

All Male Panels? Representation and Democratic Legitimacy



Amanda Clayton Vanderbilt University
Diana Z. O'Brien Texas A&M University
Jennifer M. Piscopo Occidental College

Abstract: *What does women's presence in political decision-making bodies signal to citizens? Do these signals differ based on the body's policy decisions? And do women and men respond to women's presence similarly? Though scholars have demonstrated the substantive and symbolic benefits of women's representation, little work has examined how women's presence affects citizens' perceptions of democratic legitimacy. We test the relationship between representation and legitimacy beliefs through survey experiments on a nationally representative sample of U.S. citizens. First, we find that women's equal presence legitimizes decisions that go against women's interests. We show suggestive evidence that this effect is particularly pronounced among men, who tend to hold less certain views on women's rights. Second, across decision outcomes and issue areas, women's equal presence legitimizes decision-making processes and confers institutional trust and acquiescence. These findings add new theoretical insights into how, when, and for whom inclusive representation increases perceptions of democratic legitimacy.*

Replication Materials: The data, code, and any additional materials required to replicate all analyses in this article are available on the *American Journal of Political Science* Dataverse within the Harvard Dataverse Network, at: <https://doi.org/10.7910/DVN/7190MT>

In 2017, newly inaugurated President Donald Trump sparked public outrage when he reinstated the global gag order on abortion funding while surrounded by only men. Opprobrium against groups of men making decisions concerning women is not a new phenomenon. Famously, protests erupted in 1991 when an all-male, all-white congressional committee interrogated Anita Hill—a black woman—about being sexually harassed. Nor is public outcry limited to cases that restrict women's rights. PayPal endured public shaming via social media in April 2016, when it organized a panel of “senior male leaders” to discuss pay equity.

That all-male panels confront scorn, especially when their topic addresses matters connected to women's

experiences, suggests that women's presence can affect how citizens view policy decisions and the institutions and processes that guide them. The backlash against all-male panels thus raises a central question for the study of democratic politics: Does the inclusion of representatives from historically underrepresented groups (typically called *descriptive representation*) legitimize decisions and decision-making procedures in the eyes of the general public? Democratic theorists argue that legislative outcomes, processes, and institutions cannot be legitimate when certain social groups remain systematically excluded from elected office (Dovi 2007; Mansbridge 1999; Phillips 1995). Despite these strong normative expectations, most research on symbolic representation—that is, the link

Amanda Clayton is Assistant Professor, Department of Political Science, Vanderbilt University, 230 Appleton Place, Nashville, TN 37203 (amanda.clayton@vanderbit.edu). Diana Z. O'Brien is Associate Professor, Department of Political Science, Texas A&M University, 2010 Allen Building, 4348 TAMU, College Station, TX 77843-4348 (dzobrien@tamu.edu). Jennifer M. Piscopo is Assistant Professor, Department of Politics, Occidental College, 1600 Campus Road, Los Angeles, CA 90041 (piscopo@oxy.edu).

We are grateful for feedback we received from Georgia Anderson-Nilsson, Lisa Baldez, Larry Bartels, Josh Clinton, Maria Escobar-Lemmon, Martin Gilens, Matthew Hayes, Nahomi Ichino, Cindy Kam, Dave Lewis, Kristin Michelitch, Michael Nelson, Efrén Perez, Kira Sanbonmatsu, Michelle Taylor-Robinson, Liz Zechmeister, and others. This work also benefited from comments provided by seminar participants at Birkbeck, University of London, Indiana University, the Research on Individuals, Politics, and Society (RIPS) lab at Vanderbilt University, the University of Illinois Urbana Champaign, and the University of California-Merced, as well as attendees at the 2016 Empirical Study of Gender (EGEN) conference at the University of Pennsylvania and the 2017 annual meeting of the Midwest Political Science Association in Chicago, IL. Funding for this work was generously provided by the Center on American Politics at Indiana University.

[Correction add on 10/08, after first online publication: A correction has been made in References]

American Journal of Political Science, Vol. 63, No. 1, January 2019, Pp. 113–129

between presence and citizens' attitudes and behavior—has focused on political engagement (Alexander 2012; Clayton 2015; Kittilson and Schwindt-Bayer 2012; Wolbrecht and Campbell 2007). Less work has investigated citizens' responses to the exclusion of marginalized social groups from decision-making bodies, especially when these institutions produce policies that affect the group's rights or well-being. Yet, this dimension of symbolic representation is crucial, as the legitimacy-conferring effects of descriptive representation are central to questions concerning regime stability and the acceptance (or rejection) of democratic political institutions (Atkeson and Carrillo 2007).

We explain whether, when, and for whom descriptive representation affects citizens' perceptions of democratic legitimacy. To begin, we consider the distinct ways in which descriptive representation influences citizens' attitudes toward different parts of the decision-making process. Because women's presence may signal whether the decision itself was correct or fair, descriptive representation may affect citizens' immediate reaction to the content of the decision reached (*substantive legitimacy*). Women's presence may also affect citizens' perceptions of the fairness of decision-making procedures, as captured through citizens' assessments of the decision-making process, acquiescence to the group's decisions, and trust in representative institutions (*procedural legitimacy*). We theorize that the effects of women's presence may differ across these two forms of legitimacy beliefs. Adding to existing scholarship, we further posit that the gender composition of the decision-making body interacts with the decision itself to shape citizens' perceptions of legitimacy. Because women's presence signals that women's interests were represented, we hypothesize that descriptive representation affects citizens' perceptions of substantive legitimacy more strongly when evaluating a political decision that harms women's rights. At the same time, we argue that descriptive representation communicates procedural legitimacy irrespective of the policy outcome.

Finally, we expect that the symbolic consequences of women's presence may differ for women (the historically marginalized group) as compared to men (the historically dominant group). Building on Mansbridge's (1999) notion of uncrystallized interests, we argue that women's presence sends stronger signals to men than it does to women about substantive legitimacy, as it informs men more than women about which outcome is "correct" for women as a group. In particular, we expect that women's inclusion in political decision making will affect men more strongly than women when political decisions roll back group rights. Our expectations are reversed for citizens' perceptions of procedural legitimacy.

Drawing on the broader literature on symbolic representation, which suggests that inclusion has particular meaning for the historically excluded group, we theorize that women's presence sends stronger signals about the legitimacy of decision-making processes and institutions to women than to men.

Contributing to the burgeoning body of experimental research on gender in U.S. politics (Bauer 2017; Kanthak and Woon 2015; Karpowitz, Mendelberg, and Shaker 2012; Klar 2018), we provide the first causal test examining how women's descriptive representation affects citizens' perceptions of democratic legitimacy. Our experiment varies the gender composition of a hypothetical legislative committee (all-male or gender-balanced) as well as the outcome reached (a decision that either expands or restricts women's rights). We fielded this survey experiment on a nationally representative sample in the 2016 Cooperative Congressional Election Survey and completed additional analyses using Amazon's Mechanical Turk service.

We find that citizens, both men and women, strongly prefer inclusion. Yet, we observe important differences depending on the measure of legitimacy (substantive or procedural), the decision the group reaches, and respondent gender. With respect to substantive legitimacy, women's presence does not affect the perceived legitimacy of decisions that expand women's rights, for either men or women respondents. Yet, women's presence on legislative committees does improve evaluations of decisions that harm women, and this effect appears to be somewhat stronger for men. Our work thus indicates that the inclusion of representatives from the marginalized group may only affect citizens' perceptions of a decision's content when group rights are rescinded. Moving to perceptions of procedural legitimacy, we find that citizens attach greater legitimacy to decision-making procedures when women are present, both when decisions expand and restrict group rights. Contrary to our expectations, we show that this result holds similarly for both women and men. We conclude by discussing the mixed normative implications of both sets of findings; on the one hand, that women's equal presence legitimizes anti-feminist outcomes, perhaps particularly for men, and, on the other hand, that women's equal presence confers legitimacy on democratic procedures more broadly.

The Symbolic Effects of Women in Politics

Does women's inclusion in political institutions affect citizens' perceptions of politics? This question is central not

only to the study of gender and politics, but also to scholarship on historically underrepresented groups more generally. The large body of normative and empirical work on this topic typically falls under the umbrella of “symbolic representation.” Though initially broadly defined as the meaning that a representative has for those being represented (Pitkin 1967), existing gender and politics scholarship on symbolic representation focuses primarily on citizens’ gender biases, attitudes toward traditional gender roles, and political engagement. Women citizens are more likely than men to want to be represented by women (Sanbonmatsu 2002), for example, and women who are represented by women tend to offer more positive evaluations of their members of Congress (Lawless 2004). Descriptive representation can further enhance women’s beliefs in their own ability to govern (Alexander 2012) and decrease citizens’ implicit biases against women leaders (Beaman et al. 2009), particularly among young women (Clayton 2018). Women likewise appear to be more politically engaged when represented by women (Atkeson and Carrillo 2007; Kittilson and Schwindt-Bayer 2012; Wolbrecht and Campbell 2007), though American politics research on this front remains more mixed (e.g., Lawless 2004).

