

Appendix

A. Survey information

Interviews: 1000

Field Period: April 26, 2018 - May 04, 2018

YouGov interviewed 1096 respondents who were then matched down to a sample of 1000 to produce the final dataset. The respondents were matched to a sampling frame on gender, age, race, and education. The frame was constructed by stratified sampling from the full 2016 American Community Survey (ACS) 1-year sample with selection within strata by weighted sampling with replacements (using the person weights on the public use file).

The matched cases were weighted to the sampling frame using propensity scores. The matched cases and the frame were combined and a logistic regression was estimated for inclusion in the frame. The propensity score function included age, gender, race/ethnicity, years of education, and region. The propensity scores were grouped into deciles of the estimated propensity score in the frame and post-stratified according to these deciles.

The weights were then post-stratified on 2016 Presidential vote choice, and a four-way stratification of gender, age (4-categories), race (4-categories), and education (4-categories), to produce the final weight.

B. Descriptions of Judicial Philosophies

Original Intent: Looks to the intent of the drafters and ratifiers of the Constitution to reach conclusions about its meaning.

Original Meaning: Considers what the words in the Constitution meant at the time those words were crafted.

Living Constitution: Looks to the evolving views and values of the American public when interpreting the Constitution.

Global Outlook: Looks to how public officials in other countries have interpreted similar words or phrases.

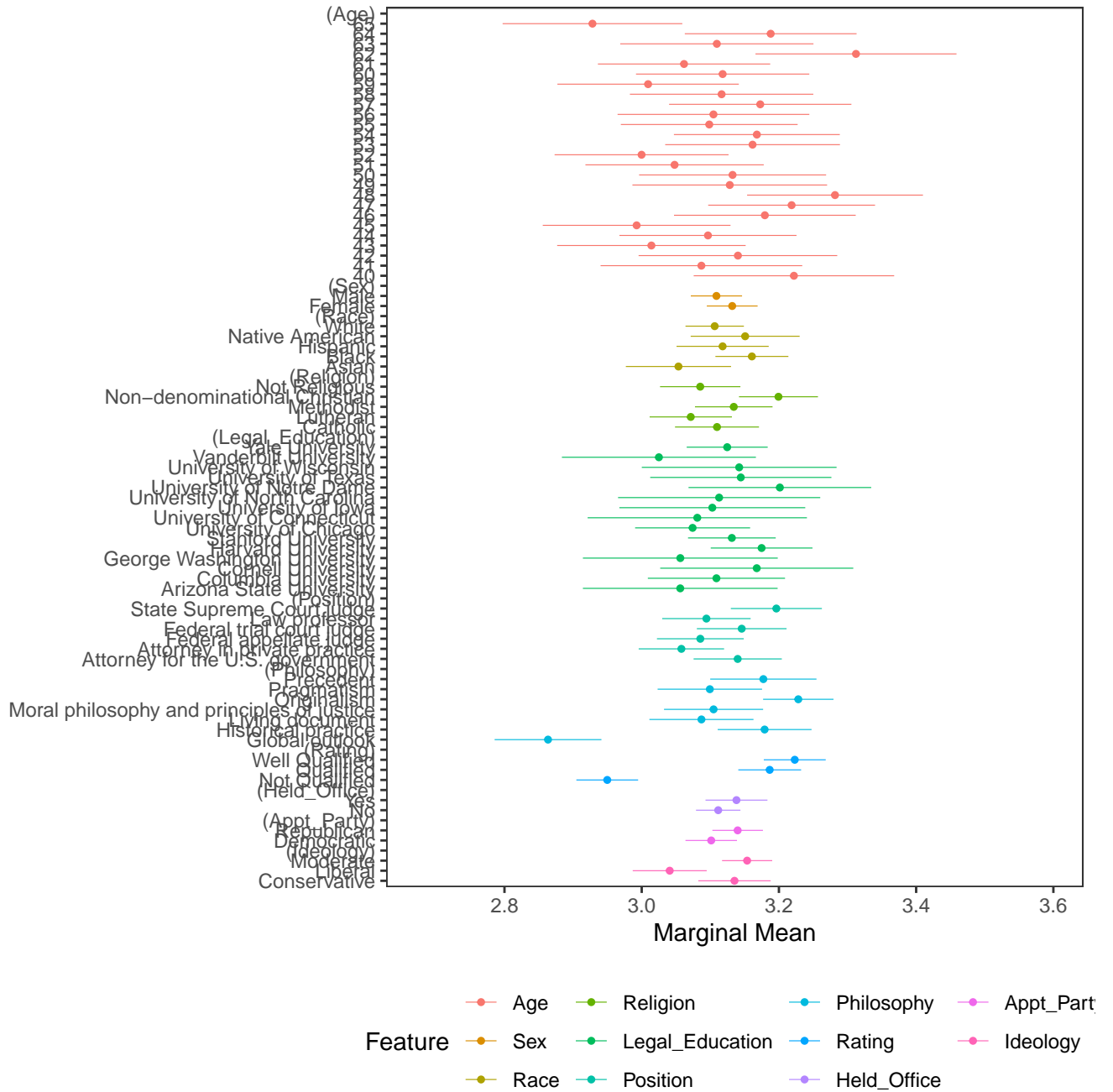
Precedent: Interprets the Constitution based on the precedents established in previous cases.

