What Have We Learned about Gender from Candidate Choice Experiments? A Meta-Analysis of Sixty-Seven Factorial Survey Experiments

Susanne Schwarz, Princeton University
Alexander Coppock, Yale University

Candidate choice survey experiments in the form of conjoint or vignette experiments have become a standard part of the political science toolkit for understanding the effects of candidate characteristics on vote choice. We collect 67 such studies from all over the world and reanalyze them using a standardized approach. We find that the average effect of being a woman (relative to a man) is a gain of approximately 2 percentage points. We find some evidence of heterogeneity across contexts, candidates, and respondents. The difference is somewhat larger for white (vs. black) candidates and among survey respondents who are women (vs. men) or, in the US context, identify as Democrats or Independents (vs. Republicans). Our results add to the growing body of experimental and observational evidence that voter preferences are not a major factor explaining the persistently low rates of women in elected office.

Do voters discriminate against women running for office? Owing to the real, pervasive, and pernicious biases against women in many areas of society, a reasonable guess would be that voters tend to prefer men over women when choosing among candidates. In the United States, the 2016 presidential election in particular confirmed for many observers that women seeking higher office face unique challenges, including gender-based discrimination (see, e.g., Burleigh 2016; Crockett 2016).

Setting aside that particular election, the empirical evidence of voter bias against women candidates (conditional on running) is surprisingly thin. Some early studies indeed reported gender gaps in electoral outcomes. In the 1960s and early 1970s, men tended to outpoll women in many Western democracies, albeit by relatively small margins (e.g., Darcy and Schramm 1977; Hills 1981; Kelley and McAllister 1984). For example, Kelley and McAllister (1984) reported that, conditional on party affiliation, vote margins for women in Britain and Australia were on average 2.5 and 4 percentage points lower, respectively. However, election returns from the late 1970s onward indicate these gender-based discrepancies in elections have by and large dissipated. When women run for political office, they are not less likely to win than men. In the United States, for example, women candidates typically garner the same level of support as men with similar characteristics, in both primary and general elections (e.g., Fox 2005; Welch, Clark, and Darcy 1987; Zipp and Plutzer 1985). Likewise, a study that tracked vote shares for parliamentary candidates between 1903 and 2004 in Australia found that gender-based discrepancies diminished drastically after the 1980s, and win rates for women today are virtually identical to those of men (King and Leigh 2010). An analysis of elections for the US House of Representatives from 1980 to 2012 comes to similar conclusions, even when comparing ideologically similar men and women (Thomsen 2020). And even in 2016, when concerns about gender discrimination and misogyny in politics were heightened in the US context, women in nonpresidential contests encountered little resistance from voters. Democratic women running for the US House of Representatives outpolled men in both primary and general elections, while Republican women performed only slightly worse than men candidates (Dittmar 2017).

Susanne Schwarz (susanneschwarz@princeton.edu) is a doctoral candidate in the Department of Politics at Princeton University, Princeton, NJ 08544. Alexander Coppock (alex.coppock@yale.edu) is an assistant professor of political science at Yale University, New Haven, CT 06520.

Replication files are available in the *JOP* Dataverse (https://dataverse.harvard.edu/dataverse/jop). The empirical analysis has been successfully replicated by the *JOP* replication analyst. An online appendix with supplementary material is available at https://doi.org/10.1086/716290.

Published online March 3, 2022.

The Journal of Politics, volume 84, number 2, April 2022. Published by The University of Chicago Press for the Southern Political Science Association. https://doi.org/10.1086/716290

These analyses of electoral returns yield the important descriptive finding that, conditional on running, men and women tend to win elected office at similar rates. It is tempting to give this finding a causal interpretation, that is, that the average causal effect of being a woman (vs. a man) on the probability of winning an election is close to zero. However, the men and women who successfully arrive on the ballot are different from each other in more ways than gender alone. Previous work shows that women who run are of higher quality than their men counterparts (Anzia and Berry 2011; Fulton 2011). Further, the women who win appear (at least by some measures) to go on to become superior elected officials (Brollo and Troiano 2016; Jeydel and Taylor 2003). The complex factors that eventuate in a person's name appearing on the ballot are surely different by gender so that the types of men and women who end up as candidates differ from one another in both observed and unobserved ways. If so, then comparisons of men and women candidates (even controlling for observables) may yield biased estimates of the true effect of gender on vote choice. Indeed, if the women who run are of higher quality than otherwise observationally equivalent men, then electoral gender parity would imply bias against women (Anzia and Berry 2011).

In response to this inferential difficulty, scholars have turned to a particular flavor of randomized experiment, the candidate choice survey experiment. In these studies, survey respondents evaluate hypothetical candidates—presented to them in either vignettes or statements of their personal characteristics in a conjoint table—and report whether they would or would not vote for a given candidate. Because the gender of the candidate (and typically, many of the other candidate characteristics) is randomized, a comparison of the support for women candidates versus men candidates yields an estimate of the effect of gender that is not biased by unobserved confounding factors that may plague many observational analyses.[1]

While survey experiments provide design-based assurances that causal effect estimates are unbiased, we may still be concerned that the question experiments answer is not quite the question we care about. We want to know the average effect of switching the genders of men and women candidates in the real world, but the question answered by the hypothetical candidate choice experiment is subtly different. First, we have some evidence that subjects evaluate hypothetical and real candidates differently (McDonald 2019); it is possible that voters would prefer a hypothetical woman in principle but would not prefer any actual real-world woman when they see her on the ballot. Second, we might be concerned that subjects spend less cognitive effort when filling out a survey questionnaire than they do when filling out a ballot. In particular, survey respondents might not pay serious attention in these experiments, but if they did, their biases against women would reveal themselves. Third, we might worry that the observed response patterns are due to experimenter demand effects (Zizzo 2010), as respondents may anticipate the research objectives of a study and adjust their answers accordingly to "confirm" the researcher's hypotheses. Respondents may also desire to appear like "good" people who behave in ways that are socially desirable, thus masking their true preferences (e.g., Krupnikov, Piston, and Bauer 2016).

Such explanations are difficult to square, however, with recent investigations that have found very little evidence for demand effects, especially in the context of studies administered online (e.g., De Quidt, Haushofer, and Roth 2018; Hopkins 2009). For example, response patterns did not differ systematically when researchers varied the amount of information that participants received about the study objectives at the beginning of an online survey (Mummolo and Peterson 2019). Another study randomly altered demographic information about the researcher in an online experiment and reported no difference in how respondents answered subsequent questions on racial resentment, gender roles, or support for women and minority presidential candidates (White et al. 2018). We agree with Clayton et al. (2019) that social desirability is not likely to be a consequential factor in conjoint experiments where the multiplicity of candidate attributes makes it difficult for the enumerator to determine which of the factors drove an individual's response (for designs that exploit this logic, see Dahl 2018; Horiuchi, Markovich, and Yamamoto 2021).

Abramson, Koçak, and Magazinnik (2019) offer a different critique of conjoint experiments. They prove that a positive average treatment effect (ATE) for the "woman" characteristic need not imply that a majority of the survey subjects prefer women candidates to men candidates. The explanation for this counterintuitive result is that it is possible the share of subjects who prefer women to men is smaller than 50%, but that share holds that preference more intensely. We therefore do not interpret these experiments as revealing majority preference but rather as revealing the average effect of gender on electoral support, the interpretation emphasized by Bansak et al. (2020) in their response to Abramson et al. (2019).

We aim in this article to document and summarize in one place what has been learned from hypothetical candidate choice

1. In this article, we will speak exclusively of candidate gender being either a "man" or a "woman." We recognize that gender can take on many values beyond these two. However, none of the experiments in our sample assign candidates a gender beyond man or woman. We also recognize that gender is socially constructed and a concept distinct from biological sex (Bittner and Goodyear-Grant 2017). Our goal is to understand the effects of the social construction, but we grant that in these experiments gender and sex are perfectly collinear in the sense that, when respondents are informed that a candidate is a woman, they likely infer both her gender and her sex.

experiments about how gender influences voters' support for political candidates. At this point, many such experiments have been conducted. We collected 67 conjoint and factorial candidate-choice experiments in which the researchers randomized candidate gender and studied respondents' vote choice. Our set of 67 includes studies conducted on six continents and in countries with widely varying levels of democratization.[2] Our meta-analytic estimate of the average effect of being a woman (vs. a man) is an approximately 2 percentage point increase in support. While effect estimates are clearly negative in some parts of the world, positive ATEs are common across the globe. We show some variation in effects associated with respondent and candidate characteristics. Effects are on average positive for both men and women respondents, but they are more positive among women. In the United States, the average effect among respondents who identify as Democrats or Independents is positive, while the average effect among Republicans is slightly negative. When we consider whether the effect of gender is similar for candidates from different racial backgrounds in the United States, we find that effects are smaller (although still positive) among nonwhite candidates compared to white candidates.