The opprobrium visited on all-male panels suggests that the exclusion (or inclusion) of marginalized groups has further consequences for how citizens view policy decisions and political institutions. Indeed, symbolic representation captures the “represented’s feelings of being fairly and effectively represented” (Schwindt-Bayer and Mishler 2005, 407), and a body of survey research has explicitly examined this dimension of representation. This work suggests that citizens view governments as more democratic when women are represented in elected office (Schwindt-Bayer and Mishler 2005). Stauffer (2018), for example, shows a strong correlation between beliefs about women’s descriptive representation and citizens’ views of government effectiveness and responsiveness. Yet, empirical gender and politics scholarship has not directly addressed the relationship between women’s presence and beliefs about the legitimacy of decisions, decision-making procedures, and decision-making institutions.

Though this core dimension of symbolic representation remains understudied in the empirical gender and politics scholarship, political scientists across subfields are deeply concerned with whether and how political systems foster legitimacy. Here, we conceptualize perceptions of democratic legitimacy as the popular belief among citizens that their government is acting competently, impartially, and in the service of the entire population (see Levi, Sacks, and Tyler 2009). Democratic theorists often associate legitimacy with procedures, such as electing

representatives through territorially based constituencies and establishing protocols for fair deliberation and debate (Rehfeld 2005; Urbinati and Warren 2008; Walsh 2010). In particular, fair procedures build up a reservoir of goodwill that endures past an individual outcome (Easton 1965). This work, for example, has inspired a vibrant body of scholarship asking whether variations in democratic procedures affect assessments of procedural fairness and decision acceptance. In particular, scholars ask whether and how citizen participation in (or influence over) public decision making—such as through direct voting, expert decision making, and deliberation—shapes these beliefs (e.g., De Fine Licht et al. 2014; Esaiasson, Gilljam, and Persson 2012; Esaiasson et al. 2016; Lind and Tyler 1988; Persson, Esaiasson, and Gilljam 2013).

Theories of symbolic representation are broadly related to procedural fairness scholarship in their conceptualizations of citizens’ legitimacy beliefs (in particular, work by Tyler [2006] and others). At the same time, the symbolic representation literature draws our attention beyond decision rules to further ask how the political inclusion or exclusion of historically marginalized groups affects citizens’ perceptions of legitimacy. To theorize about group representation, we draw from the work of political theorists specifically interested in marginalized groups, who argue that without a diverse range of perspectives, both political procedures and outcomes can lose legitimacy (Mansbridge 1999; Phillips 1995). Mansbridge (1999, 634) argues that the aggregative function of democracy “aims at producing some form of relatively legitimate decision,” but such outcomes cannot be obtained when certain groups remain systematically unheard or ignored. In particular, excluding women and ethnic minorities from political representation places these groups outside the political order, conveying their lack of full citizenship (Celis and Mazur 2012). Those concerned with legislative inclusion argue that decisions made about groups, but in the absence of group representatives capable of articulating group interests, are fundamentally unjust and may undermine the legitimacy of otherwise democratic institutions (Arnesen and Peters 2017; Mansbridge 1999).

Work on symbolic representation by scholars of race and ethnic politics supports the notion that the marginalized group’s presence cues the legitimacy of both outcomes and procedures. Black citizens perceive police officers’ actions as more legitimate when Black officers are present (Theobald and Haider-Markel 2008). Minorities who see themselves represented in decision-making bodies likewise view their institutions as more responsive (Banducci, Donovan, and Karp 2004) and their procedures as more just (Hayes and Hibbing 2017). African

Americans, for example, perceive the federal courts as more legitimate as the proportion of African American judges rises, though Whites reduce their support for courts under these conditions (Scherer and Curry 2010).

Taken together, normative and empirical scholarship indicates that representative diversity affects perceptions of legitimacy. Yet, gender and politics scholars have not clearly established whether presence conveys legitimacy, or under what circumstances and for whom. Further, the work touching on this topic has been either theoretical or based on observational research. We add to this line of inquiry by offering causal evidence explicating how citizens' perceptions of legitimacy vary based not only on the inclusion/exclusion of women in decision-making bodies, but also on the substance of deliberative decisions and the identity of the observer.

How, When, and for Whom Women's Presence Conveys Legitimacy

We draw together insights from these diverse theoretical and empirical literatures to build a theory about how, when, and for whom the inclusion of women conveys democratic legitimacy. We begin by taking seriously normative and empirical scholars' views that legitimacy attaches to both decisions and procedures, and posit that women's presence affects both substantive and procedural legitimacy. Yet, these perceptions of legitimacy vary based on the decision reached (feminist or anti-feminist) and the identity of the observer (members of the dominant or marginalized group).

Women's involvement in decision making may legitimize political decisions and political procedures for different reasons. First, with respect to substantive legitimacy, *women's presence may communicate that substantive representation has occurred*—that is, that women's interests were manifested during the policy process. Thus, women's presence signals that a political decision treats women as a group justly. Second, with respect to procedural legitimacy, *women's presence may symbolize that institutions are working properly or effectively*. Institutions may occasionally make bad decisions, but fair procedures build up a reservoir of goodwill that endures past an individual outcome (Easton 1965; Hayes and Hibbing 2017; Scherer and Curry 2010). This reservoir of goodwill speaks to the diffuse or long-term nature of procedural legitimacy. Thus, women's presence may contribute to institutional trust and acquiescence irrespective of whether any singular decision harms or benefits group rights.

We expect that women's presence will raise perceptions of both substantive and procedural legitimacy. At the same time, citizens' legitimacy beliefs are likely conditioned on the policy decision reached. Outrage over all-male panels when they limit access to abortion or question the truth of sexual harassment claims (as in our motivating examples) indicates the strength of the descriptive–substantive link in the popular imagination. On issues dealing explicitly with women's rights, women's absence implies the exclusion of the perspectives of the very citizens who have the most at stake with respect to the group's decision. Conversely, when group representatives are included, their presence signals that their interests were part of the deliberations—even if the ultimate decision does not expand group rights. *Women's inclusion may thus enhance the substantive legitimacy of anti-feminist decisions.*

We have less reason to suspect that women's presence will affect the substantive legitimacy of feminist decisions. When members of the dominant group “get it right”—that is, when they expand the rights of historically marginalized groups—we expect that an inclusive decision-making body will be less consequential in the eyes of citizens. In this case, when a group of men makes a decision that expands women's rights, citizens assume women's interests were represented even in their absence. We therefore expect that *women's presence will grant substantive legitimacy to anti-feminist decisions, but will not affect citizens' assessment of feminist decisions.*

At the same time, we anticipate that irrespective of the policy decision reached, women's presence enhances citizens' judgments vis-à-vis procedural legitimacy. Procedural legitimacy reflects the belief that decision-making procedures and institutions are working properly. Even if citizens view a feminist policy outcome made by an all-male group as legitimate, they may not wholly support the way in which the decision was made. Indeed, because we expect decision-making bodies to tackle other issues in the future, we may prefer women to be present in those deliberations irrespective of the outcome reached in a single case. We thus posit that *women's presence will grant procedural legitimacy to both anti-feminist and feminist decisions.*

Beyond the outcome of the decision, we theorize that presence conveys legitimacy differently for the dominant and the marginalized group. With respect to substantive legitimacy, the strong link between women's descriptive and substantive representation in the public imagination suggests that citizens believe women representatives are advocating for women's interests during group deliberations. At the same time, men and women sometimes differ in their reliance on cues sent by politicians' gender

(Bauer 2015). In this case, it is not women citizens who require information about issues like sexual harassment, as women's positions tend to be more crystallized on issues related to women. Even if women disagree on the appropriate policy solution, they are more likely than men to have considered whether and how these issues matter for them. That is, women tend to have feelings on what is "right" for women, even in the absence of women decision makers (Mansbridge 1999). Men, in contrast, tend to be less sure about the "correct" outcome for women. When an issue is not well established on the political agenda, women's presence signals to men that women have participated in the deliberation and that the resultant decision is therefore "right" for women. In other words, women's preferences over outcomes may be more fixed.

