Quality in Measurement Matters: Adjusted American Bar Association Ratings and Circuit Court Confirmation Hearing Word Choice

James Sieja[1]
Assistant Professor of Government
St. Lawrence University
102 Hepburn Hall
23 Romoda Drive
Canton, NY 13617
jsieja@stlaw.edu

[1] All data to reproduce the tables and figures in this publication are available on the author's website and in the Journal's Dataverse archive.

Abstract

Though widely used in studies of judicial politics, American Bar Association (ABA) ratings have a partisan bias. As a result, when researchers include ABA ratings and ideology in a model together, the results may be biased toward non-findings with respect to the effect of ideology, qualifications, or both. This study leverages new data on the ABA rating process to create a valid and reliable new measure for the qualifications of nominees to the US Courts of Appeals. In an empirical example, I test the new measure against alternative specifications to demonstrate its potential. The empirical example also presents a new data set on circuit court confirmation hearing speech. The findings contrast with well-established conclusions from previous studies.

How well do qualifications measures capture judicial nominees' "quality?" How does "quality" affect potential judges at different stages in their careers, e.g., during their selection, nomination, confirmation, and while they are on the bench? Existing research is equivocal about the effect of qualifications. Some studies find that highly qualified judges are more likely to earn promotion (Savchack, Hansford, Songer, Manning, and Carp 2006), get through the nomination process more quickly (Allison 1996, Martinek, Kemper, and Van Winkle 2002, Stratmann and Garner 2004, Hendershot 2010), and be influential (Landes, Lessig, and Solimine 1998, Black and Owens 2013). Yet, other studies find qualifications to have no effect on the length of the nomination process (Binder and Maltzman 2002), whether senators deploy institutional delay tactics against the nominee (Black, Madonna, and Owens 2014), or the likelihood of confirmation (Scherer, Bartels, and Steigerwalt 2008; Steigerwalt 2010). Moreover, highly qualified judges work harder (Choi, Gulati, and Posner 2013) and more efficiently (Christensen and Szmer 2012). But, a negative relationship exists between productivity and influence, on the one hand, and time to confirmation, on the other (Lott 2005).

The puzzling pattern of results might arise from an imperfect measure of qualifications: American Bar Association (ABA) ratings. The inadequacy of the ABA rating measure leads researchers to come up with alternative operationalizations that muddle findings and stymie the development of a coherent body of knowledge. This study aims to develop a better measure of judicial qualifications by improving ABA ratings for US Courts of Appeals, i.e., circuit court, nominees in a valid and reliable way.

Despite their well-documented problems, ABA ratings provide a useful starting point for the creation of the new measure. Indeed, the fact that researchers have documented the problems with ABA ratings so well leads to a straightforward improvement. Briefly, the study leverages

new data on the people who produce ABA ratings to indirectly minimize unwarranted biases in the ratings, especially biases stemming from partisan, gender, and racial identities. To show its benefits, I then test the adjusted measure in an empirical context: the word choices senators make at the nominee's confirmation hearing.

The results are promising. The adjusted ABA ratings measure corrects many of the problems of raw ABA ratings, most importantly their questionable construct validity and reliability. The raw measure assigns 56.5 percent of all circuit court nominees between 1958 and 2020 the highest possible rating, "well qualified." Empirically, nominees' backgrounds are more continuous than categorical, a difference that the adjusted measure captures well. It also reliably assigns similar ratings to nominees with similar backgrounds.

Turning to the empirical example, adjusted ABA ratings change the analysis of confirmation hearing word choice. The qualifications word choice analysis using raw ABA ratings finds few substantively significant differences between low- and high-quality nominees, observations consistent with the argument that the hearings are "not personal" (Dancey, Nelson, and Ringsmuth 2014, 2020). But, the analysis using the adjusted measure demonstrates that the difference in senators' use of qualifications-related words between low- and high-quality nominees can be as much as 69 percent. Thus, the study presents evidence that the hearings *are* personal in part, so long as researchers measure nominees' personal backgrounds accurately.

The main contribution of the study is the new measure of qualifications for circuit court nominees. Wide adoption of the measure will allow for comparable findings across independent studies, which will in turn lead to a coherent body of knowledge about the role of qualifications in nomination politics and a judge's behavior once on the bench. In the empirical example, I contribute a new machine-coded dataset on the content of circuit court confirmation hearings.

Additionally, the findings speak to a wider discussion about the ABA's role in vetting candidates for federal judicial office. Recent research that ties ABA ratings to performance, e.g., citation (Lott 2013) and reversal rates (Sen 2014), finds no relationship between the two. However, in muting some of the ratings' considerable "noise"—a heavily left-skewed distribution, partisan bias, and demographic bias—the study tries to recover from them a useful "signal" about the nominee's qualifications (Silver 2012). Given the exposure of ABA research in popular media (Lott 2006, Voeten 2013), the results could have real implications for the policy choices presidents make with respect to the ABA's continued involvement in the judicial nomination process.

The next section documents the problems with ABA ratings and how scholars have tried to deal with them as both independent and dependent variables. Then, I present the adjusted measure of qualifications and describe some of its beneficial qualities. I end the empirical discussion with an analysis of confirmation hearing word choice. The analysis compares ABA ratings, the adjusted measure, and a proxy indicators approach to demonstrate the problems and promise in studying the role of qualifications in judicial politics.

1. Problems with ABA Ratings: Heterogeneity in Lower Court Qualifications Measures

Unfortunately for scholars of judicial politics, the relatively minor salience of lower court nominations precludes the application of the standard method used to create an exogenous measure of qualifications for Supreme Court nominees. Cameron, Segal, and Songer (1990) analyzed editorial content discussing nominees' qualifications in four national newspapers, two liberal and two conservative, to create the standard qualifications variable for the Supreme Court. But, editorial writers for national newspapers do not, for the most part, write about specific circuit or district judge nominations. To be sure, lower court nominations generate editorial

attention at the regional level, but reliance on regional newspapers introduces a host of confounds. Most prominently, the difficulty in scaling editorial ideology for newspapers across circuits unfavorably contrasts with the simplicity of using the same four newspapers for every nominee. The difficulty notwithstanding, there should be a single measure of lower court nominees' qualifications that is both valid and reliable. The current standard measurement of qualifications in research on lower courts now, however, is that there is no standard.

Broadly, studies that use a measure of qualifications fit into three categories. First, studies in the proxy indicators approach use a series of variables designed to represent the underlying qualifications construct. Second, the modified ABA ratings approach uses simplified versions of the ordinal scale, "not qualified" to "well qualified," on which the ABA rates nominees. Finally, some studies use the raw ABA rating as a measure of qualifications. Table 1 categorizes several studies over a recent two-decade span on their approaches to measurement of qualifications.

Proxy Indicators	Modified ABA Ratings	Raw ABA Ratings
Choi, Gulati, and Posner (2011) [3]	Allison (1996)	Black et al. (2014)
Choi, Gulati, and Posner (2013) [4]	Binder and Maltzman (2002)	Dancey, Nelson, and Ringsmuth (2014)
Christensen and Szmer (2012) [1]	Binder and Maltzman (2009a, 2009b)	Lott (2005)
Dancey, Nelson, and Ringsmuth (2014) [4]	Hendershot (2010)	Scherer, Bartels, and Steigerwalt (2008)
Merrit and Brudney (2001) [1]	Hettinger, Lindquist, and Martinek (2006)	Steigerwalt (2010)
Stratmann and Garner (2004) [2]	Landes, Lessig, and Solimine (1998)	Stratmann and Garner (2004)
	Martinek, Kemper, and Van Winkle (2002)	
	Savchak, et al. (2006)	
	Sen (2014)	
	Taha (2004)	

Table 1. Existing Approaches to Qualifications Measurement. *Note*: The number of indicators used in each of the proxy indicators studies appears in brackets after the citation. The 11 studies in the modified ABA ratings approach used seven distinct modifications.