Our meta-analytic results may come as a surprise to some, as they did to us when we began the project. Indeed, many of the authors of the experiments included in our meta-analysis expressed surprise at their own results as well. Clayton et al. (2019, 620) write, "To our surprise, our experimental results did not reveal a generalized public distaste for women leaders." Kage, Rosenbluth, and Tanaka (2019, 285) remark, "We were surprised to find, based on three experimental surveys, that Japanese voters do not harbor particularly negative attitudes toward female politicians." Saha and Weeks (2020) conclude, "Contrary to the expectations of Hypothesis 1, we find no evidence that voters penalize ambitious women." Teele, Kalla, and Rosenbluth (2018, 530) summarize their main result as, "In both surveys and among most subgroups we do not find evidence that women are discriminated against *as* women. In fact, . . . female candidates actually get a *boost* over men." Our meta-analysis demonstrates that these individual findings are not flukes but instead generalize quite well to many times and contexts.

SUPPLY AND DEMAND EXPLANATIONS FOR CANDIDATE SELECTION

Women make up roughly half of the world's population but hold only 23% of the elected positions in national legislatures (Inter-Parliamentary Union 2018). The existing theory and evidence tends to group possible causes of this persistent gender gap in electoral politics into factors that shape the supply and demand for politicians who are women (e.g., Holman and Schneider 2018; Karpowitz, Monson, and Preece 2017).

Supply-side explanations consider each of the critical junctures that may shape whether and what sorts of women stand for election. Women might not aspire to run for office at the same rates as men (Kanthak and Woon 2015; Lawless and Fox 2010). They also might face higher entry costs into politics, especially in primary-based electoral systems (Lawless and Pearson 2008), and they might be more averse to highly competitive settings than men (Preece and Stoddard 2015). As many women continue to be primary caregivers for their families, they may shy away from politics when their political career adversely affects their work-life balance, for example, in the form of extended commutes to legislative offices (Silbermann 2015) or even increased risk of divorce (Folke and Rickne 2020). In addition, women are less likely to be recruited by party gatekeepers to run for office (Crowder-Meyer 2013). By contrast, when parties face high levels of electoral competition, they might be more inclined to consider women and minority candidates if it increases their chances of winning (Folke and Rickne 2016).

In this article, however, we focus on what scholars have labeled demand-side barriers. Voters simply might not have a "taste" for women candidates and politicians, or gender stereotypes about leadership abilities may disadvantage women at the ballot box (e.g., Alexander and Andersen 1993; Brescoll and Okimoto 2010; Sanbonmatsu 2002; Smith, Paul, and Paul 2007). One possibility sometimes invoked in the popular press is that women might lack political skill or might be otherwise legislatively deficient, so voters elect them at lower rates. The empirical record on this count, however, indicates that women are effective legislators (Jeydel and Taylor 2003) and are often more likely to get things done than men (Brollo and Troiano 2016), although this pattern may result from a selection process in which only the highest quality women are elected (Anzia and Berry 2011; Fulton 2011).

Hypothetical candidate choice experiments mainly shed light on the demand side of the candidate selection process, as they measure voter support for candidates of various types (one might argue that they also inform the supply side to the extent that they measure the tastes of party gatekeepers; e.g., Doherty, Dowling, and Miller 2019). Our article aims to summarize the evidence from candidate choice survey experiments on three distinct demand-side explanations. First, we explore whether individuals discriminate against women candidates on average. Next, we ask whether individuals discriminate against specific types of women candidates, with a focus on the intersections between race and gender. Finally, we test whether

2. That said, the majority of studies on gender and vote choice we revisit here were conducted in industrialized, democratic countries.

certain subgroups of respondents display stronger (or weaker) preferences for women candidates. We briefly review the theory underlying each of these three before turning to our meta-analysis of candidate choice experiments.

Voter preferences and gender discrimination

Hostility against women or a general distaste for women politicians predicts, all else equal, that the effect of being a woman on candidate support should be negative. Perceptions about gender roles may shape evaluations of political candidates, especially in the context of male-dominated arenas such as politics. Women candidates may face electoral penalties because they are perceived as defying traditional sex roles and prescriptive gender norms, even when they are as qualified for the job as men (the gender-incongruency hypothesis; e.g., Brescoll and Okimoto 2010). Indeed, in one study, respondents rated fictitious men presidential candidates as more skilled and as having more political potential than women candidates despite their having otherwise identical profiles (Smith et al. 2007). In a study of state legislative contests in the United States between 1970 and 1980, men candidates garnered more support than women candidates by an average of 2 percentage points (Welch et al. 1985).

However, more recent studies have suggested that outright discrimination against women political candidates may not be as prevalent as it once was. An observational study of candidates for US Congress did not find evidence of differential candidate evaluations by gender after conditioning on partisanship (Dolan and Lynch 2014). Similarly, a study of local media coverage of political candidates in nearly 350 congressional contests found no significant differences in the portrayal of women and men office seekers (Hayes and Lawless 2015). Indeed, our meta-analysis of candidate choice experiments finds only limited support for the gender-incongruency hypothesis as well. Across the 67 studies we summarize, three-fourths find a positive effect of being described as a woman.

Interactions with other candidate characteristics

Even if voters do not discriminate against women in general, they might evaluate certain types of women or men candidates differently. In other words, women candidates might face double standards in terms of the qualifications or attributes they need to bring to the table if they want to succeed in politics (the double-standard hypothesis; see Teele et al. 2018). Experimental research has demonstrated penalties for women who overtly "seek power" or who are as assertive as men (Brescoll and Okimoto 2010). In addition, respondents react negatively to women who show emotions like anger (Brescoll and Uhlmann 2008; Brooks 2011). However, in a recent conjoint experiment, Teele et al. (2018) found that women faced bigger electoral advantages than men when they had a larger family, and for all other characteristics—age, marital status, experience in politics, and previous occupation—men and women were not rewarded or penalized differentially. As we will show below, the balance of evidence from a large set of candidate choice experiments also does not support the double-standard hypothesis.

Intersectional theories of gender and politics predict that whatever effects candidate gender may have on candidate support, the effects are likely to be different for candidates of different racial or ethnic groups (e.g., Crenshaw 1991; Hardy-Fanta 2013; Holman and Schneider 2018; Hooks 1982; Mügge and De Jong 2013). Even if the effect of being a woman is positive for white candidates, it need not be for black candidates. Candidate choice experiments often randomize both characteristics of candidates independently, so they are well placed to evaluate this possibility. As we will demonstrate, once we aggregate across a number of studies conducted with US samples that manipulate both gender and race, we find only a modest difference in the effect of gender for white and black candidates.

Interactions with respondent characteristics

Identity-based theories of vote choice suggest that individuals favor political candidates who "look" and "think" like themselves (Besley and Coate 1997; Converse et al. 1961). We might therefore expect voters who are women to prefer candidates who are women as well (the gender affinity hypothesis; see Dolan 2008). Findings from a number of survey and experimental studies lend some support to this hypothesis (Dolan 1998; Plutzer and Zipp 1996; Sanbonmatsu 2002). However, more recent studies report no such gender affinity pattern (Dolan 2008; Teele et al. 2018). In the present meta-analysis, we find some evidence supporting a gender affinity argument: the positive effect of being a woman candidate is larger among respondents who are themselves women than among men.

Candidates' personal characteristics may also provide informational shortcuts for voters, allowing them to infer candidates' policy positions and ideological orientations in low-information environments (Downs 1957; Kirkland and Coppock 2018; Popkin 1991). Similar to candidate partisan affiliation, candidate gender may provide such a shortcut (the gender heuristic hypothesis). For example, women candidates are often believed to be more liberal than men candidates, which can advantage women Democratic candidates over their male counterparts among liberal voters (McDermott 1997, 1998). By the same logic, Republican women running for office may face additional barriers (Bucchianeri 2018) among conservative voters.

In addition, gender stereotypes may mold perceptions of issue positions that candidates hold and of their skills and leadership abilities. Women are seen as "more dedicated to honest government" (McDermott 1998) and viewed as better suited to handling issues related to women, children, the aged, and the poor (Huddy and Terkildsen 1993). By contrast, men politicians are often more trusted with issues related to national security or the economy (Holman, Merolla, and Zechmeister 2016). Both Democratic and Republican voters appear to hold these gendered stereotypes, but because these groups differ in their policy preferences and ideological positions, they may endorse women candidates to varying degrees (Sanbonmatsu and Dolan 2009). Indeed, as we will discuss in more detail below, our meta-analysis of studies conducted among American voters yields positive effects among Democrats but negative effects among Republicans.

