedures, increasing numbers of both women and minorities gained serious consideration for the first time, and Carter ushered in a new and lasting vision of the judiciary. Once President Carter opened the federal judiciary to all who were qualified, there was no turning back" (Ginsburg and Brill 1995, p 288).

Consistent with President Carter's use of the merit system and subsequent diversification of the federal judiciary, many observers expressed optimism that merit selection procedures would necessarily promote diversity on the bench more generally (Clark 2002; Crompton 2001). Mary Mullarky, the second woman selected to Colorado's state supreme court, for example, writes about the selection of women judges to the Colorado court: "[A]nother important factor in Colorado was a new method of selecting judges... Partisan political elections of judges were replaced by a merit selection system... Hopes were high that women would do better under the merit selection system than they had under partisan political elections"(Mullarky 2012). Similarly, Norman Krivosha, the chief justice of the Nebraska Supreme Court from 1978 to 1987, writes, "there is no question but that the merit selection system affords greater opportunities for women and minorities to find their way to the bench" (Krivosha 1987, p 19).

However, the belief that merit selection procedures would be net beneficial for the selection of women judges has not been born out in the several decades since Carter's use of merit selection to diversify the federal judiciary (Goldstein 2006). Merit selection did not accelerate the selection of the first woman justice to state supreme courts (Goelzhauser 2011)[4] nor is merit selection associated with a greater proportion of women judges on the bench (Reddick, Nelson and Caufield 2010). Indeed, some research suggests that merit selection disadvantages women: in Nebraska, Goelzhauser (2018b; 2019) finds that women are disadvantage at the commission nomination stage of the merit selection process, which could

---

[4]Although Goelzhauser (2011) does find that merit selection procedures are associated with the selection of the first Black judge on state supreme courts.

explain why women are more likely to fill state supreme court seats through appointment than through merit selection (Goelzhauser 2016). Williams (2007) finds no relationship between merit selection procedures and women's representation at the state trial court level but finds a negative relationship between the two for state appellate courts, a finding mostly confirmed by Reddick, Nelson and Caufield (2010) who find that merit selection has no effect on the selection of women to state high courts or trial courts but is negatively associated with the selection of women judges to intermediate state appellate courts.

In this article, I propose one explanation for why expectations about the role of merit selection in facilitating the selection of women candidates has not come to fruition: I argue that the popular belief that merit selection institutions are "fair" leads observers to be less critical of observed disparity, which allows the exclusion of women under merit selection procedures to persist unpunished. I test this expectation using evidence from a survey experiment. Comparing merit selection procedures to unilateral, gubernatorial selection procedure, I find that respondents perceive of merit selection procedures as, a priori, more fair (this pattern persists even when merit selection is referred to as the less-leading "committee-assisted selection"). Consistent with expectations, I then find that respondents are less critical of gender disparity under merit selection procedures relative to gubernatorial selection procedures. These findings contribute to the growing literature about the role of selection institutions and gender diversity in office and they suggest a potential, counterintuitive drawback to implementing institutional changes that observers find more trustworthy on the surface.

The rest of the article proceeds as follows: first I outline existing literature about the role of selection institutions in the judiciary. Then, I describe how citizens' beliefs in the fairness of institutions can undermine prospects for judicial diversity. Finally, I report evidence from a survey experiment and discuss the results.

Merit Selection Procedures, Judicial Quality, and Judicial Diversity

Debates about the best way to select judges for preferred outcomes have a long tradition, both globally and in the United States. Founding Father Alexander Hamilton, for example, celebrated judicial independence while Thomas Jefferson prioritized judicial accountability (Nejelski 1981; Webster 1995). More recently, scholars have addressed whether and how different selection institutions affect the quality of judicial qualifications and judicial performance. Patterns are inconsistent. For US state supreme courts, selection systems do not shape the probability that selected judges have prior judicial experience nor are there differences across elite appointment and merit selection systems with respect to the quality of law school education or judges' performance in law school; however, electing judges is associated with a decreased probability that a judge attended an elite law school or served on law review (Goelzhauser 2016). Cann (2006) finds that judges in states where most judges are selected through merit selection systems or executive appointment rate the quality of their state's courts higher than in states where most judges are elected.

When it comes to judicial performance, evidence suggests that elected judges respond to constituent pressure (Brace and Boyea 2008; Hall 1992). In particular, judges become more punitive as elections approach (Huber and Gordon 2004) to avoid appearing "soft" on crime (Bright and Keenan 1995). Other measures of performance are less conclusive. Using US Supreme court review and reversal as an indicator of professional criticism of a lower court ruling, Owens et al. (2015) find no relationship between selection method and the probability of being reviewed or reversed, which suggests that state supreme court justices are equally "forward-looking" across selection procedures. Choi, Gulati and Posner (2010) find that appointed state supreme court judges write better quality opinions while elected judges write more opinions; the authors suggest that appointed judges care more about their

professional reputations whereas elected judges care more their reputation in their local communities.

In addition to qualifications and performance, scholars address how selection systems may shape prospects for diversity on the bench. Evidence, though, is likewise mixed. In the cross-national context, Williams and Thames (2008) find that among OECD countries, systems in which presidents appoint judges are associated with a greater presence of women judges, and Valdini and Shortell (2016) find a positive relationship between gender diversity and selection systems in which the selectors are "exposed" to rather than "sheltered" from the electorate. Hoekstra, Kittilson and Bond (2014) find only weak support that selection method influences the presence of women on courts cross-nationally.