The heterogeneity in how researchers measure qualifications—there are almost as many ways to operationalize the construct as there are studies that include some measurement of

qualifications—makes it difficult to compare results across studies.[2] With uncomparable results, knowledge about the role of qualifications does not accumulate; that is, the field of judicial politics does not have much concrete to say about the role of qualifications in the stages of a (potential) judge's career.

Moreover, both the proxy indicators and modified ABA ratings approaches make the falsification of a qualifications-related theory a shifting target. For example, consider a situation in which the inclusion of variables measuring private practice experience, judicial experience, and elite law school produces no significant results in a multiple regression analysis, yet the inclusion of prosecutorial experience and clerkship experience variables do produce significant results. Given the pronounced publication bias in the social sciences, a researcher is far more likely to write up and publish the latter output (Franco, Malhotra, and Simonovits 2014). So, not only does knowledge not generally accumulate, at least some of the knowledge that does build up reflects a skewed sample of all possible results.

If researchers treat the raw ABA rating consistently, the epistemological problems of the proxy indicators and modified ABA ratings approaches do not arise. The raw scale is reliable from study to study. However, its validity is questionable. The hodgepodge of relationships, some diametrically opposed, between the ABA rating and various outcomes calls its face and construct validity into question. With respect to unbiasedness, the ABA rating fares no better. Political elites have asserted at different times that liberal or conservative nominees receive lower ratings because of the nominees' ideology or background (Slotnick 1983, Kamenar 1990,

---

[2] Dancey, Nelson, and Ringsmuth's (2020, 176) book-length treatment of confirmation hearings notes the lack of a "uniform way of measuring ABA ratings." It demonstrates the diversity with the inclusion of at least three different modifications to ABA ratings in separate analyses throughout the book.

ABA Watch 2006). And, the balance of social scientific research finds strong evidence for ideologically biased ratings for circuit court nominees (e.g., Smelcer, Steigerwalt, and Vining 2012, Sieja 2023; see Sen 2014 for gender and racial bias at the district court level).

Biased measures are not necessarily useless, especially if, as is true with ABA ratings, researchers know the bias's direction. Using a biased measure ordinarily induces a researcher to note that the results represent a "conservative estimate" or the "ceiling" of the actual effect size. Yet, *ideologically* biased ratings present real problems as a measure of qualifications in multiple regression analyses. Specifically, the correlation between two independent variables—ABA rating and ideology—could affect the estimation of the effect of one or both. Determining the actual presence, size, and direction of any potential effect is difficult.

As researchers presently treat it, the ABA rating both induces heterogeneity in scientific measurement of qualifications and potentially biases the results of studies that use it unaltered. Despite its problems, though, the ABA rating is a useful starting point for the creation of a reliable and valid measure of nominee qualifications. The ABA rating's wide recognition and basic reliability are positives. Another plus is that it is the product of an extensive investigation into a nominee's background (American Bar Association 2020, 1-9).[3] Given general agreement that judges ought to have traits like integrity and an even temperament, in addition to law degrees and experience in the courtroom or classroom, researchers using ABA ratings have a richer measure of qualifications. The biases remain, though, and expunging them is not a simple task. What follows is an attempt to purge the ratings of the unwarranted biases in a valid and reliable way.

[3] See also Appendix A for a brief description of the process.

2. Adjusting the Ratings

2.1 Data and Methods

To adjust ABA ratings, I used a dataset that began with demographic, background, and partisanship information on the 643 nominees to the circuit courts and the 250 ABA Standing Committee on the Federal Judiciary (SCFJ) members whose investigations produced the nominees' ratings from 1958 to 2020.[4] Using a genetic matching algorithm, I matched nominees on race, gender, age, years as a federal judge, years as a state judge, years as a government attorney, years in private practice, whether they taught law school, whether they were a federal clerk, whether they had political experience, and whether they graduated from a Top 14 law school (Diamond and Sekhon 2013; Ho, Imai, King, and Stuart 2011). Party of the appointing president was the treatment variable. The preprocessing left 562 observations, evenly split between Republican and Democratic nominees, investigated by 188 SCFJ members.

In addition to balancing the data, matching is particularly useful for adjusting ABA ratings. This is because it restricts the data for producing the adjustment coefficients to a portion of the set, yet all nominees will receive an adjusted rating based on the results of the restricted estimation. If the new measure performs better than raw ABA ratings for the whole population, including out-of-matched observations, the improved performance would be solid evidence that the adjustment is worthwhile.[5]

There are at least two ways to adjust ABA ratings. First, consider a less-valid adjustment: Taking the estimated coefficients on the partisanship variable from Smelcer, Steigerwalt, and

---

[4] See Appendix B for details on the data collection effort, including the coding and analysis of SCFJ member partisanship.

[5] In the qualifications word choice analysis subsequently presented, 28 percent of the nominees do not appear in the matched data. Their hearings account for 42 percent of the 919 total observations.

Vining's (2012) model, one could simply "zero-out" the partisan bias by generating a predicted rating for the nominees with their partisanship set at "Democratic." But, the assumption of uniform bias from the ABA is incorrect. Indeed, Republican nominees suffer from lower ABA ratings *on average*, but the estimated effect for or against Republican and Democratic nominees varies with the partisanship of the investigating member and different levels of partisanship on the SCFJ as a whole. Therefore, the naïve adjustment would unduly punish (reward) nominees who were (not) exposed to bias. In effect, the zero-out approach to adjustment would trade one form of construct invalidity for another.

Instead, the more-valid approach focuses primarily on the adjustment for the SCFJ investigator and committee composition and secondarily on the effects of race and gender, all variables over which the nominee has no control. Sieja (2015, 2023) demonstrates that the individual member primarily and the committee secondarily interact to have significant effects on the nominee's final rating, after controlling for other potential covariates. Therefore, to produce the adjusted rating, I first estimated an ordered logit model with the matched sample that included a three-way interaction term between the partisanship of the nominee, the SCFJ investigator, and the committee as a whole. The dependent variable was the nominee's ABA rating, preserving split ratings but collapsing all "well-qualified" and above ratings to "well-qualified" (Martinek, Kemper, and Van Winkle 2002). I also included experience controls for years as a federal judge, years as a state judge, years in private practice, years as a government attorney, graduation from a Top 14 law school, having taught law, holding a clerkship for a federal judge, and participating in partisan political activity; demographic controls for minority status, gender, age, and the square of age; and an "era" control for the existence of publicly available split votes. With the exception of the SCFJ-related variables, the model is the same as

Smelcer, Steigerwalt, and Vining's (2012) model.

The results in Table 2 provide estimates for the relative weight that the ABA places on each element that goes into the rating decision. Some elements, such as experience as a federal judge, state judge, in private practice, as a government attorney, and as a federal clerk, seem appropriate as positive predictors of a qualified nominee. The effects of elite law school attendance and law professorship are substantively negligible, and the ABA itself explicitly discounts prior political experience (American Bar Association 2020, 4). Minority and female identifiers have a lower probability of a high rating. Though it is not statistically significant overall, the committee proportion variable is substantively important over certain values.