DESIGN

Most of the candidate choice experiments we consider here were not designed specifically to study gender but nevertheless vary candidate gender as one of the many candidate characteristics included in the description of hypothetical candidates. Our goal is to leverage the randomization of candidate gender in many countries, time periods, and contexts in order to gain a holistic understanding of the effects of gender on vote choice. We attempted to collect all candidate choice experiments ever conducted and described in academic papers, whether published or unpublished. We used two main inclusion criteria: (1) candidate gender is randomized, and (2) the dependent variable is, or can be transformed into, a binary vote choice for or against the candidate.

We followed standard practices to locate our studies: citation chains, internet searches using the terms "factorial," "candidate choice," "voter preference experiment," "conjoint experiment," "gender, vote, experiment," and "vignette," and word of mouth using social media as well as personal conversations with scholars in the field. In total, we located 67 experiments from 49 papers. In 48 of the 67 cases, we were able to obtain replication data through either private communication or publicly available repositories. In the remaining 19 cases, we attempted to recover the necessary statistics from the article text or graphical presentation of results. We did not exclude studies on the basis of the manner in which candidate gender was signaled to the survey respondent. Some studies manipulate gender by indicating "man" or "woman," or "male" or "female," in a matrix of candidate characteristics (e.g., Kirkland and Coppock 2018), while others use pictures (e.g., Crowder-Meyer et al. 2015). We did not limit our data collection to any specific geographic context. While over half our studies were conducted in the United States, we also include samples from Afghanistan, Argentina, Australia, Brazil, Chile, Denmark, Germany, India, Japan, Jordan, Malawi, Norway, Switzerland, Tunisia, the United Kingdom, Vietnam, and Zambia. Moreover, studies were included regardless of their sampling procedures. Some use convenience samples like Amazon's Mechanical Turk, Lucid, Survey Sampling International, or student samples. Others use samples that are nominally representative of the voting-age population in a given country at the time the study was conducted. We included experiments that used a variety of designs, including standard factorial experiments in which only a few characteristics are varied, highly factorial conjoint experiments in which many characteristics are varied, and vignette experiments that embed manipulations in a larger dose of information about the candidate. Some experiments asked respondents to rate one candidate at a time; others asked respondents to choose between two at a time. We excluded several excellent studies that randomized gender but measured favorability or perceptions of competence instead of vote choice. Overall, our database of studies includes 67 experiments from six continents across three and a half decades. Table 1 provides an overview of the studies included in our analysis.

Our main focus will be a meta-analysis of the ATE estimates of being a woman candidate (vs. a man) in each study. These ATEs are typically sample average treatment effects (SATEs), although some studies target population average treatment effects (PATEs), by using either a probability sampling scheme or poststratification weights. The ATE in conjoint experiments is usually referred to as the average marginal component effect (AMCE; Hainmueller, Hopkins, and Yamamoto 2014). Describing an ATE as an AMCE emphasizes that the estimand itself depends on the distribution of the other candidate attributes included in the study. For simplicity, we will refer to all of these (SATEs, PATEs, and AMCEs) as ATEs for the remainder of the article.

Because we require candidate gender to be randomized, the difference in means will be an unbiased estimator of the ATE in each case. Where the raw data are available, we estimate robust standard errors and 95% confidence intervals using the estimatr package for R (Blair et al. 2018). We include sampling weights when provided by the original researchers in their replication data sets. When subjects rate multiple candidate profiles, we follow standard practice in the analysis of conjoint experiments and cluster our standard errors by respondent (Hainmueller et al. 2014). Where raw data were not available, we searched the original publications for estimates of the ATE as well as uncertainty estimates. Occasionally, this process involved digitally measuring coefficient plots for both point estimates and 95% confidence intervals.[3]

3. In app. sec. B, we show that this procedure yields accurate measurements by comparing digital measurement with direct computation.

Table 1. Study Manifest

	N Subjects	*N* Ratings	Raw Data	Sample Type
Aguilar, Cunow, and Desposato (2015), Brazil	3,908	27,076	Yes	Representative
Aguilar, Cunow, and Desposato (2015), Sao Paulo	608	608	Yes	Convenience
Arnesen, Duell, and Johannesson (2019), Norway 1	1,134	4,420	Yes	Representative
Arnesen, Duell, and Johannesson (2019), Norway 2	1,077	6,370	Yes	Representative
Bansak et al. (2018), USA, MTurk	2,411	144,494	Yes	Convenience
Bansak et al. (2018), USA, SSI	643	38,482	Yes	Convenience
Blackman and Jackson (2021), Tunisia, face to face	383	3,064	Yes	Representative
Blackman and Jackson (2021), Tunisia, YouGov	574	5,740	Yes	Representative
Campbell et al. (2019), UK, frequency of MP dissent	1,899	18,990	Yes	Representative
Campbell et al. (2019), UK, type of MP dissent	1,919	19,190	Yes	Representative
Carnes and Lupu (2016), Argentina	1,149	2,298	Yes	Representative
Carnes and Lupu (2016), UK	5,548	11,096	Yes	Representative
Carnes and Lupu (2016), USA	1,000	2,000	Yes	Representative
Clayton and Nyhan (2020), USA, donors	570	11,192	Yes	Convenience
Clayton and Nyhan (2020), USA, YouGov	954	19,080	Yes	Representative
Clayton et al. (2019), Malawi	604	3,624	Yes	Representative
Costa (2020), USA, Lucid	1,501	18,012	Yes	Convenience
Dahl and Nyrup (2021), Denmark	1,621	15,916	Yes	Convenience
Doherty, Dowling, and Miller (2020), USA	831	13,040	Yes	Representative
Eggers, Vivyan, and Wagner (2018), UK	1,367	2,806	Yes	Representative
Goyal (2020), India	1,664	9,984	Yes	Representative
Hainmueller, Hopkins, and Yamamoto (2014), USA	311	3,466	Yes	Convenience
Harris, Kao, and Lust (2020), Malawi	7,522	7,522	Yes	Representative
Harris, Kao, and Lust (2020), Zambia	5,508	5,508	Yes	Representative
Henderson et al. (2019), USA, CCES	2,791	22,328	Yes	Representative
Holman, Merolla, and Zechmeister (2016), USA	1,001	1,001	Yes	Representative
Hopkins (2014), USA	551	7,714	Yes	Representative
Horne (2020), UK	3,257	26,056	Yes	Representative
Kao and Benstead (2021), Jordan	1,490	2,926	Yes	Representative
Kirkland and Coppock (2018), USA, MTurk	1,204	12,032	Yes	Convenience
Kirkland and Coppock (2018), USA, YouGov	1,200	11,432	Yes	Representative
Leeper and Robison (2020), USA, SSI	743	7,430	Yes	Convenience
Lemi (2021), USA, Qualtrics	786	6,394	Yes	Convenience
Mares and Visconti (2020), Romania	502	5,020	Yes	Convenience
Martin and Blinder (2021), UK, YouGov	3,943	7,886	Yes	Representative
Mo (2015), Florida	407	5,700	Yes	Convenience
Ono and Yamada (2020), Japan	2,686	21,488	Yes	Convenience
Saha and Weeks (2020), DLABSS 1	551	3,280	Yes	Convenience
Saha and Weeks (2020), UK, Prolific	869	8,682	Yes	Convenience
Saha and Weeks (2020), USA, DLABSS 2	497	4,886	Yes	Convenience
Saha and Weeks (2020), USA, SSI	1,248	7,480	Yes	Representative
Sen (2017), USA, SSI 1	798	4,797	Yes	Convenience
Sen (2017), USA, SSI 2	765	4,594	Yes	Convenience
Senninger and Bischof (2021), Germany	993	9,930	Yes	Convenience
Shaffner and Green (2020), USA, YouGov Blue	2,953	29,530	Yes	Convenience
Simas and Murdoch (2019), US, MTurk	1,312	1,312	Yes	Convenience
Teele, Kalla, and Rosenbluth (2018), USA	1,052	6,312	Yes	Representative
Visconti (2020), Chile	210	3,360	Yes	Convenience
Armendariz, Farrer, and Martinez (2020), USA	1,495	2,990	No	Convenience
Atkeson and Hamel (2020), USA	1,500	3,000	No	Convenience
Bermeo and Bhatia (2017), Afghanistan	2,485	7,455	No	Representative

Table 1 (*Continued*)