In the US context, Carbon, Houlden and Berkson (1982) find that more female judges came to their positions on state courts via elite selection than election; Bratton and Spill (2002) find that the first woman state supreme court justice is more likely to be selected when initially appointed, and Goelzhauser (2016) finds that state supreme court vacancies filled by appointment have a higher probability of being filled by a woman justice relative to both merit selection and election. Others, however, find no effect of formal selection procedures on gender diversity in the judiciary (Alozie 1988; 1990; Slotnick 1984; Hurwitz and Lanier 2008; 2003; Goelzhauser 2011).

When it comes to the effect of merit selection proceeders on gender diversity, the empirical evidence is, at best, bleak. Merit selection at the state supreme court level not only failed to accelerate the selection of the first women justices (Goelzhauser 2011), the procedure may disadvantage women judges (Goelzhauser 2018b; Reddick, Nelson and Caufield 2010; Williams 2007). Taken as a whole the literature suggests that selection institutions play little role in judicial diversity (Goelzhauser 2016), and in the context of US state courts, it appears that if selection institutions do shape diversity, the selection institution "best" for the selection of women is executive appointment (Carbon, Houlden and Berkson 1982;

Goelzhauser 2016). These empirical findings contradict the persistent expectation that merit selection is beneficial for gender diversity in the judiciary.

Merit Selection, Beliefs in Fairness, and Unintended Consequences

One explanation for the discrepancy between the enthusiasm for merit selection procedures and the empirical reality of gender diversity under merit selection may hinge on that enthusiasm for merit selection: because people believe merit selection procedures to be "better" than gubernatorial selection, observers may be more likely to accept the outcomes of merit selection, even when those outcomes are suboptimal.

Gibson, Caldeira and Spence (2003) pose the theory of positivity bias, in which legitimizing symbols of the judiciary lead people to believe that the judiciary is a different type of institution than the other political institutions, one based on objective legal principles and qualifications rather than politics. Given these perceived differences between courts and other political institutions, courts are perceived as "worthy of more respect, deference, and obedience – in short, legitimacy" (**?**, p 142). In turn, this institutional legitimacy shapes how people perceive of outcomes and makes people more willing to acquiesce to policies they oppose (Gibson, Caldeira and Spence 2005).

Applied to the context of judicial gender diversity, I argue that merit selection procedures promote the institutional legitimacy of the judicial selection process. Relative to gubernatorial selection, merit selection prohibits unilateral appointment and insures at least some overlap in preferences between the commission and the governor, which can serve to check the power of either actor. As such, we might expect fewer opportunities for unwarranted patronage under merit selection procedures. To the extent that merit selection procedures promote institutional legitimacy, merit selection procedures may lead observers to acquiesce

to outcomes that they might otherwise be critical of, including as gender homogeneity. In other words, a belief that the selection process is procedurally fair leads observers to be less critical of outcomes they might otherwise perceive to be problematic, such as the selection of an all-male or overwhelmingly-male bench. In turn, the acceptance of gender disparity on the bench under merit selection procedures can persist un-criticized or un-punished.

In the next sections I first test whether (and why) respondents do perceive of merit selection procedures to be more fair than unilateral, gubernatorial selection procedures. Then, I test whether the selection mechanism for judges shapes the extent to which observers are critical of gender disparity on the bench. I test these two hypotheses using evidence from two survey experiments.

Empirical Evidence

Beliefs about the Fairness of Merit Selection

Do people perceive of merit selection asm ore fair than gubernatorial selection? To test perceptions of fairness across selection institutions, I conducted a survey fielded in December 2017. The survey was designed in Qualtrics and fielded through Amazon's Mechanical Turk (MTurk). The analyses presented here are restricted to residents of the United States. As is well documented, MTurk survey respondents are not randomly drawn from the population. Instead, MTurk workers tend to be younger and more politically liberal than randomly selected respondents. Even so, research suggests that MTurk samples are more representative than in person convenience samples and student samples (Berinsky, Huber and Lenz 2012). Summary statistics for the sample are listed in table 1.

Respondents were shown information about two selection procedures to a hypothetical court. In one selection procedure, the governor is tasked with selecting judges to vacancies (gubernatorial selection). In the other procedure, a commission generates a short list of

8

Table 1: Sample Characteristics

Gender	Male	Female		
	.46 (149)	.55 (173)		
Age	<25	25-34	35-49	50+
	.10 (31)	.45 (145)	.31 (99)	.15 (49)
Ideol.	Very Cons.	Smwht Cons.	Moderate	Smwht Lib.	Very Lib.
	.06 (18)	.30 (96)	.21 (69)	.33 (105)	.11 (34)
Edu.	High School	Some Cllge	BA/BS	Masters	Doctorate
	.06 (20)	.34 (110)	.45 (146)	.12 (39)	.02 (7)

Descriptive characteristics of the survey sample. n=324

candidates from which the governor selects a judge to fill the vacancy (merit selection). Figure 4 in the appendix shows an example of the survey instruments. Of course, merit and gubernatorial selection are not the only ways in which judges are selected in the United States; state court judges are also selected through partisan elections, non-partisan election, legislative election, and supreme court/chief justice selection (Reddick, Nelson and Caufield 2010), although classifying selection systems is not always straightforward (Goelzhauser 2018a).