Federal Judge (Years)	0.108**
	(0.021)
State Judge (Years)	0.058*
	(0.023)
Private Practice (Years)	0.034
	(0.018)
Government Attorney (Years)	0.013
	(0.024)
Top 14 Law School	0.208
	(0.179)
Law Professor	0.040
	(0.206)
Federal Clerk	0.907**
	(0.227)
Political Experience	-0.672*
	(0.243)
Republican President	-0.729
	(0.485)
Republican Investigator	0.348
	(0.876)
Republican Committee (Proportion)	-0.470
	(1.290)
R President * R Investigator	-0.413
	(0.962)
R President * R Committee	-0.406
	(1.480)
R Investigator * R Committee	-1.770
	(1.740)

R President * R Investigator * R Committee	2.184 (2.034)
Minority Status	-0.315 (0.222)
Female	-0.132 (0.205)
Age	0.479* (0.182)
Age^2	-0.005* (0.002)
No Split Vote Era	0.982* (0.393)
N	**558**
Log-pseudolikelihood	**-646.092**

Notes: *: p<0.05 **: p<0.01 (two-tailed)

Table 2. *Ordered Logit Results Used for Adjusted ABA Measure Creation.* The dependent variable is a six-point ordinal measure of the nominee's ABA rating.

After model estimation, I duplicated all observations in the dataset and then set the duplicated observations' values for SCFJ investigator partisanship to 1 (Republican), minority status, gender, and "no split vote" era to 0 (white, male, and exposed to split vote possibility, respectively), and proportion of Republicans on the SCFJ to 0.647. Empirically, this proportion is the point at which the predicted probability of a "well-qualified" rating for Republican SCFJ investigators is equal for both Republican and Democratic nominees (0.54). In effect, this made all the duplicated observations white male nominees investigated by a Republican SCFJ member when the committee was about two-thirds Republicans and one-third Democrats. The observed values for all other variables, such as party of the nominating president, years of judicial or litigation experience, and whether the nominee was a law professor, remained the same in the duplicated observations.

Finally, I created the linear prediction from the model as each nominee's adjusted ABA rating, rounded it to the nearest 0.1, and standardized the resulting distribution, i.e., subtracted the mean and divided by the standard deviation. Predictions to derive a single measure of a

complex underlying construct are, of course, not unknown to scholars of judicial politics. For example, Black and Boyd (2011) used a model of certiorari grants to derive a case-level measure of certworthiness, an approached followed by others in the field (e.g., Benesh, Armstrong, and Wallander 2020).

2.2 Properties of Adjusted ABA Ratings: Less Bias and Increased Face Validity

The adjusted ABA ratings are attractive for several reasons. First, they adjust for the partisan and demographic biases in the existing ratings. In the case of partisan bias, the measure reduces the connection between nominee partisanship and ABA ratings *without changing the partisanship of any nominee*. The adjustment altered only SFCJ investigators' partisanship, the level of partisanship on the SCFJ committee, and nominees' race and gender. Nevertheless, the bivariate correlation between the party of appointing president and qualifications went down from the raw ABA rating's r=-0.129 to the adjusted rating's r=-0.056, a 57 percent reduction. At the same time, the magnitudes of the bivariate correlations between the adjusted measure and other non-adjusted variables increased for all but one covariate (law professor). In most cases, the increases were well over 100 percent.

Second, in addition to its increased unbiasedness, the adjusted ABA rating is a more facially valid measure of qualifications. That is, it generates similar values for nominees with similar backgrounds. Consider the case of Goodwin Liu, a University of California law professor nominated to the Ninth Circuit by President Obama in 2010. The ABA rated Professor Liu unanimously "well-qualified," even though he had no judicial experience and only two years of experience in private practice, a full decade less than the ABA recommended at the time (American Bar Association 2009, 3). Critics railed against the "ridiculous" rating, comparing it unfavorably to the split "qualified-not qualified" rating given to Reagan Seventh Circuit nominee

Frank Easterbrook, who had a similar résumé to Liu at the time of his appointment (Whelan 2010). The adjusted ABA rating for Liu is -1.30 and -1.00 for Easterbrook. While these ratings are both substantively low—they reflect nominees with qualifications about one standard deviation below the mean—they are nevertheless consistent.

The adjusted measure's improved consistency applies to more than the high-profile Liu and Easterbrook pairing. Consider further 39-year-old Clinton DC Circuit nominee Elena Kagan, "qualified-well qualified" according to the ABA, and 35-year-old Reagan Ninth Circuit nominee Alex Kozinski, whom the ABA rated "qualified-not qualified." Both lacked previous Article III federal or state judicial experience, did not meet the 12-year benchmark as a practicing attorney (three years and six years of experience, respectively), yet clerked for federal judges. Kagan's law degree came from a Top 14 institution, Harvard, but Kozinski's came from UCLA, just outside the upper echelon. Of course, Kagan was female and Kozinski male, which advantaged Kozinski. The adjusted measure assigns both nominees essentially the same rating, -1.42 for Kagan and -1.41 for Kozinski, again reflecting their similarly thin legal résumés at the time of their nominations.

The adjusted measure benefits not just Republican nominees. Consider finally Obama Second Circuit nominee Susan Carney—a Harvard Law School graduate and former federal clerk with 2.5 times the ABA's suggested amount of experience as a lawyer—whom the ABA nevertheless rated only "qualified." Carney's background compares well with Trump Fifth Circuit nominee Halil Ozerden—a Stanford Law School graduate and former federal clerk with a decade less overall legal experience than Carney. But, Ozerden served 12 years as a federal district court judge. The ABA rated Ozerden "well qualified." The adjusted measure, though, assigns Carney and Ozerden the same 1.08 score—both are one standard deviation above the

mean in overall qualifications—which suggests the ABA may have unduly discounted, relative to other nominees, Carney's prodigious private practice experience or unfairly downgraded her due to her gender.

[FIGURE 1 ABOUT HERE]

Figure 1. *Distributions of qualifications measures.* The top panel shows the distribution of raw ABA ratings. The bottom panel displays the distribution of adjusted ABA ratings.

Third, as Figures 1 shows, the adjusted measure is more fine-grained than the categories of the raw ABA ratings. ABA ratings skew drastically to the left, with over 56 percent of the rated nominees receiving the top possible rating. The adjusted ratings approximate a normal curve, which is a function not of model estimation, but rather the combinations of over 600 unique backgrounds. Practically, the distribution comports with expectations for circuit court nominees.[6] Most are professionally competent, and there are a few outstanding nominees. And, there are others whose résumés show promise but lack much objectively measurable experience. Table 3 compares the summary statistics for raw and adjusted ABA ratings.

	Raw ABA Ratings	Adjusted ABA Ratings
Mean	3.89	-0.01
Median	5	-0.003
Mode	5	---
1st Quarter	2	-0.647
3rd Quarter	5	0.594
Range	5	5.85
Variance	2.22	1.01
Observations	640	646
Correlation	r=0.38	

[6] Standardization does not produce a normal distribution, but rather a distribution with mean 0 and a standard deviation of 1. If applied to raw ABA ratings with whole-number values 0 to 5, standardization will produce the same six groupings of observations. The groupings will simply shift left along the x-axis.

Table 3. Summary Statistics of Qualifications Measures. *Note*: The number of observations for adjusted ABA ratings includes nominees whom the ABA did not officially rate. The method of adjustment creates ratings for these nominees. None of the nominees who did not have an official ABA rating appears in the data for the empirical example.

In sum, the adjusted measure is less biased as a measurement of the underlying qualifications construct, valid on its face, and reliably assigns values nominees with similar backgrounds. In short, adjusted ABA ratings correct many of the ABA ratings' measurement problems. However, empirical demonstration shows the ultimate usefulness of the adjustment. The three attractive qualities of adjusted ABA scores are not much help if the measure does not have a fourth quality: improving a researcher's ability to model real world political behavior.