	N Subjects	*N* Ratings	Raw Data	Sample Type
Crowder-Meyer et al. (2015), MTurk	430	1,290	No	Convenience
Crowder-Meyer et al. (2015), UC Merced	350	1,050	No	Student
Fox and Smith (1998), UCSB	650	2,600	No	Student
Fox and Smith (1998), University of Wyoming	990	3,960	No	Student
Hobolt and Rodon (2020), UK, YouGov	1,936	19,360	No	Representative
Horiuchi, Smith, and Yamamoto (2020), Japan	2,200	22,000	No	Convenience
Kage, Rosenbluth, and Tanaka (2019), Japan	1,611	9,666	No	Representative
Kang et al. (2018), Australia	2,290	4,580	No	Representative
Ono and Burden (2019), USA, SSI	1,583	15,830	No	Representative
Piliavin (1987), USA	245	245	No	Student
Schuler (2020), Vietnam	13,576	27,152	No	Representative
Sigelman and Sigelman (1982), USA	227	227	No	Student
Tomz and Van Houweling (2016), USA	4,200	25,200	No	Representative
Vivyan and Wagner (2015), UK, YouGov 2012	1,899	1,899	No	Representative
Vivyan and Wagner (2015), UK, YouGov 2013	1,919	1,919	No	Representative
Wuest and Pontusson (2017), Switzerland	4,500	9,000	No	Representative
Total	120,601	774,971		

Note. CCES = Cooperative Congressional Election Study; DLABSS = Digital Lab for the Social Sciences; MP = Member of Parliament; MTurk = Mechanical Turk; SSI = Survey Sampling International.

Our second analysis will present estimates of the conditional average treatment effect (CATE) of candidate gender depending on other (randomly assigned) candidate characteristics. For example, to estimate the CATE given that candidates are black, we condition the data set to only include black candidates, then estimate the difference in means using the same procedure described above. We estimate CATEs for all candidate dimensions for which we have data. These dimensions are overlapping, but we do not estimate CATEs at the intersection of candidate characteristics (e.g., among 55-year-old Democratic former police officers) because we run out of data too quickly.

Our third and final analysis estimates the CATE of candidate gender conditional on respondent characteristics. Because the space of possible respondent characteristics is very large, we limit ourselves to the evaluation of the gender affinity and the (partisan) gender heuristic hypotheses described above.

When averaging across studies, we employ random effects meta-analysis. Random effects (rather than fixed effects) is appropriate in this setting because we do not assume that the true average effect of gender is exactly the same across contexts. Instead, we assume that these effects vary from context to context but are nevertheless drawn from a common distribution. The estimand in the random effects meta-analysis is the expectation (or average) of this distribution.

RESULTS

Main finding: Positive electoral effects for women on average

Our main result is presented in figure 1. Using random effects meta-analysis, we estimate the average ATE to be 1.8 percentage points, with a confidence interval stretching from 1.1 to 2.5 points. We see a very consistent pattern in favor of women candidates on average: 47 estimates are positive (23 significant) and 18 are negative (11 significant). We find clearly positive estimates among the American, European, and South American samples. Our lone entry from South Asia (Goyal 2020) is positive as well. While this strongly positive pattern is widespread, it is not universal. In sub-Saharan Africa, we report mixed results: while in some countries, we see a clear positive effect of gender on candidate choice (Clayton et al. 2019), the estimates are mildly negative in other contexts (Harris, Kao, and Lust 2020). We find negative effects in Afghanistan (Bermeo and Bhatia 2017), Jordan (Kao and Benstead 2021), Tunisia (Blackman and Jackson 2021), Vietnam (Schuler 2020), and among students at the University of Wyoming in 1998 (Fox and Smith 1998). The estimates from Japan are truly mixed: one estimate is positive and significant (Kage et al. 2019), another is negative and significant (Ono and Yamada 2020), and a third comes in precisely at 0.0 (Horiuchi, Smith, and Yamamoto 2020). Our summary read of this collection of studies

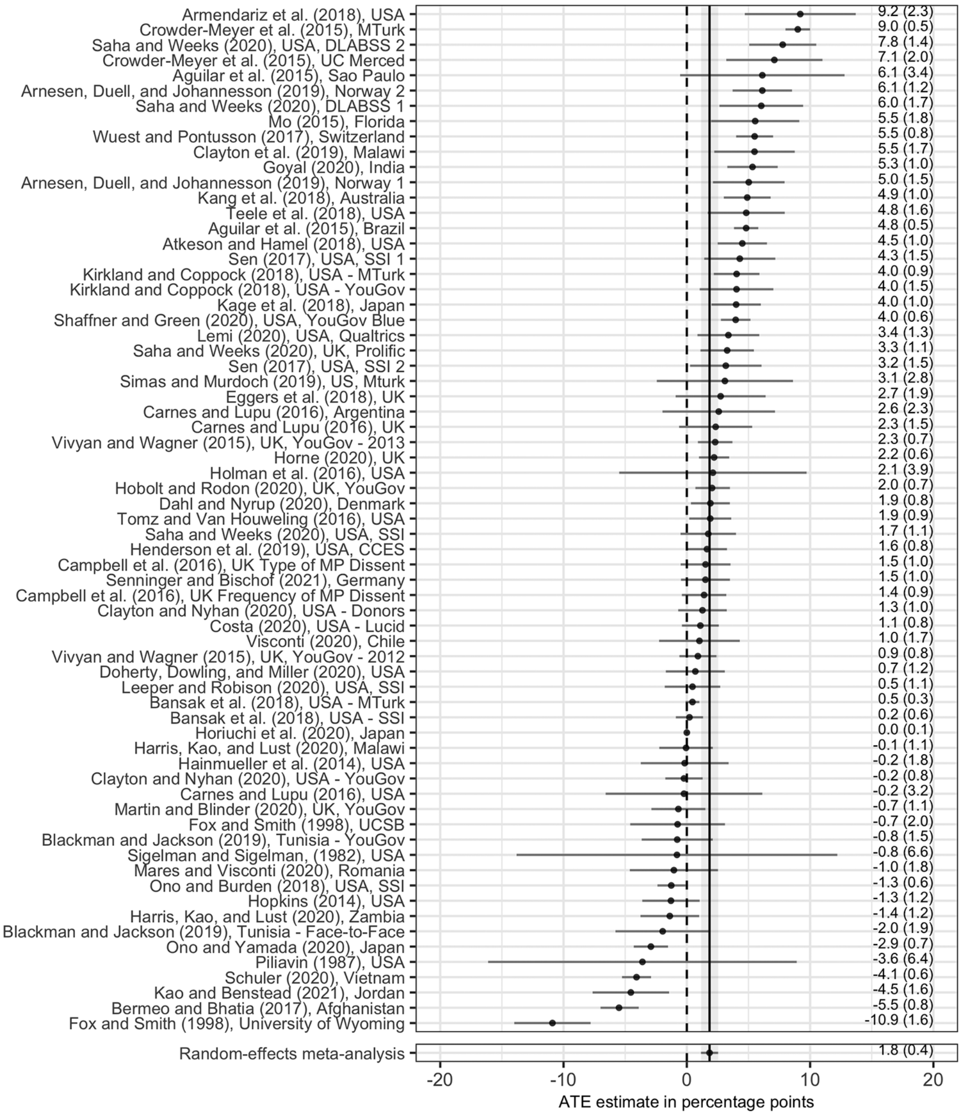

Figure 1. Results of 67 candidate choice experiments on the effect of candidate gender

is that the average effect of being described as a woman candidate is an approximately 2 percentage point gain in vote margin and that this positive effect applies in most, but emphatically not all, contexts.

In addition to the overall results, we calculated meta-analytic estimates for different subsets of studies as a robustness check (table 2). We find that the effect of gender is slightly more positive among studies conducted post-2014, whereas it appears to be negative for samples collected before 1998. This difference could be due to changing gender norms over time, although it is difficult to be sure because we were only able to include four studies from this period.[4] Effects appear to be somewhat larger, on average, among convenience or student samples as compared to representative samples (2.3 vs. 1.5 percentage points, respectively). Average effects are slightly larger among American samples than non-American samples (2.3 vs. 1.4 percentage points, respectively). However, all estimates for these subtypes, with the exception of the pre-1998 studies, are positive and statistically significant, and none of the differences across types are statistically significant. Importantly, even when we trim the bottom and top 5% of point estimates from our analysis, our meta-analytic estimate remains robust, suggesting that our findings are not merely attributable to a handful of outlying studies.