Pragmatism: When assessing outcomes, selects the one that has the best consequences and rejects the ones that have the worst.

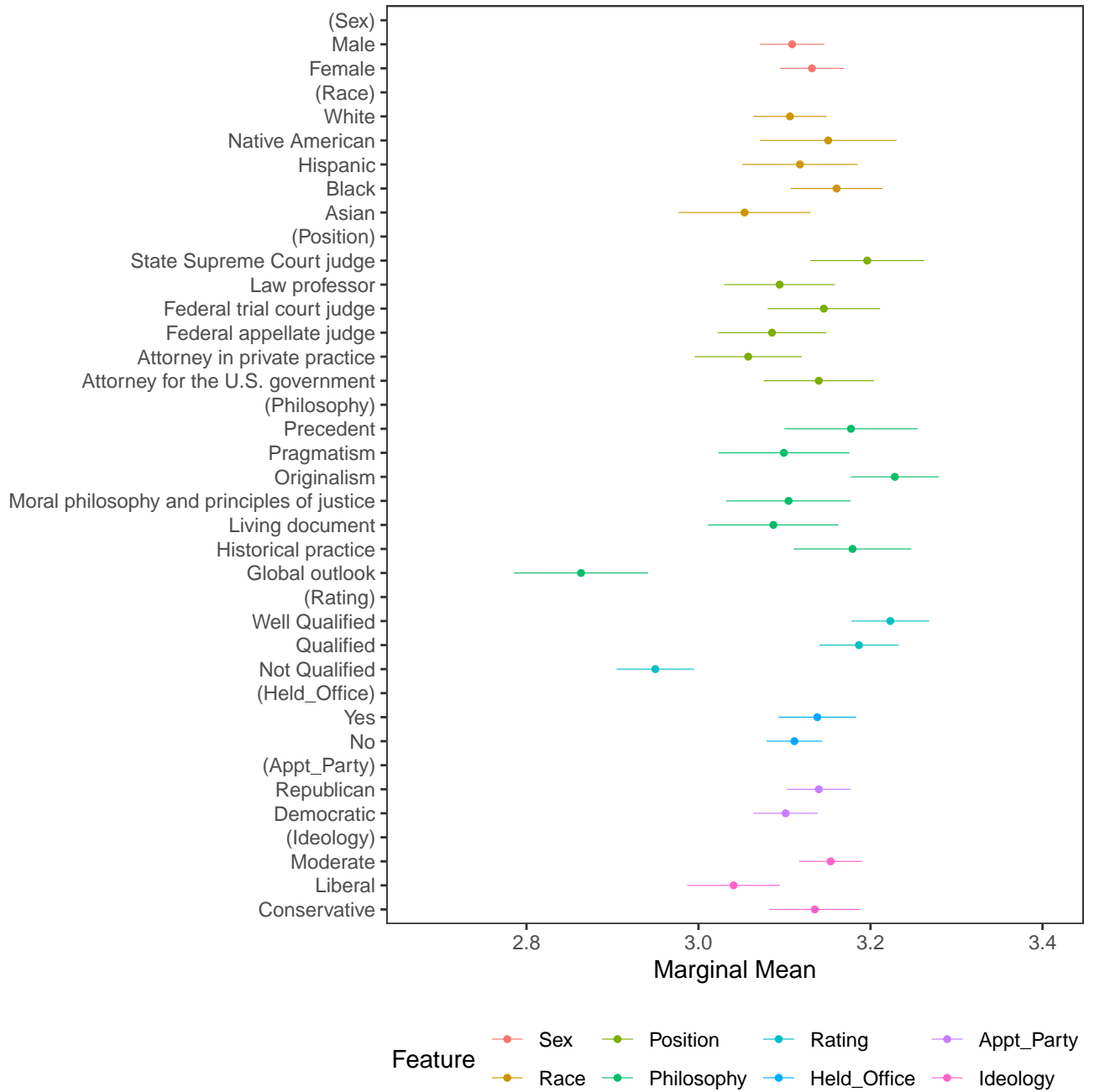
Historical Practice: Looks to longstanding historical practices, especially of the federal, executive, and legislative branches.

Moral Philosophy: Looks to moral philosophy and principles of justice when interpreting the Constitution.