The comparison in this study is limited to merit and gubernatorial selection for a few reasons. First, research indicates a pattern between gubernatorial selection and greater gender diversity in the judiciary (Carbon, Houlden and Berkson 1982; Goelzhauser 2016); for those interested in identifying which selection institution is more likely to lead to increased gender diversity, gubernatorial selection may be relevant alternative. Second, both merit and gubernatorial selection are elite selection procedures (Bratton and Spill 2002), so comparing gubernatorial selection to merit selection, therefore, holds constant some of the key actors and types of actors involved in selection. Finally, merit selection and gubernatorial selection are sometimes grouped together as appointment systems in contrast to election based selec-

Figure 1: Merit vs. Gubernatorial Selection

Gub. vs. Merit Selection?
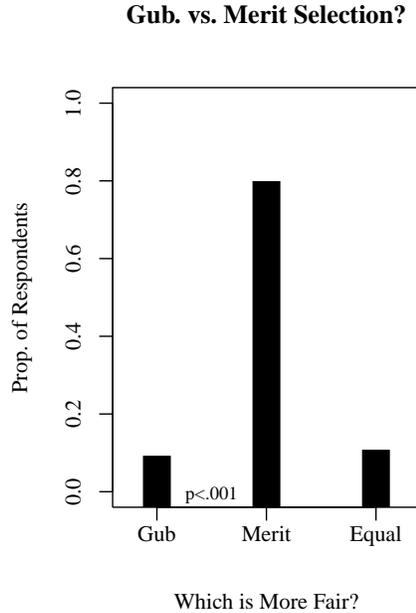
Proportion of respondents who indicated each process as more fair. Respondents overwhelmingly indicated that the merit selection process was more fair than the gubernatorial selection process (p<.001; n=324).

tion systems (Arrington 2018; Bratton and Spill 2002); comparing unilateral, gubernatorial selection to merit selection helps to disaggregate our understanding of judicial appointment systems.

After reading brief summaries of the selection procedures, respondents were asked "Which selection procedure do you think will be more fair?" If selection institutions do not affect respondents' prior beliefs in the fairness of institutions, respondents should select each type equally or indicate that they two systems are equally fair (a third option). In contrast, if beliefs about fairness vary across institution, respondents should favor one institution over the other. Given the debate surrounding merit selection and the argument that the commission will decrease the role of politics in selection, I expect respondents to perceive merit selection as more fair than gubernatorial selection.

Figure 1 reports the proportion of respondents who gave each answer. Just under 80% of respondents indicated that Merit Selection is more fair; 10.8% of respondents indicated that the two processes were equally fair, and only 9.3% of respondents indicated that gubernatorial selection was more fair. Moreover, it is highly unlikely that this difference in preference between Merit and Gubernatorial selection is due to chance (p<.001). Overwhelmingly, respondents indicated that merit selection procedures were more fair than gubernatorial selection procedures, absent any information about outcomes.

Unpacking the Preference for Merit Selection

After respondents indicated which institution they perceived as more fair, they had an opportunity to describe the reason they made their choice. Assessing these qualitative responses can shed light on why respondents perceived merit selection as more fair than gubernatorial selection. Qualitative explanations for respondents' choices were classified into seven categories representing the most common themes in the responses:

1. merit selection provides a check on the power of the governor,
2. merit selection allows for more input in selection (having more people involved),
3. merit selection will limit corruption, the role of politics, nepotism, etc,
4. in merit selection, *experts* choose judges,
5. merit selection will result in better qualified judges,
6. gubernatorial selection has more accountability, and
7. gubernatorial selection allows the governor to get input from the broad population.

Of course, there is some overlap in these responses. For example, for many answers classified as "providing a check on the governor," the mechanism through which the governor is limited is the presence of multiple actors, which overlaps with the idea of having "more input" in selection. Table 2 gives examples of typical qualitative responses for each category for those who chose merit selection and those who chose gubernatorial selection.

Table 2: Qualitative Response Examples

Institution Selected	Category	Response
Merit	Check on Gov	"There is less of a chance of a Governor simply appointing one of his friends to the position."
Merit	Check on Gov	"With merit selection, the governor gets the final say but since he can only select from a short list, he doesn't have too much control over the pick."
Merit	More Input	"Because more people are involved in the decision so more opinions can be heard about the candidates"
Merit	More Input	"More eyes on the matter; less reliance on a single individual and the biases they possess."
Merit	Less Corrupt	"Depending on the committee, it could limit cronyism and political appointments."
Merit	Experts Select	"The committee is comprised of experts and will have a good list of qualified candidates."
Merit	Better Qualified	"Because it is more fair to be judged based on actual merits rather than just someones opinion."
Gov	Accountability	"People elect the governor to make choices like that. I have no clue who the committee is and probably didn't vote for them. I prefer the choice be made by the guy who was elected over some random people with no accountability to the voters."
Gov	Accountability	"[B]ecause the Governor gets to appoint judges who align with his or her political views, with the committee they could nominate people who don't align with the public's choice of politics."
Gov	Popular Input	"I like that they get recommendations from various segments of the population."
Gov	Popular Input	"Because he get advice from different people"

Qualitative explanations for why respondents chose the institution they selected as more fair.