The expectation that descriptive representation may send a stronger signal to men than women is consistent with findings from procedural fairness scholars. This work shows that individuals with stronger preexisting policy preferences and group attachments are less moved by variations in legislative features. For instance, Esaiasson (2010) demonstrates that good behavior on the part of a public official increases decision acceptance, but that this effect is attenuated among citizens who are morally disappointed in the outcome (see also Mullen and Skitka 2006; Skitka and Mullen 2008). Relatedly, Leung, Tong, and Lind (2007) find that among citizens with strong national attachments, procedural justice matters less—and collective outcome favorability matters more—in determining support for government policies. In line with this research, we theorize that those with stronger preexisting preferences and group attachments (in our case, women) will be less moved by variations in legislative features. Therefore, *with respect to substantive legitimacy, the presence of women decision makers provides a more meaningful cue to men than to women.*

Though descriptive representation sends a stronger signal to men about policy outcomes, we expect that it offers a stronger cue to women with respect to decision-making procedures. Previous literature has identified the symbolic benefits of women's presence in political institutions, particularly for women citizens (e.g. Alexander 2012; Wolbrecht and Campbell 2007). Indeed, Atkeson and Carrillo (2007, 79) find that women's presence in legislatures has "important benefits to a democratic society" because "higher levels of collective female descriptive representation promote higher values of external efficacy for female citizens," though, strikingly, not for men. Thus, we expect that *women's presence will signal greater procedural legitimacy to women than to men.*

Experimental Design: Measuring the Causal Effects of Presence on Legitimacy Beliefs

We test our theory using a series of survey experiments, which allow us to control both the makeup of the decision-making body and the outcome reached. We can thus isolate the effects of women's descriptive representation when the decision either expands or retracts protections for women as a group. This provides a distinct advantage over observational work. Electorates who choose high levels of descriptive representation systematically differ from those who exclude women, making it difficult to make causal claims based on observational group differences. Additionally, because women's representation is likely correlated with progressive policies toward women's rights, it is difficult to observe the legitimacy-conferring effects of descriptive representation when legislative bodies harm group interests. Our experimental design avoids these concerns.

Our design varies two treatment dimensions. The vignettes we use modify the gender composition of an eight-member state legislative committee to consist of either only men or four men and four women. We also vary the type of outcome the group reaches: increasing or decreasing penalties for workplace sexual harassment. Of course, given the diversity among women as a group, not all women will support the feminist outcome of increasing penalties for offenders (Celis and Childs 2012). At the same time, increasing penalties is consistent with feminism's commitment to protect women from abuse. Further, focusing on sexual harassment allows us to test how women's presence affects substantive and procedural legitimacy on an issue that, at the time, was not strongly identified with the major parties' platforms. Our theory suggests the presence of marginalized groups should be especially important when the policy is uncrystallized on the political agenda.

In our main series of experimental vignettes, respondents read a newspaper article describing the gender composition of the committee and the decision reached. The mocked-up article includes a headline and photos of the eight legislators (see Appendix 1 in the supporting information [SI]). Because legislative groups always have a gender composition (i.e., there are no genderless committees), we opt not to include a control condition that provides no information about the gender makeup of the panel. Due to men's overrepresentation in state and national legislatures, a prompt that did not signal the gender composition of the group would likely elicit assumptions of predominantly male membership, especially when

lawmakers rescind group rights. We test this assumption through a separate experiment conducted via Amazon's Mechanical Turk service. When we do not note the gender composition of the panel, respondents believe that an eight-member committee making an anti-feminist decision has 1.6 women members on average, and that an eight-member committee making a feminist decision has four women on average ($n = 150$, two-tailed t -test, difference significant at $p \leq .001$). As we cannot use a genderless prompt to test whether gender balance cues legitimacy or whether male dominance rescinds it, we remain agnostic empirically and refer to our results in both ways.

We collect data from two sources. First, we randomized our four treatment conditions and included our main response questions on the Cooperative Congressional Election Study (CCES), a nationally representative survey with a stratified sample of U.S. adult citizens administered by the sampling firm YouGov. Our survey experiment was fielded through CCES to 847 respondents in the post-election wave in November 2016.¹

Second, we simultaneously ran the same survey experiment with an additional 993 adult U.S. citizens via Amazon's survey platform, Mechanical Turk (MTurk).² We removed respondents with missing data on our variables of interest and those who failed to pass our manipulation checks (correctly identifying the number of women on the committee and the decision reached), reducing our sample size to 881 respondents.³ Our MTurk survey included a more extensive battery of questions related to our outcomes of interest as well as supplementary questions on respondents' political attitudes and behavior. We also used MTurk to test additional treatment conditions in order to evaluate our theoretically motivated scope conditions and to run several robustness checks of our main results.

¹The pre-election wave was fielded to 1,000 respondents, but due to attrition, the full post-election wave was fielded to 847 respondents.

²Between November 20 and November 23, 2016, we recruited U.S. residents to participate in our experiment who were over the age of 18, whose prior approval rate for previous MTurk work exceeded 98%, and who had completed over 1,000 previous MTurk tasks. The survey took, on average, 9 minutes, and we paid participants \$1.30 for their completed response.

³To maintain consistency across data sets, we list-wise deleted observations with missing values on our variables of interest: our battery of legitimacy questions, respondent gender, partisan identification, and opinion on preventing sexual harassment. We exclude 60 observations because of missing values and 52 observations for failing to pass the manipulation checks. Our results are robust and substantively similar when we include respondents who did not pass the manipulation checks. We report these intent-to-treat (ITT) effects in Appendix 9 in the SI. Ability to pass the manipulation checks is not systematic across treatment conditions (chi-squared test, p -value = .262). Missingness is also not systematic across conditions (chi-squared test, p -value = .961).

We remove observations with missing data from the CCES sample, resulting in a sample size of 827.⁴ Our balance diagnostics reveal no statistical differences on observable characteristics—such as partisanship and a range of demographic variables—across treatment conditions in both the CCES and MTurk samples (see balance tables in Appendix 2 in the SI). Comparing the two samples, we observe that the MTurk sample has respondents who are younger, more educated, more likely to be white, and more likely to be Democrats than the CCES sample. Our ability to test for treatment effects across the two samples strengthens our confidence in the generalizability of our findings.

Experimental Results

We first present our results related to citizens' perceptions of substantive legitimacy and then turn to procedural legitimacy.⁵ In both cases, we begin by considering men and women respondents together and then examine heterogeneous treatment effects by respondent gender. As a theoretically motivated scope condition, we then compare our main results to a series of treatment conditions using a non-gendered policy area. Together, our results demonstrate that women's descriptive representation enhances citizens' legitimacy beliefs. Importantly, however, they also reveal that substantive effect sizes vary based on the measure of legitimacy, the decision the group reaches, and respondent gender.

Perceptions of Substantive Legitimacy

We begin by investigating how the panel's gender composition affects respondents' perceptions of the decision reached. We measure substantive legitimacy beliefs by prompting respondents to consider the decision's rightness and fairness for the polity broadly as well as for women specifically. We measure these perceptions through the following statements and question:

The committee made the right decision for all the state's citizens.

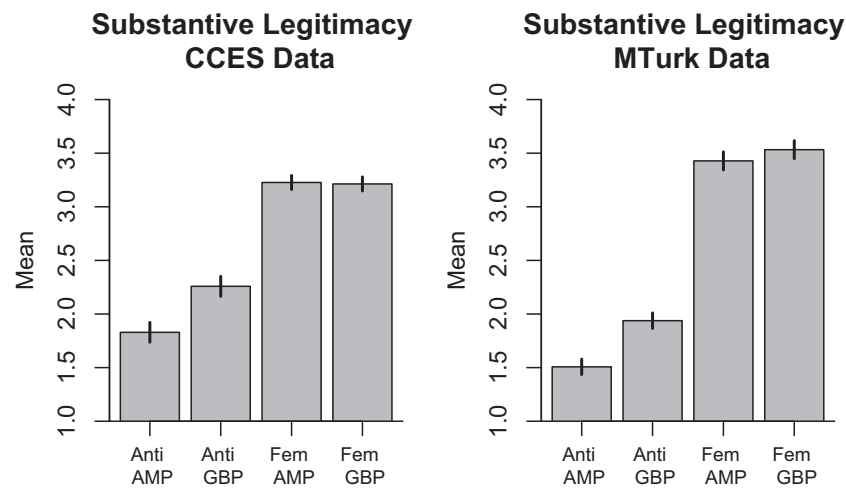
The committee made the right decision for women.

How fair was this decision to women?

⁴Missingness is also not systematic across treatment conditions for the CCES sample (chi-squared test, p -value = .491).

⁵In Appendix 3 in the SI, we conduct an empirical test to verify that substantive and procedural legitimacy are empirically distinct concepts. Moreover, the significant treatment effects we present also hold when we examine each constituent question separately (see Appendix 4 in the SI).

FIGURE 1 Perceptions of Substantive Legitimacy



Note: Error bars at 95% confidence intervals.

TABLE 1 Respondents’ Perceptions of Substantive Legitimacy

	Anti-Fem. All-Male Mean	Anti-Fem. Balanced Mean	ATE (95% CI)	p-value (two- tailed)	Feminist All-Male Mean	Feminist Balanced Mean	ATE (95% CI)	p-value (two- tailed)
CCES								
Substantive Legitimacy Scale	1.829	2.258	0.429 [0.246, 0.611]	≤ .001	3.227	3.213	0.014 [-0.142, 0.115]	.835
MTurk								
Substantive Legitimacy Scale	1.507	1.938	0.431 [0.283, 0.578]	≤ .001	3.428	3.533	0.105 [-0.028, 0.237]	.122

Note: Group means and differences by treatment condition. n = 827 CCES respondents; n = 881 MTurk respondents. ATEs with significance of p < .05 are indicated in boldface.

