I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17376

Do Americans Favor Female or Male Politicians? Evidence from Experimental Elections

Panu Poutvaara
Andreas Graefe

OCTOBER 2024

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 17376

Do Americans Favor Female or Male Politicians? Evidence from Experimental Elections

Panu Poutvaara
LMU Munich, ifo Institute, CESifo and IZA

Andreas Graefe
Macromedia University of Applied Sciences

OCTOBER 2024

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Do Americans Favor Female or Male Politicians? Evidence from Experimental Elections

Women are severely underrepresented in American politics, especially among Republicans. This underrepresentation may result from women being less willing to run for office, from voter bias against women, or from political structures that make it more difficult for women to compete. Here we show how support for female candidates varies by voters' party affiliation and gender. We conducted experimental elections in which participants made their vote choices based solely on politicians' faces. When choosing between female and male candidates, Democrats, and especially Democratic women, preferred female candidates, while Republicans were equally likely to choose female and male candidates. These patterns held after controlling for respondents' education, age, and political knowledge, and for candidates' age, attractiveness, and perceived conservatism. Our findings suggest that voter bias against women cannot explain women's underrepresentation. On the contrary, American voters appear ready to further narrow the gender gap in politics.

JEL Classification: D72, J16, H23

Keywords: gender, elections, gender discrimination, political candidates, redistribution

Corresponding author:
Panu Poutvaara
Ifo Institute
Poschingerstr. 5
81679 Munich
Germany
E-mail: poutvaara@ifo.de

1 Introduction

Female politicians are more likely to implement policies that benefit women and children (Chattopadhyay & Duflo, 2004; Miller, 2008; Baskaran & Hessami, 2023). Yet, women remain severely underrepresented in American politics. Although the percentage of elected women has steadily increased over the past four decades, women account for only 28% of House and Senate members in the 118th Congress, seated in January 2023, 33% of state legislatures, and 30% of statewide elective offices (Fig. 1A). Supply side explanations conclude that women are less willing to run for office, which could result from lack of encouragement or self-confidence (Fox & Lawless, 2004; Fox & Lawless, 2010; Kanthak & Woon, 2015). Demand side explanations suggest voter bias against women (Baltrunaite et al., 2019; Baskaran & Hessami, 2018), who must be more qualified than men to even compete (Pearson & McGhee, 2013). Female disadvantage could also follow from women being held back by party leaders (Besley et al., 2017; Casas-Arce & Saiz, 2015). In addition, the incumbency advantage (Gelman & King, 1990) hurts women given their historically low representation.

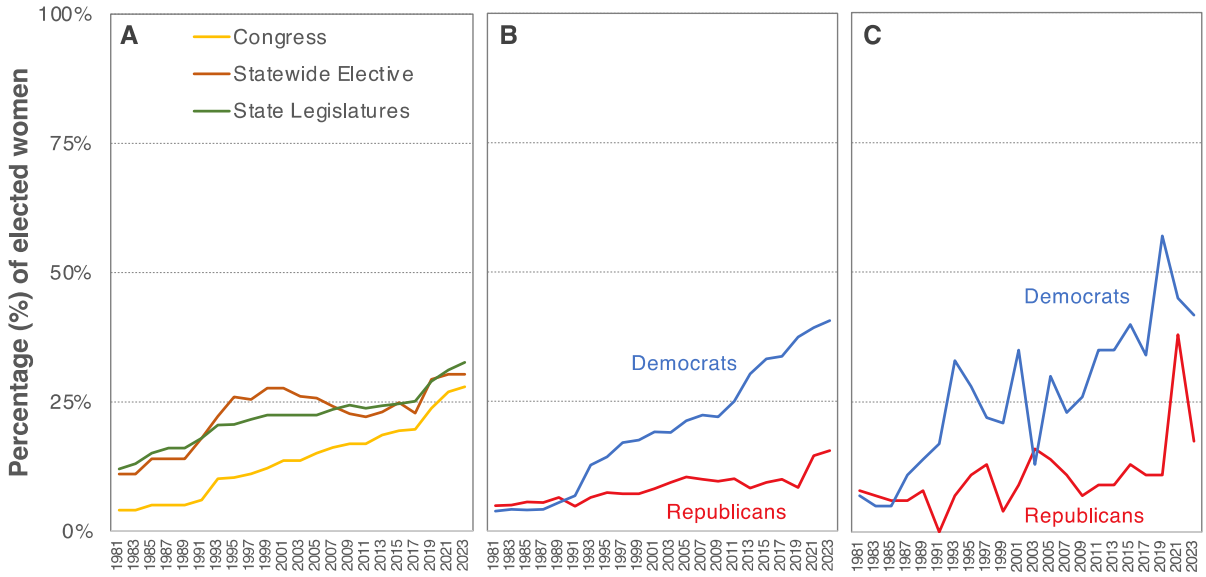

Fig. 1. Percentage of elected female politicians in the United States.
(A) In Congress, statewide elective offices (e.g., governor or lieutenant governor, attorney general, and secretary of state), and state legislatures. (B) Among Democrats and Republicans in Congress. (C) Among newly elected Democrats and Republicans in Congress. Data sources: Center for American Women and Politics, Eagleton Institute of Politics, Rutgers University; congress.gov; house.gov; senate.gov.

However, these forces do not seem to affect Democrats and Republicans equally, as a large gap has opened up between the two parties in terms of female representation (Fig. 1B). In

the 118th Congress, the percentage of women in the Democratic caucus (41%) is nearly three times that of the Republican caucus (15%), despite the surge of newly elected female Republicans in 2021 (Fig. 1C). Research offers several explanations for this partisan gap. On the supply side, compared to their Democratic counterparts, Republican women are less likely to run in districts favorable to the party (Pearson & McGhee, 2013), and are at a disadvantage in raising campaign funds (Bucchianeri, 2018; Thomsen & Swers, 2017). On the demand side, female politicians may simply appeal more to Democratic voters (Dolan, 2004; Sanbonmatsu, 2002), perhaps because Americans perceive women as more liberal than men, even within the same party (McDermott, 1997; Sanbonmatsu & Dolan, 2009). Another explanation is that voters tend to prefer candidates of their own gender (Sanbonmatsu, 2002). This should help female politicians among Democrats but hurt them among Republicans, as women identify more as Democrats than Republicans, while the opposite is true for men (Edlund & Pande, 2002; Gillion et al., 2020). A meta-analysis of vote choice experiments found that, on average, being described as a woman increased a candidate's vote share by four percentage points among Democrats, but decreased it by one point among Republicans (Schwarz & Coppock, 2021). However, the estimated effects of gender vary widely, with some studies finding no effect of gender (Hainmueller et al., 2014) or a small disadvantage for women (Ono & Burden, 2019).

To disentangle the potentially interacting effects of voter gender and partisanship, we conducted experimental elections in which participants made vote choices based solely on politicians' faces. American respondents, recruited using Amazon Mechanical Turk (MTurk), were shown 99 pairs of randomly selected headshot photographs of politicians they did not know, and asked each time whom they would vote for as their representative (see Figs. A1 and A2). Respondents were told beforehand that the task involved voting, but not that the aim of the study was to test the influence of candidate gender on vote choice. To ensure that respondents did guess the purpose of the study, each experimental election was equally likely to depict two women, two men, a woman on the left and a man on the right, or a man on the left and a woman on the right. The same set of photographs of 736 elected Members of the European Parliament (MEPs), was used in two surveys, one in 2016 (N=293) and one in 2020 (N=436). The experimental vote tasks were followed by background questions, one of which was the respondent's preferred candidate in that year's presidential election. Answers to this question were used to classify Democratic and Republican voters.

We find that Democrats strongly prefer female candidates. The pattern is strongest among Democratic women, who are two to three times more likely to prefer a female candidate

to a male candidate in mixed-gender experimental elections. Republicans, on the other hand, are about equally likely to vote for men and women in experimental elections. The pattern that Democrats, and especially Democratic women, are much more likely to support female candidates than male candidates holds after controlling for perceptions of candidates' attractiveness and conservatism. Both Democrats and Republicans favor attractive-looking candidates. Republicans are more likely, while Democrats are less likely, to support candidates who look more conservative.

We also analyzed the heterogeneity in voting in experimental elections according to respondent characteristics. Older respondents are slightly more likely to support female candidates in experimental elections, while the effects of education are not statistically significant. However, there is an interesting difference according to political knowledge. Among Republicans, political knowledge is uncorrelated with gender preference. Among Democrats, however, higher political knowledge is correlated with a stronger preference for female candidates. This could be explained by Democrats with higher political knowledge being more aware that women are, on average, more liberal. Another interpretation relates to the ideal of group representation: Democrats may favor women in order to reduce the gender imbalance in politics.

The remainder of the paper is organized as follows. Section 2 presents a theoretical framework that distinguishes the role of candidate gender from two other thin slices of information, candidate attractiveness and perceived conservatism, in a low-information election. We use this framework to derive testable predictions about how support for female candidates differs between female and male Democrats and Republicans, and how differences in candidate attractiveness and perceived conservatism are related to their support. Section 3 presents our data, and Section 4 presents the empirical results. Section 5 analyzes possible explanations for our findings. Section 6 concludes.

2 Theoretical framework

We analyze voting in a low-information setting where voters and candidates differ in their gender, ideology, and non-ideological characteristics. Voters belong to one of two political parties, Democrats on the left and Republicans on the right. When choosing whom to vote for, Democrats would prefer to vote for a liberal candidate and Republicans would prefer to vote for a conservative candidate, but they do not observe the candidates' true ideology. Instead,

voters infer ideology from candidate photographs, particularly from the candidate's gender, the candidate's attractiveness, and how conservative the candidate looks. Voters may also value attractive looks either because of taste-based discrimination, consistent with the beauty premium in the labor market (Hamermesh & Biddle, 1994), or because of a halo effect, whereby good looks are associated with, for example, competence or intelligence (Langlois et al., 2000).[1]

We denote gender by g, which takes the value f for females and m for males. An indicator variable I_K for candidate K's gender obtains value 0 for males and 1 for females. Party is denoted by P, which can also take two values, D for Democrat and R for Republican. Voters' gender preference may depend on their party preference and their own gender. More specifically, we denote valuation of the candidate being female by $\varphi_{P,g}$ for supporters of party P of gender g. For voters who are indifferent to the politician's gender, $\varphi_{P,g}$ would be equal to zero and any advantage of female or male candidates would arise through differences in their attractiveness or perceived conservatism. If voters have an inherent gender preference, $\varphi_{P,g} < 0$ for groups favoring men and $\varphi_{P,g} > 0$ for groups favoring women. Attractiveness of candidate K is denoted by a_K. We denote the valuation of attractiveness by supporters of party P by β_P, with our prior assuming that this beauty premium is positive for both parties (i.e., voters generally prefer better-looking candidates). Perceived conservatism of candidate K is denoted by c_K. Its valuation by supporters of party P is denoted by κ_P, with $\kappa_D < 0 < \kappa_R$.

We analyze probabilistic voting with two candidates, X and Y. As in Lindbeck and Weibull (1987, 1993) and Dixit and Londregan (1996), voters differ in a continuously distributed term whose ex-ante distribution is known but whose realization is unknown to parties or candidates. We analyze voting by individual i who belongs to group J, in which J is defined by a combination of party P, $P \in \{D, R\}$, and gender g, $g \in \{f, g\}$. As Persson and Tabellini (2002), we model the random component as a voter-specific popularity parameter γ_{iJ} that measures voter i's who belongs to group J idiosyncratic preference for candidate X, relative to candidate Y.[2] We analyze choices of voter i who belongs to party P and is of gender g. Voter i's expected utility from voting for candidate X is

$$EU_X^i = \varphi_{P,g} I_X + \beta_P a_X + \kappa_P c_X + \gamma_{iJ}.$$

Voter i's expected utility from voting for candidate Y is

$$EU_Y^i = \varphi_{P,g} I_Y + \beta_P a_Y + \kappa_P c_Y .$$

[1] Berggren et al. (2010) show that political candidates with more attractive looks are also generally perceived to be more competent, intelligent, likeable, and trustworthy in ratings based solely on photographs.
[2] Theoretical predictions would remain qualitatively similar if we assumed that there is a separate random term associated with each candidate.

Voter i maximizes his or her expected utility, and votes for X if $EU_X^i > EU_Y^i$. We introduce notation $\tilde{a} = a_X - a_Y$ for the extent to which candidate X looks more attractive than candidate Y, with negative values indicating the extent to which Y looks more attractive, and $\tilde{c} = c_X - c_Y$ for the extent to which candidate X looks more conservative than candidate Y, with negative values indicating the extent to which Y looks more conservative. $EU_X^i > EU_Y^i$ is equivalent to

$$\gamma_{iJ} > \varphi_{P,g}(I_Y - I_X) - \beta_P \tilde{a} - \kappa_P \tilde{c} = \hat{\gamma}_{iJ}(I_X, I_Y, \tilde{a}, \tilde{c}).$$

The ex-ante probability that voter i votes for candidate X, measured before the realization of γ_{iJ} is revealed, is denoted by q_{iJ}. As Persson and Tabellini (2002), we assume that it follows a symmetric uniform distribution around zero, which is sufficiently wide so that $0 < q_{iJ} < 1$. This assumption rules out corner solutions and allows deriving closed-form solutions. Denoting the range of the distribution by $[-\frac{1}{2\Gamma}, \frac{1}{2\Gamma}]$ with Γ being the density function for γ_{iJ}, the ex-ante probability that voter i in group J votes for candidate X is given by

$$(1)\ q_{iJ}(I_X, I_Y, \tilde{a}, \tilde{c}) = \int_{\hat{\gamma}_{iJ}}^{\frac{1}{2\Gamma}} \Gamma d\gamma_{iJ} = \frac{1}{2} + \Gamma\left[\varphi_{P,g}(I_X - I_Y) + \beta_P \tilde{a} + \kappa_P \tilde{c}\right].$$

We can now derive

Proposition 1. $\forall I_X, I_Y, \tilde{a}, \tilde{c}$:

(i) $q_{iJ}(1, I_Y, \tilde{a}, \tilde{c}) - q_{iJ}(0, I_Y, \tilde{a}, \tilde{c}) = q_{iJ}(I_X, 0, \tilde{a}, \tilde{c}) - q_{iJ}(I_X, 1, \tilde{a}, \tilde{c}) = \Gamma\varphi_{P,g}$;

(ii) $\frac{\partial q_{iJ}(I_X, I_Y, \tilde{a}, \tilde{c})}{\partial \tilde{a}} = \Gamma\beta_P$;

(iii) $\frac{\partial q_{iJ}(I_X, I_Y, \tilde{a}, \tilde{c})}{\partial \tilde{c}} = \Gamma\kappa_P$.

Proof. (i) follows by inserting equation (1) with different values of the gender indices and (ii) and (iii) follow by differentiating equation (1).

Proposition 1 shows how voters of a given type respond to changes in candidate characteristics. Part (i) establishes how preference for a candidate being female among voters of gender g among supporters of party P translates into difference in the probability of supporting a given candidate, once controlling for candidates' perceived attractiveness and conservatism. Part (ii) derives how party-specific attractiveness premium translates into votes,

and part (iii) how party-specific preference for or against conservative-looking candidates translates into votes.

Previous research has concluded that female politicians appeal more to the Democratic electorate (McDermott, 1997; Sanbonmatsu & Dolan, 2009), suggesting that $\varphi_{D,0} > \varphi_{R,0}$ and $\varphi_{D,1} > \varphi_{R,1}$. Also, some research has concluded that voters tend to prefer candidates of the same gender (Sanbonmatsu, 2002), suggesting $\varphi_{D,1} > \varphi_{D,0}$ and $\varphi_{R,1} > \varphi_{R,0}$. Therefore, we expect Democrats of either gender to be more likely to vote for a female candidate in a mixed-gender race than Republicans of the same gender. Furthermore, we expect Democratic women to be more likely to select female candidates than Democratic men, and Republican women to be more likely to select female candidates than Republican men. Importantly, our model leaves it open whether a given group of voters is more likely to select female or male candidates and presents predictions only on how support for female candidates differs by voter gender and partisanship. Evaluating the overall gender preference is left to the empirical analysis. Our main testable hypotheses are as follows[3]:

Hypothesis 1. Democrats of either gender are more likely to vote for women than Republicans of the same gender.

Hypothesis 2. Men tend to support male candidates more often than women.