Table 3 shows a summary of the distribution of types of qualitative responses. The most common explanation among those who indicated merit selection was more fair was that merit selection allowed for more input in the selection process. Respondents with responses in this category identified the presence of multiple actors as a justification for preferring merit selection over gubernatorial selection. Closely related were those who indicated that their preference stemmed from the committee's check on the power of the governor. These respondents generally mistrusted the (seemingly) unconstrained power of the governor in selecting judges. These two categories – more input in the selection process and a check on the power of the governor – together comprise 48% of explanations for those respondents who chose merit selection as more fair than gubernatorial selection.

Among those who indicated gubernatorial selection was more fair, responses were split rather evenly between emphasizing the accountability of the governor, the ability of the governor to get broad input from the population, and other or unclear reasons.

Table 3: Qualitative Responses: Merit Selection

Response Category	Number Responses	Proportion Responses
Merit Selection		
More Input in Selection	79	29%
Check on Governor	53	19%
Less Corruption/Politics/Nepotism/etc	36	13%
Better Qualified Judges	35	13%
Experts are Selecting	9	3%
Other, Unclear	65	23%
Gubernatorial Selection		
Electoral Accountability	9	35%
Broad Input	8	30%
Other, Unclear	9	35%

A summary of the qualitative explanations for why respondents indicated that merit or gubernatorial selection was more fair than the other.

In summary, respondents overwhelmingly indicated that they perceived merit selection procedures as more fair than gubernatorial selection procedures. Those who indicated that merit selection was more fair generally identified the presence of additional actors and constraints on the governor as their justification. For the few who indicated that gubernatorial selection was more fair, justifications included accountability of the governor and the governor's ability to get advice from a broad swath of the population.

The Importance of "Merit" in Merit Selection

Although most respondents provided substantive explanations for their perception of merit selection as more fair than gubernatorial selection, it is possible that respondents were swayed by the the normatively laden term "merit" in the name.

Jones (2012) describes the modern trend of naming short titles for federal statutes with "evocative language" to attract support for the bills (p. 455). He notes how "many short titles imply that measures will be successful (for example, that they will "prevent" certain actions or "protect" certain populations) or contain various subjective characteristics (such as "responsibility" or "accountability")" (p. 458).[5] These names help garner popular support for bills and make voting against the bills potentially more costly for legislators. The term "merit" in "merit selection" could serve a similar function by implying that the process *is necessarily* meritorious, regardless of the actual procedures or outcomes of the process. As Dimino (2003, p. 803) writes, "merit selection [is] purely, so far as I can tell, a propagandistic

---

[5]Examples include, "Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT) Act; the No Child Left Behind Act of 2001; the Prosecutorial Remedies and Other Tools to End the Exploitation of Children Today (PROTECT) Act; the Unborn Victims of Violence Act 2004; the Prison Rape Elimination Act of 2003; the Partial-Birth Abortion Ban Act of 2003; the Adam Walsh Child Protection and Safety Act of 2006; and the Emergency Economic Stabilization Act of 2008... the Lilly Ledbetter Fair Pay Act of 2009; the American Recovery and Reinvestment Act of 2009; the Serve America Act; the Helping Families Save Their Homes Act of 2009; the Credit Card Accountability Responsibility and Disclosure (Credit CARD) Act of 2009; the Patient Protection and Affordable Care Act; and the Dodd-Frank Wall Street Reform and Consumer Protection Act"(Jones 2012, p. 457-458)

misnomer: nothing ensures that judges chosen under that plan will be better than judges under any other system."[6] If this is the case, the respondents who indicated a preference for "merit" selection procedures may have been persuaded by the name of the procedure rather than the process of the procedure.

To test whether respondents' beliefs in the fairness of institutions depend on the name of the institutions or the process, I add a treatment group in which respondents are shown the exact same information, but the "merit" selection process is referred to as "commission assisted" selection. If the term "merit" is responsible for differences in perceptions of fairness across merit and gubernatorial selection, then the difference between gubernatorial selection and "merit" selection should disappear when merit selection is referred to as "commission assisted" selection. In contrast, if the process of the selection procedure is more important than the label, the pattern of responses between merit and gubernatorial selection should be the same or similar to the pattern of responses between commission assisted and gubernatorial selection.

It is worth noting, however, that prior beliefs in the fairness of selection institutions should affect how respondents interpret observed information, regardless of how those prior beliefs were formed. Even so, having a better understanding of why observers believe certain institutions are more fair than others can inform the policy discourse and recommendations for selection procedures that might promote inclusion.

Table 4 reports summary statistics of the treatment group and control group (the control group is the same group of respondents reported on above). Respondents were randomly assigned into the two groups.