All question responses are on 4-point Likert-type scales, with higher values indicating greater levels of agreement or perceived fairness. As noted above, because these responses are highly correlated and load together onto a single factor, we generate a composite score of substantive legitimacy, which closely mirrors the 1–4 range of the individual response questions.⁶

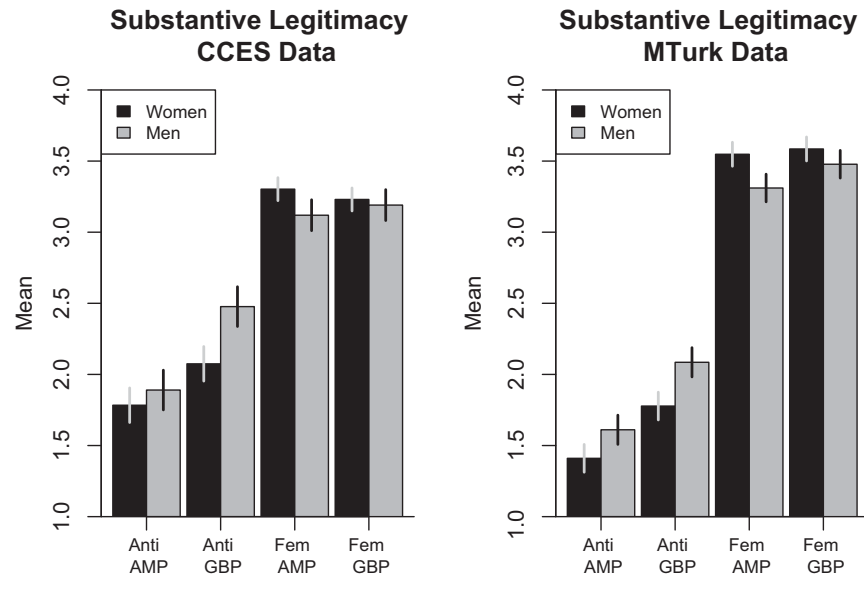
Figure 1 shows our results from the CCES sample and the MTurk sample, respectively, for each treatment condition: an anti-feminist decision made by an all-male panel (labeled “Anti AMP”), an anti-feminist decision

made by a gender-balanced panel (“Anti GBP”), a feminist decision made by an all-male panel (“Fem AMP”), and a feminist decision made by a gender-balanced panel (“Fem GBP”). Although CCES respondents show slightly more support for the anti-feminist outcome in general (as expected, given the composition of the samples), our average treatment effects (ATEs) across the two samples are very similar. Table 1 displays group averages for each treatment condition. Differences in group means between the all-male panel and the gender-balanced panel can be interpreted as the average treatment effect of women’s equal presence for both the anti-feminist decision and the feminist decision, respectively.

Figure 1 and Table 1 reveal two important findings. First, they show that gender balance improves perceptions of substantive legitimacy when the panel reaches an anti-feminist decision. This finding is replicated in both the CCES and MTurk samples, and the effect sizes are

⁶Due to space constraints, the CCES data exclude the first question, and we create a composite score from the second and third questions only. Scales in both samples are highly reliable. For the MTurk sample: Cronbach’s $\alpha = .96$. The individual item factor loadings land in a narrow range, from 0.92 to 0.97. For the CCES sample: Cronbach’s $\alpha = .93$. Individual loadings of the two response questions are 0.94 each. Correlation matrices across outcome measures are included in Appendix 2 in the SI.

FIGURE 2 Perceptions of Substantive Legitimacy by Respondent Gender



Note: Error bars at 95% confidence intervals.

remarkably similar in both. The effect size of women’s equal presence (as compared to an all-male panel) is a 0.43-point increase on the 4-point substantive legitimacy scale. This difference is approximately one-half of a standard deviation, an effect size typically characterized as moderate in the experimental literature (Cohen 1992). Second, the results indicate that when the panel reaches a feminist decision, the gender composition of the committee does not affect respondents’ perceptions of substantive legitimacy. This result holds in both the CCES and MTurk samples. That our findings are so consistent suggests that the MTurk sample behaves similarly to the nationally representative CCES sample (see Berinsky, Huber, and Lenz 2012).

We also separately evaluated whether respondents’ perceptions of substantive legitimacy are related to their level of policy support. We asked:

Now turning to your own personal opinion about this decision, do you agree or disagree with the following statement: Personally, I think the committee made the right decision.

Personal agreement is highly correlated with our substantive legitimacy scale ($r = .87$ for the CCES sample) and loads onto the same factor (see Appendix 3 in the SI). Importantly, we find that policy support also responds to women’s equal presence. Respondents express higher levels of agreement with anti-feminist outcomes

reached by a gender-balanced panel. Again, women’s representation increases personal agreement when the panel restricts women’s rights, but not when it expands them. Thus, although citizens may be expected to hold relatively fixed preferences on sexual harassment policy, for at least some respondents, women’s presence affects their level of support.

Gender Differences in Perceptions of Substantive Legitimacy

We find that gender balance only cues substantive legitimacy for anti-feminist outcomes. These results raise questions about which groups may be especially affected by women’s presence. We theorized that respondent gender should interact with perceptions of legitimacy because the dominant group (men) depends more on the cues sent by women’s presence than the marginalized group (women). Figure 2 illustrates mean responses to our substantive legitimacy scale for men and women respondents. Table 2 displays the associated conditional average treatment effects (CATEs) for both samples.

Across both samples, we see that gender balance significantly improves both men’s and women’s perceptions of substantive legitimacy when the group reaches an anti-feminist decision. At the same time, we observe larger treatment effects for men. In the CCES sample, men have an effect size twice the magnitude of women’s. We do note,

TABLE 2 Respondents' Perceptions of Substantive Legitimacy

	Anti-Fem All-Male Mean	Anti-Fem Balanced Mean	CATE (95% CI)	p-value (two- tailed)	Feminist All-Male Mean	Feminist Balanced Mean	CATE (95% CI)	p-value (two- tailed)
Men (CCES)								
Substantive Legitimacy Scale	1.890	2.477	0.587 [0.311, 0.862]	≤ .001	3.119	3.191	0.071 [−0.143, 0.285]	.513
Men (MTurk)								
Substantive Legitimacy Scale	1.612	2.085	0.474 [0.265, 0.684]	≤ .001	3.311	3.478	0.167 [−0.033, 0.367]	.101
Women (CCES)								
Substantive Legitimacy Scale	1.783	2.074	0.291 [0.051, 0.531]	.018	3.302	3.230	−0.072 [−0.232, 0.087]	.372
Women (MTurk)								
Substantive Legitimacy Scale	1.410	1.778	0.367 [0.163, 0.572]	≤ .001	3.548	3.585	0.037 [−0.135, 0.209]	.671

Note: Group means and differences by treatment condition and respondent gender. $n = 357$ (CCES) and 440 (MTurk) for men respondents and $n = 470$ (CCES) and 441 (MTurk) for women respondents. CATEs with significance of $p < .05$ are indicated in boldface.

however, that when we interact the treatment—gender balance on the panel—with respondent gender, we find that the difference in effect sizes between men and women fails to reach conventional statistical significance levels ($p = .11$; see SI Table 5 in Appendix 5). Similarly, though we also observe a larger effect size for men than for women in the MTurk sample, the difference between the two groups is substantively smaller than in the CCES sample and is not statistically significant ($p = .44$; see SI Table 6 in Appendix 5). When we combine the samples and control for the respondent pool (CCES or MTurk), the interaction between the treatment and respondent gender also does not reach conventional statistical significance levels ($p = .09$; see SI Table 7 in Appendix 5). We thus treat the possibility of a heightened effect among men as suggestive in the following discussion, and we believe that future work should further interrogate the possibility of significant gender differences in effect sizes.

Though just outside conventional statistical significance, the substantive differences in effect sizes suggest that the dominant group may indeed be more moveable when a decision rolls back the rights of marginalized groups. Above, we theorized that effects would be larger for men because men have less crystallized thoughts about sexual harassment. Notwithstanding the recent media attention that surrounds high-profile men accused of harassment, policies addressing sexual harassment have traditionally been marginal to the policy agenda, and men are less likely to have been compelled to form a strong opinion about the issue. Women,

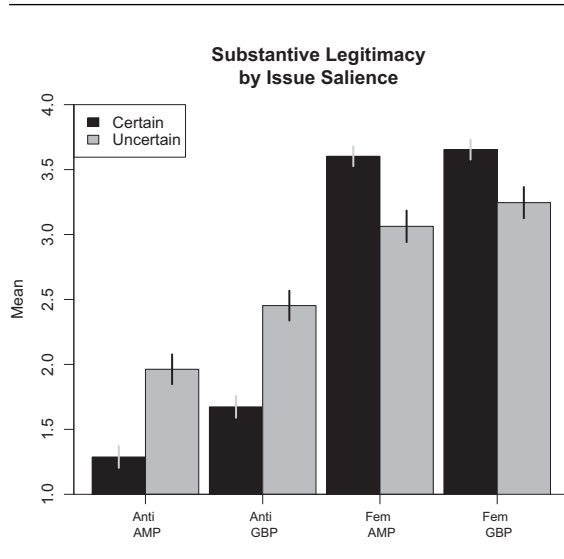
in contrast, are more likely to have crystallized views about sexual harassment, suggesting that women's presence is less relevant in shaping their perceptions of the outcome.