Hypothesis 3. Women tend to support female candidates more often than men.

Previous research also found that both conservative and liberal voters prefer more attractive-looking candidates but that the beauty premium is larger for candidates on the right in both low-information real elections and in experimental elections (Berggren et al., 2017). Therefore, our prior was that $\beta_R > \beta_D > 0$. However, we do not assume that in our model, and leave the sign and the magnitude of β_D and β_R to be estimated in the econometric analysis. Since Olivola et al. (2018) found that Republicans are more likely to support candidates who appear more conservative, we further expect that Republicans are more likely and Democrats

[3] We pre-registered these hypotheses regarding mixed-gender elections in AEA RCT Registry, with RCT ID AEARCTR-0006653. In the pre-registration plan, we operationalized Hypothesis 1 in two ways: with the current formulation, which relies on party identification, and in terms of presidential vote preference as "those supporting Biden of either gender are more likely to vote for a female candidate than those supporting Trump of the same gender)". We present evidence for both formulations in the empirical results section.

less likely to support candidates who appear more conservative. Therefore, our prior is that $\kappa_D < 0 < \kappa_R$.

The most appropriate application of our model to US politics is in primary elections, where candidates do not differ in their party labels. In general elections, it is plausible that voters would also use partisanship as an additional cue to ideology which we do not include in our model. Given that most House districts tend to repeatedly elect either a Democrat or a Republican, it may well be that the gender gap in American politics arises primarily from gender preferences in intra-party nomination contests, especially in primaries. At the cost of making the model more complicated, it would be possible to extend our model to also include party labels. In this case, strong partisans would likely always or almost always vote for a candidate from their preferred party, regardless of other factors. On the other hand, independents and voters with weak party preferences might consider candidates from different parties. They may lean toward one party, but could be swayed to support a candidate from the other party if that candidate is perceived as favorable in terms of gender, attractiveness, and perceived conservatism – enough to outweigh the partisan disadvantage.

3 Data

Our analysis is based on two online surveys conducted in April 2016 and October 2020, respectively. The Ethics Commission of the Department of Economics at the University of Munich approved the research, with decision numbers 2015-03 (Study I) and 2020-09 (Study II). Study II was pre-registered at the American Economic Association's registry for randomized trials on 22 October 2020 (AEARCTR-0006653). Respondents in both studies gave their informed consent prior to participation.

3.1 Stimulus material

Both studies used the same database of photographs, which consisted of official headshots of all 736 elected Members of the European Parliament (MEPs) from the 2009–2014 parliamentary term, collected from the web page of the European Parliament. Thereby, 256 (35%) were female and 480 were male politicians. This database had already been used in prior research (Berggren et al., 2017), which obtained ratings of the MEPs' perceived attractiveness from 296, and ratings of perceived conservatism from 292, American respondents using MTurk. Table 1 presents descriptive statistics of the MEPs in the photographs. On average, female

MEPs are evaluated as more attractive than male MEPs, and male MEPs are perceived as more conservative. The average age of male MEPs is 52 years and of female MEPs 50 years.

Table 1. Members of European Parliament (MEPs): descriptive statistics.

	Male (N=480)	**Female (N=256)**
Attractiveness	-0.17	0.33
	(0.90)	(1.10)
Perceived conservatism	0.38	-0.71
	(0.92)	(0.71)
Age	52	50
	(11)	(10)

Table shows average standardized ratings for attractiveness and perceived conservatism as well as age (standard deviation in parentheses) per MEP gender.

In study I, respondents were asked whether they recognized any of the persons shown in the photographs. As in prior research (Berggren et al., 2017), the likelihood that American respondents would recognize any of the MEPs was virtually zero. Only one respondent recognized any of the politicians (and only one). We therefore did not ask this question in study II. One study II respondent nevertheless wrote in the write-in comments that one person looked familiar and guessed correctly that the photographs could depict members of the European Parliament.

3.2 Questionnaire

The questionnaires in both studies were programmed using SoSci Survey software. After a short introduction, respondents were asked to complete the experimental voting task. Respondents in both studies saw 99 pairs of candidates, randomly selected from all 736 photographs. Thereby, each pair was equally likely to depict two males, two females, a female on the left and a male on the right, and a male on the left and a female on the right. For each pair, respondents were asked for which person they would vote as a member of the House of Representatives if they would have to decide based on photographs only. In study I (Fig. A1), respondents were asked if they would vote for the person on the left or on the right. In study II

(Fig. A2), respondents answered the question using a scale from 1 ("definitely the person on the left") to 6 ("definitely the person on the right"). In both studies, respondents could opt to abstain or prefer not to answer.

After making 99 vote choices, respondents were asked three policy questions, followed by which presidential candidate they would prefer to see elected as President. Then, respondents answered four political knowledge questions, and were asked to predict the presidential candidates' national vote shares. Finally, respondents were asked for gender, age, nationality, and education. In study II, respondents were also asked whether they think of themselves as a Republican, a Democrat, or an Independent, if they are registered to vote in the 2020 U.S. presidential election and, if so, in which state. Respondents could also provide comments on the survey or request a summary of the results before they received their MTurk completion code. The full questionnaires are available in the Online Appendix.

3.3 Participant recruiting

Participants in both studies were recruited through the Amazon Mechanical Turk (MTurk) platform.

Study I. We obtained 293 questionnaires 2-11 April 2016. Respondents had to be located in the U.S., had to have at least 1,000 MTurk tasks approved, as well as an approval rate of higher than 95%. Participants were paid $2.14 for completing the questionnaire, which took 15 minutes (SD = 5 minutes) on average. This would translate to an hourly wage of $8.7, which was 20% above the minimum wage in the U.S. at the time of the survey. Including fees paid to MTurk, we paid $3.07 per respondent.

Study II. Our plan was to collect 360 questionnaires. Respondents had to be located in the U.S. and needed an approval rate of 90% or higher for previously completed MTurk tasks. We paid for an additional qualification to make sure that our sample contained 120 respondents who registered their political affiliation at MTurk as conservative, as well as another 120 respondents who registered their political affiliation at MTurk as liberal, with the total being 436 questionnaires obtained 23-27 October 2020.[4] Our online simulation tool,

[4] We rejected approval from one respondent who did not provide a completion code, six respondents who provided an incorrect code, and 19 respondents due to speeding. Speeders were identified based on a score calculated directly by the Sosci Survey software. This score, called DEG_TIME, is normalized so that values greater than 100 indicate low quality data, although the software recommends lower scores for stricter filtering.

described below in subsection 4.2, allows readers to verify that the estimated effects change only marginally if implementing any combination of possible exclusion criteria. Participants were paid $2.1 for completing the questionnaire, which took 11 minutes (SD = 4 minutes) on average. This would translate to an hourly wage of $11.5, which was 59% above the minimum wage in the U.S. at the time of the survey. Including the fees paid to MTurk, we paid $2.65 per respondent.

Six respondents reported a nationality other than American, but we did not exclude them as they could be naturalized citizens who feel more attached to another country. A handful of people guessed in the comments that we are interested in how gender, age and/or appearance affects vote choice, but none guessed that we are interested in how respondents' political views affect their choice between men and women. We did not exclude any respondents based on write-in comments.

3.4 Participants

Table 2 presents summary statistics of all respondents in the two surveys. Of the 293 respondents in *study I*, 49% (N=144) were female. Respondents' average age was 40 years (SD = 12.6). A total of 62% said that they would vote for a Democratic candidate (either Hillary Clinton or Bernie Sanders) in the 2016 presidential election, while 33% said that they would vote for one of three Republican candidates (Donald Trump, Ted Cruz, or John Kasich). The share of Democratic respondents in our survey was somewhat higher compared to the general population. For example, a Gallup poll conducted April 6-10, 2016, around the same time as our survey, found that 49% of respondents identified as Democrats/Democratic leaners and 41% as Republicans/Republican leaners (Gallup Organization, 2016).

We rejected respondents with a score greater than 100. Another 12 (38) respondents did not meet our sampling criteria, because MTurk listed them as liberal (conservative) while they answered that they identified as Republican (Democrat). We initially decided to replace these respondents, due to concerns about the reliability of their reported political affiliation. However, because we did not specify exclusion criteria for respondents in our pre-registration, we later decided to keep all respondents in the analysis. This resulted in an oversampling of 76 questionnaires.

Table 2. Summary statistics for respondents

	Study I 2016	Study II 2020
Number of respondents	293	436
Gender		
Female	49.1%	44.3%
Male	50.9%	55.7%
Age	40.3 (sd=12.6)	38.5 (sd=12.3)
Presidential preference		
Democrat	61.8%	51.4%
Republican	33.1%	46.3%
Other	5.1%	2.3%
Party identification		
Democrat	na	50.7%
Republican	na	38.3%
Independent	na	9.2%
Other	na	1.8%
Education		
High school or less	15.7%	7.6%
Associate's degree or some college	33.4%	20.4%
Bachelor's degree	36.9%	53.4%
Master's degree or higher	14.0%	18.6%
Political knowledge	1.8 (sd=1.2)	2 (sd=1.2)

Table displays the distribution of respondents along demographic and political variables in studies I and II. Distributions of gender, presidential preference, party identification, and education are presented as column percentages, and age and political knowledge as average (standard deviation in parentheses). Distributions of presidential preference and party identification are shown excluding abstentions.

Of the 436 respondents in *study II*, 44% (N=193) were female. Respondents' average age was 39 years (SD = 12.3). A total of 51.4% said that they would vote for Joe Biden in the 2020 presidential election, while 46.3% said that they would vote for Donald Trump. This almost perfectly matched the final election outcome (Biden 51.3% vs. Trump 46.9%). With respect to political affiliation, 51% identified as Democrats (including Independent leaners), 38% as Republicans (including Independent leaners), and 9% as Independents. In comparison,

a Gallup poll conducted around the same time of our survey found that 49% of respondents identified as Democrats/Democratic leaners and 45% as Republicans/Republican leaners (Gallup Organization, 2020).

4 Results

4.1 Descriptive evidence on gender preference

We start by presenting descriptive evidence on female vote by male and female Republicans and Democrats. As respondents in each group were presented large numbers of randomly chosen pairs of photographs using the same randomization rule and set of photographs, the share of female vote should not vary systematically if the groups would not differ in their gender preference. Therefore, any systematic differences would be indicative of group-level differences in tendency to choose a female candidate in experimental elections.

As shown in Fig. 2, Democrats favored female candidates in mixed-gender races, and the effect was particularly strong for female Democrats. Republicans, regardless of their gender, chose female candidates about half of the time, thereby not favoring either gender. Average support for female candidates remained stable across male and female voters of both parties from 2016 to 2020, ranging from 42% (95%-confidence: 35–49%) among Republican men in 2016 to 75% (95%-confidence: 71–79%) among Democratic women in 2020. The picture is similar when comparing respondents based on their stated party identification (Fig. A3) as well as when accounting for the strength of the reported choice (Fig. A4), each of which was collected in 2020 only. Taken together, these figures confirm Hypothesis 1 that Democrats of either gender are more likely to vote for women than Republicans of the same gender, both when using presidential voting intentions and when using stated party identification.

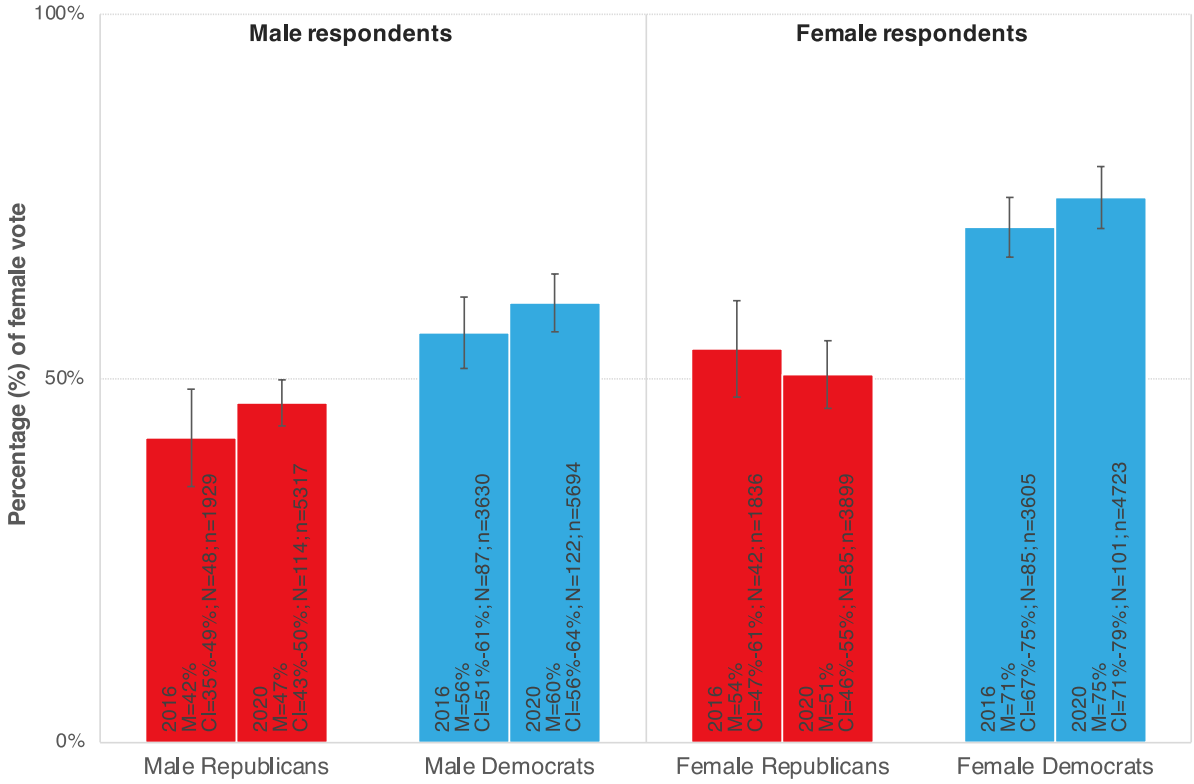

Fig. 2. Percentage of female vote by respondent gender and partisanship.
Female vote shown separately for studies I and II. Partisanship measured based on presidential preference. Results include only experimental elections with one female and one male candidate, excluding abstentions. N: number of respondents; n: number of experimental elections. Error bars depict 95% CI of the mean.

Fig. 3 compares female voting by respondent gender, without controlling for respondent partisanship. Female respondents chose female candidates almost twice as often as male candidates, while male respondents chose female and male candidates about equally often. This figure confirms Hypothesis 2 that men tend to support male candidates more often than women and Hypothesis 3 that women tend to support female candidates more often than men. The gap in support for female candidates between female and male respondents is similar in size to the gap in support for female candidates between Democratic and Republican men, but smaller than the gap in support for female candidates between Democratic and Republican women. Taken together, our descriptive evidence suggests that women and Democrats strongly favor female candidates in low-information experimental elections, while men and Republicans do not exhibit strong gender preference in low-information experimental elections, on average.

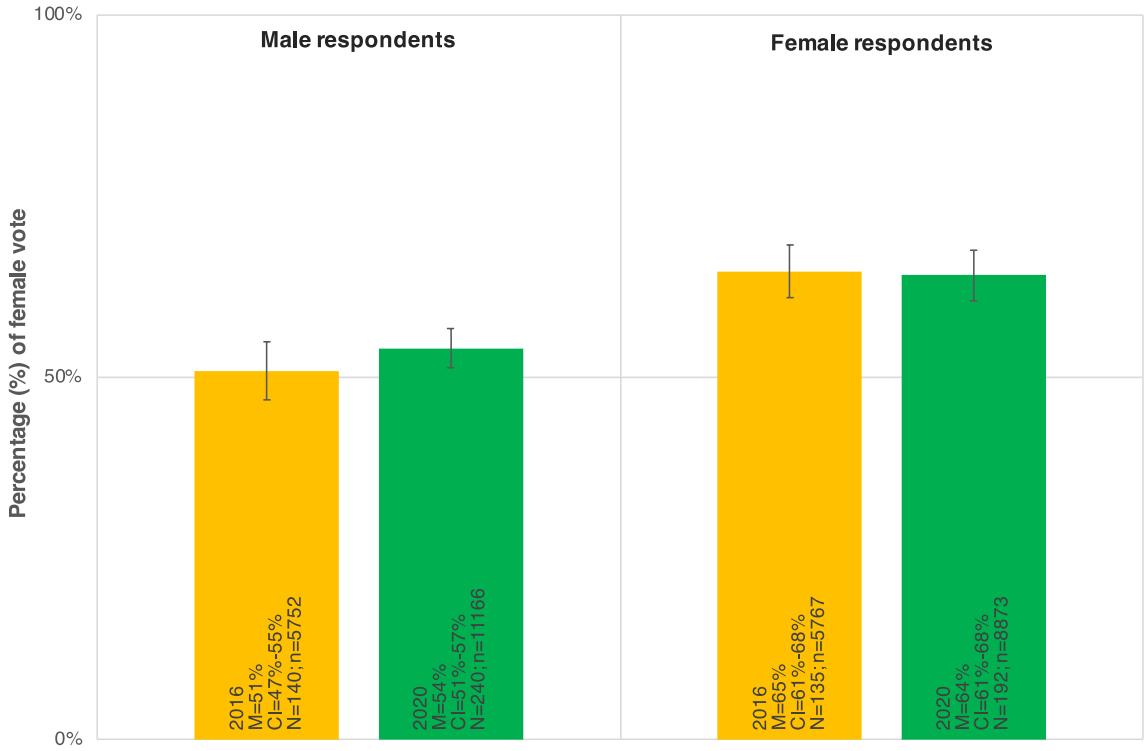

Fig. 3. Percentage of female vote by respondent gender.