C1 Full Results



C2 Truncated Results



C3 Robustness Check

To test our claim further, we sought to examine the respondents' length of exposure to the appointing or electing selection method. If our theory is correct, people exposed to a selection method for a lengthy period of time will behave consistent with our hypotheses while people exposed to the method for only a short period of time may not. While we were unable to determine how long each respondent was exposed to various judicial selection methods, we were able to compare the behavior of older respondents—those likely exposed to a method for a long time—to younger respondents. (We categorized young respondents as those below the mean age in the sample, which was 49.01 years, and those at or over the mean as older.) The results are consistent with our expectations. Comparing older individuals in appointing states with older individuals in electing states results in statistically significant differences, both descriptively and causally. Most importantly, the difference in causal effects of shifting from low to high qualification ratings (lower-right portion of Figure 1) is statistically significant, with older individuals in appointing states responding more strongly to qualification ratings than older individuals in electing states. We do not observe the same results among younger respondents.

To be sure, we do not know how long our respondents actually lived in their states. It could be, for example, that some of our respondents recently moved to their states. Nevertheless, we think that we can assume an *overall* amount of stability among our respondents. Indeed, the Census Bureau reports that Americans today are moving at “historically low rates.”¹ Further, even if *some* of our respondents recently moved from one state to another, there is no guarantee that they moved from one type of selection method to a different one in any systematically biased way.²

¹<https://www.census.gov/newsroom/press-releases/2016/cb16-189.html>.

²Similarly, if exposure to selection methods influences respondents, we should expect to see results more consistent with our hypotheses among states that have retained their selection

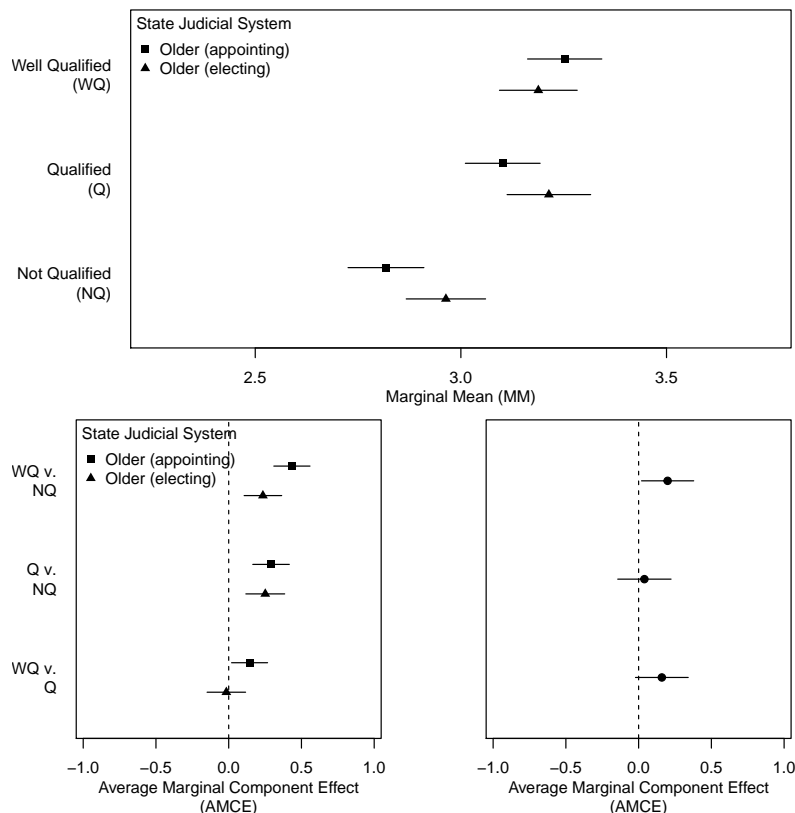


Figure 1: Upper Portion: Marginal means of nominee evaluations based on ABA Ratings for older respondents from states that appoint their judges (squares) and older respondents from states that elect their judges (triangles). Horizontal lines indicate 95 percent confidence intervals. NQ=Not Qualified; Q=Qualified; WQ=Well Qualified. Lower-left: Average Marginal Component Effects (AMCEs). Estimates the causal effect of a shift from one category to another category (pairwise comparison). Lower-right: Difference in AMCEs (Appointing - Electing) for each type of pairwise comparison. Tests hypothesis that older respondents in appointing states respond more strongly to ABA ratings than older respondents in electing states.

methods over long periods of time and less consistent results among states that recently changed their selection methods. Not enough states have recently changed their selection methods, however, to perform an adequate analysis. Future scholars should consider ways in which they might leverage judicial selection reforms.