To determine if the term "merit" explains why respondents favored merit selection, I compare the proportion of respondents who indicated that "merit" selection is more fair

---

[6]In this project, I am agnostic to whether or not "merit" selection odes or does not produce better quality candidates (whatever better quality might be). Rather I am interested in how beliefs about the quality of merit selection procedures shape inferences about the outcomes

Table 4: Sample Characteristics

Gender	Male	Female			
Merit	.46 (149)	.55 (173)			
Commission Assisted	.43 (140)	.57 (183)			

Age	<25	25-34	35-49	50+	
Merit	.10 (31)	.45 (145)	.31 (99)	.15 (49)	
Commission Assisted	.06 (19)	.44 (144)	.35 (113)	.15 (48)	

Ideology	Very Conservative	Smwht Conservative	Moderate	Smwht Liberal	Very Liberal
Merit	.06 (18)	.30 (96)	.21 (69)	.33 (105)	.11 (34)
Commission Assisted	.06 (18)	.24 (78)	.29 (94)	.29 (94)	.12 (39)

Education	High School	Some College	BA/BS	Masters	Doctorate
Merit	.06 (20)	.34 (110)	.45 (146)	.12 (39)	.02 (7)
Comm-Astd	.10 (34)	.31 (101)	.39 (125)	.17 (56)	.01 (4)

Descriptive characteristics of the survey sample. There were 324 respondents in each treatment group.

than gubernatorial selection to a treatment group in which respondents were given the exact same information except "merit" selection was labeled "commission assisted" selection. Figure 2 plots responses for both groups of respondents. The black bars show responses for those who saw "merit" and gubernatorial selection, and the grey bars show responses for those who saw "commission assisted" and gubernatorial selection. The p-values along the x-axis are for the difference in proportion across treatment groups. Respondents were more likely to choose the two-stage selection method as more fair when the process was labeled as "merit" selection than when it was labeled "commission assisted" ($p=.013$), but even with this difference, respondents still overwhelmingly chose "commission assisted" selection as fair rather than gubernatorial selection: 71.6% of respondents chose commission assisted (compared to 79.9% under "merit"); 14.8% chose gubernatorial selection, and 13.6% indicated that the two systems were equally fair.

These findings indicate that the term "merit" may affect some respondents' beliefs about the fairness of that selection system ($p=.013$). However, most respondents (71.6%) in the "commission assisted" group still indicated that the selection procedure was more fair than

16

Figure 2: Perceptions of Fairness: "Merit" vs "Commission Assisted" Selection

Merit v. Commission−Assisted Selection
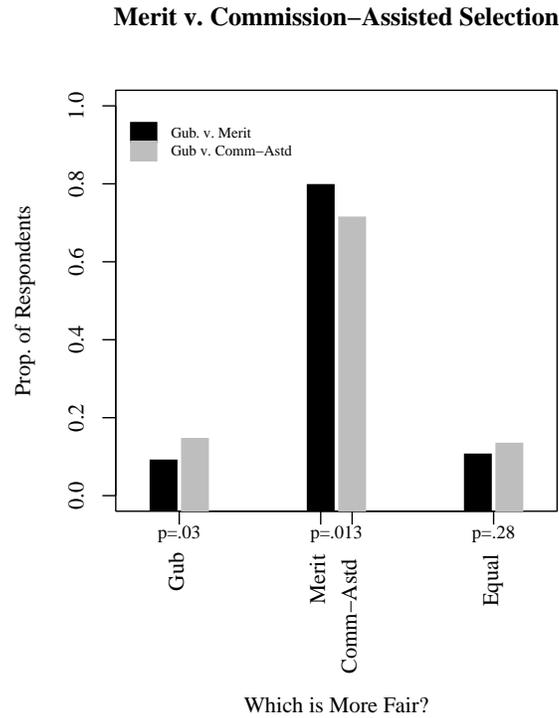
The proportion of respondents who indicated each answer in response to the question "Which selection procedure do you think will be more fair?" The black bars show respondents who were given information about "merit" and gubernatorial selection. The grey bars shows respondents who were given the same information, but were told the processes were called "commission assisted" and gubernatorial selection.

gubernatorial selection, absent the normative/subjective language. This suggests that the procedure – that is, the presence of additional actors who serve as a check on the governor – is what leads (most) people to perceive merit selection as more fair than gubernatorial selection. In the next section, I test whether gubernatorial or merit selection procedures affect how respondents interpret information about the outcome of judicial selection.

The Effect of Selection Institutions on Perceptions of Gender Disparity

To test whether and how prior beliefs in the fairness of merit selection procedures may lead respondents to be less critical of observed gender disparity, I employ another survey experiment fielded in the summer of 2017. This experiment was also developed and hosted on Qualtrics, and was completed by respondents recruited from Amazon's MTurk. Table 5 presents summary characteristics of the sample. Respondents for the analysis presented were required to reside in the U.S.A at the time they completed the survey.

Table 5: Survey Respondent Characteristics

Gender:	Male	Female		
Merit	.42 (197)	.58 (273)		
Gubernatorial	.42 (194)	.58 (270)		
Age:	<25	25-34	35-49	50+
Merit	.075 (36)	.34 (162)	.37 (176)	.22 (105)
Gubernatorial	.1 (48)	.35 (165)	.33 (155)	.21 (97)
Ideology:	Very Conservative	Smwht Conservative	Moderate	Smwht Liberal	Very Liberal
Merit	.06 (27)	.22 (106)	.225 (107)	.32 (152)	.17 (83)
Gubernatorial	.07 (31)	.22 (102)	.23 (107)	.30 (141)	.18 (82)
Education:	High School	Some Cllge	BA/BS	Masters	Phd/JD/MD
Merit	.08 (37)	.31 (148)	.40 (189)	.17 (80)	.04 (19)
Gub	.06 (29)	.32 (150)	.42 (195)	.15 (68)	.045 (21)

Summary characteristics for the MTurk survey respondents.