3. Empirical Example: Senators' Confirmation Hearing Word Choice

I now turn to an empirical example in which qualifications may influence the behavior of institutional actors in the nomination and confirmation process. The example, a quantitative analysis of senators' word choice at confirmation hearings for circuit court nominees over a 15-year period, builds directly on the foundation that Dancey, Nelson, and Ringsmuth (2014, 2020) laid. Using adjusted ABA ratings leads to a materially different understanding of the role nominees' qualifications play in the hearings.

In the example, I estimate three models. The first model uses raw ABA ratings as its measure of qualifications. The second model replaces raw ABA ratings with adjusted ABA ratings. The third model drops any form of ABA rating in favor of a proxy indicators approach. The estimation strategy generates a working hypothesis to test whether the adjusted measure works well. If adjusted ABA ratings perform as expected, the model with adjusted ratings should improve the overall model fit: Concretely, *the second model should have a lower AIC and BIC than the other models*.

AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) allow for

the comparison of non-nested models, so they are useful for swapping out variables with others. BIC tends to favor more parsimonious models, whereas AIC values reducing the log-likelihood. For both information criteria, a difference in value less than two between models suggests their equality, i.e., evidence for the null hypothesis. A difference greater than four suggests the model with the lower value is "considerably" better (Fabozzi, Focardi, Rachev, and Arshanapalli 2014, 401-403).

3.1 Senatorial Word Choice: The Qualifications and Non-Qualifications Frames

Do senators speak differently in confirmation hearings because of who the nominee is? Dancey, Nelson, and Ringsmuth's (2014, 2020) work on lower court hearings suggests the answer is no: *It's Not Personal*; it's politics. While altering some of Dancey, Nelson, and Ringsmuth's research design, I retain their basic approach, including their theoretical logic and independent variables of interest. The results point to senators' adoption of more personalized approaches to nominees than Dancey, Nelson, and Ringsmuth's findings suggest.

Why do senators mention a particular topic more often in a confirmation hearing as opposed to another? I contend that there are two frames—alternative sets of "terms in which equivalent choices are described" (Iyengar 1987, 11; Kahneman and Tversky 1984)—that senators can use in their hearing speech. One frame more heavily emphasizes nominee qualifications—even if qualifications are not the main thrust of the questioning. The other frame de-emphasizes qualifications talk in favor of other politically salient terms. The contention is broadly consistent with Dancey, Nelson, and Ringsmuth's (2014, 2020) and Collins and Ringhand's (2013) question coding schemes.

First, in the *qualifications frame*, senators exercise their advice and consent role by mentioning the nominees' professional background and temperament. In both their preliminary

comments and direct questions, the senators will talk about the nominees' experiences, making references to their legal training, academics, intellect, compassion, and courtesy. The ostensible motivation behind this type of talk is to establish the nominees' fitness for judicial office, independent of the decisions they will make once on the bench. Given that every nominee to the federal bench has a law degree and some further legal experience, senators will more likely use qualifications words to bolster nominees than denigrate them, all else equal. Indeed, Dancey, Nelson, and Ringsmuth (2020, 81) observe, "senators from the president's party can facilitate the confirmation of … policy allies by highlighting nominees' qualifications."[7] Senators thus use the frame to achieve indirectly their goal of good public policy (Fenno 1978). Hypothetically, *senators ideologically proximate to the nominee will use more qualifications frame words than senators who are ideologically distant*.

Second, in the *non-qualifications frame*, senators talk to nominees about anything but qualifications. Typically, they focus on hot-button judicial issues, e.g., the death penalty or the Violence Against Women Act in the late-1990s, to the rights of detainees or disadvantaged groups in the 2000s. Through the non-qualifications frame, senators view the confirmation hearing as both a platform to espouse their reading of the law, i.e., position-taking, and a last-ditch effort to paint nominees as "outside the mainstream," claiming credit if nominees

[7] Based on their intensive experience of hand-coding nearly 1,000 hearing transcripts, Dancey, Nelson, and Ringsmuth (2020, 81-83) argue, "While [q]ualifications questions occasionally critique a nominee's preparedness to serve on the bench, these questions more typically allow nominees to showcase their résumés and credentials. … Unlike senators from the president's party, who have an incentive to create an impression that the president is nominating qualified candidates to the federal bench, the opposition party has incentives to paint the president's nominees as ideologues." Though the present study uses machine counts of word mentions, the process of personally cleaning each hearing transcript to prepare it for machine coding leads me to the same conclusions.

ultimately do not move forward in the process (Mayhew 2004, Steigerwalt 2010). Thus, *senators ideologically farthest away from the nominee will use fewer qualifications words*.

While the sincere qualifications and non-qualifications frame hypotheses are straightforward, they ignore the potential influence nominees' actual qualifications might play in altering senators' word choices. Put simply, highly qualified nominees' résumés might force senators to take notice of them, thereby drawing senators from the non-qualifications frame into the qualifications frame. Senators already opposed to confirmation can easily ignore less-qualified nominees' résumés. Informing this logic, Dancey, Nelson and Ringsmuth (2020, 105) present results "consistent with a staffer who told us that objections to lower court nominees are 'very rarely on the merits,'" i.e., qualifications-driven. Ideologically opposed senators may also look at nominees with thin credentials as especially good targets for issue-based questioning. Not providing adequate answers to salient legal questions embarrasses nominees who lack extensive legal backgrounds to bolster their *prima facie* case for confirmation. So, there is good reason to believe that ideological distance and nominee qualifications interact in a specific way: *Ideologically distant senators will use fewer qualifications words with less-qualified nominees than they will with highly qualified nominees*.

Note that the theory does not imply the converse conditional effect. Nominee qualifications should not affect ideologically close senators' word choice. Consider the upside of talking about technical legal matters with less-qualified nominees if senators want the nominees confirmed: At best, the nominees show that they are conversant on legal topics. But, given that an ideologically proximate senator engages the nominees, the conversation will not likely impress any undecided senators necessary for confirmation. The downside is steep: Nominees flub "softball" questions on the law from friendly senators, effectively "striking out" at their

confirmation chance. The safest play is for ideologically close senators to mention qualifications no matter the nominees' backgrounds, reiterating the *prima facie* confirmation case for highly qualified nominees and bolstering less-qualified nominees. Therefore, the null hypothesis—*no difference in the number of qualifications words for ideologically close senators across levels of nominee qualifications*—should find support.

Though they were for district court seats, two recent SJC hearings, both featuring Senator John Kennedy (R-LA), illustrate the different expectations well. In 2017, faced with a panel of five Trump nominees, Senator Kennedy began his questioning by probing their backgrounds. Matthew Petersen was the only nominee who answered that he had not tried a case to verdict. After 12 more questions about Petersen's background, Kennedy then asked three questions about judicial decision making—what the *Daubert* standard for expert witnesses was, what a motion *in limine* was, and what the *Younger* and *Pullman* abstention doctrines were—that Petersen utterly failed to address.

In contrast, Kennedy began a 2023 hearing with five Biden nominees by bypassing any questions about their qualifications. After perfunctorily addressing her as "judge," Kennedy asked Washington state court Judge Charnelle Bjelkengren what Articles V and II of the Constitution covered and what "purposivism" in judging was. Judge Bjelkengren, like Petersen, did not attempt to explain or define the passages and concepts Kennedy asked about.