Female vote shown separately for studies I and II. Results include only experimental elections with one female and one male candidate, excluding abstentions. N: number of respondents; n: number of experimental elections. Error bars depict 95% CI of the mean.

Prior research suggests that voters tend to vote for more attractive candidates (Ahler et al., 2017; Berggren et al., 2010, 2017), and use candidate appearance as a cue to ideology (Berggren et al., 2017; Olivola & Todorov, 2010; Rule & Ambady, 2010). To address potential concerns that Democrats may prefer female candidates simply because women may look better, or more liberal, we took into account evaluations of each MEP's attractiveness and perceived conservatism, obtained in prior research (Berggren et al., 2017). The partisan gender gap prevailed when analyzing separately experimental elections in which the female candidate looked more attractive (Fig. A5) or less attractive (Fig. A6). Similarly, partisan gender gap remains when analyzing separately experimental elections in which the female (Fig. A7) or the male (Fig. A8) candidate looked more conservative. In each case, female Democrats clearly favored female candidates.

4.2 Econometric evidence on gender preference

We used ordinary least squares (OLS) regression to analyze how vote choice relates to respondent and candidate characteristics, as recommended when estimating treatment effects on binary outcomes (Gomila, 2020). All variables and their coding are described in Table A1. Our dependent variable *vote choice* is a dummy variable, coded as 1 if the respondent chose the female candidate in experimental elections with a male and a female candidate, and 0 otherwise. Experimental elections in which the respondent abstained were excluded. Although the assumption that errors are normally distributed is violated because of a binary outcome variable, this should not be a large problem given our sample size (Hellevik, 2009). To account for heteroscedasticity and correlation of the outcome variable within respondents, we used robust standard errors clustered at the respondent level (N= 684).

Table 3 presents the results for the support for female candidate in the mixed-gender elections, using four different specifications. In model 1, we used only respondent characteristics related to gender, partisanship defined as Democrat for those supporting Democratic presidential candidate and Republican for those supporting Republican presidential candidate in the upcoming presidential election, age, level of education, and time dummy for those who responded in 2016. In model 2, we added a measure of the two candidates' attractiveness difference, measured as average attractiveness score of the female candidate minus the average attractiveness score of the male candidate, and its interaction with Democrat. In model 3, we added a measure of the two candidates' difference in perceived conservatism, measured as average perceived conservatism of the female candidate minus the average perceived conservatism of the male candidate, as well as the age difference between the two candidates, defined as the age of the female candidate minus the age of the male candidate, and the interactions of both measures with the respondent being Democrat. In the model 4, we added a measure of respondent's political knowledge and its interaction with respondent being Democrat. Regressions were run using STATA/MP 16.0. Pre-registered conjecture that Democrats of either gender are more likely to vote for women than Republicans of the same gender was tested against one-sided null hypothesis that this is not the case with an F-test, separately for females and males.

Table 3. OLS regression analysis on the determinants of voting for the female candidate in mixed-gender elections (Party classification based on Presidential voting intention)

	1	2	3	4
Democrat	0.146***	0.149***	0.108***	0.100***
	(0.023)	(0.023)	(0.023)	(0.022)
Female	0.064*	0.066*	0.066*	0.067*
	(0.026)	(0.026)	(0.026)	(0.026)
Democrat x Female	0.076*	0.073*	0.073*	0.080*
	(0.034)	(0.034)	(0.034)	(0.033)
Age	0.002**	0.002**	0.002**	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
High school or less	-0.023	-0.026	-0.026	-0.015
	(0.028)	(0.028)	(0.028)	(0.028)
Associate degree or some college	-0.011	-0.012	-0.012	0.000
	(0.022)	(0.022)	(0.022)	(0.022)
Master's degree or higher	-0.001	-0.002	-0.001	0.002
	(0.021)	(0.021)	(0.021)	(0.021)
Responded in 2016 (Study I)	-0.027	-0.027	-0.026	-0.018
	(0.018)	(0.018)	(0.018)	(0.018)
Female more attractive		0.061***	0.062***	0.063***
		(0.005)	(0.005)	(0.005)
Female more attractive x Democrat		-0.005	-0.006	-0.006
		(0.006)	(0.006)	(0.006)
Female looks more conservative			0.025***	0.025***
			(0.004)	(0.004)
Female looks more conservative x Democrat			-0.037***	-0.038***
			(0.006)	(0.006)
Female older			-0.000	-0.000
			(0.000)	(0.000)
Female older x Democrat			0.000	0.000
			(0.001)	(0.001)
Political knowledge				-0.007
				(0.011)
Political knowledge x Democrat				0.047**
				(0.014)
Constant	0.461***	0.429***	0.456***	0.449***
	(0.018)	(0.018)	(0.018)	(0.018)
R squared	0.05	0.08	0.08	0.08
Number of respondents	683	683	683	683
Number of experimental elections	30633	30633	30633	30633

Each observation is an experimental election between a female and a male candidate, excluding experimental elections in which the respondent abstained. Dependent variable equals 1 if the respondent chose the female candidate and 0 if the respondent chose the male candidate. Robust standard errors are clustered at the respondent level. Respondent reference categories (not shown) are Republican, Male, Bachelor's degree and Responded in 2020 (study II).

The conjecture that Democrats of either gender are more likely to support female candidates than Republicans of the same gender received strong support in all models. According to model 1, Democratic men were 15 percentage points more likely than Republican men to vote for female candidates (estimated change in probability 0.146; $P < 0.001$; $F(1,683)$

= 42.27) and Democratic women were 22 percentage points more likely than Republican women to vote for female candidates (estimated change in probability 0.222; $P < 0.001$; $F(1,683) = 73.84$). Column 2 shows that the results remain when controlling for the difference in candidate attractiveness. Respondent education has no statistically significant effects. Older respondents are somewhat likelier to support female candidates than younger respondents, although the magnitude of this effect is small. Moving from column 1 to column 2 confirms findings from previous research: respondents were more likely to select more attractive candidates (Ahler et al., 2017; Berggren et al., 2010, 2017). The regressions also confirm that Republicans tended to vote for women about as often as for men, as already suggested by Fig. 2. The reference person (male Republican with bachelor's degree in 2020) voted for female candidate in 46% of mixed-gender races in column 1, and female Republicans in 52% of mixed-gender races.

Column 3 adds perceived conservatism and age difference between candidates. Republicans tended to support candidates who looked more conservative as in Olivola et al. (2018), while Democrats tended to support candidates who looked more liberal. Candidate age difference has no effect on average vote shares. Controlling for perceived conservatism somewhat reduces the effect of gender, but most of the gender gap remains: male Democrats are 11 percentage point more likely to support female candidate than male Republicans, and female Democrats 18 percentage point more likely than female Republicans. Finally, column 4 shows that the effects of political knowledge on gender preference varied by party. While political knowledge did not affect gender preference of Republicans, Democrats who were more knowledgeable about US politics were more likely to support female candidates.

In our pre-registration, we did not list hypotheses with respect to perceived conservatism. Therefore, the estimates in columns 1 and 2 are the best test of our pre-registered hypotheses. Importantly, controlling for attractiveness leaves the estimated effect of gender almost unchanged. This is in line with Proposition 1 deriving the effects of gender and attractiveness that are independent of each other. Comparing columns 2 and 3 allows evaluating how our testable hypotheses can be expected to change if voters have access to only two thin slices of information, candidate gender and attractiveness. In that case, we would expect that both could be used as a cue for candidate ideology. The empirical results confirm this clearly for gender: if perceived conservatism is not controlled for, the estimate for differential support for female candidates among Democrats compared with among Republicans is somewhat higher. The difference in gender preference between male Democrats and male Republicans in

our model, $\Gamma(\varphi_{D,0} - \varphi_{R,0})$, would correspond to the estimated coefficient for Democrat in Table 3, and the difference in gender preference between female Democrats and female Republicans in our model, $\Gamma(\varphi_{D,1} - \varphi_{R,1})$, would correspond to the sum of the estimated coefficients for Democrat and Democrat x Female in Table 3. Our theoretical model corresponds most closely to column 3, which includes perceived conservatism, and has a value of 0.11 for Democrat (gap among men) and 0.18 for Democrat + Democrat x Female (gap among women). Omitting perceived conservatism is equivalent to moving to column 2, and increases the partisan gap to 0.15 for men and 0.22 for women. Thus, omitting conservatism as an additional thin slice of information increases gender preference for female candidates by about four percentage points both for male Democrats relative to male Republicans and for female Democrats relative to female Republicans. This can be explained by women being generally perceived as less conservative, as shown in Table 1 for evaluations of MEPs in our photographs, and McDermott (1997) and Sanbonmatsu & Dolan (2009) for perceptions of female and male politicians in the United States.

As Berggren et al. (2017) showed that more attractive-looking candidates are perceived as more conservative, we also expected that not controlling for perceived conservatism could increase the estimated rewards for attractiveness among Republicans, but reduce it among Democrats. This conjecture was not supported, as the estimated returns to attractiveness are similar in columns 2 and 3, and remain similar for Democrats and Republicans independently of whether perceived conservatism is controlled for.

The effects of respondent gender and partisanship are robust with respect to researcher decisions. In Table A2, we present corresponding analysis based on self-reported party identification, available only in 2020. If anything, the effects of respondent ideology are even stronger, at least for women. In columns 1 and 2, female Democrats are 28 percentage points more likely to vote for female candidate in an experimental election than female Republicans. In columns 3 and 4, the gap is still about 25 percentage points. While female Democrats are 13 percentage points more likely to support a female candidate than male Democrats, there is no statistically significant difference in support for female candidates between female and male Republicans.

In an effort to provide full transparency, we created an online simulation tool that allows readers to interact with our complete data (tinyurl.com/femalevote). For example, readers can estimate the likelihood of a female vote based on voter and candidate characteristics, and test

the effects of introducing data cleaning rules to reject respondents who were identified as speeders or using different regression model specifications.

4.3 Gender voting and ideological positions

We also carried out exploratory analyses on gender preference according to ideological positions in three central dimensions of contemporary American politics: support for redistribution, environmental policies, and social issues. We asked respondents whether they support increasing taxes on the rich and redistributing money to those with low incomes, whether they support increasing taxes on gasoline and using the money to protect the environment, as well as whether they support same-sex marriage or registered partnership. In each issue, Republicans were more conservative than Democrats (Fig. A9), and respondents who took liberal positions were more likely to vote for female candidates in experimental elections (Fig. A10). These patterns suggest that the partisan gap in support for female candidates coincides with a gap between conservatives and liberals in economic, environmental, and social issues.

4.4 Are Republicans more confident in voting choices based on candidate looks?

Previous research has suggested that conservatives tend to rely more heavily on stereotypes than liberals (Olivola et al., 2012; Olivola et al., 2018). To study whether this is the case also in our setting, we use the feature that we collected vote choice in experimental elections in 2020 using a six-point scale, from voting definitely for the candidate on the left to voting definitely for the candidate on the right. In our pre-registration, we listed two hypotheses related to the strength of respondent choices in experimental elections: first that Republicans are, on average, more certain in their choice than Democrats, and second that Trump supporters are, on average, more certain in their choice than Biden supporters, at least when it comes to choices between two candidates of the same gender. In Figs A11 and A12, we present the evidence on these hypotheses, separately for mixed-gender elections, elections between two males, and elections between two females. In each case, a considerably higher share of Republicans than of Democrats makes the choice to vote definitely for the candidate on the left or vote definitely for the candidate on the right, whether Democrats and Republicans are identified based on self-reported party identification or based on their intention to vote for Biden or Trump. Therefore, also these two additional pre-registered hypotheses are confirmed.

5 Mechanisms

5.1 Insights from same-gender elections

Although our primary focus is on mixed-gender elections, we also present analyses on same-gender elections, to shed further light on the use of attractiveness and perceived conservatism as voting cues. Table A3 analyzes the probability of voting for the more attractive-looking candidate in experimental elections between two male candidates and table A4 in experimental elections between two female candidates, when partisanship is defined based on presidential voting preference. In column 1, constant presents the baseline probability that the reference person (who is male Republican with a bachelor's degree who responded in 2020) votes for the more attractive-looking candidate. In column 2, we include as an additional explanatory variable the marginal effect of attractiveness gap between the more and the less attractive-looking candidates. Although the constant then drops below 0.5, more attractive-looking candidates still have a clear electoral advantage, which is now partly captured by *Attractiveness advantage*. The third column again adds difference in perceived conservatism and the fourth column political knowledge.

In our pre-registration, we also hypothesized that Republicans would, on average, be more likely to choose better-looking candidates than Democrats, and that Trump supporters would, on average, be more likely to choose better-looking candidate than Biden supporters, at least when choosing between two candidates of the same gender. The empirical results refute both of these hypotheses. Instead, we find that better-looking candidates have a similar advantage among Democrats and Republicans. This differs from the finding in Berggren et al. (2017) that politicians on the right benefit more from good looks in experimental elections in both Europe and the United States, whether using MTurk or student respondents, and also in low-information real elections, using performance in terms of the number of personal votes within the party list. Since Berggren et al. (2017) was written and uses data collected before the Trump presidency, which has profoundly changed the Republican Party in many ways, it is left to future research to evaluate whether the attractiveness premium has converged between Democrats and Republicans more generally since the Trump presidency.

As in mixed-gender elections, Republicans are more likely and Democrats less likely to vote for the more conservative-looking candidate. This supports the idea underlying our theoretical model that voters use both attractiveness and perceived conservatism as cues in their

vote choice. These results are also confirmed when party is defined based on self-identification, which is only available in 2020 (Tables A5 and A6).

A comparison of Tables 3, A2, A3, A4, A5, and A6 shows that the effects of attractiveness and perceived conservatism are quite similar in both mixed-gender and same-gender elections, regardless of the other controls included. The marginal returns to attractiveness advantage in columns 2 to 4 in these six tables ranges between 0.050 and 0.076, and interaction with Democrat is always small and statistically insignificant. The marginal returns to more conservative look among Republicans range from 0.022 to 0.040, and among Democrats (after adding interaction term to the baseline estimate for Republicans) from -0.010 to -0.033. The finding that the effects of attractiveness and perceived conservatism are fairly similar regardless of candidate gender is consistent with our theoretical model, which treats attractiveness and perceived conservatism independently of candidate gender.

5.2 The role of political knowledge

In our experimental elections, all voters are by construction uninformed about the true party affiliation or politics of the competing candidates. Nevertheless, we find an interesting partisan difference in the use of gender as a voting heuristic and how it relates to general levels of political knowledge in column 4 of Table 3. Among Republicans, political knowledge is uncorrelated with the likelihood of voting for a female candidate in a mixed-gender experimental election. Among Democrats, however, support for female candidates increases with political knowledge. To analyze potential nonlinearities in the effects of political knowledge, Figure 4 shows the vote share of female candidates in mixed-gender experimental elections among Democrats and Republicans, categorized by levels of political knowledge. The results show that the female vote share increases consistently with higher levels of political knowledge among Democrats, while it remains essentially unchanged among Republicans. These patterns are not an artifact of a small number of observations or clustering at one end of the knowledge spectrum. Fig. A13 shows that the median and modal respondent for both Democrats and Republicans correctly answered to two of the four knowledge questions. Although Democrats are slightly more politically informed (Democrats answered 2.0 questions correctly on average and Republicans answered 1.9 questions correctly), the difference is small and does not explain our findings.