Effects conditional on candidate characteristics

Most of the studies in our sample randomized other candidate features beyond gender, allowing us to study whether the positive effects we observe on average hold for candidates

4. Surprisingly, our sample includes no studies conducted between 1998 and 2014, but this drought is ended by the explosion of interest in candidate choice experiments following the publication of Hainmueller et al. (2014).

Table 2. Meta-analytic Estimates by Study Subset

	Estimate (SE)	95% CI	N Studies
Post-2014 studies	.021 (.004)	[.014, .028]	63
Pre-1998 studies	−.047 (.037)	[−.118, .025]	4
Convenience/student sample	.023 (.006)	[.012, .034]	31
Representative sample	.015 (.005)	[.004, .025]	36
American sample	.023 (.006)	[.011, .035]	35
Non-American sample	.014 (.005)	[.004, .023]	32
Trimmed (middle 90% of estimates)	.018 (.003)	[.012, .023]	61
All studies	.018 (.004)	[.011, .025]	67

of different parties, ages, races, professions, marital status, or political experiences, among other attributes. In the appendix, we show 954 separate CATEs, the majority of which are positive.

Here we focus on candidate race in the US context, since we have a large number of studies (16) over which to pool. As shown in figure 2, the effects are slightly larger among white candidates (2.2 percentage points) than black candidates (0.9 points), although the difference between these two estimates is not statistically significant ($p = .104$). This finding provides only modest support for the intersectional hypothesis that the effects of candidate gender depend on candidate race.

Effects conditional on respondent characteristics

Finally, we consider the effects of gender, conditional on respondent characteristics, in particular by gender and partisans affiliations. Figure 3 shows the CATEs of candidate gender, conditional on respondent gender, for the 36 experiments where we were able to identify respondent gender. The effect is positive for both groups: 3.0 points among women and 0.9 points among men. The difference between the two estimates is itself statistically significant ($p = .004$), providing some support for the gender affinity hypothesis. Women tend to prefer women candidates more so than men do, even though both groups on average respond positively to women running for office.

In figure 4, we summarize the results of 20 studies conducted in the US context, for which the partisan identification of respondents was available. The effect is negative among Republicans (−1.4 points) but 3.3 percentage points for Independents and 4.2 points for Democrats. The Republican estimate is statistically significantly different from the Democratic and Independent estimates. These estimates by party underline a general difficulty in interpretation. The differential effects could represent a gender heuristic whereby people infer the sorts of policies women are likely to pursue when elected, or

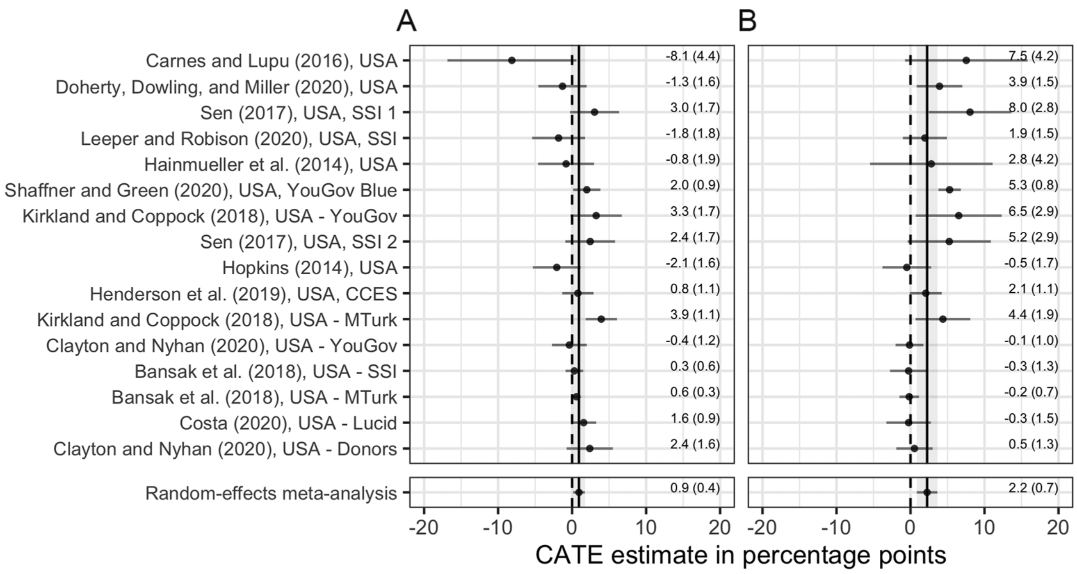

Figure 2. Conditional average effect of candidate gender, conditional on candidate race: *A*, black candidate; *B*, white candidate

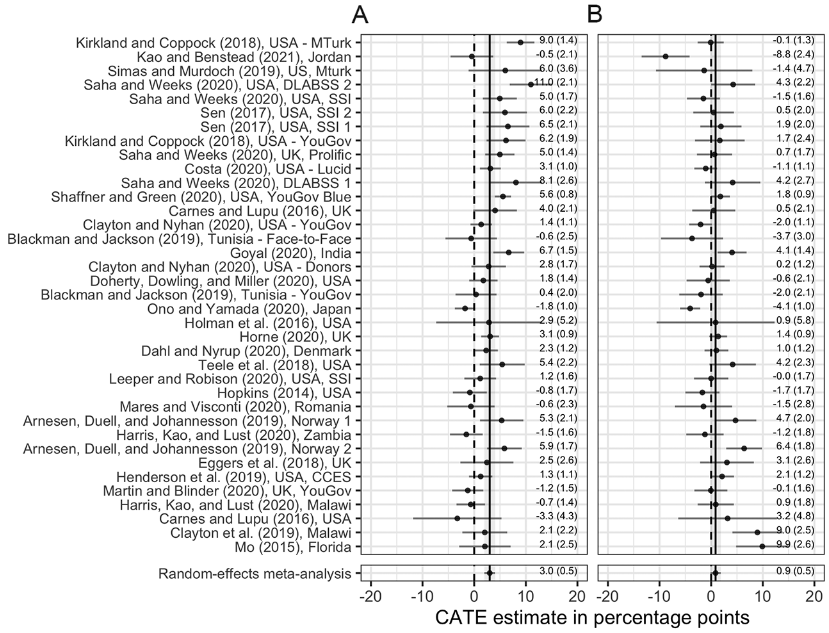

Figure 3. Conditional average treatment effects of candidate gender by respondent gender: *A*, female respondents; *B*, male respondents

they could reflect difference taste-based preferences for women by partisan group. Yet, even when disaggregating by party, we cannot disentangle the two mechanisms because Democrats may both prefer the sorts of policies typically championed by women and also prefer women on taste-based grounds.

DISCUSSION

We have summarized the statistical evidence on gender from 67 candidate choice experiments. Our main finding is that the average effect of being a woman candidate is a 1.8 percentage point increase in support. We observe considerable study-to-study variation, although more than three-fourths of the studies show a positive effect. Even in studies that estimated negative treatment effects, vote margins for women are much closer to 50% than might be expected given the clear evidence of sexism in many sectors of society.

We further investigated whether these average effects mask important heterogeneities. We find suggestive evidence that the effect is stronger among white candidates than among black candidates in the United States. However, on the whole, our results do not depend on other candidate characteristics such as experience, age, or occupation as we detail in the

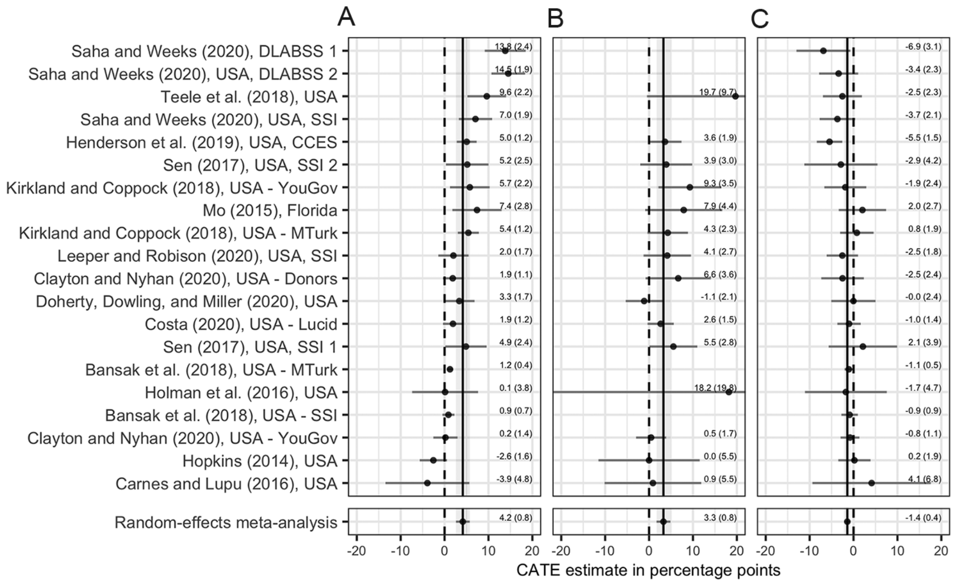

Figure 4. Conditional average treatment effects of candidate gender by respondent partisanship: *A*, Democratic respondents; *B*, Independent respondents; *C*, Republican respondents.

appendix (see figs. A.2–A.45, available online). Therefore, in line with Teele et al. (2018), our findings do not support the hypothesis that voters systematically apply double standards when they evaluate women candidates.