Several important facets of the two hearings deserve highlighting. First, consistent with the theoretical expectations, Kennedy spent ample time talking about the qualifications of the eventually embattled Trump nominee Petersen. Only after firmly establishing the nominee's background did Kennedy ask the procedural questions, and Petersen could have coasted to confirmation if he had been able to answer them. Second, and also consistent with expectations,

Kennedy did not spend any time on the Biden nominees' backgrounds; rather, he presented them with questions about conservative judicial priorities right away, including the difference between purposivism and originalism. Finally, all nominees in the two hearings who engaged with Kennedy—Petersen, Bjelkengren, Matthew Brookman, and Oriela Eleta Merchant—consistently referred to their own qualifications in their responses, again demonstrating that, in SJC hearings, qualifications talk is overwhelmingly more often a shield than a sword.

3.2 Data and Measures

To determine why senators mention qualifications-related words, I machine coded all of the confirmation hearings for circuit court nominees held between 1997 and 2012.[8] The dependent variable is a count of the qualifications-related words a senator spoke in a hearing. I identified terms related to judicial qualifications and experience in part by taking the beginning fragments of words from the ABA's backgrounder on how it rates nominees (American Bar Association, 2020). I included some word fragments inductively after reading several presentations of nominees.[9]

---

[8] See Appendix C for the details of the machine coding process. I excluded one entire hearing, Michael Wallace in 2006, because he was the only nominee to receive a "not qualified" ABA rating. The time frame in this study does not go beyond 2012 because the study attempts to hew as closely as possible to Dancey, Nelson, and Ringsmuth's (2020) research design, which covers 1993 to 2012. If the study included post-nuclear option hearings from 2013 to present, the change in Senate rules could easily confound the finding with respect to senators' use of different frames.

[9] The word fragments were "exper," "quali," "profess," "competen," "integri," "temperam," "charac," "reputa," "indust," "dilige," "intell," "bias," "compass," "civic," "serv," "talen," "scholar," "school," "teach," "lawy," "prosecuto," "attor," "universi," "clerk," "colleg," "laude," "honors," "alumn," "solicitor," "credent," "coif," and "kappa." I also identified "aba" and "edito" as key qualification-related terms. These words would fall under the "experience" and "temperament" question categories Dancey, Nelson, and Ringsmuth (2014, 2020) identify, which follows the ABA's assertion that it rates "strictly on professional qualifications: integrity, professional competence and judicial temperament" (American Bar Association 2020, 1).

Following Dancey, Nelson, and Ringsmuth (2020), I split the independent variables of theoretical interest into two groups: nominee-specific factors and contextual factors.[10] Nominee-specific factors begin with *ideological distance*, which is the absolute value of distance between the speaking senator's Common Space score (Poole 1998) and the nominee's Common Space score, calculated using the method Giles, Hettinger, and Peppers (2001) suggest. The GHP method leverages the norm of senatorial courtesy in nominations by assigning nominees the Common Space of their home-state senators—or the average of the two home-state senators—if the home-state senator(s) are members of the same party as the president. If both home-state senators are of the party opposite the president, then the nominee receives the president's Common Space score. Because theory suggests that senators will use fewer qualifications words with any ideologically distant nominee, regardless of direction of the distance, the absolute value is appropriate.

The second nominee specific factor, qualifications, includes three distinct measurements, each estimated in a separate model. The first is *raw ABA rating*, which ranges from 1 ("qualified-not qualified") to 5 ("well-qualified") in the analysis because I exclude the Wallace hearing. Second is the *adjusted ABA rating*, which ranges from -2.24 to 3.23. Third, *proxy indicators* can represent qualifications. For this approach, I include the same experience variables—years as a federal judge, state judge, private practice, and government attorney; experience as a law professor, federal clerk, and in partisan politics; and graduation from a Top 14 law school—used to derive the adjusted ABA ratings. To test the conditional hypotheses, I interact each of these qualifications measures with ideological distance.

[10] Appendix D presents the summary statistics for the dependent and independent variables.

The remaining nominee-specific and contextual factors are of less importance to demonstrating the usefulness of adjusted ABA ratings, so I condense their discussion. The third nominee-specific factor is whether there is at least one senator of the opposition party to the president representing the nominee's home state. Dancey, Nelson, and Ringsmuth (2020) argue that the presence of an opposition party home state senator can moderate the nominee's ideology; therefore, senators will use more qualifications frame words with such nominees. Demographic variables round out the nominee-specific factors. I include indicators for *non-white* or *female* identity, as well as nominees' *age* and the *square of age*.

The contextual factors are the same four variables that Dancey, Nelson, and Ringsmuth (2020) use, plus an additional one due to this study's slightly wider scope. Whether the hearing occurs during a *presidential election year* or *divided government* should reduce the number of qualifications words, given that both contexts heighten senators' ideological concerns. *Panel balance*—the existing difference between the fraction of active judges on the circuit appointed by Democratic presidents and the fraction appointed Republican presidents (Binder and Maltzman 2009a)—can affect the number of qualifications words. As panel balance approaches zero, i.e., a 50-50 split, a new judge will tip the panel in one party's favor. Hence, in this ideologically charged context, qualifications talk will likely recede. Dancey, Nelson, and Ringsmuth (2020) also include the interaction between divided government and panel balance to capture particularly intense ideological contexts. The two other contextual variables are indicators for nominees to the *DC* and *Federal Circuits*, both due to the unique nature of their circuits: the "prestige and salience" of the DC Circuit (Dancey, Nelson, and Ringsmuth 2020, 85) and the technical backgrounds of Federal Circuit judges (Handler 2022).

For pure controls, I include first whether the speaker is in the *majority party* of the Senate at the time of the hearing. The *presenter*, often a senator, introduces the nominee to the committee by recounting the nominee's biography, thus increasing the number of qualifications-related words used. The *chair of the hearing*, as identified by the hearing transcript, presides over the hearing, recognizing the order of speakers and keeping time. I include the *log of the total words* spoken by a senator during a hearing and whether it was the *second hearing* for the nominee. The former variable ensures that verbose senators, who in their logorrhea happen to use some qualifications-related words as well, do not produce spurious results. The latter variable is important because the senators may dispense with many qualifications references in a second hearing, having already heard them in the first.

3.3 Methods and Results

Because the dependent variable is a count, I estimate three negative binominal regressions, replacing different measures of qualifications in each. The first model uses raw ABA ratings. The second uses adjusted ABA ratings, and the third uses proxy indicators. In all models, I clustered standard errors on the date of the hearing. In some cases, the SJC considered and questioned multiple nominees in the same hearing, though each nominee counted as a separate set of observations. Thus, the errors in qualifications-related words are likely correlated within the hearing itself. Table 4 displays the regression estimates for the different qualifications measures and ideological distance.[11]

	I	**II**	**III**
	Raw ABA	**Adj. ABA**	**Proxy Ind.**
Ideological Distance	-0.86**	-0.55**	-0.34
	(0.22)	(0.07)	(0.24)

[11] See Appendix E for the full regression results.