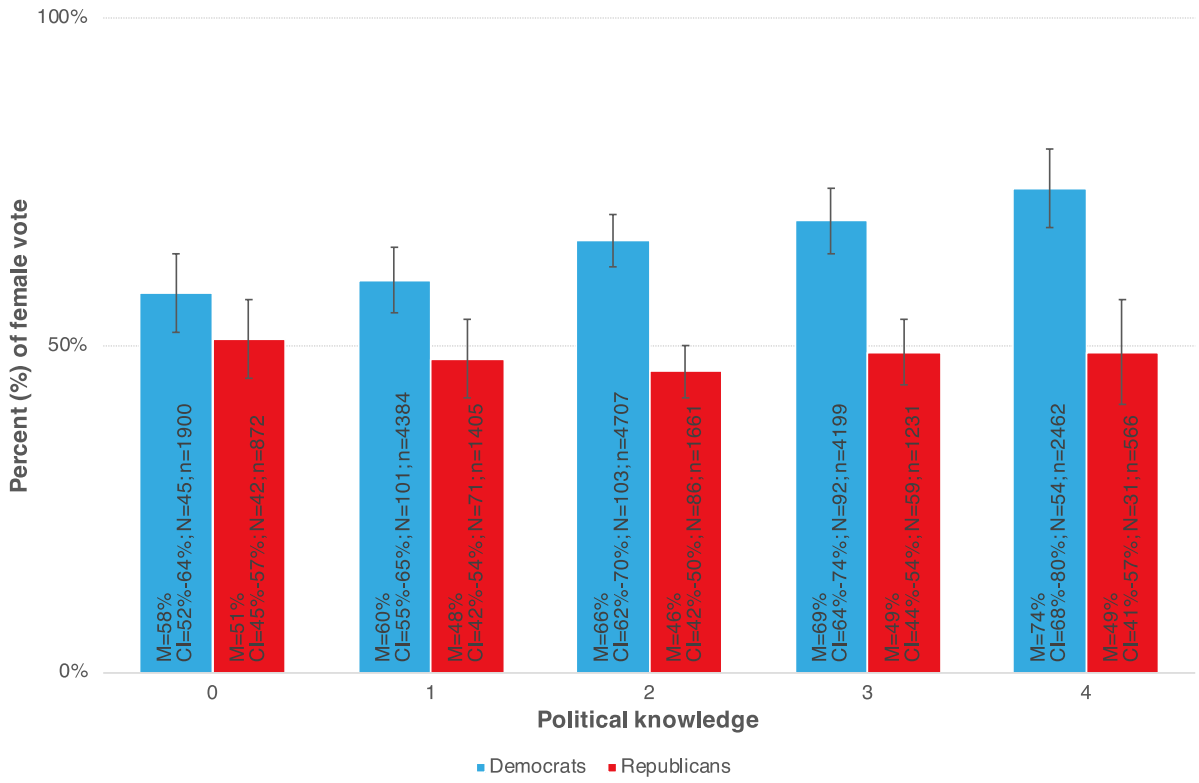

Fig. 4. Percentage of female vote by respondents' partisanship and political knowledge. Female vote shown across studies I and II. Political knowledge categorized as the number of correctly answered political knowledge questions. Results include only experimental elections with one female and one male candidate, excluding abstentions. N: number of respondents; n: number of experimental elections. Error bars depict 95% CI of the mean.

Why would support for female candidates increase with political knowledge among Democrats? One possible explanation is that Democrats with more political knowledge are more likely to know that women are more likely to be liberal than men, and thus use gender more as a cue to ideology in low-information settings. Another possible explanation concerns group representation. If Democrats with more political knowledge are more concerned about the underrepresentation of women in politics, they might be more likely to support female candidates as a matter of principle to correct this. A third possible explanation concerns the expected competence of candidates. Given that women face greater difficulties in entering politics, politically well-informed respondents may form a prior that women who succeed in becoming political candidates are, on average, of higher quality than the average male candidate, and thus favor them. Nevertheless, this explanation should hold for both Democrats and Republicans. However, it is possible that among Democrats, the three motivations reinforce

each other, while among Republicans, the first and third motivations cancel each other out, resulting in no differential gender preference by candidate gender.

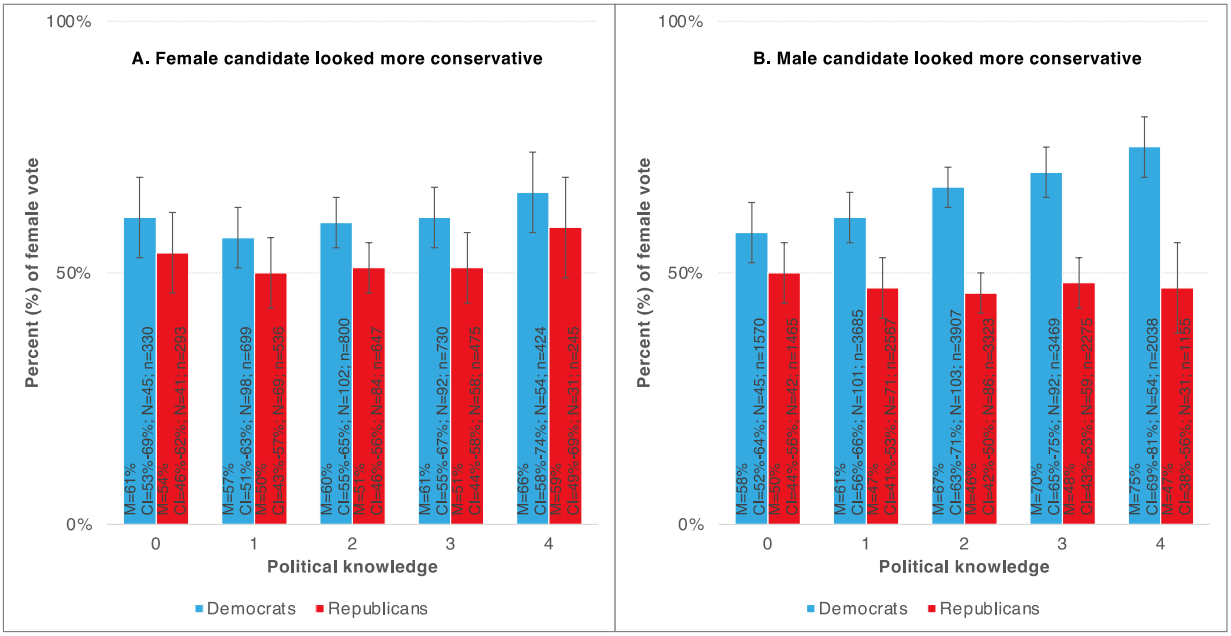

Fig. 5. Percentage of female vote by respondents' partisanship and political knowledge, categorized by candidates' relative perceived conservatism.

(A) Experimental elections in which the female candidate looked more conservative than her male counterpart. (B) Experimental elections in which the male candidate looked more conservative than his female counterpart. Female vote shown across studies I and II. Political knowledge categorized as the number of correctly answered political knowledge questions. Results include only experimental elections with one female and one male candidate, excluding abstentions. N: number of respondents; n: number of experimental elections. Error bars depict 95% CI of the mean.

To test the relative performance of these alternative explanations, we examined the extent to which support for female candidates in mixed-gender elections depend on political knowledge, depending on whether the female or male candidate looked more conservative. Figure 5 compares the relationship between the vote share of female candidates and political knowledge in experimental elections in which (A) the female candidate looked more conservative versus those in which (B) the male candidate looked more conservative. First, we observe that, regardless of candidates' perceived conservatism, the vote share of female candidates is higher among Democrats. Second, both Democrats and Republicans respond to

perceived conservatism. Among Democrats, the vote share of female candidates is higher in experimental elections in which the male candidate looks more conservative while among Republicans the vote share of female candidates is higher in experimental elections in which the female candidate appears more conservative, although these differences are small. Third, among Democrats, women's vote share increases strongly with political knowledge in experimental elections in which the male candidate looks more conservative. Fourth, the female vote among Republicans does not vary with political knowledge, even when experimental elections in which the female candidate looks more conservative and those in which the male candidate looks more conservative are analyzed separately. Taken together, our findings suggest that the partisan gap in the effect of political knowledge is driven primarily by Democrats in experimental elections in which the male candidate looks more conservative.

To assess the effect of political knowledge on appearance-based stereotyping in the absence of gender differences, we also analyzed same-gender races. Figure A14 shows that in experimental elections between two male candidates, the vote share of the more conservative-looking candidate increases with political knowledge among Republicans and decreases with political knowledge among Democrats, suggesting that the use of appearance-based stereotyping in low-information settings increases in political knowledge. However, this result is only suggestive as the pattern is weak and the differences are not statistically significant in experimental elections between two female candidates.

6 Conclusion

In recent decades, large gender gaps have opened up in American politics. Women are more likely than men to support Democrats (Gillion et al., 2020), Democratic voters are more likely than Republicans to support female candidates (Schwarz & Coppock, 2021), and the female share of congressional Democrats is nearly three times that of congressional Republicans (Fig. 1B). We conducted experimental elections in 2016 and 2020 to disentangle how voter gender and partisanship interact in support for female candidates. Our results show that Democrats generally favored female candidates, and that preference for female candidates was particularly strong among Democratic women. In our 2020 survey, Democratic women chose the female candidate three times as often as the male candidate. Republican respondents, on the other hand, chose female and male candidates about equally often. Our findings suggest that voter bias against women cannot explain the underrepresentation of women in American politics, even among Republicans. If anything, voters, on average, prefer women to men.

Our approach to studying gender discrimination in voting complements vignette and conjoint survey experiments, which have become a staple of political science research (Hainmueller et al., 2015; Hainmueller et al., 2014). In these studies, respondents indicate their preferences based on brief, standardized descriptions of hypothetical candidates. Vignette and conjoint survey experiments allow for the simultaneous investigation of the effects of different cues, such as gender, age, and reported experience. However, this comes at the cost of researchers defining the characteristics that are presented to respondents, and how they are presented. Our approach of asking respondents to make vote choices based on candidate photographs does not require researchers to specify which textual cues are presented to respondents and in what order. Instead, we collected vote choices for experimental elections among all 736 Members of the European Parliament. An advantage of using MEPs was that they are real and elected politicians. Therefore, the photographs are likely to contain cues that are relevant in politics, which may not be the case when using stock photographs. Another advantage of using MEPs was that American respondents are unlikely to recognize the candidates, which could have introduced bias. Finally, previous research has shown that ratings of politicians' photographs help to predict election outcomes around the world, providing external validity for using photographs (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Berggren et al., 2010; Lawson et al., 2010; Todorov et al., 2005).

A major concern in all research is that subjects may change their behavior based on cues about what constitutes appropriate behavior (Zizzo, 2010). In our setting, the concern is that respondents would perceive supporting female candidates in experimental elections as the appropriate choice, even if they would not vote for the female candidate in a real election. Our study design mitigates this concern by randomizing the gender combinations in the experimental elections. In addition, we did not refer to gender in our task description, but only to voting with very little information. Moreover, recent research has found that experimenter demand effects are small in online surveys, even when respondents are given cues about the hypothesis the researchers are testing (de Quidt et al., 2018; Mummolo & Peterson, 2019). A comparison of conjoint and vignette experiments with real referendums in Switzerland also suggests that estimates from survey experiments are remarkably good at predicting actual voting outcomes (Hainmueller et al., 2014).

Our results highlight the critical role of supply side factors as remaining barriers to closing the gender gap in political representation, such as women's reluctance to enter politics and discrimination by party elites and donors, as well as the weight of historical

underrepresentation of women through incumbency advantage. Given that voters with prior exposure to female leaders are more likely to vote for women (Baskaran & Hessami, 2018; Beaman et al., 2009; Bhavnani, 2009), and that prominent women in politics serve as role models that encourage other women to enter politics (Ladam et al. 2018), the recent increase in the proportion of elected female politicians, and the potential shattering of the highest remaining glass ceiling if Kamala Harris is elected as the first female President of the United States, may foreshadow a narrowing of the gender gap in the coming years.

References

Ahler, D. J., Citrin, J., Dougal, M. C., & Lenz, G. S. (2017). Face Value? Experimental Evidence that Candidate Appearance Influences Electoral Choice. *Political Behavior*, *39*(1), 77-102. https://doi.org/10.1007/s11109-016-9348-6

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science*, *323*(5918), 1183-1183. https://doi.org/10.1126/science.1167748

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948-17953. https://doi.org/10.1073/pnas.0705435104

Baltrunaite, A., Casarico, A., Profeta, P., & Savio, G. (2019). Let the voters choose women. *Journal of Public Economics*, *180*, 104085. https://doi.org/https://doi.org/10.1016/j.jpubeco.2019.104085

Baskaran, T., & Hessami, Z. (2018). Does the Election of a Female Leader Clear the Way for More Women in Politics? *American Economic Journal: Economic Policy*, *10*(3), 95-121. https://doi.org/10.1257/pol.20170045

Baskaran, T., & Hessami, Z. (2023). Women in Political Bodies as Policymakers. *Review of Economics and Statistics.* In press. https://doi.org/10.1162/rest_a_01352

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., & Topalova, P. (2009). Powerful Women: Does Exposure Reduce Bias? *The Quarterly Journal of Economics*, *124*(4), 1497-1540. https://doi.org/10.1162/qjec.2009.124.4.1497

Berggren, N., Jordahl, H., & Poutvaara, P. (2010). The looks of a winner: Beauty and electoral success. *Journal of Public Economics*, *94*(1), 8-15. https://doi.org/http://dx.doi.org/10.1016/j.jpubeco.2009.11.002

Berggren, N., Jordahl, H., & Poutvaara, P. (2017). The right look: Conservative politicians look better and voters reward it. *Journal of Public Economics*, *146*, 79-86. https://doi.org/https://doi.org/10.1016/j.jpubeco.2016.12.008

Besley, T., Folke, O., Persson, T., & Rickne, J. (2017). Gender Quotas and the Crisis of the Mediocre Man: Theory and Evidence from Sweden. *American Economic Review*, *107*(8), 2204-2242. https://doi.org/10.1257/aer.20160080

Bhavnani, R. R. (2009). Do Electoral Quotas Work after They Are Withdrawn? Evidence from a Natural Experiment in India. *The American Political Science Review*, *103*(1), 23-35. https://doi.org/10.2307/27798484

Bucchianeri, P. (2018). Is Running Enough? Reconsidering the Conventional Wisdom about Women Candidates. *Political Behavior*, *40*(2), 435-466.