Respondents were shown information about a hypothetical court with seven members. They were told that judges were drawn from large and deep candidate pool that included

many women. The names of seven judges were listed along with an arbitrary bar association score. There were two treatment variables. First, the number of women's names listed varied from zero to three. The names were drawn from a list of past and current state supreme court judges in which gender-ambiguous names were removed.[7] Second, half the respondents were told that judges were selected by the governor while the other half of respondents were told that judges were selected through a merit procedure in which a commission generated a short list of three names from which the governor chose one judge to fill a vacancy. Respondents were not given any information about the composition of the short-lists generated by the commission.

Based on the information given, respondents were asked if the selection of judges appeared fair or not using a five-point scale: Definitely Unfair, Probably Unfair, Neither Fair nor Unfair, Probably Fair, and Definitely Fair. Figure 5 in the Appendix shows the survey instrument for the treatment group with merit selection procedures and two women justices. If selection institutions have no effect on how respondents interpret information, perceptions of bias across different levels of gender diversity will be the same across both institutional treatments.

Figure 3 shows the proportion of respondents who indicated that the process seemed unfair across the number of women judges *and* the selection institution type. There is one notable difference in the interpretation of information across institution type: when there were no women on the bench, almost half of respondents in the gubernatorial selection group interpreted the gender disparity as evidence of unfairness. In contrast, only a third of respondents in the merit selection group indicated the all-male court was unfair. This difference in perceptions of unfairness of an all-male court across institution type is statistically significant (p=.004), suggesting that observers are indeed less critical of gender homogeneity when it is the outcome of a process they believe to be more fair.

---

[7] I removed gender-ambiguous names from a list of all state supreme court judges selected between 1960 and 2010 from the State High Court Justice Database (*https* : //*www.lsu.edu*/*faculty*/*bratton*/*research.htm*) and then randomly drew male and female names from the list.

Figure 3: Survey Results, Perceptions of Unfairness Across Institutions Type and Diversity

Prop. Unfair Responses by # Women
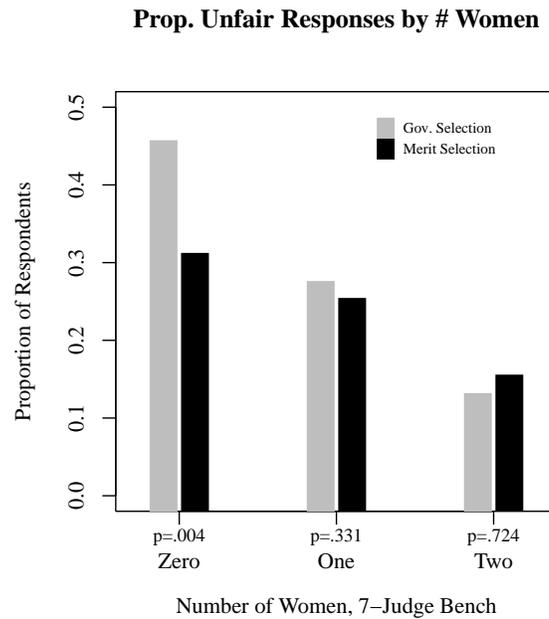
Perceptions of unfairness across levels of gender diversity and institution type. The bars show the proportion of respondents in each gender diversity treatment group who indicated that the process appeared "definitely unfair" or "probably unfair" disaggregated by institution type. The grey bars show the responses for those who saw gubernatorial selection institutions. The black bars show responses for those who saw merit selection procedures. When there were no women on the court, more respondents in the gubernatorial group perceived the process as biased relative to those in the merit group (p=.004).

Discussion and Conclusion

Despite enthusiasm for the ways in which merit selection procedures could/would facilitate the selection of women and minority judges, empirical evidence has been inconclusive – at best – about the role of merit selection in promoting gender diversity on the bench. One explanation for this discrepancy hinges on observers' enthusiasm or trust for merit selection. To the extent that observers believe merit selection procedures are more fair than gubernatorial selection, observers may be more willing to accept suboptimal outcomes of merit selection.

Evidence from a survey experiment confirms that respondents overwhelmingly consider merit selection procedures to be more fair than gubernatorial selection, absent information about outcomes. This preference for the two-stage process persists even when "merit selection" is referred to as "commission assisted" selection. Qualitative responses suggest that respondents' preferences for merit selection stem from the check it provides on the governor and the formal requirement of having more people give input into selection.

Given that respondents perceive merit selection procedures to be more fair than gubernatorial selection procedures, I hypothesized that observers would, in turn, be more accepting of suboptimal outcomes under merit selection. Indeed, respondents were less critical of an all-male bench when they were told that judges were selected through merit selection. This finding corroborates the idea that trust in institutions leads to acquiescence over outcomes.

This is not to say that merit selection is "bad" or should be avoided. This lesson of this project is not that merit selection is worse for women than executive selection; the takeaway is more subtle: the *belief* that a system is better for women can undermine the extent to which observers are critical of all-male selections. Even if merit selection were worse for women, though, merit selection need not be "bad." The prospect for diversification is just one of myriad implications of judicial selection institutions. The choice of selection institutions requires an assessment of the tradeoffs across several criteria. The contribution

of this project is to note how a belief that a particular institution is fair can undermine accountability by making people less critical of suboptimal outcomes.