To measure this expectation, we posit that the crystallization of citizens' views on the issue of sexual harassment—that is, how much they have considered its personal and political relevance—will correspond with whether they have firm feelings about the importance of addressing the issue. Measured in this way, our data support our expectation concerning the relative crystallization of men's versus women's views. In our MTurk sample, we ask respondents whether they agree that preventing sexual harassment is an important issue. Women are significantly more likely than men to strongly agree: 75% compared to 55%.⁷

To test whether the average treatment effects above are moderated by the degree to which respondents have crystallized views about sexual harassment, we split our sample based on responses to this question of issue salience. We place all respondents who either strongly agree or strongly disagree that sexual harassment is an important issue into the "certain" group, as their feelings on the issue are fixed, albeit in different directions ($n = 598$). Respondents who either somewhat agree or somewhat disagree make up the "uncertain" group

⁷This question was asked at the end of the survey, several question blocks below the treatment. Responses to the question of sexual harassment salience are not statistically differentiable across treatment conditions ($p = 1$). See SI Table 2 in Appendix 2.

FIGURE 3 Perceptions of Substantive Legitimacy by Respondents' Attitudes toward Sexual Harassment (MTurk Sample)



Note: Error bars indicate 95% confidence intervals.

($n = 283$). We then compare the conditional average treatment effects for these two groups, as shown in Figure 3.

As expected, though we continue to observe significant treatment effects for both sets of respondents when the panel reaches an anti-feminist decision, the magnitude of the effect size is greater for those with less certain feelings about sexual harassment. The group with more crystallized views is less moved. We observe a conditional average treatment effect of 0.388 points on our substantive legitimacy scale for those who feel certain about the issue of sexual harassment prevention and a CATE of half of a point (0.490) on our substantive legitimacy scale for those who feel less certain. We do find, however, that when we interact the treatment of gender balance with respondent certainty on the issue of sexual harassment, the interaction term does not reach statistical significance ($p = .461$). As above, we view these differences in effect size as suggestive.

In combination, our analyses suggest that respondents who have weaker feelings about preventing sexual harassment are more persuaded that an outcome is just when women are present in decision making. We theorized that because men tend to have less crystallized views on sexual harassment, the effect is particularly strong for the dominant group. Our findings add tentative support to this theoretical expectation and provide a foundation for subsequent research on variable perceptions of legitimacy based on dominant versus marginalized group membership.

Perceptions of Procedural Legitimacy

Beyond citizens' perceptions of policy outcomes, we are also interested in how women's presence affects citizens' views of processes and institutions. We measure procedural legitimacy through respondents' direct assessment of procedural fairness, as well as their institutional trust and acquiescence. Scholars have used these measures to capture the diffuse consequences of descriptive representation on public attitudes, as these perceptions relate to attitudes beyond the decision itself (Hayes and Hibbing 2017; Scherer and Curry 2010). We measure citizens' feelings about procedural legitimacy through responses to the following question and statements, each measured on a 4-point Likert-type scale:

Thinking for a moment about the gender composition of the committee, how fair was the decision-making process?

Thinking about the gender composition of the committee, the committee's decision should be overturned. (Reverse coded.)

Thinking about the gender composition of the committee, the committee can be trusted to make decisions that are right for the state's citizens.

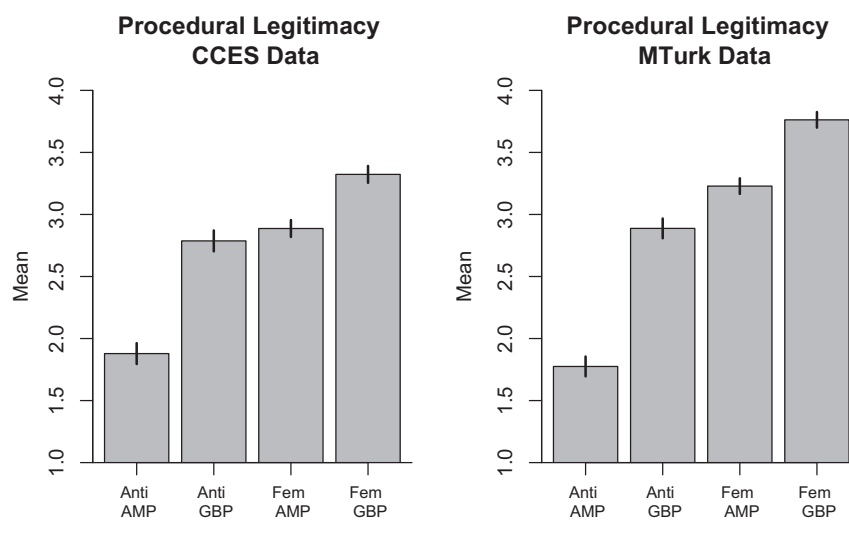
The State Legislature can be trusted to make decisions that are right for the state.

Because these measures load onto a single factor, we generate a single procedural legitimacy scale, which closely mirrors the 1–4 range of the individual response questions.⁸ Average treatment effects for individual response questions, which also achieve statistical significance separately, are included in Appendix 4 in the SI. Figure 4 shows results from both the MTurk and CCES samples, and Table 3 displays group means and differences. Again, we find very similar treatment effect sizes across samples.

Gender balance significantly and substantially improves perceptions of procedural legitimacy for both the anti-feminist and feminist outcomes. Respondents who received the gender-balanced anti-feminist treatment scored 0.91 points higher on the 4-point procedural legitimacy scale in the CCES sample and 1.11 points higher in the MTurk sample, as compared to those who received the all-male anti-feminist treatment (a very large

⁸Due to space constraints, we only include the first three questions on the CCES survey, and we use these measures to construct our scale for our CCES respondents. Again, both scales are reliable. MTurk sample: Cronbach's $\alpha = .92$. The individual item factor loadings land in a narrow range, from 0.80 to 0.91. CCES sample: Cronbach's $\alpha = 0.73$, and individual factor loadings are each 0.76.

FIGURE 4 Perceptions of Procedural Legitimacy



Note: Error bars at 95% confidence intervals.

TABLE 3 Procedural Legitimacy: Group Means and Differences by Treatment Condition

	Anti-Fem All-Male Mean	Anti-Fem Balanced Mean	ATE (95% CI)	p-value (two-tailed)	Feminist All-Male Mean	Feminist Balanced Mean	ATE (95% CI)	p-value (two-tailed)
CCES								
Procedural Legitimacy Scale	1.879	2.787	0.908 [0.742, 1.074]	≤ .001	2.887	3.323	0.436 [0.302, 0.570]	≤ .001
MTurk								
Procedural Legitimacy Scale	1.775	2.888	1.113 [0.956, 1.270]	≤ .001	3.229	3.763	0.534 [0.410, 0.658]	≤ .001

Note: n = 827 CCES respondents; n = 881 MTurk respondents. ATEs with significance of p < .05 are indicated in boldface.

standardized effect size of 1.1 and 1.5 standard deviations, respectively). Here, we also observe treatment effects when respondents read about a group making a feminist decision. Respondents who viewed the gender-balanced feminist decision scored 0.44 points higher on the procedural legitimacy scale in the CCES sample (and 0.53 points higher in the MTurk sample) than those who received the all-male feminist condition (an improvement of 0.56 and 0.72 standard deviations, respectively). We thus observe meaningful treatment effects for both decision conditions, and especially strong effects when the group reaches an anti-feminist decision.