Casas-Arce, P., & Saiz, A. (2015). Women and Power: Unpopular, Unwilling, or Held Back? *Journal of Political Economy*, *123*(3), 641-669. https://doi.org/10.1086/680686

Chattopadhyay, R., & Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica*, *72*(5), 1409-1443. https://doi.org/https://doi.org/10.1111/j.1468-0262.2004.00539.x
de Quidt, J., Haushofer, J., & Roth, C. (2018). Measuring and Bounding Experimenter Demand. *American Economic Review*, *108*(11), 3266-3302. https://doi.org/10.1257/aer.20171330
Dixit, A., & Londregan, J. (1996). The Determinants of Success of Special Interests in Redistributive Politics. *The Journal of Politics*, *58*(4), 1132-1155. https://doi.org/10.2307/2960152
Dolan, K. (2004). *Voting for Women: How the Public Evaluates Women Candidates*. Westview Press.
Edlund, L., & Pande, R. (2002). Why Have Women Become Left-Wing? The Political Gender Gap and the Decline in Marriage*. *The Quarterly Journal of Economics*, *117*(3), 917-961. https://doi.org/10.1162/003355302760193922
Fox, R. L., & Lawless, J. L. (2004). Entering the Arena? Gender and the Decision to Run for Office. *American Journal of Political Science*, *48*(2), 264-280. https://doi.org/https://doi.org/10.1111/j.0092-5853.2004.00069.x
Fox, R. L., & Lawless, J. L. (2010). If only they'd ask: Gender, recruitment, and political ambition. *The Journal of Politics*, *72*(2), 310-326.
Gallup Organization. (2016). *Gallup Poll: April 2016*.
Gallup Organization. (2020). *Gallup Poll: October 2020*.
Gelman, A., & King, G. (1990). Estimating Incumbency Advantage without Bias. *American Journal of Political Science*, *34*(4), 1142-1164. https://doi.org/10.2307/2111475
Gillion, D. Q., Ladd, J. M., & Meredith, M. (2020). Party Polarization, Ideological Sorting and the Emergence of the US Partisan Gender Gap. *British Journal of Political Science*, *50*(4), 1217-1243. https://doi.org/10.1017/S0007123418000285
Gomila, R. (2020). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *Journal of Experimental Psychology: General*, Advance online publication. https://doi.org/10.1037/xge0000920
Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, *112*(8), 2395. https://doi.org/10.1073/pnas.1416587112
Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments. *Political Analysis*, *22*(1), 1-30. https://doi.org/10.1093/pan/mpt024
Hamermesh, D. S., & Biddle, J. E. (1994). Beauty and the Labor Market. *The American Economic Review*, *84*(5), 1174-1194. http://www.jstor.org/stable/2117767
Hellevik, O. (2009). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & Quantity*, *43*(1), 59-74.
Kanthak, K., & Woon, J. (2015). Women Don't Run? Election Aversion and Candidate Entry. *American Journal of Political Science*, *59*(3), 595-612. https://doi.org/https://doi.org/10.1111/ajps.12158
Ladam, C., Harden, J. J. & Windett, J. H. (2018). Prominent Role Models: High-Profile Female Politicians and the Emergence of Women as Candidates for Public Office. *American Journal of Political Science*; 62(2), 369-381. https://doi.org/10.1111/ajps.12351
Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychol Bull*, *126*(3), 390-423. https://doi.org/10.1037/0033-2909.126.3.390

Lawson, C., Lenz, G. S., Baker, A., & Myers, M. (2010). Looking Like a Winner: Candidate Appearance and Electoral Success in New Democracies. *World Politics*, *62*(4), 561-593. https://doi.org/10.1017/S0043887110000195

Lindbeck, A., & Weibull, J. W. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, *52*(3), 273-297. https://doi.org/10.1007/BF00116710

Lindbeck, A., & Weibull, J. W. (1993). A model of political equilibrium in a representative democracy. *Journal of Public Economics*, *51*(2), 195-209. https://doi.org/https://doi.org/10.1016/0047-2727(93)90084-7

McDermott, M. 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), 270-283. https://doi.org/10.2307/2111716

Miller, G. (2008). Women's Suffrage, Political Responsiveness, and Child Survival in American History. *The Quarterly Journal of Economics*, *123*(3), 1287-1327. https://doi.org/10.1162/qjec.2008.123.3.1287

Mummolo, J., & Peterson, E. (2019). Demand Effects in Survey Experiments: An Empirical Assessment. *American Political Science Review*, *113*(2), 517-529. https://doi.org/10.1017/S0003055418000837

Olivola, C. Y., Sussman, A. B., Tsetsos, K., Kang, O. E., & Todorov, A. (2012). Republicans Prefer Republican-Looking Leaders: Political Facial Stereotypes Predict Candidate Electoral Success Among Right-Leaning Voters. *Social Psychological and Personality Science*, *3*(5), 605-613. https://doi.org/10.1177/1948550611432770

Olivola, C. Y., Tingley, D., & Todorov, A. (2018). Republican Voters Prefer Candidates Who Have Conservative-Looking Faces: New Evidence From Exit Polls. *Political Psychology*, *39*(5), 1157-1171. https://doi.org/https://doi.org/10.1111/pops.12489

Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Reexamining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, *46*(2), 315-324. https://doi.org/https://doi.org/10.1016/j.jesp.2009.12.002

Ono, Y., & Burden, B. C. (2019). The Contingent Effects of Candidate Sex on Voter Choice. *Political Behavior*, *41*(3), 583-607. https://doi.org/10.1007/s11109-018-9464-6

Pearson, K., & McGhee, E. (2013). What it takes to win: Questioning "gender neutral" outcomes in US House elections. *Politics & Gender*, *9*(4), 439-462.

Persson, T., & Tabellini, G. (2002). *Political Economics: Explaining Economic Policy*. MIT press.

Rule, N. O., & Ambady, N. (2010). Democrats and Republicans Can Be Differentiated from Their Faces. *PLOS ONE*, *5*(1), e8733. https://doi.org/10.1371/journal.pone.0008733

Sanbonmatsu, K. (2002). Gender Stereotypes and Vote Choice. *American Journal of Political Science*, *46*(1), 20-34. https://doi.org/10.2307/3088412

Sanbonmatsu, K., & Dolan, K. (2009). Do gender stereotypes transcend party? *Political Research Quarterly*, *62*(3), 485-494.

Schwarz, S., & Coppock, A. (2021). What Have We Learned About Gender From Candidate Choice Experiments? A Meta-analysis of 67 Factorial Survey Experiments. *The Journal of Politics*, forthcoming.

Thomsen, D. M., & Swers, M. L. (2017). Which women can run? Gender, partisanship, and candidate donor networks. *Political Research Quarterly*, *70*(2), 449-463.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*(5728), 1623-1626. https://doi.org/10.1126/science.1110589

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*(1), 75-98. https://doi.org/10.1007/s10683-009-9230-z

Acknowledgments: We thank Joop Adema for outstanding research assistance and Thushyanthan Baskaran, Niclas Berggren, Zohal Hessami, Henrik Jordahl, Patrick Kaiser, Helmut Küchenhoff, Chris Olivola, and Paola Profeta as well as the participants at the Silvaplana Political Economy Workshop in 2021 and the Meeting of the European Public Choice Society, the CESifo Area Conference on Public Economics, and the Journées Louis-André Gérard-Varet in 2022 for helpful comments.

Funding: This report did not receive any funding.

Author contributions: PP conceived the study and collected the photographs used in the study. PP and AG planned the survey. AG programmed and implemented the survey and visualized the results. PP and AG analyzed the data and co-wrote the paper.

Competing interests: Authors declare that they have no competing interests.

Data and materials availability: All data, code, and materials used in the analysis will be made available via Harvard Dataverse.

Appendix

Table A1. Variables and their coding

MEASURE	STUDY	QUESTION(S)	ANSWER OPTIONS	CODING IN ANALYSIS
Party identification	II	Generally speaking, do you usually think of yourself as a Republican, a Democrat, an Independent, or what?	a strong Democrat	1
			not a strong Democrat	1
			an Independent, leaning Democratic	1
			an Independent	0
			an Independent, leaning Republican	-1
			not a strong Republican	-1
			a strong Republican	-1
			Other	-9
			I am not interested in politics	-9
Democrat	I	Whom would you like to see elected as the next President of the United States?	Hillary Clinton (Democrat)	1
			Bernie Sanders (Democrat)	1
			Donald Trump (Republican)	0
			Ted Cruz (Republican)	0
			John Kasich (Republican)	0
			Don't know / prefer not to answer	-1
			(Candidates rotated)	
	II	Whom would you like to see elected as the next President of the United States?	Joe Biden (Democrat)	1
			Donald Trump (Republican)	0
			Don't know / prefer not to answer	-1
			(Candidates rotated)	
Female	I&II	What is your gender?	Male	0
			Female	1
			Other/Prefer not to answer	-1
Age	I&II	What is your year of birth?	Free text (4-digit number), converted to age and mean-centered	
Education	I&II	What is the highest degree or level of school you have completed?	No schooling completed	1
			Nursery school to 8th grade	1
			Some high school, no diploma	1
			High school graduate, diploma or the equivalent	1
			Some college credit, no degree	2
			Trade/technical/vocational training	2
			Associate degree	2
			Bachelor's degree	3
			Master's degree	4
			Professional degree	4
			Doctorate degree	4
Knowledge	I&II	Index from 0 to 4, calculated as the number of correct answers to the four knowledge questions KQ1 to KQ4		
KQ1	I	How many female justices are there currently on the Supreme Court of the United States?	None	0
			One	0
			Two	0
			Three	1
			Four	0
			Don't know	0
	II	How many male justices are there currently on the Supreme Court of the United States?	3	0
			4	0
			5	0
			6	1
			7	0
			Don't know	0
KQ2	I&II	What is the number of representatives with full voting rights in the United	100	0
			235	0
			435	1

MEASURE	STUDY	QUESTION(S)	ANSWER OPTIONS	CODING IN ANALYSIS
		States House of Representatives?	501	0
			603	0
			Don't know	0
KQ3	I&II	Which of the following statements is correct? (answer options rotated)	*Currently, the United States Senate has a Republican and the House has a Democratic majority.*	*1 in study II, 0 in study I*
			Currently, the United States Senate and House both have Republican majorities.	*1 in study I, 0 in study II*
			Currently, the United States Senate and House both have Democratic majorities.	0
			Currently, the United States Senate has a Democratic and the House has a Republican majority.	0
			Currently, the United States Senate is evenly divided between Democrats and Republicans, while the House has a Republican majority.	0
			Don't know	0
KQ4	I	Who is the current Secretary of Defense? (answer options rotated)	Dick Cheney	0
			Ashton Carter	1
			Chuck Hagel	0
			Leon Panetta	0
			John Kerry	0
			Don't know	0
	II	Who is the current Secretary of Defense? (answer options rotated)	Mark Esper	1
			Ashton Carter	0
			Chuck Hagel	0
			Leon Panetta	0
			John Kerry	0
			Don't know	0
Vote choice (recoded so that left refers to the male candidate and right to the female candidate)	I	For which person would you vote?	Person on the left	0
			Person on the right	1
			Abstain from voting	-9
			Prefer not to answer	-9
	II	Which person would you vote for?	Definitely the person on the left	0
			The person on the left	0
			Most probably the person on the left	0
			Most probably the person on the right	1
			The person on the right	1
			Definitely the person on the right	1
			Abstain from voting	-9
			Prefer not to answer	-9
Strength of vote choice (recoded so that left refers to the male candidate and right to the female candidate)	II	Which person would you vote for?	Definitely the person on the left	1
			The person on the left	2
			Most probably the person on the left	3
			Most probably the person on the right	4
			The person on the right	5
			Definitely the person on the right	6
			Abstain from voting	-9
			Prefer not to answer	-9
Female more attractive	Data obtained from Berggren et al. (2017)	Defined as the average attractiveness rating of the female candidate minus the average attractiveness rating of the male candidate. Individual attractiveness ratings were originally collected on a scale from 1 (very unattractive) to 5 (very handsome or beautiful). We standardized those ratings as follows: After excluding abstentions, we standardized each individual respondent's ratings to a mean of 0 and a standard deviation of unity. Standardized scores per MEP were then obtained by averaging standardized individual respondents' ratings of each politician, and standardizing those averages. See Table S1 for descriptive statistics.		
Female looks more conservative		Defined as the average perceived conservatism of the female candidate minus the average perceived conservatism of the male candidate. Perceived conservatism ratings were originally collected using scale from 1 (farthest to the left) to 10 (farthest to the right). We standardized those ratings as follows: After excluding abstentions, we standardized each individual respondent's ratings to a mean of 0 and a standard deviation of		

MEASURE	STUDY	QUESTION(S)	ANSWER OPTIONS	CODING IN ANALYSIS
		unity. Standardized scores per MEP were then obtained by averaging standardized individual respondents' ratings of each politician, and standardizing those averages. See Table S1 for descriptive statistics.		
Female older		Defined as the age of the female candidate minus the age of the male candidate. Candidate age information obtained from the European Parliament's web site. See Table S1 for descriptive statistics.		

Table A2. OLS regression analysis on the determinants of voting for the female candidate in mixed-gender elections in 2020 (Party classification based on self-identification)

	1	2	3	4
Democrat	0.149***	0.151***	0.115***	0.115***
	(0.029)	(0.029)	(0.029)	(0.029)
Female	0.003	0.004	0.004	0.004
	(0.037)	(0.037)	(0.037)	(0.037)
Democrat x Female	0.131**	0.130**	0.130**	0.130**
	(0.046)	(0.046)	(0.046)	(0.046)
Age	0.002*	0.002*	0.002*	0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
High school or less	-0.023	-0.025	-0.025	-0.025
	(0.039)	(0.038)	(0.038)	(0.038)
Associate degree or some college	-0.010	-0.011	-0.011	-0.011
	(0.032)	(0.032)	(0.032)	(0.032)
Master's degree or higher	-0.005	-0.006	-0.006	-0.006
	(0.028)	(0.028)	(0.028)	(0.028)
Female more attractive		0.050***	0.050***	0.050***
		(0.006)	(0.007)	(0.007)
Female more attractive x Democrat		-0.007	-0.006	-0.005
		(0.007)	(0.008)	(0.008)
Female looks more conservative			0.024***	0.024***
			(0.006)	(0.006)
Female looks more conservative x Democrat			-0.034***	-0.034***
			(0.008)	(0.008)
Female looks older			-0.000	-0.000
			(0.001)	(0.001)
Female looks older x Democrat			0.001	0.001
			(0.001)	(0.001)
Political knowledge				-0.004
				(0.005)
Political knowledge x Democrat				0.003
				(0.006)
Constant	0.467***	0.441***	0.467***	0.467***
	(0.022)	(0.022)	(0.022)	(0.022)
R squared	0.06	0.08	0.08	0.08
Number of respondents	383	383	383	383
Number of experimental elections	17880	17880	17880	17880

Each observation is an experimental election between a female and a male candidate in 2020, excluding experimental elections in which the respondent abstained. Dependent variable equals 1 if the respondent chose the female candidate and 0 if the respondent chose the male candidate. Robust standard errors are clustered at the respondent level. Respondent reference categories (not shown) are Republican, Male, and Bachelor's degree. Party identification is based on self-identification. Respondents who do not identify as Democrats or Republicans or at least leaning towards one of these parties are excluded.

Table A3. OLS regression analysis on the determinants of voting for the more attractive candidate in elections between two males (Party classification based on Presidential voting intention)

	1	2	3	4
Democrat	0.021	0.006	0.004	0.003
	(0.013)	(0.016)	(0.016)	(0.016)
Female	0.007	0.005	0.006	0.007
	(0.015)	(0.015)	(0.015)	(0.015)
Democrat x Female	0.002	0.004	0.002	0.001
	(0.019)	(0.019)	(0.019)	(0.019)
Age	0.001	0.001	0.001	0.001
	(0.000)	(0.000)	(0.000)	(0.000)
High school or less	0.011	0.010	0.009	0.012
	(0.018)	(0.018)	(0.018)	(0.018)
Associate degree or some college	0.015	0.018	0.017	0.019
	(0.012)	(0.012)	(0.012)	(0.012)
Master's degree or higher	0.005	0.007	0.008	0.008
	(0.013)	(0.013)	(0.013)	(0.012)
Responded in 2016 (Study I)	0.042***	0.041***	0.041***	0.041***
	(0.010)	(0.010)	(0.010)	(0.010)
Attractiveness advantage		0.065***	0.075***	0.076***
		(0.008)	(0.009)	(0.009)
Attractiveness advantage x Democrat		0.015	0.004	0.003
		(0.011)	(0.011)	(0.011)
More attractive looks more conservative			0.040***	0.040***
			(0.006)	(0.006)
More attractive looks more conservative x Democrat			-0.069***	-0.069***
			(0.008)	(0.008)
More attractive looks older			0.001*	0.001*
			(0.000)	(0.000)
More attractive looks older x Democrat			-0.001	-0.001
			(0.001)	(0.001)
Political knowledge				0.011*
				(0.006)
Political knowledge x Democrat				-0.011
				(0.008)
Constant	0.552***	0.486***	0.488***	0.487***
	(0.011)	(0.013)	(0.013)	(0.013)
R squared	0.00	0.02	0.03	0.03
Number of respondents	675	675	675	675
Number of experimental elections	14414	14414	14414	14414

Each observation is an experimental election between two male candidates, excluding experimental elections in which the respondent abstained. Dependent variable equals 1 if the respondent chose the candidate who looks more attractive and 0 if the respondent chose the candidate who looks less attractive. Robust standard errors are clustered at the respondent level. Respondent reference categories (not shown) are Republican, Bachelor's degree and Responded in 2020 (study II).