Raw ABA Rating	-0.02 (0.02)	---	---
Adjusted ABA Rating	---	-0.02 (0.04)	---
Ideological Distance * Raw ABA Rating	0.08 (0.05)	---	---
Ideological Distance * Adjusted ABA Rating	---	0.14* (0.06)	---
Federal Judge (Years)	---	---	-0.02* (0.01)
Ideological Distance * Federal Judge (Years)	---	---	0.01 (0.01)
State Judge (Years)	---	---	-0.01 (0.01)
Ideological Distance * State Judge (Years)	---	---	0.00 (0.01)
Private Practice (Years)	---	---	-0.01 (0.01)
Ideological Distance * Private Practice (Years)	---	---	-0.01 (0.01)
Government Attorney (Years)	---	---	0.01 (0.01)
Ideological Distance * Government Attorney (Years)	---	---	-0.02 (0.011)
Law Professor	---	---	0.06 (0.08)
Ideological Distance * Law Professor	---	---	-0.28* (0.13)
Federal Clerk	---	---	0.22** (0.08)
Ideological Distance * Federal Clerk	---	---	0.04 (0.13)
Top 14 Law School	---	---	-0.04 (0.08)
Ideological Distance * Top 14 Law School	---	---	0.20 (0.12)
Political Experience	---	---	0.02 (0.01)
Ideological Distance * Political Experience	---	---	-0.32* (0.15)
N	919	919	919
Log-pseudolikelihood	-2380.63	-2378.53	-2349.44
AIC	4803.26	4799.07	4768.89
BIC	4904.55	4900.36	4937.70

Notes: *: p<0.05 **: p<0.01 (two-tailed)

Table 4. *Negative Binomial Regression Results, Limited to Qualifications Measures.* The dependent variable is a count of the number of qualifications-related words senators used in a nominee's hearing. Model I uses unadjusted ABA ratings. Model II uses adjusted ABA ratings. Model III uses multiple measures of nominee qualifications.

Before discussion of the independent variables' effects, a comparison of the models is necessary to see which approach to modeling qualifications best fits the data. The proxy indicators model (Model III) has the lowest AIC, suggesting the best model fit. However, the adjusted ABA ratings model (Model II) fares best under BIC, the information criterion that favors more parsimonious models. Indeed, with respect to BIC, Model III is the worst-fitting model. Both the AIC and BIC difference between Models I and II is greater than four, which is "considerabl[e]" and positive evidence that Model II fits better than Model I (Fabozzi, et al. 2014, 401-403). Model II is not the worst fitting model under either criterion. It is distinguishably the best under the more exacting BIC.

Though they fit the data best, do adjusted ABA ratings change anything substantive about the interpretation of qualifications' role in the confirmation process? They do, considering the difference in effects displayed in Figure 2. While the sincere qualifications frame and non-qualifications frame hypotheses find support in Model I—the effect of ideological distance is negative and statistically significant—the conditional hypothesis finds barely marginal support. The interaction between qualifications and ideological distance is not statistically significant.[12] Yet, in Model II, both the sincere and conditional hypotheses find clear and strong support. The effect of ideological distance is effectively the same in both models; however, the interaction between ideological distance and qualifications is positive and statistically significant in the

[12] The interactive term results in a statistically significant difference over less than a quarter of the observed range of ideological distance.

adjusted ABA ratings model.[13] The results suggest that the background of highly qualified nominees can, in fact, influence senators' speech in the confirmation hearing. The finding contrasts with Dancey, Nelson, and Ringsmuth's (2014, 2020) findings that individual nominees' characteristics do not influence senators' questioning.

Figure 2 shows how the introduction of adjusted ABA ratings changes the substantive interpretation of qualifications' role in SJC hearings. In both panels, the predicted number of qualifications-related terms are plotted against increasing ideological distance between the senator and the nominee. To provide some context for the substantive discussion, the model predicts that nominee presenters—ordinarily home-state senators whose one job is to provide the SJC with background information on the nominee—will use 12.8 qualifications words. The ideologically closest presenters, senators whom the GHP measure assumes have essentially made the nomination themselves, will use 16.5 qualifications words, while the ideologically farthest presenters, senators from the nominee's home state but not of the president's party, will use 8.7 qualifications words.[14]

In the top panel of Figure 2, the dashed line represents the predictions for a "well qualified" nominee, and the solid line represents a "qualified-not qualified" nominee, the lowest ABA category used in Model I's estimation. A senator at minimal ideological distance is expected to use 10.31 qualifications-related words for a "well qualified" nominee and 11.08 for a "qualified-not qualified" nominee. The difference is not statistically significant ($X^2_{(1)}$ = 0.51,

[13] Moreover, the switch from ABA ratings to adjusted ABA ratings does not materially alter the estimated effects other variables. The consistent estimation suggests that the adjusted measure's main benefit—making the distinction between qualifications and politics clearer—does not bias the understanding of unrelated variables.

[14] The most distant presenter observed in these data is Senator John Corzine, who introduced Bush Third Circuit nominee Michael Chertoff at his May 7, 2003, SJC hearing.

p>0.47). At maximum ideological distance, a senator will use 5.57 qualifications words for a "well qualified" nominees and 3.72 for "qualified-not qualified" nominees. The 40 percent difference reaches conventional levels of statistical significance ($X^2_{(1)}$ = 4.12, p<0.05). The differences between minimum and maximum levels of ideological distance within each ABA rating are statistically significant ($X^2_{(1)}$ = 30.13, p<0.01 for "well qualified"; $X^2_{(1)}$ = 29.50, p<0.01 for "qualified-not qualified"). Thus, individual nominees matter mostly to the extent that their ideology diverges from the senator, a finding consistent with the *It's Not Personal* argument.

[FIGUR 2 ABOUT HERE]

Figure 2. *Comparison of the substantive effect of qualifications and ideological distance*. The top panel displays the substantive change in ideological distance for "well qualified" (dashed line) and "qualified-not qualified" (solid line) qualified nominees from Model I. The bottom panel shows the same for high (+2 standard deviations) and low (-1 standard deviation) qualified nominees from Model II. The shaded regions represent the predictions' 95% confidence intervals.

In contrast, the bottom panel, created using estimates from Model II, shows strong support for the conditional hypothesis. Because the adjusted ABA ratings do not fall into discrete categories, the dashed line represents a nominee one standard deviation below the mean qualifications. The solid line represents a nominee two standard deviations above the mean.[15] At the minimum of ideological distance, there is still no statistically significant difference between higher- and lower-qualified nominees. With highly qualified nominees, senators use 10.16 qualifications-related words, and they use 10.85 words with low-qualified nominees ($X^2_{(1)}$ = 0.29,

[15] I chose these values for the adjusted ABA rating to reflect "low qualifications" and "high qualifications," respectively. The values capture a rough analogue to the ABA's split "qualified-not qualified" and unanimously "well qualified" ratings. The results are substantially similar for alternative values for "low" and "high," though the point estimates of qualifications words narrow predictably as the values converge.

p>0.58). Like Model I, ideologically close senators mention qualifications almost as much as the average presenter does.

But, at maximum ideological distance, there is a statistically and substantively significant difference between high and low qualifications. While an ideologically distant senator uses 4.13 qualifications-related words—about 47.5 percent of an ideologically-distant presenter—with nominees who have relatively thin résumés, the same senator uses 7.00 qualifications words with highly qualified nominees. The 69 percent increase in qualifications word usage is statistically significant ($X^2_{(1)}$ = 7.74, p<0.01). Substantively, the 2.87-word increase makes the ideologically distant senator's speech resemble 80.5 percent of an ideologically distant presenter's speech in terms of qualifications mentions.[16] The differences between highly- and low-qualified nominees are statistically significant for all values of ideological distance greater than 0.60, which account for 47 percent of all observations.

For highly qualified nominees, the 3.18-word decline in qualifications frame mentions between an ideologically proximate senator and an ideologically distant senator is statistically significant ($X^2_{(1)}$ = 4.86, p<0.05).[17] For low-qualified nominees, the difference is also statistically significant ($X^2_{(1)}$ = 60.88, p<0.01), but the magnitude of the decline is over 110 percent larger:

[16] The average number of questioning senators per hearing in these data is four. If all senators in the hearing were maximally ideologically hostile to the nominee, the cumulative effect of the increase would be about 11.5 words over the course of the hearing. Thus, the cumulative effect would be nearly the same as having another average presenter in the room.