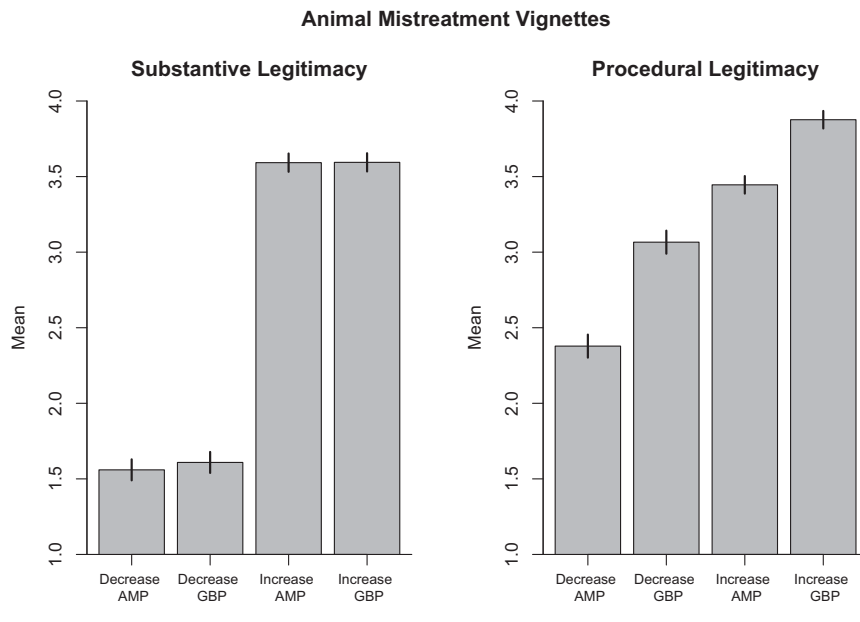
Gender Differences in Perceptions of Procedural Legitimacy

As above, we are interested in heterogeneous treatment effects by respondent gender. We theorized that women’s

presence would confer legitimacy to decision-making procedures to a greater degree for women than men. Contrary to our expectations, we find statistically significant and substantively large effects for both men and women, indicating that both groups prefer inclusion when thinking about how decisions are made (see Appendix 6 in the SI). Yet, looking at the magnitude of treatment effects, this is the only area in which our two data sources show slightly different patterns. In the CCES data, gender balance affects men’s and women’s perceptions of procedural legitimacy similarly, whereas our MTurk data suggest that women are more strongly affected than are men.⁹

⁹We are agnostic about why the two data sets differ on this finding. It is possible that this mixed result reflects differences in the partisan composition of the two samples. The MTurk sample skews Democratic, and it may be that Democratic women are more susceptible to the legitimacy-conferring effects of women’s presence vis-à-vis procedural legitimacy, particularly for the feminist

FIGURE 5 Perceptions of Substantive and Procedural Legitimacy: Non-Gendered Issue Area (MTurk Sample)



Note: Error bars at 95% confidence intervals.

Women's Presence and Legitimacy in Non-Gendered Issues

Normative theorists posit that descriptive representation matters not only for giving voice to women's interests, but also for signaling that governments are more inclusive and therefore more legitimate (Dovi 2007; Mansbridge 1999). If citizens perceive women's equal presence as more just, then women's inclusion may matter for perceptions of procedural legitimacy even when group rights are not at stake. Yet, women's presence should have no effect on beliefs about the substantive legitimacy of these decisions.

As a scope condition, we alter the vignette to consider whether women's presence shapes legitimacy beliefs in a non-gendered issue area. After extensive pretesting, we chose a scenario related to the mistreatment of animals on commercial farms.¹⁰ We closely mirrored our

outcome, than are Republican women. Yet, when we run a formal difference-in-difference test of different effect sizes by gender in the MTurk sample, the interaction term does not reach statistical significance for the anti-feminist outcome ($p = .204$), and it is just barely outside the traditional threshold for statistical significance for the feminist outcome ($p = .055$). Gender differences in effect size in the MTurk sample are therefore only suggestive.

¹⁰We pretested several control scenarios, including texting while driving, workplace bullying, and issues around homelessness. Farm animal mistreatment had balance with our main sexual harassment scenario on several key dimensions, namely, perceived liberal/conservative divide, issue salience, and potential to agree or disagree with the outcome.

sexual harassment scenarios, with the committee again either increasing or decreasing penalties for offenders. The wording of our response questions is identical, except we replace the question that asks "the committee made the right decision for women" with "the committee made the right decision for the treatment of animals." These treatment vignettes were shown to a separate respondent pool.

Figure 5 displays the results from our MTurk sample. As expected, gender balance does not affect citizens' perceptions of substantive legitimacy when the policy issue does not address women's rights. When comparing the all-male and gender-balanced conditions, we observe no difference in perceptions of whether the decision was right or fair. Importantly, however, gender balance continues to improve citizens' perceptions of procedural legitimacy. These findings support claims that inclusion matters for broader reasons of justice. Women's equal presence affects citizens' perceptions of procedural legitimacy even when women's rights are not at stake.

Robustness Checks

We conduct a series of robustness checks to eliminate three potential alternative explanations for our treatment effects. First, we address the concern that the panel's gender composition sends partisan cues—that is, reactions to

all-male panels may actually reflect Democratic respondents' distaste for what they perceive as an all-Republican panel. Second, we run three checks to examine whether priming or respondent social desirability bias affects our results. Third, we reran our survey in July 2017 to assess whether treatment effects were specific to the timing of the original survey immediately after the 2016 presidential election. Our results replicate across specifications.

Perceptions of Partisanship

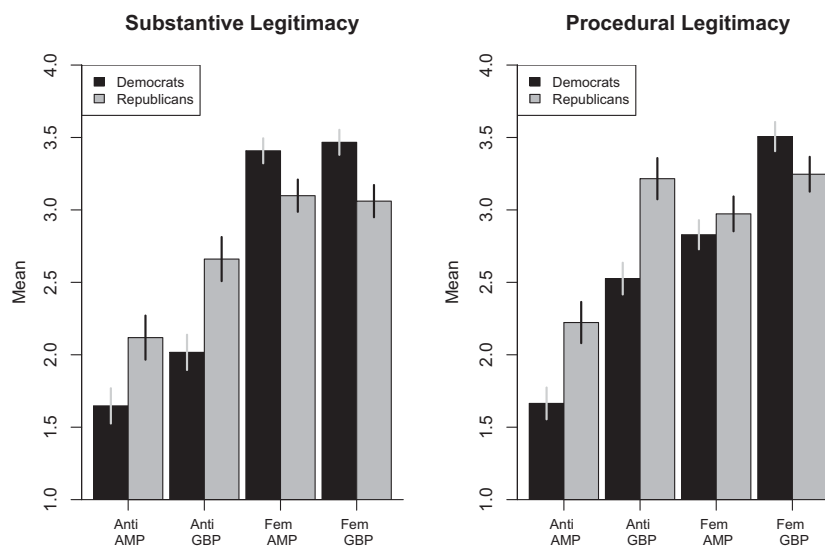
Our vignette explicitly notes that the legislative committee is bipartisan. Nonetheless, women are better represented within the Democratic Party and are often viewed as more left-leaning than men (King and Matland 2003). Respondents may therefore view our all-male panel as an all-Republican panel and our gender-balanced panel as having more Democratic members. In this case, opprobrium against all-male panels may actually reflect Democratic respondents' distaste for what they perceive as an all-Republican panel. Our treatment effects would therefore be driven by the Democratic respondents in our sample, who are responding to partisanship cues and not to women's presence per se.

To address these concerns, we test for conditional average treatment effects by partisan identification. We split our CCES data into those who self-identify as Republicans and those who self-identify as Democrats (including strong partisans and partisan leaners). Figure 6 shows conditional average treatment effects for both

groups. Our results hold among both Republicans and Democrats (n = 399 self-identified Democrats; n = 282 self-identified Republicans). Across party identification, respondents on average rate anti-feminist outcomes as less substantively legitimate when made by an all-male panel. They also report lower average perceptions of procedural legitimacy for both anti-feminist and feminist decisions reached by an all-male panel. That our main findings hold among members of both parties suggests that our results are not driven by perceptions of the panel's partisanship.

Though the results hold for both Republicans and Democrats, there are some notable partisan differences in effect size. First, in the feminist condition, we find significantly larger effect sizes for Democratic than Republican respondents with respect to procedural legitimacy (a CATE of 0.678 for Democrats vs. 0.273 for Republicans, a difference in effect size significant at p = .01). That is, compared to Republicans, Democrats attach greater procedural importance to women's presence in political decision making when the group makes a feminist decision. Interestingly, we do not see the same pattern in the anti-feminist condition. To the contrary, if there is any partisan effect, it is that Republican respondents may be more influenced by gender balance than Democrats. Republicans have a CATE size of 0.993, and Democratic respondents have a CATE size of 0.861 on the 4-point procedural legitimacy scale, though this difference in effect size is not statistically significant (p = .46). With respect to substantive legitimacy, we also find

FIGURE 6 Perceptions of Substantive and Procedural Legitimacy by Respondent Party (CCES Sample)



Note: Error bars at 95% confidence intervals.

larger effect sizes for Republican respondents in the anti-feminist outcome (a CATE of 0.542 for Republicans vs. 0.369 for Democrats), although again, this difference is not statistically significant ($p = .39$).

Social Desirability Bias and Priming

We use a series of robustness checks to examine whether our respondents are simply reporting socially desirable answers or responding to explicit primes about the committee's gender composition. We use MTurk to rerun our sexual harassment vignettes with modified headlines and question wording. In the experiments presented above, we include the gender composition of the group in the vignette headline (e.g., "Gender-Balanced (All-Male) Committee Supports Decreased (Increased) Penalties for Sexual Harassment"; see Appendix 1 in the SI). Our alternative treatments removed the gender makeup cue from the headlines and relied solely on pictures of the group members and the vignette text. Even without the explicit headline prompt, our main results hold and have substantively similar effect sizes (see Appendix 7 in the SI).