We found some interactions of candidate gender with respondent characteristics, however. While the effect is positive for both men and women respondents, it is somewhat larger among women, lending some support to the gender affinity hypothesis. We also observed a stronger effect among Democrats and Independents compared with Republicans, for whom the average effect is in fact negative. It is unclear, however, whether this difference is due to a gender heuristic, whereby partisans infer the sorts of policies women are likely to pursue when elected, or whether it arises from a taste-based preference among Democrats and Independents for women candidates in general.

Overall, our findings offer evidence against demand-side explanations of the gender gap in politics. Rather than discriminating against women who run for office, voters on average appear to reward women. What then explains the persistent gender gap in politics across the globe? For us, the findings we discuss here suggest that supply-side factors that include gendered differences in political ambition, party structures, donor preferences, candidate recruitment, and differences in opportunity costs are correctly coming under deeper scrutiny by political scientists (e.g., Crowder-Meyer 2013; Lawless and Fox 2010; Preece, Stoddard, and Fisher 2016; Silbermann 2015; Thomsen and Swers 2017). Of course, evaluating the causal influence of such supply-side factors on women's representation is inherently more difficult as candidate nomination and selection are complex, often opaque processes. Nevertheless, some recent scholarship has made important progress in this area. Foos and Gilardi (2020) show that a randomized invitation to meet with women politicians did not increase self-reported political ambition among women university students in Switzerland. By contrast, Kalla and Porter (2020) show that female high school students who receive political skill training show higher levels of political efficacy, even when the program does not explicitly emphasize gender and political ambition. Similarly, Karpowitz et al. (2017) randomly induced leaders of precinct-level caucus meetings to read statements encouraging their membership to elect more women delegates to the statewide convention, increasing the fraction of precincts electing at least one woman by more than 5 percentage points. We hope that future work will continue to push forward this promising research agenda.

ACKNOWLEDGMENTS

We would like to thank William Hunt for his research assistance, as well as Dawn Teele and Amanda Clayton for their support and comments. We also thank our anonymous reviewers for their helpful feedback. Previous versions of this manuscript were presented at the Midwest Political Science Association annual conference in 2018 as well as the Yale University Gender in the Social Sciences Workshop. Finally, this project would not have been possible without the many researchers who generously shared their data with us.

REFERENCES

Abramson, Scott F., Korhan Koçak, and Asya Magazinnik. 2019. "What Do We Learn about Voter Preferences from Conjoint Experiments." Unpublished manuscript, University of Rochester. https://scholar.princeton.edu/sites/default/files/kkocak/files/conjoint_draft.pdf.

Aguilar, Rosario, Saul Cunow, and Scott Desposato. 2015. "Choice Sets, Gender, and Candidate Choice in Brazil." *Electoral Studies* 39 (Suppl. C): 230–42.

Alexander, Deborah, and Kristi Andersen. 1993. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46 (3): 527–45.

Anzia, Sarah, and Christopher Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55 (3): 478–93.

Armendariz, Karen Caballero, Ben Farrer, and Monica Martinez. 2020. "Badge of Courage or Sign of Criminality? Experimental Evidence for How Voters Respond to Candidates with a Political Arrest in Their Past." *Social Science Quarterly* 101:2203–19.

Arnesen, Sveinung, Dominik Duell, and Mikael Poul Johannesson. 2019. "Do Citizens Make Inferences from Political Candidate Characteristics When Aiming for Substantive Representation?" *Electoral Studies* 57:46–60.

Atkeson, Lonna Rae, and Brian T. Hamel. 2020. "Fit for the Job: Candidate Qualifications and Vote Choice in Low Information Elections." *Political Behavior* 42 (1): 59–82.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2018. "Beyond the Breaking Point? Survey Satisficing in Conjoint Experiments." *Political Analysis* 26:112–19.

Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins, and Teppei Yamamoto. 2020. "Using Conjoint Experiments to Analyze Elections: The Essential Role of the Average Marginal Component Effect (AMCE)." Unpublished manuscript, SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3588941.

Bermeo, Nancy, and Jasmine Bhatia. 2017. "Strongmen or Technocrats? Experimental Evidence Testing Leadership Preferences in Afghanistan." Technical report, International Growth Center.

Besley, Timothy, and Stephen Coate. 1997. "An Economic Model of Representative Democracy." *Quarterly Journal of Economics* 112 (1): 85–114.

Bittner, Amanda, and Elizabeth Goodyear-Grant. 2017. "Sex Isn't Gender: Reforming Concepts and Measurements in the Study of Public Opinion." *Political Behavior* 39 (4): 1019–41.

Blackman, Alexandra Domike, and Marlette Jackson. 2021. "Gender Stereotypes, Political Leadership, and Voting Behavior in Tunisia." *Political Behavior* 43:1037–66.

Blair, Graeme, Jasper Cooper, Alexander Coppock, Macartan Humphreys, and Luke Sonnet. 2018. estimatr: Fast Estimators for Design-Based Inference. https://declaredesign.org/r/estimatr/.

Brescoll, Victoria L., and Tyler G. Okimoto. 2010. "The Price of Power: Power Seeking and Backlash against Female Politicians." *Personality and Social Psychology Bulletin* 36 (7): 923–36.

Brescoll, Victoria L., and Eric Luis Uhlmann. 2008. "Can an Angry Woman Get Ahead? Status Conferral, Gender, and Expression of Emotion in the Workplace." *Psychological Science* 19 (3): 268–75.

Brollo, Fernanda, and Ugo Troiano. 2016. "What Happens When a Woman Wins an Election? Evidence from Close Races in Brazil." *Journal of Development Economics* 122:28–45.

Brooks, Deborah Jordan. 2011. "Testing the Double Standard for Candidate Emotionality: Voter Reactions to the Tears and Anger of Male and Female Politicians." *Journal of Politics* 73:597–615.

Bucchianeri, Peter. 2018. "Is Running Enough? Reconsidering the Conventional Wisdom about Women Candidates." *Political Behavior* 40 (2): 435–66.

Burleigh, Nina. 2016. "The Presidential Election Was a Referendum on Gender and Women Lost." *Newsweek*, November 14. https://www.newsweek.com/2016/11/18/hillary-clinton-presidential-election-voter-gender-gap-520579.html.

Campbell, Rosie, Philip Cowley, Nick Vivyan, and Markus Wagner. 2019. "Legislator Dissent as a Valence Signal." *British Journal of Political Science* 49 (1): 105–28.

Carnes, Nicholas, and Noam Lupu. 2016. "Do Voters Dislike Working-Class Candidates? Voter Biases and the Descriptive Underrepresentation of the Working Class." *American Political Science Review* 110 (4): 832–44.

Clayton, Amanda, Amanda Lea Robinson, Martha C. Johnson, and Ragnhild Muriaas. 2019. "(How) Do Voters Discriminate against Women Candidates? Experimental and Qualitative Evidence from Malawi." *Comparative Political Studies* 53 (3–4): 601–30.

Clayton, Katherine, and Brendan Nyhan. 2020. "Who Will Defend Democracy? Evaluating Tradeoffs in Candidate Support among Partisan Donors and Voters." Unpublished manuscript. https://preprints.apsanet.org/engage/api-gateway/apsa/assets/orp/resource/item/5d02becd39ef0400184a49fd/original/who-will-defend-democracy-evaluating-tradeoffs-in-candidate-support-among-partisan-donors-and-voters.pdf.

Converse, Philip E., Angus Campbell, Warren E. Miller, and Donald E. Stokes. 1961. "Stability and Change in 1960: A Reinstating Election." *American Political Science Review* 55 (2): 269–80.

Costa, Mia. 2020. "Ideology, Not Affect: What Americans Want from Political Representation." Unpublished manuscript.

Crenshaw, Kimberle. 1991. "Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Color." *Stanford Law Review* 43 (6): 1241–99.

Crockett, Emily. 2016. "Why Misogyny Won." Vox, November 15. https://www.vox.com/identities/2016/11/15/13571478/trump-president-sexual-assault-sexism-misogyny-won.