[17] The decline is *not* statistically significant for nominees 2.15 standard deviations above the mean of the adjusted ABA ratings. While an insignificant decline would present the strongest evidence supporting the argument that nominee résumés matter—every senator uses the qualifications frame at the same rate, regardless of ideological distance—I maintain the two standard deviations above the mean threshold for consistency.

6.72 words.[18] Substantively, the decline represents moving from 85 percent to under one-third of the average presenter's speech. In short, the Model II results demonstrate that nominees' backgrounds can influence the frame senators use, a clear contrast to Dancey, Nelson, and Ringsmuth's (2020) conclusions.[19]

4. Discussion

This study makes two substantive contributions to the study of lower courts and the lower court confirmation process. First, it created a single valid and reliable measure of qualifications for nominees to the circuit courts. It then tested the measure against two alternative operationalizations of qualifications: the existing raw ABA ratings and a proxy indicators approach. The adjusted ABA ratings fit the data better on balance than either of the other two approaches. In future studies, when appropriate, researchers can use the adjusted ABA ratings as a common, consistent measure of qualifications, which will allow studies to build on each other directly.

Second, the study contributes the data and findings with respect to the content of confirmation hearings. It showed that ideology is a driving force behind the choice senators make to mention qualifications. It further showed that nominees' backgrounds, if sufficiently impressive, could induce senators to mention their qualifications. Moving from a low-qualified nominee to a highly qualified nominee results in a 69 percent increase in the number of qualifications-related words used *by each senator* in the hearing. A finding of this magnitude

---

[18] The same comparison in the raw ABA model produces only a 55 percent larger magnitude—4.75 words and 7.36 words, respectively.

[19] In the interests of space, I do not offer a full discussion of Modell III's results. They are somewhat inconsistent, with more federal judicial and private practice experience leading to less qualifications talk, but federal clerk experience leading to more. Two interactions with ideological distance (law professor and partisan political experience) are negative, but the interaction with Top 14 law school graduation is positive.

does not appear using raw ABA ratings. The significant left-skew and blunt categorization of ABA ratings obscure the actual information transmitted to senators. The adjusted ratings strip the noise related to ideology and skewness, leaving a signal that is closer to the information gathered by senators and their staff while evaluating nominees. The findings also contrast with the notion that lower court confirmation hearings are not personal (Dancey, Nelson, and Ringsmuth 2014, 2020).

On a higher level, the research highlights the close connection between qualifications and politics. Political considerations characterize the path to a seat on the lower federal courts. The equivocal findings of research into qualification effects in the confirmation process and on judicial behavior reinforce the link between ideology and qualifications. Disentangling qualifications and politics is key to doing methodologically sound social science to answer questions about the federal judiciary. While a disentangled measure exists for Supreme Court justices and nominees (Cameron, Segal, and Songer 1990), none existed for judges on the circuit courts. None still exists for the district court.

Adjusted ABA ratings do not reveal the content of the confidential interviews that apparently can cause members of the SCFJ to switch their votes on nominees (Tober 2006). Knowing the factors and personalities that enter the ratings' formulation can lead to their better consumption by journalists, senators, and presidents. Because it appears that the original investigator, along with the committee, influence the rating, the research can lead to better procedures within the SCFJ itself, which are already under discussion (Little 2001). If "the quality of our judges ... is ... the quality of our civilization," then we should know "the significance of" the quality "measures which are taken" (Leflar 1960, 305; Mott 1948, 262).

Acknowledgements
I thank Brad M. Jones for his programming expertise. I thank Ronnie Olesker, Nancy Scherer, Justin Wedeking, Steve Wasby, Logan Dancey, Ken Mayer, Alex Tahk, Ryan Owens, Ellie Powell, Sida Liu, and anonymous reviewers at this and other journals for their encouragement, insight, and suggestions. I presented earlier versions of this research at the 2014 and 2015 Midwest Political Science Association Meetings and the 2014 American Political Science Association Meeting.

Funding Declaration
The author received no financial support for the research, authorship, and/or publication of this article.

Competing Interests Declaration
The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability Statement
All data to reproduce the tables and figures in this publication will be available on the author's website and the Journal's Dataverse archive.

References

ABA Watch. 2006. A Focus On: The ABA Standing Committee on Federal Judiciary. *ABA Watch* (August). Washington, DC: The Federalist Society for Law and Public Policy Studies.

Allison, Garland. 1996. "Delay in Senate Confirmation of Federal Judicial Nominees." *Judicature* 80: 8-15.

American Bar Association. 2009. "Standing Committee on the Federal Judiciary: What It Is and How It Works." Online at http://www.americanbar.org/content/dam/aba/uncategorized/GAO/Backgrounder.authcheckdam.pdf.

American Bar Association. 2020. "Standing Committee on the Federal Judiciary: What It Is and How It Works." Online at https://www.americanbar.org/content/dam/aba/administrative/government_affairs_office/backgrounder-9-21-2020.pdf.

Benesh, Sarah, David Armstrong, and Zachary Wallander. 2020. "Advisors to Elites: Untangling Their Effect." *Journal of Law and Courts* 8 (1): 51-73.

Binder, Sarah, and Forrest Maltzman. 2002. "Senatorial Delay in Confirming Federal Judges, 1947-1998." *American Journal of Political Science* 46: 190-199.

Binder, Sarah, and Forrest Maltzman. 2009a. "Advice and Consent During the Bush Years: The Politics of Confirming Federal Judges." *Judicature* 92: 320-329.

Binder, Sarah, and Forrest Maltzman. 2009b. *Advice and Dissent: The Struggle to Shape the Federal Judiciary*. Washington, DC: Brookings.

Black, Ryan, and Christina Boyd. 2011. "The Role of Law Clerks in the U.S. Supreme Court's Agenda-Setting Process." *American Politics Research* 40: 147–173.

Black, Ryan, Anthony Madonna, and Ryan Owens. 2014. "Qualifications or Philosophy? The Use of Blue Slips in a Polarized Era." *Presidential Studies Quarterly* 44 (2): 290-308.

Black, Ryan, and Ryan Owens. 2013. "Bargaining and Legal Development in the United States Courts of Appeals." *American Politics Research* 41: 1071-1101.

Cameron, Charles, Albert Cover, and Jeffrey Segal. 1990. "Senate Voting on Supreme Court Nominees: A Neoinstitutional Model." *American Political Science Review* 84: 525-534.

Choi, Stephen, Mitu Gulati, and Eric Posner. 2011. "What Do Federal District Judges Want? An Analysis of Publications, Citations, and Reversals." *Journal of Law, Economics, and Organization* 28 (3): 518-549.

Choi, Stephen, Mitu Gulati, and Eric Posner. 2013. "How Well Do Measures of Judicial Ability Predict Judicial Performance?: A Case Study Using Securities Class Actions." *International Review of Law and Economics* 33: 37-53.

Christensen, Robert, and John Szmer. 2012. "Examining the Efficiency of the U.S. Courts of Appeals: Pathologies and Prescriptions." *International Review of Law and Economics* 32: 30-37.

Collins, Paul, and Lori Ringhand. 2013. *Supreme Court Confirmation Hearings and Constitutional Change*. New York: Cambridge University Press.

Dancey, Logan, Kjersten Nelson, and Eve Ringsmuth. 2014. "Individual Scrutiny or Politics as Usual? Senatorial Assessment of U.S. District Court Nominees." *American Politics Research* 42 (5): 784-814.

Dancey, Logan, Kjersten Nelson, and Eve Ringsmuth. 2020. *It's Not Personal: Politics and Policy in Lower Court Confirmation Hearings*. Ann Arbor, MI: University of Michigan Press.

Diamond, Alexis, and Jasjeet Sekhon. 2013. "Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies." *Review of Economics and Statistics* 95 (3): 932-945.