Our findings also hold when we exclude the phrase "thinking about the gender composition of the committee" in our procedural legitimacy question wording. We continue to find that gender balance significantly increases citizens' perceptions of procedural legitimacy in both the anti-feminist and feminist conditions at the $p \leq .10$ level or lower. We do note that the magnitudes of our effects are reduced here. Above, we reported that gender balance improves perceptions of procedural legitimacy by about 1 point on the 4-point scale when the committee makes an anti-feminist decision. The effect size drops to about 0.59 points when we do not explicitly mention gender in the question wording. When the committee makes a feminist decision, the effect size drops from 0.53 (in the original MTurk sample) to about 0.19 and only maintains significance at the $p = .08$ level. We find this a particularly important—and theoretically motivated—check of our main results, as it relates to respondent priming. The difference in effect size with and without the gender-composition prompt can be interpreted as the significance citizens place on women's presence in the absence of other cues. Even when respondents are not explicitly asked to consider the gender composition of the decision-making body, gender balance confers legitimacy on political procedures.¹¹

Finally, we include a host of survey questions typically used in social psychology to assess respondents'

propensity to report socially desirable responses. When we restrict our sample only to respondents for whom social desirability is less likely to be a confounding factor, our results hold (see Appendix 7 in the SI for details). In combination, these robustness checks suggest that our results are not driven by the vignette explicitly referencing the exclusion (or inclusion) of women or by explicitly priming respondents on the committee's gender composition in our response question wording.

Survey Timing

Finally, to assuage concerns that our treatment effects are specific to the timing of the CCES survey immediately following the 2016 U.S. presidential election, in which gender and sexual harassment were both particularly salient, we reran our original survey in July 2017 on MTurk.¹² We continue to find strong, statistically significant treatment effects for substantive legitimacy in the anti-feminist condition and for procedural legitimacy in both the anti-feminist and feminist conditions (see Appendix 8 in the SI). These results give us confidence that our original experiment was not merely capturing a response that was specific to the U.S. election, but rather reflects the deleterious effects of male overrepresentation more generally.

The Corrosive Effects of Male Overrepresentation

This article addresses a fundamental question concerning citizens' perceptions of democratic legitimacy: Does descriptive representation matter? Political theorists posit a strong link between presence and legitimacy, and empirical research exploring the consequences of women's representation often draws heavily on this normative literature. At the same time, gender and politics scholarship largely focuses on how descriptive representation leads to substantive representation or shapes citizens' political engagement. Less attention has been dedicated to examining the signals that women's inclusion sends about policy decisions and political institutions. Our work fills this gap, providing a theoretical framework and empirical tests that illuminate the link between women's presence in decision making and citizens' legitimacy beliefs.

We find that women's equal presence grants legitimacy to political decisions and democratic procedures.

¹¹As we did not use this prime in the original substantive legitimacy question wording, this is not a concern vis-à-vis that finding.

¹²Importantly, these survey experiments were also conducted before the advent of the #MeToo movement.

To begin, we show that citizens, on average, see anti-feminist decisions as more legitimate when women are included in the decision-making process. Further, we find some suggestive evidence that this effect is especially pronounced among men, those with less crystallized views on sexual harassment, and self-identified Republicans. The conclusion that group members' presence can legitimize outcomes that violate group rights is in line with recent work on race and legitimacy beliefs (Hayes and Hibbing 2017). On the one hand, this outcome is deeply troubling, suggesting that actors looking to roll back group rights could manipulate public opinion by placing group members on decision-making bodies. On the other hand, this outcome signals the profound importance of inclusion. Assuming good-faith deliberations, the voices of marginalized groups matter even when those groups lose benefits or protections.

Moving to perceptions of procedural legitimacy, we find that, on average, gender balance improves citizens' attitudes regardless of the decision the panel makes. Even in cases in which all-male panels advance feminist policies, citizens report lower average levels of procedural fairness, institutional trust, and acquiescence (though we observe larger effects when the group makes an anti-feminist decision, which further supports our finding that women's presence especially matters when decisions counter group interests). Women's presence even communicates procedural legitimacy when the issue under discussion is unrelated to gender. More generally, these results speak to broader efforts to document the causes and consequences of male overrepresentation (see Besley et al. 2017; Murray 2014), and our findings provide the first causal evidence connecting men's overrepresentation to diminished perceptions of democratic legitimacy.

That citizens view gender-balanced institutions as more legitimate—and as producing more equitable outcomes for women—is important both normatively and instrumentally. Scholars are concerned with the ways in which governments propagate legitimating beliefs because these beliefs foster citizens' voluntary obedience to authority and compliance with existing laws and regulations (Levi, Sacks, and Tyler 2009). Thus, when citizens believe political institutions are legitimate, governments operate more effectively. Our finding that male-dominated decision-making bodies delegitimize political decisions and democratic procedures suggests that homogenous institutions make effective governance more difficult.

Our theory and results also serve as a point of departure for a much broader research agenda. To begin, we intentionally focus on sexual harassment policy because the normative literature suggests that descriptive repre-

sentation is especially important in issue areas without sharp partisan divisions. Future work can examine gendered issues that are more partisan and for which both citizens and political elites likely have more crystallized views. In particular, we expect that women's presence has a weaker effect on substantive legitimacy when opinions are more crystallized, such as support for abortion rights. In these cases, even men may know the "right" policy stance for women because of their partisan leanings. Yet, in such instances, we may see stronger effects for procedural legitimacy, as women's exclusion from deliberation on well-established gendered issues may be seen as especially egregious. Drawing on our motivating example, Trump signing abortion restrictions while surrounded by only men speaks to this scope condition. Our work suggests that this action degrades both men's and women's perceptions of the decision and the procedures surrounding it, but future research could test how respondent gender and partisanship mitigate or exacerbate our treatment effects on this polarized issue.

Additional extensions also entail varying the presence conditions such that women hold just one, two, or three seats. To advance this agenda, we conducted a preliminary examination of one of the most interesting alternative conditions: tokenism. Here, we aimed to establish whether the presence of a single woman legislator is sufficient to confer legitimacy on decisions, decision-making processes, and decision-making institutions. To do so, we conducted a preliminary analysis in which we modified our eight-member committee to contain just one woman. We find that in this case, the legitimizing effects of women's presence no longer hold. As compared to the gender-balanced panel, our results either lose significance completely or are greatly reduced in effect size (see Appendix 10 in the SI). This suggests that citizens cannot be swayed by the presence of a token group representative. Future work should test whether there is a tipping point at which presence confers legitimacy. For instance, does the threshold of 30%, often considered a critical mass of group representation, generate the same legitimacy beliefs as parity?

In addition, future research should also examine how variations in the decision rule used within the deliberative body—or cues about the nature of the interactions within the group (e.g., the presence or absence of conflict, a robust discussion with different perspectives)—might change respondents' perceptions of legitimacy (see Karpowitz and Mendelberg 2014). This line of inquiry presents an important bridge to traditional theories of procedural fairness, which emphasize how the nature of citizen involvement in, and influence over, political decisions affects legitimacy beliefs (see De Fine Licht et al.

2014; Esaiasson, Gilljam, and Persson 2012; Esaiasson et al. 2016; Persson, Esaiasson, and Gilljam 2013). For instance, scholarship in this vein could examine whether the link between women's presence and substantive and procedural legitimacy is affected by the decision rule used by the committee (e.g., majority vs. unanimity rule).

Likewise, we should establish whether the legitimizing effects of descriptive representation found in a bipartisan legislative committee also hold for expert panels. Drawing on anecdotal evidence, the public outrage that followed in 2012 after Rep. Darrell Issa convened a hearing on contraceptive coverage with only male panelists suggests that women's presence likely confers legitimacy across institutional settings. Yet, more work is needed to explicitly investigate this claim. Likewise, future research should consider whether and how women's presence affects legitimacy beliefs across other branches of government or even in nongovernmental positions, such as women's presence on corporate boards. With respect to political bodies like the U.S. Supreme Court, observational work based on time-series data may even allow us to assess how citizens' expectations of descriptive representation have evolved as institutions become more diverse over time.

Finally, our work has implications for the significant body of scholarship on gender in comparative politics. There is a great deal of variation in women's representation worldwide, and in some countries, women are approaching equal representation in national legislative seats (e.g., Mexico, Sweden, and South Africa). This trend raises clear questions as to whether the relationship between women's presence and citizens' perceptions of democratic legitimacy varies in countries that have experienced different levels of—and debates around—women's descriptive representation. In this way, our work lays the foundation for additional research in the United States and beyond.