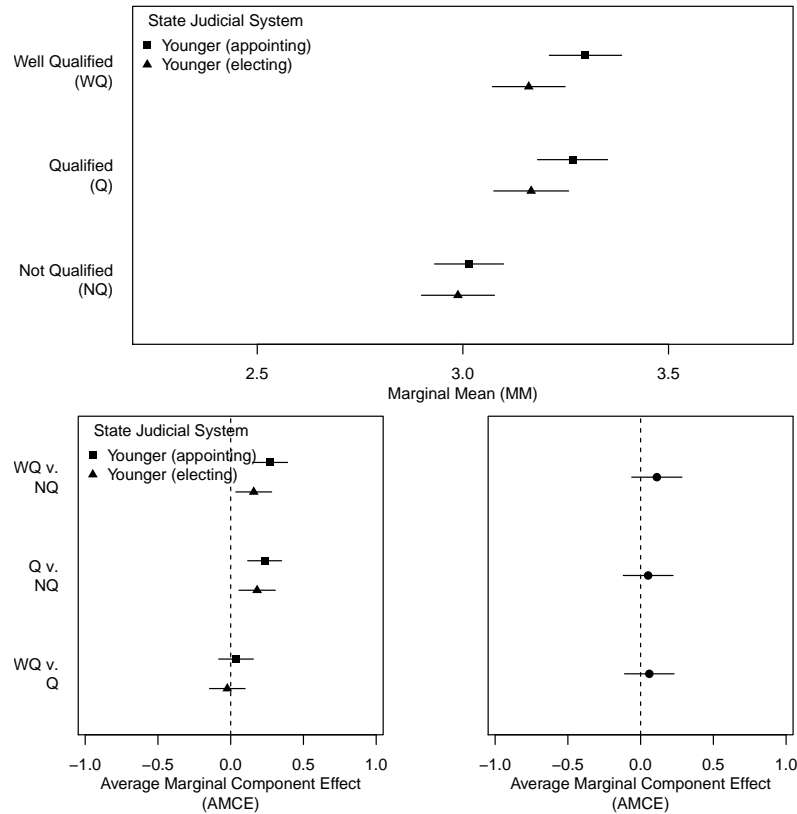


Figure 2: Upper Portion: Marginal means of nominee evaluations based on ABA Ratings for younger respondents from states that appoint their judges (squares) and younger respondents from states that elect their judges (triangles). Horizontal lines indicate 90 (thick line) and 95 (thin line) percent confidence intervals. NQ=Not Qualified; Q=Qualified; WQ=Well Qualified. Lower-left: Average Marginal Component Effects (AMCEs). Estimates the causal effect of a shift from one category to another category (pairwise comparison). Lower-right: Difference in AMCEs (Appointing - Electing) for each type of pairwise comparison. Tests hypothesis that younger respondents in appointing states respond more strongly to ABA ratings than younger respondents in electing states.

Stability and Carryover Effects

We re-estimated our analyses while only considering results obtained from the respondents' first tasks. One important result is that ABA ratings no longer have a meaningful influence on evaluations. Individuals from appointing states do not value qualifications more than individuals from electing states. It appears that individuals from appointing states realize their preference for highly qualified individuals over the course of their evaluations, while individuals from electing states consistently failed to differentiate between nominees of different quality throughout the conjoint.

What are the implications of finding no significant effects on preference for ABA ratings alone, at least for the first conjoint comparison? This could be because the preference is "learned" during the survey (i.e., it does not exist outside of the survey experience). At the same time, there must be some reason that respondents from appointing states begin to focus more on ABA ratings over the course of making multiple conjoint comparisons while respondents from electing states do not. We think our expectancy theory can explain this.

Still, we must grapple with the fact that respondents in appointing states did not come into the survey with a greater preference for legal qualifications than did people from electing states. Fortunately, our argument does not depend on a single nominee characteristic. We analyze the effect of state judicial system more globally using the omnibus F-test; and we still find that state judicial selection system structures evaluations of nominees using only the first conjoint pairing. Furthermore, we find stronger evidence of differentiation for non-legal factors when using only the first conjoint pairing, which counter-balances the disappearance of an effect for legal qualifications.

Looking only at data from the first profile comparison, we continue to find that state judicial selection method structures how individuals evaluate judicial philosophy. An *F*-test (see Footnote 14) shows that interacting the effect of judicial philosophy with our selection method variable improves model fit ($p \approx .03$). We also find strong effects for partisanship, with Democrat respondents evaluating Republican nominees more harshly ($p \approx$

.16), Independent respondents evaluating Republican nominees more favorably ($p \approx .02$), and Republican respondents evaluating Republican nominees more favorably ($p \approx .08$) in electing states compared to appointing states.

Finally, while an omnibus F-test incorporating all of the conjoint attribute levels indicates that judicial selection method does not structure the evaluation of every conjoint characteristic, we do find evidence based on F-tests that it jointly structured how respondents evaluated nominee judicial philosophy, nominee gender, and nominee partisanship ($p \approx .02$).