Fabozzi, Frank, Sergio Focardi, Svetlozar Rachev, and Bala Arshanapalli. 2014. *The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications*. Hoboken, NJ: Wiley.

Farganis, Dion, and Justin Wedeking. 2014. *Supreme Court Confirmation Hearings in the U.S. Senate: Reconsidering the Charade*. Ann Arbor, MI: University of Michigan Press.

Fenno, Richard. 1978. *Home Style: House Members in Their Districts*. Boston: Little, Brown, and Company.

Franco, Annie, Neil Malhotra, and Gabor Simonovits. 2014. "Publication Bias in the Social Sciences: Unlocking the File Drawer." *Science* 345 (6203): 1502-1505.

Giles, Michael, Virginia Hettinger, and Todd Peppers. 2001. "Picking Federal Judges: A Note on Policy and Partisan Selection Agendas." *Political Research Quarterly* 54: 623-641.

Handler, Samantha. 2022. "Black Women Lawyers Still Sparsely Seen at Federal Circuit." *Bloomberg Law.* Online at https://news.bloomberglaw.com/ip-law/black-women-lawyers-still-sparsely-seen-at-federal-circuit.

Hendershot, Marcus. 2010. "From Consent to Advice and Consent: Cyclical Constraints within the District Court Appointment Process." *Political Research Quarterly* 63: 328-342.

Hettinger, Virginia, Stefanie Lindquist, and Wendy Martinek. 2006. *Judging on a Collegial Court: Influences on Federal Appellate Decision-Making*. Charlottesville, VA: University of Virginia Press.

Ho, Daniel, Kosuke Imai, Gary King, and Elizabeth Stuart. 2011. "MatchIt: Nonparametric Preprocessing for Parametric Causal Inference." Online at r.iq.harvard.edu/docs/matchit/2.4-20/matchit.pdf.

Iyengar, Shanto. 1987. *Is Anyone Responsible? How Television Frames Political Issues*.

Chicago: University of Chicago Press.

Kahneman, Daniel, and Amos Tversky. 1984. "Choices, Values, and Frames." *American Psychologist* 39: 341-350.

Kamenar, Paul. 1990. "The Role of the ABA in the Judicial Selection Process." In *Judicial Selection: Merit, Ideology, and Politics.* Henry Abraham, Griffin Bell, Charles Grassley, Eugene Hickok, John Kern III, Stephen Markman, and Wm. Bradford Reynolds, eds. Washington, DC: National Legal Center for the Public Interest, 93-102.

Landes, William, Lawrence Lessig, and Michael Solimine. 1998. "Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges." *The Journal of Legal Studies* 27: 271-332.

Leflar, Robert. 1960. "The Quality of Judges." *Indiana Law Journal* 35: 289-305.

Little, Laura. 2001. "The ABA's Role in Prescreening Federal Judicial Candidates: Are We Ready to Give up on the Lawyers?" *William and Mary Bill of Rights Journal* 10: 37-73.

Lott, John. 2005. "The Judicial Confirmation Process: The Difficulty with Being Smart." *Journal of Empirical Legal Studies* 2: 407-447.

Lott, John. 2006. "Pulling Rank." *New York Times*, January 25. Online at http://www.nytimes.com/2006/01/25/opinion/25Lott.html

Lott, John. 2013. "ABA Ratings: What Do They Really Measure." *Public Choice* 156 (1-2): 139-161.

Martinek, Wendy, Mark Kemper, and Steven Van Winkle. 2002. "To Advise and Consent: The Senate and Lower Federal Court Nominations, 1977-1998." *Journal of Politics* 64 (2): 337-361.

Mayhew, David. 2004. *Congress: The Electoral Connection*. New Haven, CT: Yale University Press.

Merritt, Deborah Jones, and James Brudney. 2001. "Stalking Secret Law: What Predicts Publication in the United States Courts of Appeals." *Vanderbilt Law Review* 54: 71-121.

Mott, Rodney. 1948. "Measurement of Judicial Personnel." *New York University Quarterly Law Review* 23: 262-277.

NOLA.com. 2017. "Watch Trump Judicial Pick Stumble over Sen. John Kennedy's Basic Legal Questions." *YouTube,* December 15. Online at https://www.youtube.com/watch?v=wd-eYl-LiPs.

Poole, Keith. 1998. "Recovering a Basic Space from a Set of Issue Scales." *American Journal of Political Science* 42: 954-993.

Savchak, Elisha, Thomas Hansford, Donald Songer, Kenneth Manning, and Robert Carp. 2006. "Taking It to the Next Level: The Elevation of District Court Judges to the U.S. Courts of Appeals." *American Journal of Political Science* 50: 478-493.

Scherer, Nancy. 2005. *Scoring Points: Politicians, Activists, and the Lower Federal Court Appointment Process*. Stanford, CA: Stanford University Press.

Scherer, Nancy, Brandon Bartels, and Amy Steigerwalt. 2008. "Sounding the Fire Alarm: The Role of Interest Groups in the Lower Federal Court Confirmation Process." *Journal of Politics* 70: 1026-1039.

Sen, Maya. 2014. "How Judicial Qualification Ratings May Disadvantage Minority and Female Candidates." *Journal of Law and Courts* 2 (1): 33-65.

Sieja, James. 2015. Do They Know Them When They See Them? The ABA Standing Committee on the Federal Judiciary and Qualifications for the Federal Bench. PhD diss., University of Wisconsin-Madison.

Sieja, James. 2023. "How You Rate Depends on Who Investigates: Partisan Bias in ABA Ratings

of US Courts of Appeals Nominees, 1958-2020." *Political Research Quarterly* 76 (4): 1723-1735.

Silver, Nate. 2012. *The Signal and the Noise: Why So Many Predictions Fail—But Some Don't.* New York: Penguin.

Slotnick, Elliot. 1983. "The ABA Standing Committee on Federal Judiciary: A Contemporary Assessment—Part 1." *Judicature* 66 (8): 349-362.

Smelcer, Susan, Amy Steigerwalt, and Richard Vining. 2012. "Bias and the Bar: Evaluating the ABA Ratings of Federal Judicial Nominees." *Political Research Quarterly* 65 (4): 827-840.

Steigerwalt, Amy. 2010. *Battle over the Bench: Senators, Interest Groups, and Lower Court Confirmations*. Charlottesville, VA: University of Virginia Press.

Stratmann, Thomas, and Jared Garner. 2004. "Judicial Selection: Politics, Biases, and Constituency Demands." *Public Choice* 118: 251-270.

Taha, Ahmed. 2004. "Publish or Paris? Evidence of How Judges Allocate Their Time." *American Law and Economics Review* 6: 1-27.

Tober, Stephen. 2006. "Statement of Stephen L. Tober on Behalf of the Standing Committee on the Federal Judiciary of the American Bar Association Concerning the Nomination of Brett Michael Kavanaugh to be Judge of the United States Court of Appeals for the District of Columbia Circuit before the Committee on the Judiciary, United States Senate." May 8. Online at http://www.americanbar.org/content/dam/aba/migrated/scfedjud/statements/ kavanaugh. authcheckdam.pdf

Voeten, Eric. 2013. "The ABA Prefers White Men." *The Monkey Cage*, February 28. Online at http://themonkeycage.org/2013/02/28/the-aba-prefers-white-men/

Whelan, Ed. 2010. "Re: The ABA and Ninth Circuit nominee Goodwin Liu." *Bench Memos*,

February 25. Online at http://www.nationalreview.com/bench-memos/49265/re-aba-and-ninth-circuit-nominee-goodwin-liu-ed-whelan.