References

- Alexander, Amy C. 2012. "Change in Women's Descriptive Representation and the Belief in Women's Ability to Govern: A Virtuous Cycle." *Politics & Gender* 8(4): 437–64.
- Arnesen, Sveinung, and Yvette Peters. 2017. "The Legitimacy of Representation: How Descriptive, Formal, and Responsiveness Representation Affect the Acceptability of Political Decisions." *Comparative Political Studies* 51(7): 868–99.
- Atkeson, Lonna Rae, and Nancy Carrillo. 2007. "More Is Better: The Influence of Collective Female Descriptive Representation on External Efficacy." *Politics & Gender* 3(1): 79–101.
- Banducci, Susan A., Todd Donovan, and Jeffrey A. Karp. 2004. "Minority Representation, Empowerment, and Participation." *Journal of Politics* 66(2): 534–56.
- Bauer, Nichole M. 2015. "Who Stereotypes Female Candidates? Identifying Individual Differences in Feminine Stereotype Reliance." *Politics, Groups, and Identities* 3(1): 94–110.
- Bauer, Nichole M. 2017. "The Effects of Counterstereotypic Gender Strategies on Candidate Evaluations." *Political Psychology* 38(2): 279–95.
- Beaman, Lori, Raghendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *Quarterly Journal of Economics* 124(4): 1497–549.
- Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3): 351–68.
- Besley, Timothy, Olle Folke, Torsten Persson, and Johanna Rickne. 2017. "Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden." *American Economic Review* 107(8): 2204–42.
- Celis, Karen, and Sarah Childs. 2012. "The Substantive Representation of Women: What to Do with Conservative Claims?" *Political Studies* 60(1): 213–25.
- Celis, Karen, and Amy G. Mazur. 2012. "Hanna Pitkin's 'Concept of Representation' Revisited." *Politics & Gender* 8(4): 508–12.
- Clayton, Amanda. 2015. "Women's Political Engagement under Quota-Mandated Female Representation: Evidence from a Randomized Policy Experiment." *Comparative Political Studies* 48(3): 333–69.
- Clayton, Amanda. 2018. "Do Gender Quotas Really Reduce Bias? Evidence from a Policy Experiment in Southern Africa." *Journal of Experimental Political Science* 1–13.
- Cohen, Jacob. 1992. "A Power Primer." *Psychological Bulletin* 112(1): 155–59.
- De Fine Licht, Jenny, Daniel Naurin, Peter Esaiasson, and Mikael Gilljam. 2014. "When Does Transparency Generate Legitimacy? Experimenting on a Context-Bound Relationship." *Governance* 27(1): 111–34.
- Dovi, Suzanne. 2007. "Theorizing Women's Representation in the United States." *Politics & Gender* 3(3): 297–319.
- Easton, David. 1965. *A Framework for Political Analysis*. Vol. 25. Englewood Cliffs, NJ: Prentice-Hall.
- Esaiasson, Peter. 2010. "Will Citizens Take No for an Answer? What Government Officials Can Do to Enhance Decision Acceptance." *European Political Science Review* 2(3): 351–71.
- Esaiasson, Peter, Mikael Gilljam, and Mikael Persson. 2012. "Which Decision-Making Arrangements Generate the Strongest Legitimacy Beliefs? Evidence from a Randomised Field Experiment." *European Journal of Political Research* 51(6): 785–808.
- Esaiasson, Peter, Mikael Persson, Mikael Gilljam, and Torun Lindholm. 2016. "Reconsidering the Role of Procedures for Decision Acceptance." *British Journal of Political Science* 1–24.
- Hayes, Matthew, and Matthew V. Hibbing. 2017. "The Symbolic Benefits of Descriptive and Substantive Representation." *Political Behavior* 39(1): 31–50.

- Kanthak, Kristin, and Jonathan Woon. 2015. "Women Don't Run? Election Aversion and Candidate Entry." *American Journal of Political Science* 59(3): 595–612.
- Karpowitz, Christopher F., and Tali Mendelberg. 2014. *The Silent Sex: Gender, Deliberation, and Institutions*. Princeton, NJ: Princeton University Press.
- Karpowitz, Christopher F., Tali Mendelberg, and Lee Shaker. 2012. "Gender Inequality in Deliberative Participation." *American Political Science Review* 106(3): 533–47.
- King, David C., and Richard E. Matland. 2003. "Sex and the Grand Old Party: An Experimental Investigation of the Effect of Candidate Sex on Support for a Republican Candidate." *American Politics Research* 31(6): 595–612.
- Kittilson, Miki Caul, and Leslie A. Schwindt-Bayer. 2012. *The Gendered Effects of Electoral Institutions: Political Engagement and Participation*. New York: Oxford University Press.
- Klar, Samara. 2018. "When Common Identities Decrease Trust: An Experimental Study of Partisan Women." *American Journal of Political Science* 62(3): 610–22.
- Lawless, Jennifer. 2004. "Politics of Presence? Congresswomen and Symbolic Representation." *Political Research Quarterly* 57(1): 81–99.
- Leung, Kwok, Kwok-Kit Tong, and E. Allan Lind. 2007. "Re-alpolitik versus Fair Process: Moderating Effects of Group Identification on Acceptance of Political Decisions." *Journal of Personality and Social Psychology* 92(3): 476–89.
- Levi, Margaret, Audrey Sacks, and Tom Tyler. 2009. "Conceptualizing Legitimacy, Measuring Legitimizing Beliefs." *American Behavioral Scientist* 53(3): 354–75.
- Lind, E. Allan, and Tom R. Tyler. 1988. *The Social Psychology of Procedural Justice*. New York: Springer Science & Business Media.
- Mansbridge, Jane. 1999. "Should Blacks Represent Blacks and Women Represent Women? A Contingent 'Yes'." *Journal of Politics* 61(3): 628–57.
- Mullen, Elizabeth, and Linda J. Skitka. 2006. "Exploring the Psychological Underpinnings of the Moral Mandate Effect: Motivated Reasoning, Group Differentiation, or Anger?" *Journal of Personality and Social Psychology* 90(4): 629–43.
- Murray, Rainbow. 2014. "Quotas for Men: Reframing Gender Quotas as a Means of Improving Representation for All." *American Political Science Review* 108(5): 520–32.
- Persson, Mikael, Peter Esaiasson, and Mikael Gilljam. 2013. "The Effects of Direct Voting and Deliberation on Legitimacy Beliefs: An Experimental Study of Small Group Decision-Making." *European Political Science Review* 5(3): 381–99.
- Phillips, Anne. 1995. *The Politics of Presence*. New York: Oxford University Press.
- Pitkin, Hanna. 1967. *The Concept of Representation*. Berkeley: University of California Press.
- Rehfeld, Andrew. 2005. *The Concept of Constituency: Political Representation, Democratic Legitimacy, and Institutional Design*. New York: Cambridge University Press.
- Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science*, 46(1): 20–34.
- Scherer, Nancy, and Brett Curry. 2010. "Does Descriptive Race Representation Enhance Institutional Legitimacy? The Case of the U.S. Courts." *Journal of Politics* 72(1): 90–104.
- Schwindt-Bayer, Leslie, and William Mishler. 2005. "An Integrated Model of Women's Representation." *Journal of Politics* 67(2): 407–28.
- Skitka, Linda J., and Elizabeth Mullen. 2008. "Moral Convictions Often Override Concerns about Procedural Fairness: A Reply to Napier and Tyler." *Social Justice Research* 21(4): 529–46.
- Stauffer, Katelyn E. 2018. *The Sum of Its Parts? Women's Collective Representation and American Evaluations of Political Institutions*. Ph.D. dissertation, Indiana University.
- Theobald, Nick A., and Donald P. Haider-Markel. 2008. "Race, Bureaucracy, and Symbolic Representation: Interactions between Citizens and Police." *Journal of Public Administration Research and Theory* 19(2): 409–26.
- Tyler, Tom R. 2006. "Psychological Perspectives on Legitimacy and Legitimation." *Annual Review of Psychology* 57, 375–400.
- Urbinati, Nadia, and Mark E. Warren. 2008. "The Concept of Representation in Contemporary Democratic Theory." *Annual Review of Political Science* 11, 387–412.
- Walsh, Denise M. 2010. *Women's Rights in Democratizing States: Just Debate and Gender Justice in the Public Sphere*. New York: Cambridge University Press.
- Wolbrecht, Christina, and David E. Campbell. 2007. "Leading by Example: Female Members of Parliament as Political Role Models." *American Journal of Political Science* 51(4): 921–39.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix 1: Pretesting Description

Appendix 2: Descriptive Statistics, Balance Diagnostics, and Correlation of Outcome Variables

Appendix 3: Verifying Dependent Variable Scales

Appendix 4: Individual Question Plots

Appendix 5: Substantive Legitimacy: Difference-in-Difference Estimate

Appendix 6: Gender Differences in Perceptions of Procedural Legitimacy

Appendix 7: Social Desirability Checks

Appendix 8: Survey Timing

Appendix 9: MTurk Results: Intent-to-Treat (ITT) Effects

Appendix 10: All-Male Panels vs. One-Woman Panels