Crowder-Meyer, Melody. 2013. "Gendered Recruitment without Trying: How Local Party Recruiters Affect Women's Representation." *Politics and Gender* 9 (4): 390–413.

Crowder-Meyer, Melody, Shana Kushner Gadarian, Jessica Trounstine, and Kau Vue. 2015. "Complex Interactions: Candidate Race, Sex, Electoral Institutions and Voter Choice." Unpublished manuscript. https://bit.ly/3e2YB6p.

Dahl, Malte. 2018. "Social Desirability Bias in Conjoint Experiments: What Is the Optimal Design When Studying Sensitive Topics?" Unpublished manuscript. https://pdfs.semanticscholar.org/9d22/5de39131b4d7b303ea6316530d11b45c37d6.pdf?_ga = 2.47096548.270470920.1589566831-1671928746.1589566831.

Dahl, Malte, and Jacob Nyrup. 2021. "Confident and Cautious Candidates: Explaining Underrepresentation of Women in Danish Municipal Politics." *European Journal of Political Research* 60 (1): 199–224.

Darcy, R., and Sarah Slavin Schramm. 1977. "When Women Run against Men." *Public Opinion Quarterly* 41 (1): 1–12.

De Quidt, Jonathan, Johannes Haushofer, and Christopher Roth. 2018. "Measuring and Bounding Experimenter Demand." *American Economic Review* 108 (11): 3266–302.

Dittmar, Kelly. 2017. "Candidates Matter: Gender Differences in Election 2016." Center for American Women in Politics, Rutgers. https://www.cawp.rutgers.edu/sites/default/files/resources/closer_look_candidates_matter_2.14.17.pdf.

Doherty, David, Conor M. Dowling, and Michael G. Miller. 2019. "Do Local Party Chairs Think Women and Minority Candidates Can Win? Evidence from a Conjoint Experiment." *Journal of Politics* 81 (4): 1282–97.

Doherty, David, Conor M. Dowling, and Michael G. Miller. 2020. "Small Power: How Local Parties Shape Elections." Unpublished manuscript.

Dolan, Kathleen. 1998. "Voting for Women in the 'Year of the Woman.'" *American Journal of Political Science* 42 (1): 272–93.

Dolan, Kathleen. 2008. "Is There a 'Gender Affinity Effect' in American Politics?" *Political Research Quarterly* 61 (1): 79–89.

Dolan, Kathleen, and Timothy Lynch. 2014. "It Takes a Survey: Understanding Gender Stereotypes, Abstract Attitudes, and Voting for Women Candidates." *American Politics Research* 42 (4): 656–76.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper.

Eggers, Andrew C., Nick Vivyan, and Markus Wagner. 2018. "Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?" *Journal of Politics* 80 (1): 321–26.

Folke, Olle, and Johanna Rickne. 2016. "Electoral Competition and Gender Differences in Political Careers." *Quarterly Journal of Political Science* 11 (1): 59–102.

Folke, Olle, and Johanna Rickne. 2020. "All the Single Ladies: Job Promotions and the Durability of Marriage." *American Economic Journal: Applied Economics* 12 (1): 260–87.

Foos, Florian, and Fabrizio Gilardi. 2020. "Does Exposure to Gender Role Models Increase Women's Political Ambition? A Field Experiment with Politicians." *Journal of Experimental Political Science* 7:157–66.

Fox, Richard L. 2005. "Congressional Elections: Where Are We on the Road to Gender Parity?" In Susan J. Carroll and Richard L. Fox, eds., *Gender and Elections: Shaping the Future of American Politics*. New York: Cambridge University Press.

Fox, Richard L., and Eric R. A. N. Smith. 1998. "The Role of Candidate Sex in Voter Decision-Making." *Political Psychology* 19 (2): 405–19.

Fulton, Sarah A. 2011. "Running Backwards and in High Heels." *Political Research Quarterly* 65 (2): 303–14.

Goyal, Tanushree. 2020. "Local Female Representation as a Pathway to Power: A Natural Experiment in India." Unpublished manuscript, SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id = 3590118.

Hainmueller, Jens, Daniel J. Hopkins, and Teppei Yamamoto. 2014. "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments." *Political Analysis* 22 (1): 1–30.

Hardy-Fanta, Carol. 2013. *Intersectionality and Politics: Recent Research on Gender, Race, and Political Representation in the United States*. London: Routledge.

Harris, Adam, Kristen Kao, and Ellen Lust. 2020. "Research Design for 'Levels Experiment.'" Unpublished manuscript.

Hayes, Danny, and Jennifer L. Lawless. 2015. "A Non-gendered Lens? Media, Voters, and Female Candidates in Contemporary Congressional Elections." *Perspectives on Politics* 13 (1): 95–118.

Henderson, John A., Logan Dancey, Stephen N. Goggin, Geoffrey Sheagley, and Alexander G. Theodoridis. 2019. "Primary Divisions: How Voters Evaluate Policy and Non-policy Differences between Co-party Politicians." https://clas.uiowa.edu/polisci/sites/clas.uiowa.edu.polisci/files/PAPER_Theodoridis.pdf.

Hills, Jill. 1981. "Candidates: The Impact of Gender." *Parliamentary Affairs* 34 (2): 221–28.

Hobolt, Sara B., and Toni Rodon. 2020. "Cross-Cutting Issues and Electoral Choice: EU Issue Voting in the Aftermath of the Brexit Referendum." *Journal of European Public Policy* 27 (2): 227–45.

Holman, Mirya R., Jennifer L. Merolla, and Elizabeth J. Zechmeister. 2016. "Terrorist Threat, Male Stereotypes, and Candidate Evaluations." *Political Research Quarterly* 69 (1): 134–47.

Holman, Mirya R., and Monica C. Schneider. 2018. "Gender, Race, and Political Ambition: How Intersectionality and Frames Influence Interest in Political Office." *Politics, Groups, and Identities* 6 (2): 264–80.

Hooks, Bell. 1982. "Ain't I a Woman?" *Das Argument* 24 (131): 134.

Hopkins, Daniel. 2009. "No More Wilder Effect, Never a Whitman Effect: When and Why Polls Mislead about Black and Female Candidates." *Journal of Politics* 71 (3): 769–81.

Hopkins, Daniel J. 2014. "Do Voting Criteria Differ across Levels of the U.S. Federal System? A Conjoint Experiment to Understand Voter Decision-Making." https://tessexperiments.org/study/hopkins365.

Horiuchi, Yusaku, Zachary D. Markovich, and Teppei Yamamoto. 2021. "Does Conjoint Analysis Mitigate Social Desirability Bias?" *Political Analysis* (forthcoming).

Horiuchi, Yusaku, Daniel M. Smith, and Teppei Yamamoto. 2020. "Identifying Voter Preferences for Politicians' Personal Attributes: A Conjoint Experiment in Japan." *Political Science Research and Methods* 8 (1): 75–91.

Horne, Will. 2020. "Do Voters Really Want Working Class Candidates?" Unpublished manuscript.

Huddy, Leonie, and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37 (1): 119–47.

Inter-Parliamentary Union. 2018. "Women in Parliaments: World and Regional Averages." https://www.ipu.org/.

Jeydel, Alana, and Andrew J. Taylor. 2003. "Are Women Legislators Less Effective? Evidence from the U.S. House in the 103rd–105th Congress." *Political Research Quarterly* 56 (1): 19–27.

Kage, Rieko, Frances Rosenbluth, and Seiki Tanaka. 2019. "What Explains Low Female Political Representation? Evidence from Survey Experiments in Japan." *Politics and Gender* 15 (2): 285–309.

Kalla, Joshua, and Ethan Porter. 2020. "Can the Political Ambition of Young Women Be Increased? Evidence from U.S. High School Students." Unpublished manuscript.

Kang, Woo Chang, Jill Sheppard, Feodor Snagovsky, and Nicholas Biddle. 2018. "Ethnic Identities and Descriptive Representation: A Conjoint Analysis of Vote Choice." Unpublished manuscript.

Kanthak, Kristin, and Jonathan Woon. 2015. "Women Don't Run? Election Aversion and Candidate Entry." *American Journal of Political Science* 59 (3): 595–612.

Kao, Kristen, and Lindsay J. Benstead. 2021. "Female Electability in the Arab World: The Advantages of Intersectionality." *Comparative Politics* 53 (3): 427–64.

Karpowitz, Christopher F., J. Quin Monson, and Jessica Robinson Preece. 2017. "How to Elect More Women: Gender and Candidate Success in a Field Experiment." *American Journal of Political Science* 61 (4): 927–43.