Table A4. OLS regression analysis on the determinants of voting for the more attractive candidate in elections between two females (Party classification based on Presidential voting intention)

	1	2	3	4
Democrat	0.000	-0.005	-0.005	-0.006
	(0.014)	(0.016)	(0.016)	(0.016)
Female	-0.038*	-0.037*	-0.037*	-0.033*
	(0.016)	(0.016)	(0.016)	(0.016)
Democrat x Female	-0.009	-0.008	-0.008	-0.011
	(0.021)	(0.021)	(0.021)	(0.021)
Age	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High school or less	0.011	0.010	0.009	0.018
	(0.020)	(0.019)	(0.019)	(0.019)
Associate degree or some college	0.024	0.025*	0.024*	0.029*
	(0.013)	(0.012)	(0.012)	(0.012)
Master's degree or higher	-0.009	-0.008	-0.008	-0.008
	(0.014)	(0.014)	(0.014)	(0.014)
Responded in 2016 (Study I)	0.049***	0.048***	0.048***	0.048***
	(0.011)	(0.011)	(0.011)	(0.011)
Attractiveness advantage		0.058***	0.058***	0.058***
		(0.007)	(0.008)	(0.008)
Attractiveness advantage x Democrat		0.004	0.004	0.003
		(0.009)	(0.010)	(0.010)
More attractive looks more conservative			0.022**	0.022**
			(0.007)	(0.007)
More attractive looks more conservative x Democrat			-0.044***	-0.044***
			(0.010)	(0.010)
More attractive looks older			-0.000	-0.000
			(0.001)	(0.001)
More attractive looks older x Democrat			-0.000	-0.000
			(0.001)	(0.001)
Political knowledge				0.027***
				(0.006)
Political knowledge x Democrat				-0.026**
				(0.008)
Constant	0.597***	0.524***	0.525***	0.524***
	(0.012)	(0.013)	(0.013)	(0.013)
R squared	0.01	0.02	0.02	0.02
Number of respondents	681	681	681	681
Number of experimental elections	15379	15379	15379	15379

Each observation is an experimental election between two female candidates, excluding experimental elections in which the respondent abstained. Dependent variable equals 1 if the respondent chose the candidate who looks more attractive and 0 if the respondent chose the candidate who looks less attractive. Robust standard errors are clustered at the respondent level. Respondent reference categories (not shown) are Republican, Bachelor's degree and Responded in 2020 (study II).

Table A5. OLS regression analysis on the determinants of voting for the more attractive candidate in elections between two males in 2020 (Party classification based on self-identification)

	1	2	3	4
Democrat	0.011	-0.006	-0.007	-0.008
	(0.017)	(0.020)	(0.020)	(0.020)
Female	-0.001	-0.003	0.002	0.002
	(0.020)	(0.020)	(0.020)	(0.020)
Democrat x Female	0.016	0.017	0.012	0.013
	(0.025)	(0.025)	(0.025)	(0.025)
Age	-0.000	-0.000	-0.000	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
High school or less	0.036	0.033	0.034	0.035
	(0.027)	(0.027)	(0.027)	(0.027)
Associate degree or some college	0.019	0.021	0.022	0.023
	(0.017)	(0.017)	(0.017)	(0.017)
Master's degree or higher	-0.011	-0.011	-0.009	-0.009
	(0.016)	(0.017)	(0.017)	(0.017)
Attractiveness advantage		0.055***	0.063***	0.063***
		(0.011)	(0.012)	(0.012)
Attractiveness advantage x Democrat		0.016	0.006	0.006
		(0.014)	(0.015)	(0.015)
More attractive looks more conservative			0.037***	0.037***
			(0.008)	(0.008)
More attractive looks more conservative x Democrat			-0.070***	-0.070***
			(0.010)	(0.010)
More attractive looks older			0.001	0.001
			(0.001)	(0.001)
More attractive looks older x Democrat			-0.001	-0.001
			(0.001)	(0.001)
Political knowledge				0.001
				(0.008)
Political knowledge x Democrat				0.000
				(0.011)
Constant	0.559***	0.503***	0.503***	0.503***
	(0.013)	(0.016)	(0.016)	(0.016)
R squared	0.00	0.01	0.02	0.02
Number of respondents	382	382	382	382
Number of experimental elections	8526	8526	8526	8526

Each observation is an experimental election between two male candidates in 2020, excluding experimental elections in which the respondent abstained. Dependent variable equals 1 if the respondent chose the candidate who looks more attractive and 0 if the respondent chose the candidate who looks less attractive. Robust standard errors are clustered at the respondent level. Respondent reference categories (not shown) are Republican and Bachelor's degree. Party identification is based on self-identification. Respondents who do not identify as Democrats or Republicans or at least leaning towards one of these parties are excluded.

Table A6. OLS regression analysis on the determinants of voting for the more attractive candidate in elections between two females in 2020 (Party classification based on self-identification)

	1	2	3	4
Democrat	0.012	0.005	0.004	0.006
	(0.017)	(0.021)	(0.021)	(0.020)
Female	-0.026	-0.026	-0.027	-0.026
	(0.024)	(0.024)	(0.024)	(0.023)
Democrat x Female	-0.008	-0.007	-0.007	-0.008
	(0.029)	(0.029)	(0.029)	(0.028)
Age	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)
High school or less	0.033	0.031	0.030	0.041
	(0.031)	(0.030)	(0.030)	(0.030)
Associate degree or some college	0.039*	0.040*	0.040*	0.043*
	(0.017)	(0.017)	(0.017)	(0.017)
Master's degree or higher	-0.015	-0.014	-0.014	-0.017
	(0.019)	(0.019)	(0.020)	(0.019)
Attractiveness advantage		0.055***	0.051***	0.052***
		(0.010)	(0.010)	(0.010)
Attractiveness advantage x Democrat		0.005	0.004	0.003
		(0.012)	(0.013)	(0.013)
More attractive looks more conservative			0.035***	0.036***
			(0.009)	(0.009)
More attractive looks more conservative x Democrat			-0.064***	-0.065***
			(0.013)	(0.013)
More attractive looks older			-0.001	-0.001
			(0.001)	(0.001)
More attractive looks older x Democrat			-0.000	-0.000
			(0.001)	(0.001)
Political knowledge				0.029**
				(0.010)
Political knowledge x Democrat				-0.027*
				(0.012)
Constant	0.582***	0.514***	0.516***	0.512***
	(0.014)	(0.017)	(0.017)	(0.016)
R squared	0.00	0.01	0.02	0.02
Number of respondents	383	383	383	383
Number of experimental elections	9003	9003	9003	9003

Each observation is an experimental election between two female candidates in 2020, excluding experimental elections in which the respondent abstained. Dependent variable equals 1 if the respondent chose the candidate who looks more attractive and 0 if the respondent chose the candidate who looks less attractive. Robust standard errors are clustered at the respondent level. Respondent reference categories (not shown) are Republican and Bachelor's degree. Party identification is based on self-identification. Respondents who do not identify as Democrats or Republicans or at least leaning towards one of these parties are excluded.

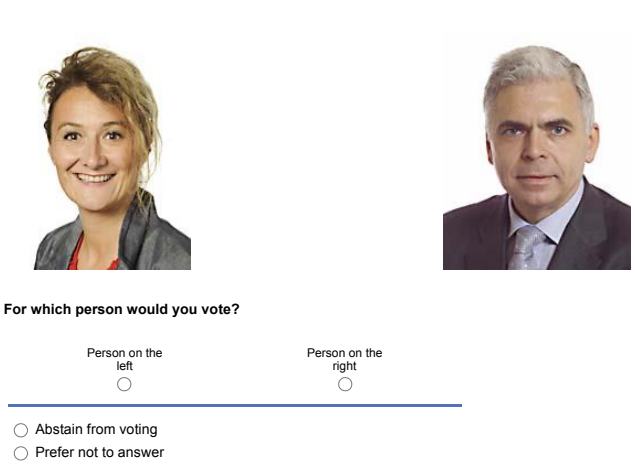

Fig. A1. Experimental voting task in study I. Participants made a binary choice between the candidate on the left and the candidate on the right. Candidates were randomly selected, and 50% of the pairs were mixed-gender races.

Fig. A2. Experimental voting task in study II. Participants chose between the candidate on the left and the candidate on the right using a 6-point scale from 1 ("definitely the person on the left") to 6 ("definitely the person on the right"). Candidates were randomly selected, and 50% of the pairs were mixed-gender races.

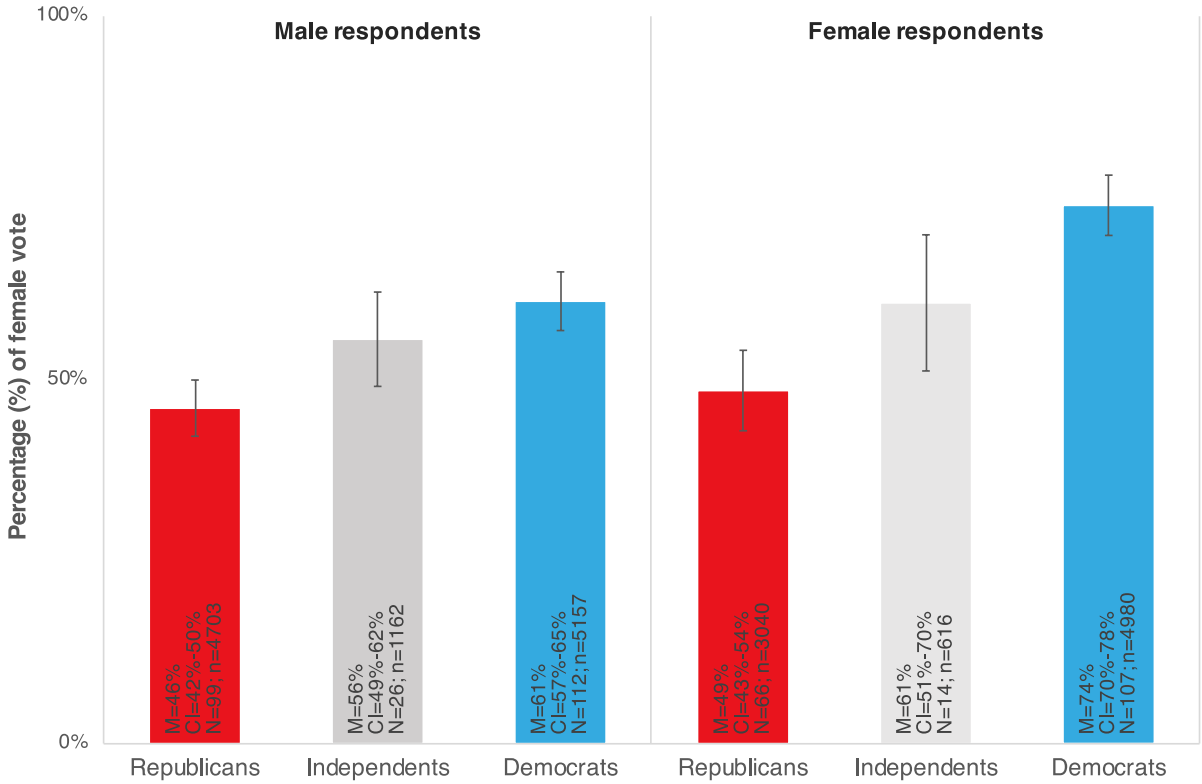

Fig. A3. Respondents who identify as Democrats support female candidates most of the time. The share of female candidates being elected in experimental elections by male and female respondents is shown according to their reported *party identification*. Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Only available in study II. Error bars depict 95% CI of the mean. N: number of respondents; n: number of experimental elections.

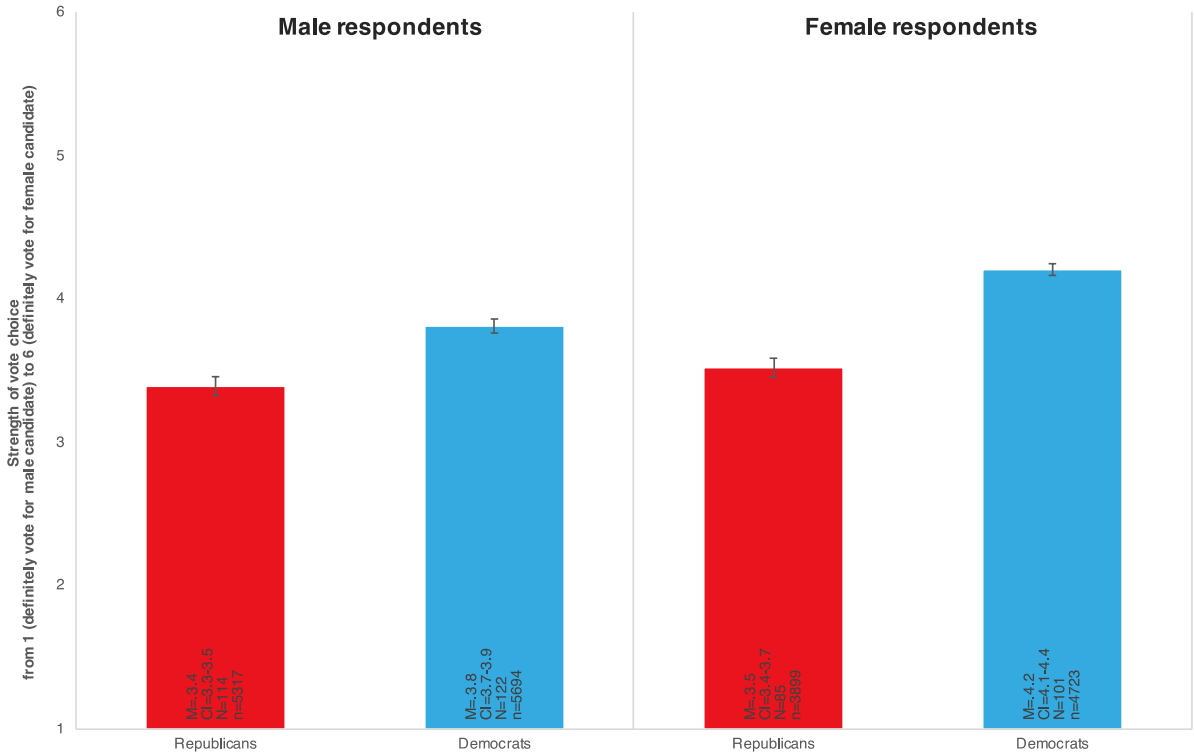

Fig. A4. Democrats support female candidates more also when accounting for the intensity of support. Average support for female candidates is measured using scale 1 to 6, in which 1 is maximum support for male candidate and 6 is maximum support for female candidates, and shown by respondent gender and partisanship. Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Only available in study II. Error bars depict 95% CI of the mean. N: number of respondents; n: number of experimental elections.

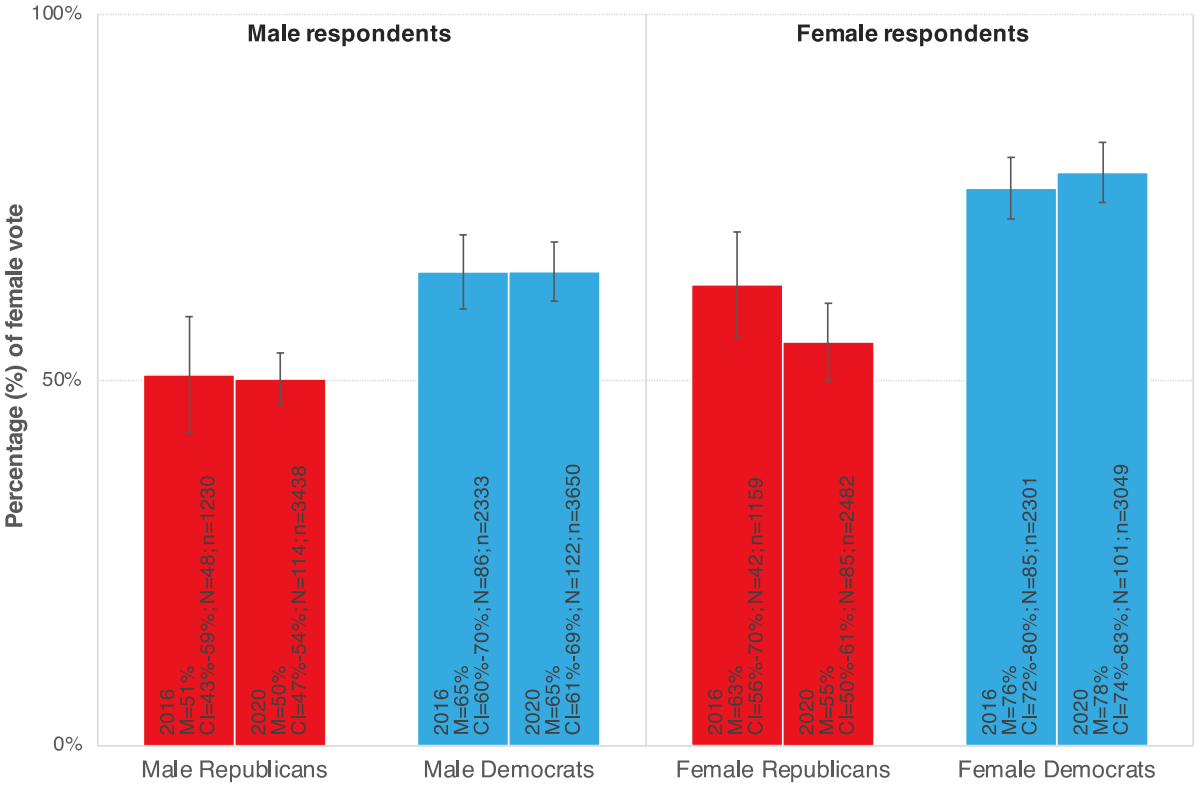

Fig. A5. Percentage of female vote for experimental elections in which the female candidate looked more attractive than the male candidate, by respondent gender and partisanship, shown separately for studies I and II. Partisanship measured based on presidential preference. Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Error bars depict 95% CI of the mean. N: number of respondents; n: number of experimental elections.