Beyond the judiciary, this project contributes to our understanding of how selection procedures may or may not lead to desired outcomes. Organizations across contexts are increasingly concerned with diversity and debate how different policies may promote diversity and/or hinder discrimination. While some of these policies may lead to the desired outcomes, it is possible that some policies actual undermine the organization's goals by indirectly assuring observers that the "problem" has been solved.

Appendix

Figure 4: Survey Instrument, Beliefs in Fairness.

We are studying the fairness of different ways of selecting judges. Below, you will be shown information about two different ways of selecting judges. Then, you will be asked which selection procedure you think is more likely to be fair.

The two selection procedures you are comparing are **(1) Gubernatorial selection** and **(2) Merit Selection.**

For **gubernatorial selection procedures,** judges are appointed by the Governor. The Governor gets advice or recommendations from different segments of the population, such as the Bar Association, but the Governor can choose to follow that advice or not.

Under **merit selection,** a judicial nominating committee searches for candidates and accepts applications for vacant posts. Next, the commission generates a short list of three judicial candidates for each vacancy. Then, the Governor chooses one candidate from the shortlist of three judges to fill a vacancy.

Which selection procedure do you think will be more fair?

Gubernatorial Selection

Merit Selection

They are equally fair

Why?

This is the survey instrument used to test prior beliefs in the fairness of merit versus gubernatorial selection procedures. To compare how the term "merit" shapes responses, the treatment group saw the same exact text but with "Merit Selection" replaced with "Commission-Assisted Selection"

Figure 5: Survey Instrument, Effect of Institutions on Perceptions of Fairness

Please read the following information carefully.

We are studying the fairness of the selection and composition of some US state supreme courts. The court you will be asked to assess has **seven judges** on the court, and all of the judges are selected through a process referred to as "merit selection." This selection process works in the following way: first a **judicial nominating committee searches** for candidates and accepts applications for vacant posts. Next, the commission generates a short list of three judicial candidates for each vacancy. Then, the **Governor chooses** one candidate from the shortlist of three judges to fill a vacancy.

This state has a very large and deep pool of qualified judicial candidates. Experts predict that about half of the qualified candidates are female. The Bar Association generates a score for potential judges that incorporates their academic and professional qualifications. This score varies from 1 to 5. A score of 5 is very rare, and state supreme court judges average a score of 4.6.

After you are shown information about the judges on the court, you will be asked some questions about your impression of the court.

Does this information make sense to you?

○ Yes
○ No

The judges currently on this state's supreme court bench are listed below, along with their ages and Bar Association scores.

Perry Hooper
Age: 67 , Score: 4.5

Herbert Meschke
Age: 58 , Score: 4.8

Jeanette Knoll
Age: 58 , Score: 4.7

Matthew Durrant
Age: 46 , Score: 4.4

Joseph Walsh
Age: 66, Score: 4.7

Linda Dalianis
Age: 65 , Score: 4.5

Benjamin Kaplan
Age: 48 , Score: 4.6

Given this information about the members of the court, do you think judges are selected in a procedure that is fair or unfair? Please indicate your response below.

Definitely Unfair	Probably Unfair	Neither Fair nor Unfair	Probably Fair	Definitely Fair
○	○	○	○	○

To test how selection institutions shape perceptions of fairness across different levels of gender diversity in the judiciary, there were six treatment groups for whom the selection institution and the number of women were manipulated. This shows the survey instrument for Merit Selection and two women.

References

Alozie, Nicholas O. 1988. "Black Representation on State Judiciaries." *Social Science Quarterly* 69:979–986.

Alozie, Nicholas O. 1990. "Distribution of Women and Minority Judges: The Effects of Judicial Selection Methods." *Social Science Quarterly* 71(2):315–325.

Arrington, Nancy B. 2018. "Gender and Judicial Replacement: The Case of US State Supreme Courts." *Journal of Law and Courts* 6(2).

Babcock, Charles R. 1980. "Carter Names Record Number of Minorities to Federal Bench.".

Berinsky, Adam, Gregory Huber and Gabriel Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351–368.
URL: *http://www.jstor.org/stable/23260322*

Berkson, Larry. 1980. "Carter's Judicial Selection System: How Well Does It Work." *Judges Journal* 19:4.

Berkson, Larry C, Susan Carbon and Alan Neff. 1979. "A Study of the US Circuit Judge Nominating Commission: Findings, Conclusions and Recommendations." *Judicature* 63:104.

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.

Bratton, K. A. and Rorie Law Spill. 2002. "Existing Diversity and Judicial Selection: The Role of the Appointment Method in Establishing Gender Diversity in State Supreme Courts." *Social Science Quarterly* 83(2):504–518.

Bright, Stephen B and Patrick J Keenan. 1995. "Judges and the politics of death: Deciding between the bill of rights and the next election in capital cases." *Boston University Law Review* 75:759.

Cann, Damon. 2006. "Beyond Accountability and Independence-Judicial Selection and State Court Performance." *Judicature* 90:226.

Carbon, Susan, Pauline Houlden and Larry Berkson. 1982. "Women on the State Bench: Their Characteristics and Attitudes about Judicial Selection." *Judicature* 65(6):294–305.