Kelley, Jonathan, and Ian McAllister. 1984. "Ballot Paper Cues and the Vote in Australia and Britain: Alphabetic Voting, Sex, and Title." *Public Opinion Quarterly* 48 (2): 452–66.

King, Amy, and Andrew Leigh. 2010. "Bias at the Ballot Box? Testing Whether Candidates' Gender Affects Their Vote." *Social Science Quarterly* 91 (2): 324–43.

Kirkland, Patricia A., and Alexander Coppock. 2018. "Candidate Choice without Party Labels: New Insights from Conjoint Survey Experiments." *Political Behavior* 40 (3): 571–91.

Krupnikov, Yanna, Spencer Piston, and Nichole M. Bauer. 2016. "Saving Face: Identifying Voter Responses to Black Candidates and Female Candidates." *Political Psychology* 37 (2): 253–73.

Lawless, Jennifer L., and Richard L. Fox. 2010. *It Still Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.

Lawless, Jennifer L., and Kathryn Pearson. 2008. "The Primary Reason for Women's Underrepresentation? Reevaluating the Conventional Wisdom." *Journal of Politics* 70 (1): 67–82.

Leeper, Thomas J., and Joshua Robison. 2020. "More Important, but for What Exactly? The Insignificant Role of Subjective Issue Importance in Vote Decisions." *Political Behavior* 42 (1): 239–59.

Lemi, Danielle Casarez. 2021. "Do Voters Prefer Just Any Descriptive Representative? The Case of Multiracial Candidates." *Perspectives on Politics* 19 (4): 1061–81.

Martin, Nicole S., and Scott Blinder. 2021. "Biases at the Ballot Box: How Multiple Forms of Voter Discrimination Impede the Descriptive and Substantive Representation of Ethnic Minority Groups." *Political Behavior* 43:1487–510.

McDermott, Monika L. 1997. "Voting Cues in Low-Information Elections: Candidate Gender as a Social Information Variable in Contemporary United States Elections." *American Journal of Political Science* 41 (1): 270–83.

McDermott, Monika L. 1998. "Race and Gender Cues in Low-Information Elections." *Political Research Quarterly* 51 (4): 895–918.

McDonald, Jared. 2019. "Avoiding the Hypothetical: Why 'Mirror Experiments' Are an Essential Part of Survey Research." *International Journal of Public Opinion Research* 32 (2): 266–83.

Mo, Cecilia Hyunjung. 2015. "The Consequences of Explicit and Implicit Gender Attitudes and Candidate Quality in the Calculations of Voters." *Political Behavior* 37 (2): 357–95.

Mügge, Liza, and Sara De Jong. 2013. "Intersectionalizing European Politics: Bridging Gender and Ethnicity." *Politics, Groups, and Identities* 1 (3): 380–89.

Mummolo, Jonathan, and Erik Peterson. 2019. "Demand Effects in Survey Experiments: An Empirical Assessment." *American Political Science Review* 113 (2): 517–29.

Ono, Yoshikuni, and Barry C. Burden. 2019. "The Contingent Effects of Candidate Sex on Voter Choice." *Political Behavior* 41:583–607.

Ono, Yoshikuni, and Masahiro Yamada. 2020. "Do Voters Prefer Gender Stereotypic Candidates? Evidence from a Conjoint Survey Experiment in Japan." *Political Science Research and Methods* 8 (3): 477–92.

Piliavin, Jane Allyn. 1987. "Age, Race, and Sex Similarity to Candidates and Voting Preference." *Journal of Applied Social Psychology* 17 (4): 351–68.

Plutzer, Eric, and John F. Zipp. 1996. "Identity Politics, Partisanship, and Voting for Women Candidates." *Public Opinion Quarterly* 60 (1): 30–57.

Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. Chicago: University of Chicago Press.

Preece, Jessica, and Olga Stoddard. 2015. "Why Women Don't Run: Experimental Evidence on Gender Differences in Political Competition Aversion." *Journal of Economic Behavior and Organization* 117:296–308.

Preece, Jessica Robinson, Olga Bogach Stoddard, and Rachel Fisher. 2016. "Run, Jane, Run! Gendered Responses to Political Party Recruitment." *Political Behavior* 38 (3): 561–77.

Saha, Sparsha, and Ana Catalano Weeks. 2020. "Ambitious Women: Gender and Voter Perceptions of Candidate Ambition." *Political Behavior* (forthcoming).

Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science* 46 (1): 20–34.

Sanbonmatsu, Kira, and Kathleen Dolan. 2009. "Do Gender Stereotypes Transcend Party?" *Political Research Quarterly* 62 (3): 485–94.

Schuler, Paul. 2020. "Gender Bias and Position Salience: A Theory and Test of Variable Political Gender Bias in a Single-Party Regime." Unpublished manuscript.

Sen, Maya. 2017. "How Political Signals Affect Public Support for Judicial Nominations: Evidence from a Conjoint Experiment." *Political Research Quarterly* 70:374–93.

Senninger, Roman, and Daniel Bischof. 2021. "Do Voters Want Domestic Politicians to Scrutinize the European Union?" *Political Science Research and Methods* (forthcoming).

Shaffner, Brian, and Jon Green. 2020. "What Attributes Do Democratic Primary Voters Value?" Data for Progress. https://www.dataforprogress.org/blog/2019/7/11/what-attributes-do-democratic-primary-voters-value.

Sigelman, Lee, and Carol Sigelman. 1982. "Sexism, Racism, and Ageism in Voting Behavior: An Experimental Analysis." *Social Psychology Quarterly* 45 (4): 263–69.

Silbermann, Rachel. 2015. "Gender Roles, Work-Life Balance, and Running for Office." *Quarterly Journal of Political Science* 10 (2): 123–53.

Simas, Elizabeth N., and Doug Murdoch. 2019. "'I Didn't Lie, I Misspoke': Voters' Responses to Questionable Campaign Claims." *Journal of Experimental Political Science* 7 (2): 75–88.

Smith, Jessi L., David Paul, and Rachel Paul. 2007. "No Place for a Woman: Evidence for Gender Bias in Evaluations of Presidential Candidates." *Basic and Applied Social Psychology* 29 (3): 225–33.

Teele, Dawn Langan, Joshua Kalla, and Frances Rosenbluth. 2018. "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* 112 (3): 525–41.

Thomsen, Danielle M. 2020. "Ideology and Gender in U.S. House Elections." *Political Behavior* 42:415–42.

Thomsen, Danielle M., and Michele L. Swers. 2017. "Which Women Can Run? Gender, Partisanship, and Candidate Donor Networks." *Political Research Quarterly* 70 (2): 449–63.

Tomz, Michael, and Robert P. Van Houweling. 2016. "Political Repositioning: A Conjoint Analysis." https://tomz.people.stanford.edu/sites/g/files/sbiybj4711/f/tomz-vh-repositioning-conjoint-2016-04.pdf.

Visconti, Giancarlo. 2020. "Ideological Preferences after Natural Disasters: Evidence from Chile." http://www.giancarlovisconti.com/pdfs/Visconti.2020.Floods.pdf.

Visconti, Giancarlo, and Isabela Mares. 2020. "Voting for the Lesser Evil: Evidence from a Conjoint Experiment in Romania." *Political Science Research and Methods* 8 (2): 315–28.

Vivyan, Nick, and Markus Wagner. 2015. "What Do Voters Want from Their Local MP?" *Political Quarterly* 86 (1): 33–40.

Welch, Susan, Margery M. Ambrosius, Janet Clark, and Robert Darcy. 1985. "The Effect of Candidate Gender on Electoral Outcomes in State Legislative Races." *Western Political Quarterly* 38 (3): 464–75.

Welch, Susan, Janet Clark, and R. Darcy. 1987. *Women, Elections, and Representation*. New York: Longman.

White, Ariel, Anton Strezhnev, Christopher Lucas, Dominika Kruszewska, and Connor Huff. 2018. "Investigator Characteristics and Respondent Behavior in Online Surveys." *Journal of Experimental Political Science* 5 (1): 56–67.

Wuest, Reto, and Jonas Pontusson. 2017. "Do Citizens Prefer Affluent Representatives? Evidence from a Survey Experiment in Switzerland." Unpublished manuscript, SSRN. https://www.ssrn.com/abstract=3077598.

Zipp, John F., and Eric Plutzer. 1985. "Gender Differences in Voting for Female Candidates: Evidence from the 1982 Election." *Public Opinion Quarterly* 49 (2): 179–97.

Zizzo, Daniel John. 2010. "Experimenter Demand Effects in Economic Experiments." *Experimental Economics* 13 (1): 75–98.