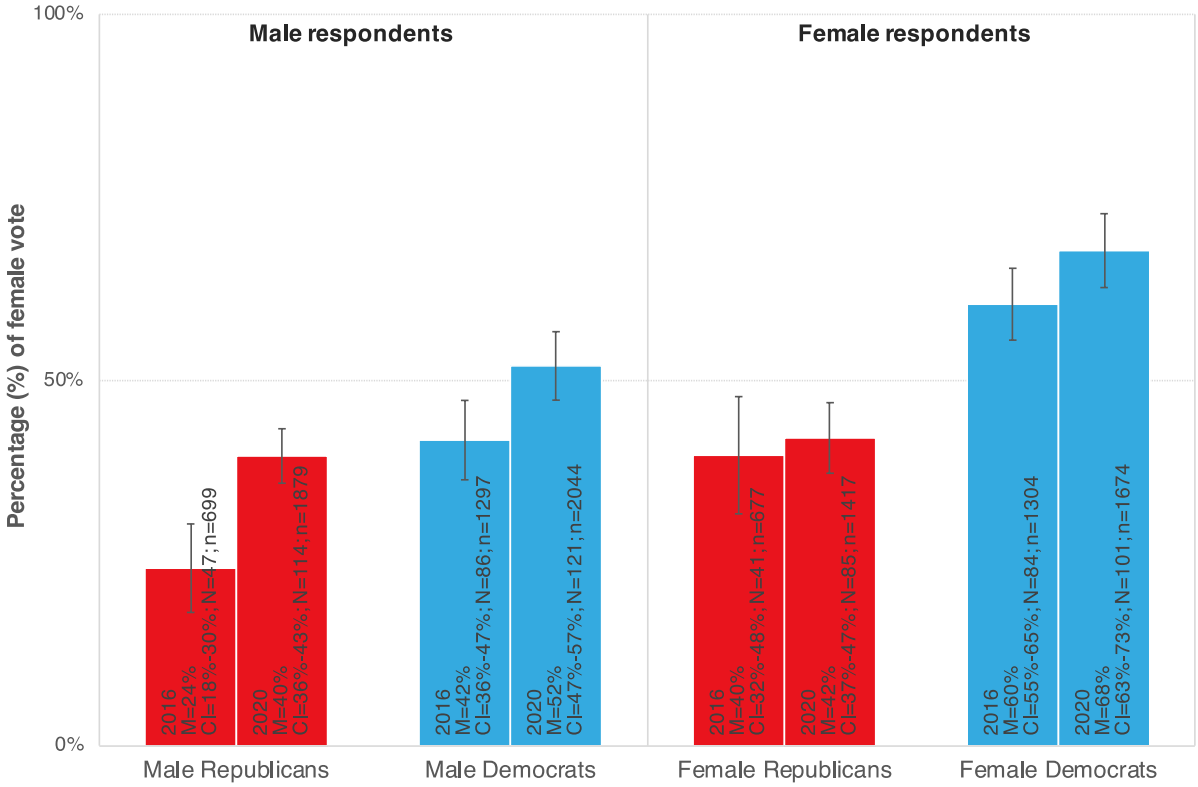

Fig. A6. Percentage of female vote for experimental elections in which the male candidate looked more attractive than the female candidate, by respondent gender and partisanship, shown separately for studies I and II. Partisanship measured based on presidential preference. Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Error bars depict 95% CI of the mean. N: number of respondents; n: number of experimental elections.

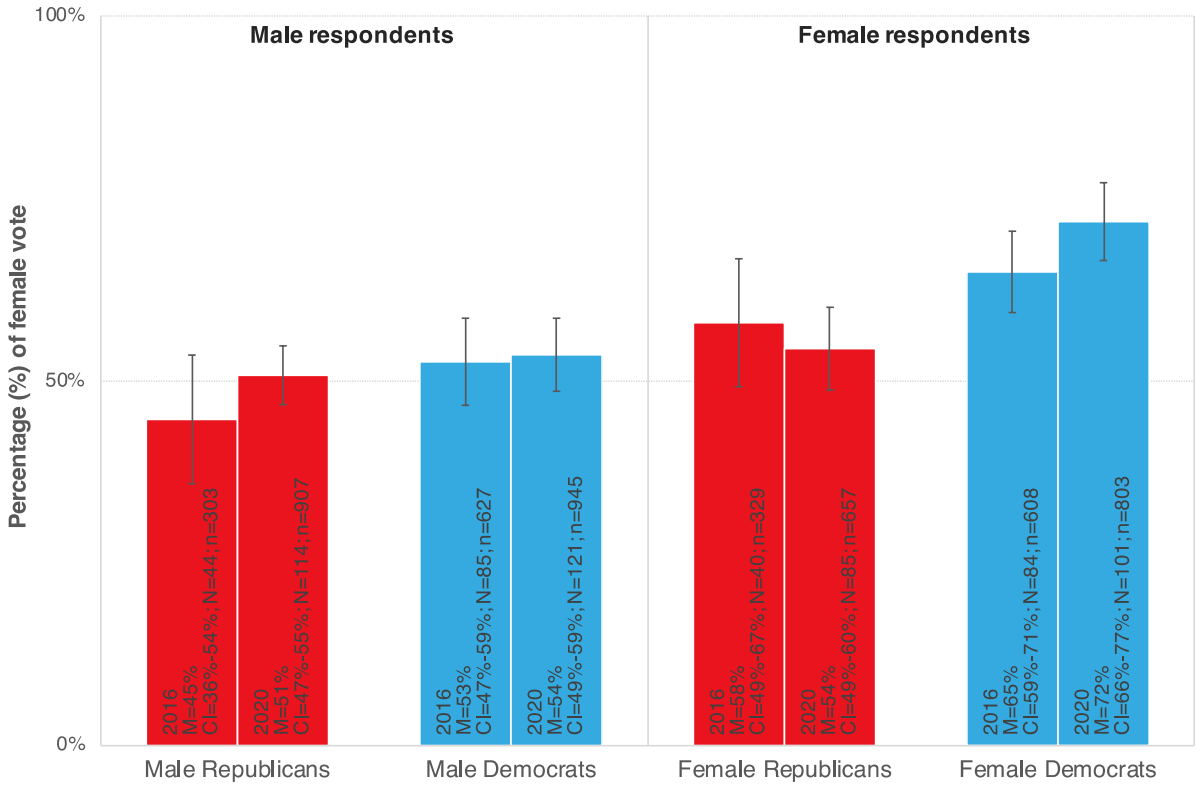

Fig. A7. Percentage of female vote for experimental elections in which the female candidate looked more conservative than the male candidate, by respondent gender and partisanship, shown separately for studies I and II. Partisanship measured based on presidential preference. Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Error bars depict 95% CI of the mean. N: number of respondents; n: number of experimental elections.

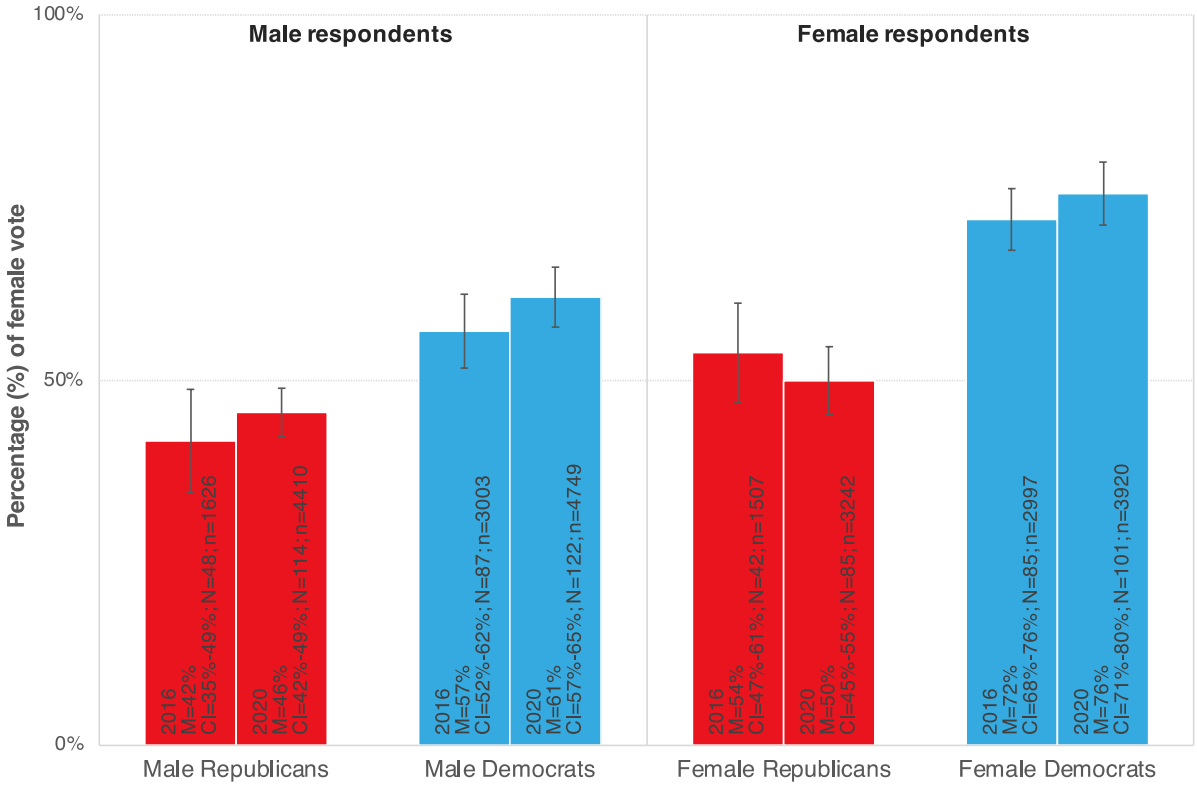

Fig. A8. Percentage of female vote for experimental elections in which the male candidate looked more conservative than the female candidate, by respondent gender and partisanship, shown separately for studies I and II. Partisanship measured based on presidential preference. Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Error bars depict 95% CI of the mean. N: number of respondents; n: number of experimental elections.

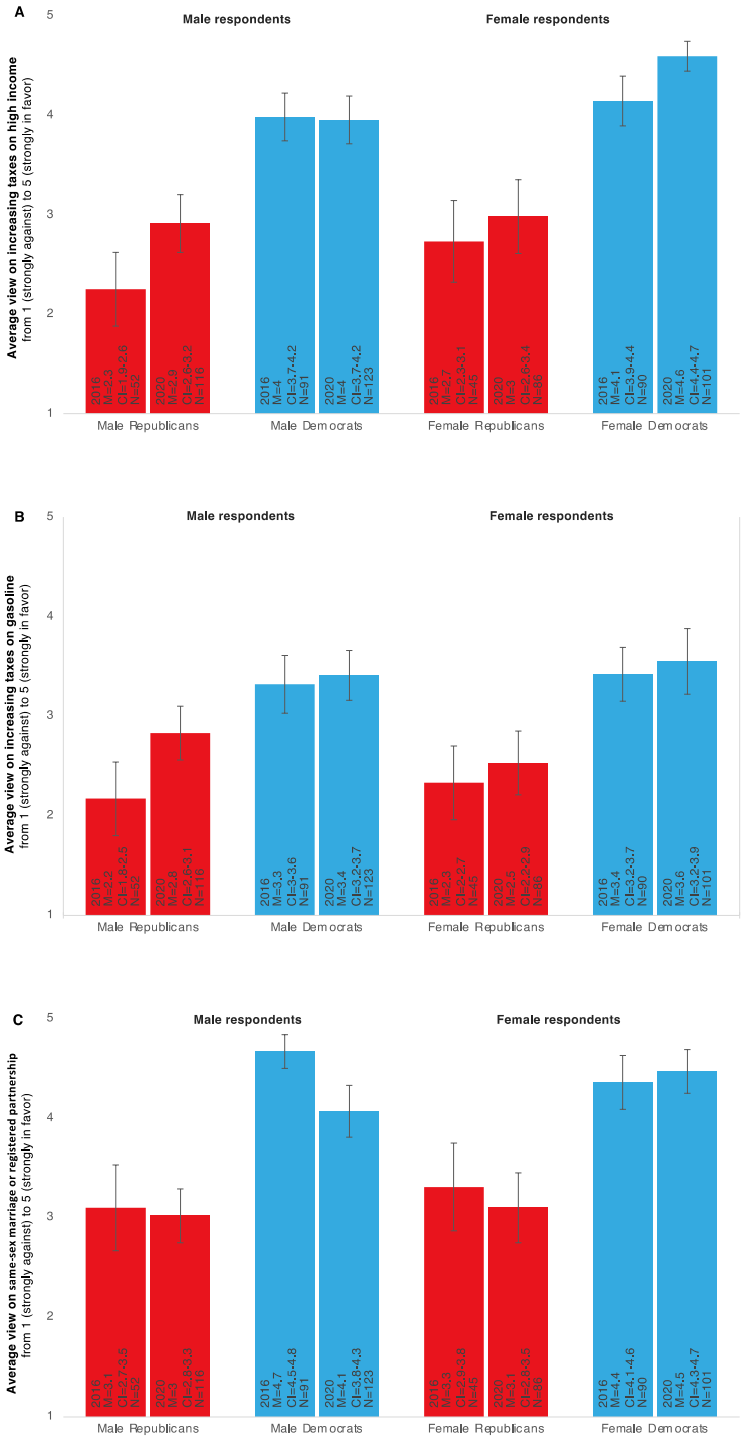

Fig. A9. Respondents' ideological views in three central dimensions of contemporary American politics, by respondent gender and partisanship, shown separately for studies I and II. Respondents were asked whether they support (A) increasing taxes on the rich and redistributing money to those with low incomes, (B) increasing taxes on gasoline and using the money to protect the environment, as well as (C) support same-sex marriage or registered partnership. Answer scale from 1 (strongly against) to 5 (strongly in favor). Error bars depict 95% CI of the mean. N: number of respondents.