Choi, Stephen J, G Mitu Gulati and Eric A Posner. 2010. "Professionals or politicians: The uncertain empirical case for an elected rather than appointed judiciary." *The Journal of Law, Economics, and Organization* 26(2):290–336.

Clark, Mary Law. 2002. "Carters Groundbreaking Appointment of Women to the Federal Bench: His Other "Human Rights" Record." *American University Journal of Gender, Social Policy and the Law'* 11(3):1131–1163.

Crompton, J Andrew. 2001. "Commentary: Pennsylvanian's Should Adopt a Merit Selection System for State Appellate Court Judges." *Dickinson Law Review* 106:755.

Dimino, Michael R. 2003. "The Futile Quest for a System of Judicial Merit Selection." *Albany Law Review* 67:803.

Gibson, James L, Gregory A Caldeira and Lester Kenyatta Spence. 2003. "The Supreme Court and the US presidential election of 2000: Wounds, self-inflicted or otherwise?" *British Journal of Political Science* 33(4):535–556.

Gibson, James L, Gregory A Caldeira and Lester Kenyatta Spence. 2005. "Why do people accept public policies they oppose? Testing legitimacy theory with a survey-based experiment." *Political Research Quarterly* 58(2):187–201.

Ginsburg, Ruth Bader and Laura W Brill. 1995. "Women in the Federal Judiciary: Three Way Pavers and the Exhilarating Change President Carter Wrought." *Fordham Law Review* 64:281.

Goelzhauser, Greg. 2011. "Diversifying State Supreme Courts." *Law and Society Review* 45(3):761–781.

Goelzhauser, Greg. 2016. *Choosing state supreme court justices: Merit selection and the consequences of institutional reform.* Temple University Press.

Goelzhauser, Greg. 2018*a*. "Classifying Judicial Selection Institutions." *State Politics and Policy Quarterly* 18(2):174–192.

Goelzhauser, Greg. 2018*b*. "Does Merit Selection Work?: Evidence from Commission and Gubernatorial Choices." *Journal of Law and Courts* 1(6):155–187.

Goelzhauser, Greg. 2019. *Judicial Merit Selection: Institutional Design and Performance for State Courts.* Temple University Press.

Goldstein, Adam. 2006. "Judicial selection as it relates to gender equality on the bench." *Cardozo Journal of Law and Gender* 13:369.

Hall, Melinda Gann. 1992. "Electoral politics and strategic voting in state supreme courts." *The Journal of Politics* 54(2):427–446.

Hoekstra, Valerie, Miki Caul Kittilson and Elizabeth Andrews Bond. 2014. Gender, High Courts, and Ideas about Representation in Western Europe. In *Representation: The Case of Women*, ed. Maria C. Escobar-Lemmon and Michelle M. Taylor-Robinson. New York: Oxford University Press book section 3, pp. 103–117.

Huber, Gregory and Sanford 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 & 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.

Jones, Brian Christopher. 2012. "Drafting Proper Short Bill Titles: Do States Have the Answer Feature." *Stanford Law and Policy Review* 23:455.

Krivosha, Norman. 1987. "Acquiring Judges by the Merit Selection Method: The Case for Adopting Such a Method." *Southwestern Law Journal* 40:15.

Martin, Elaine. 1981. "Women on the Federal Bench: A Comparative Profile." *Judicature* 65:306–313.

Mullarky, Mary J. 2012. "A Brief History of Women of the Colorado Supreme Court." *The Colorado Lawyer* 41(27).

Nejelski, Paul. 1981. "The Jeffersonian-Hamiltonian Duality: A Framework for Understanding Reforms in the Administration of Justice." *Judicature* 64:450.

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

Reddick, Malia, Michael Nelson and Rachel Caufield. 2010. "Examining diversity on state courts: How does the judicial selection environment advance–and inhibit–judicial diversity." *American Judicature Society. Retrieved February* 21:2012.

Slotnick, Elliot E. 1982. "Lowering the bench or raising it higher: Affirmative action and judicial selection during the carter administration." *Yale Law & Pol'y Review* 1:270.

Slotnick, Elliot E. 1984. "Judicial Selection Systems and Nomination Outcomes Does the Process Make a Difference?" *American Politics Quarterly* 12(2):225–240.

Solberg, R. Law Spill and K. A. Bratton. 2005. "Diversifying the Federal Bench: Presidential Patterns." *The Justice System Journal* 26(2):119–133.

Valdini, Melody E. and Christopher Shortell. 2016. "Women's Representation in the Highest Courts: A Comparative Analysis of the Appointment of Female Justices." *Political Research Quarterly* 69(4):865–876.

Walker, Thomas G and Deborah J Barrow. 1985. "The diversification of the federal bench: Policy and process ramifications." *The Journal of Politics* 47(2):596–617.

Webster, Peter D. 1995. "Selection and retention of judges: is there one Best Method." *Florida State University Law Review* 23:1.

Williams, Margaret. 2007. "Women's representation on state trial and appellate courts." *Social Science Quarterly* 88(5):1192–1204.

Williams, Margaret S. and Frank C. Thames. 2008. "Women's Representation on High Courts in Advanced Industrialized Countries." *Politics and Gender* 4(03).