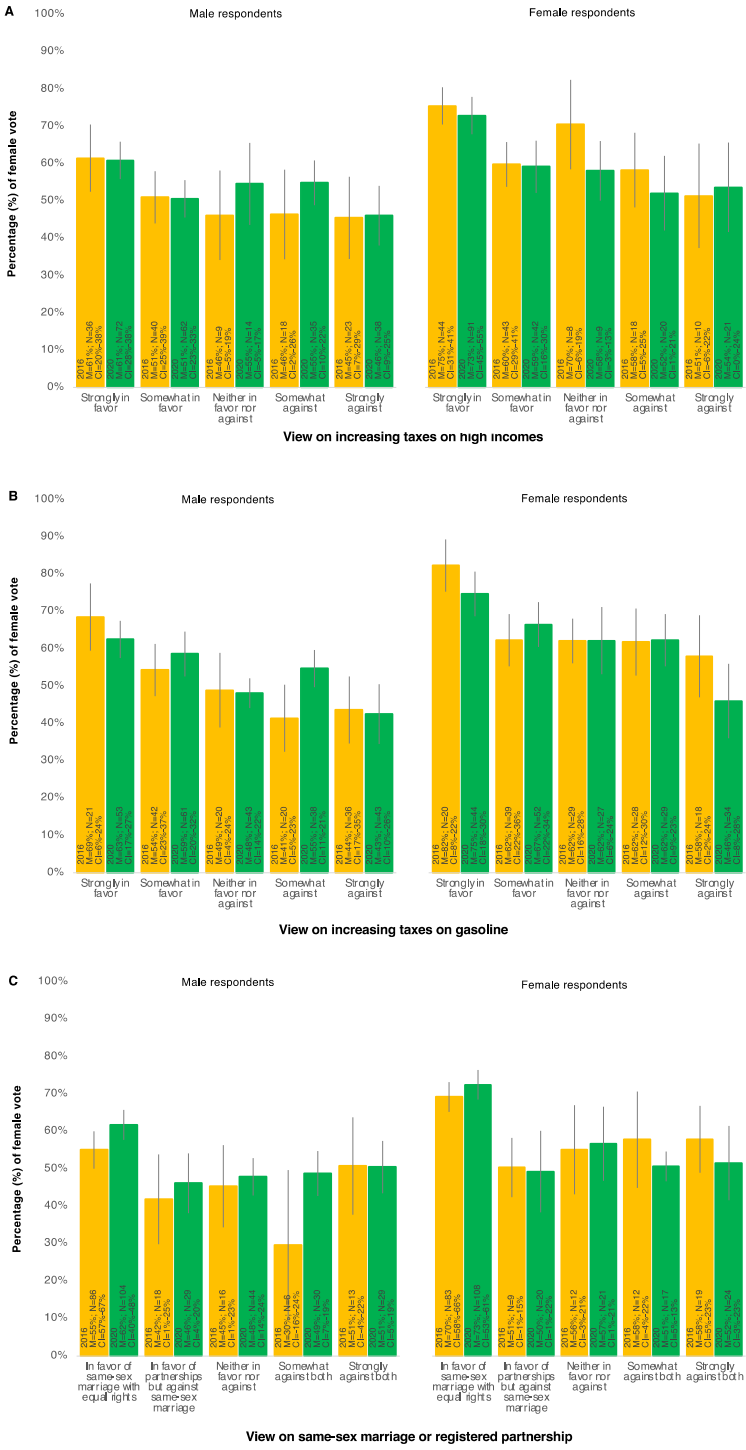

Fig. A10. Percentage of female vote depending on respondents' ideological views in three central dimensions of contemporary American politics: (A) support for redistribution, (B) environmental policies, and (C) social issues (see Fig. A9). Results shown separately for studies I (orange) and II (green). Only experimental elections with one female and one male candidate are included. Experimental elections with abstentions are excluded. Error bars depict 95% CI of the mean. N: number of respondents.

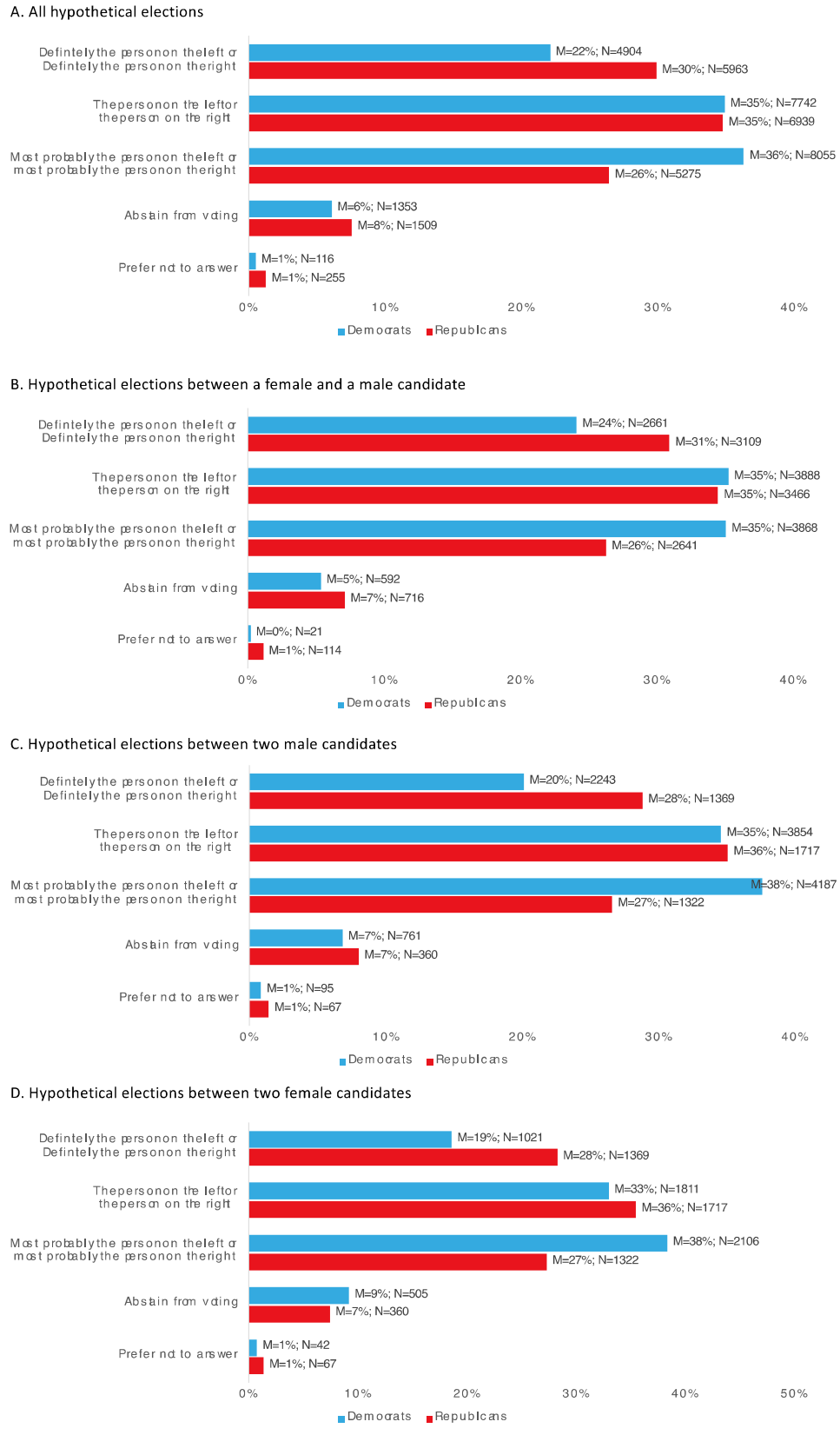

Fig. A11. Confidence in vote choice among Democrats and Republicans. (Party classification based on Presidential voting intention)

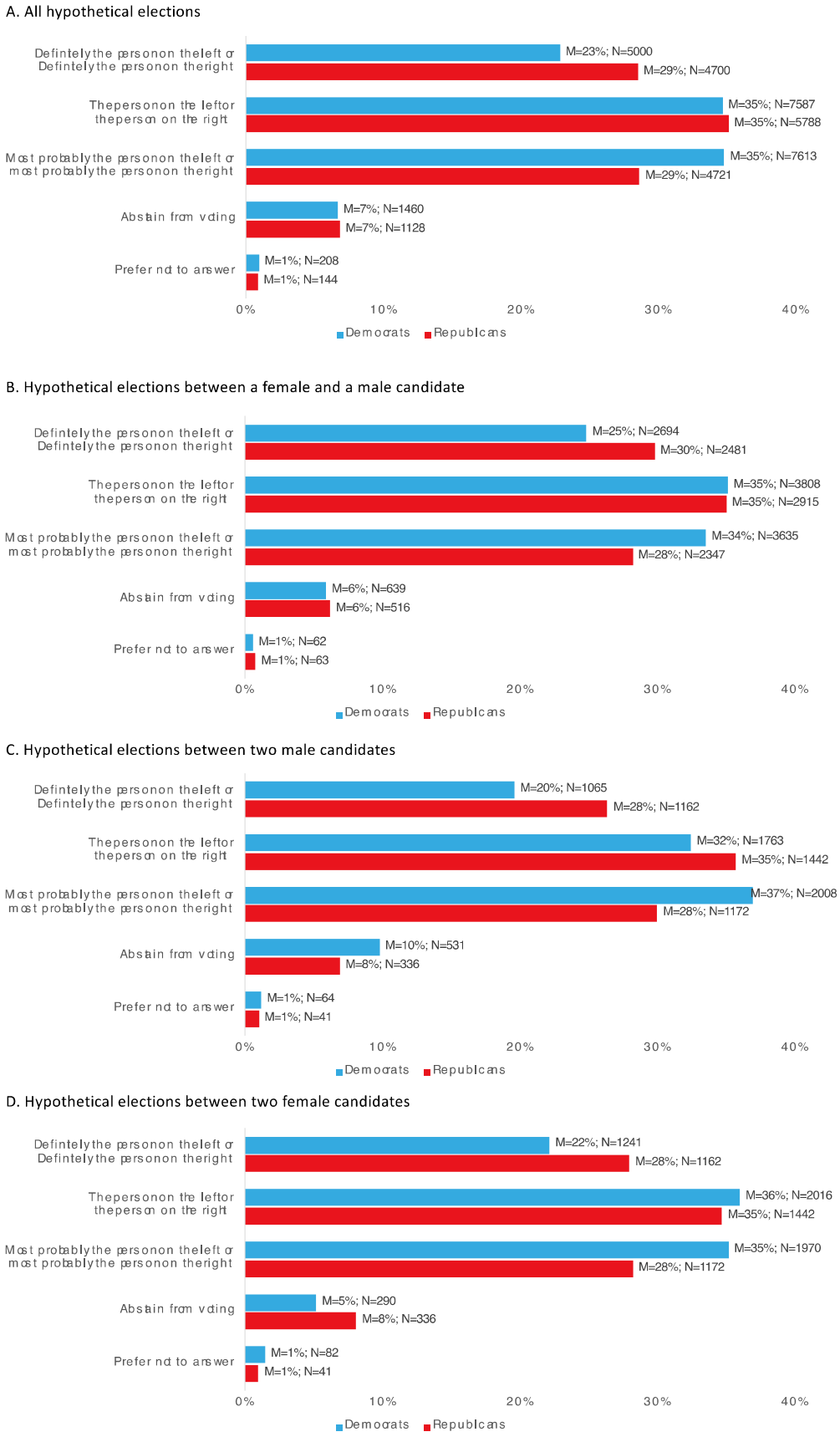

Fig. A12. Confidence in vote choice among Democrats and Republicans. (Party classification based on self-identification)

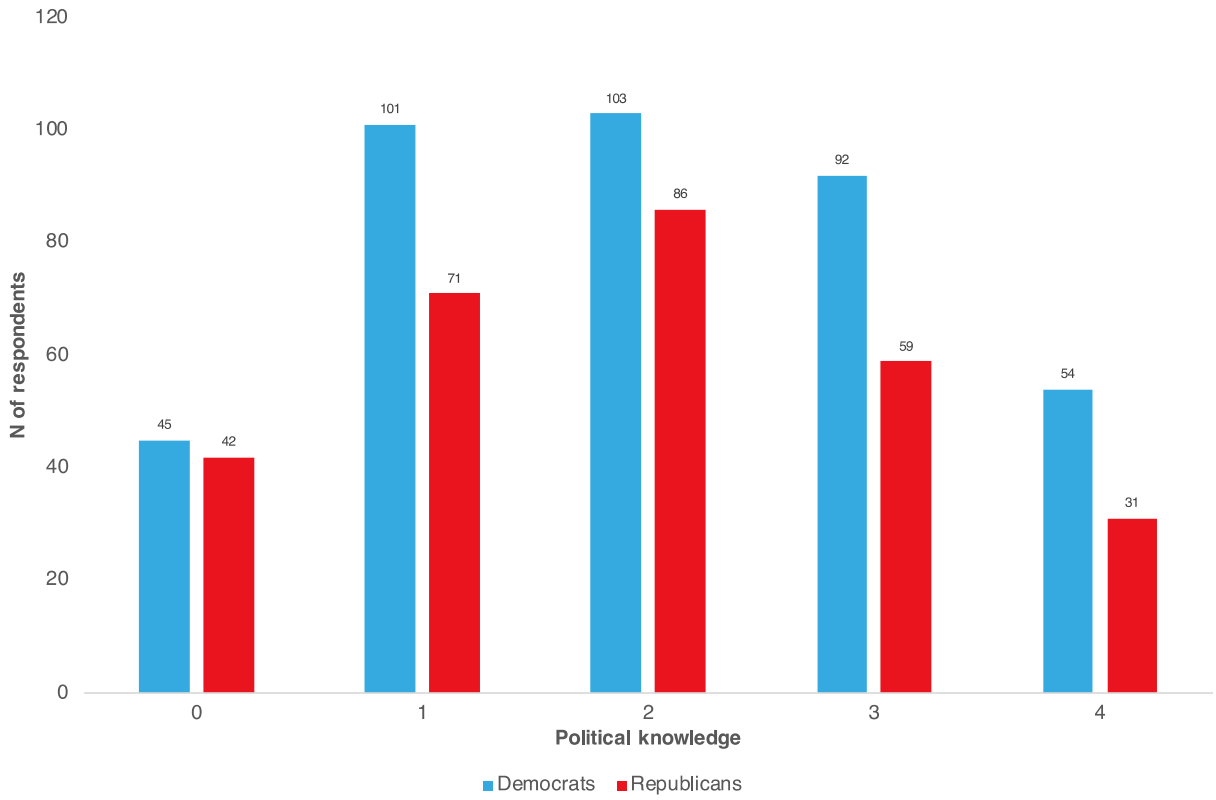

Fig. A13. Number of respondents in experimental elections by partisanship and political knowledge

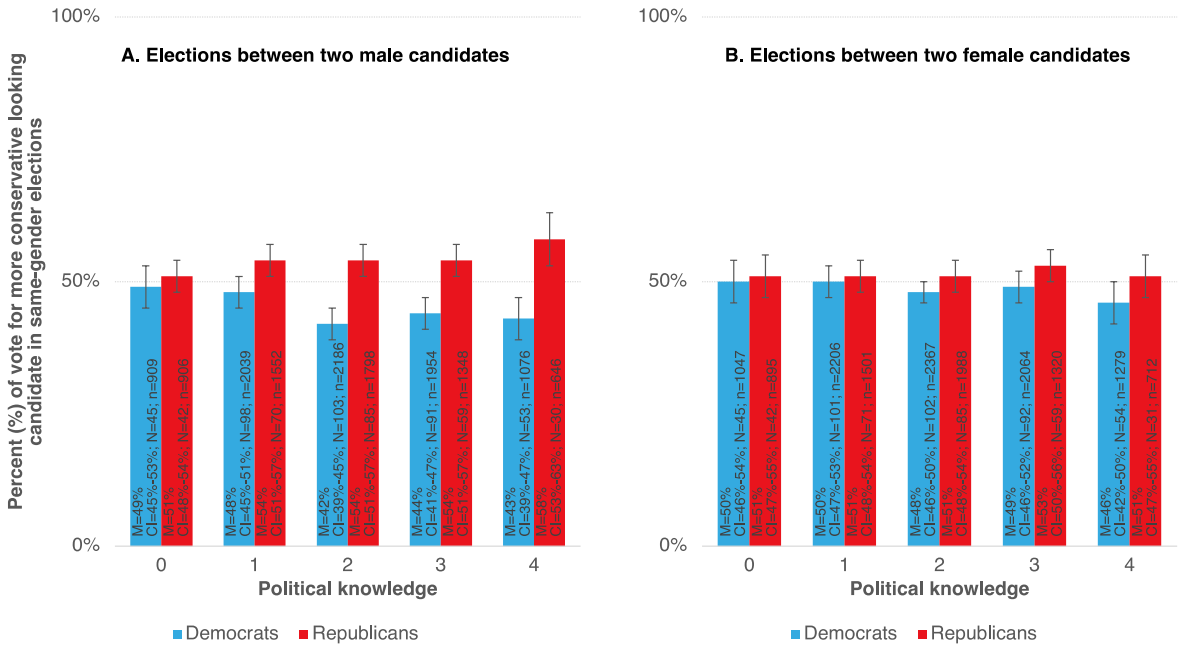

Fig. A14. Percent (%) of conservative vote in same-gender elections by respondent partisanship and political knowledge