Title
Still Not Equal: Heuristics and Gender Stereotyping in American Elections

Author
Bernhard, Rachel Velázquez

Publication Date
2018

Peer reviewed|Thesis/dissertation

Still Not Equal: Heuristics and Gender Stereotyping in American Elections

By

Rachel Velázquez Bernhard

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Political Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Gabriel Lenz, Co-Chair
Associate Professor Laura Stoker, Co-Chair
Associate Professor Sarah Anzia
Professor Ruth Collier
Associate Professor Amy Lerman

Summer 2018

Abstract

Still Not Equal: Heuristics and Gender Stereotyping in American Elections

by

Rachel Velázquez Bernhard

Doctor of Philosophy in Political Science

University of California, Berkeley

Associate Professor Gabriel Lenz, Co-Chair

Associate Professor Laura Stoker, Co-Chair

Research on gender in politics has long emphasized the gender gap—the number or proportion of women in office—as its chosen measure of women's attainment. When women win elections at rates equal to men, as they do in the U.S., many scholars thus conclude that candidate sex must play little or no role in voters' decisions. Yet the question of *which* women win should interest scholars of gender as much as *how many* women win—and tells us a great deal more about how politics is gendered.

How then does candidate gender enter into voters' decisions? Modern theories of voting behavior, observing that acquiring information about candidates is costly, often conclude that supplying voters with more information about the candidates would obviate reliance on cognitive shortcuts with known biases. This dissertation stakes a different claim: while heuristics and stereotypes do enable voters to "fill in" incomplete information, they also allow voters to simplify overwhelming information. Gender, as one of the simplest and most ubiquitous organizing principles of human behavior, helps us address the difficult task of voting in a candidate-centered electoral system by reducing a multidimensional task to a simple set of guidelines. This ensures biases in how we vote not only in low-information settings, but in high-information settings as well.

In this dissertation, I use experiments and election data to test whether our decision-making process affects the descriptive representation of women in the U.S. I present evidence that voter reliance on gender as a cue affects the type of women elected at every level of office. In Chapter 1, I examine state legislative elections in Oregon and find that voters winnow the field by selecting women with attractive, stereotype-congruent appearances, even though voters have a great deal of information and partisan cues in such races. In Chapter 2, I show that in local elections in California, where voters often know little about candidates, women fare better when they run for offices that fit feminine stereotypes (city council and school board) and worse in stereotype-incongruent offices (mayor). In Chapter 3, I show that even in high-information and high-salience national elections, partisan voters feel more favorable towards women who exhibit gendered leadership styles congruent with their party's preferences: feminine for Democrats, masculine for Republicans.

The familiarity and simplicity of gender as an organizing principle suggests the use of gendered heuristics and stereotypes is inherent to any political system in which voters evaluate individual candidates. This holds whether voters have a great deal of information about candidates or only a little, in elections of high- and low-salience, and across both partisan and non-partisan races. Even when women win elections at rates equal to men, women who run counter to stereotype—in appearance, in office type, in party—are at a disadvantage compared to women who fit stereotypes.

To Miguel, who reminds me every day to have courage and to be kind.

Table of Contents

List of Tables

List of Figures

Acknowledgements

"I set out to become the greatest lover in Vienna, the greatest horseman in Austria, and the greatest economist in the world. Alas, for the illusions of youth: as a horseman, I was never really first-rate." –Joseph A. Schumpeter

For one brief and yet endless summer in my childhood, my dad signed my siblings and me up for tennis lessons. I was terrible at tennis, which made this an unwelcome development. It was made worse by the fact that, as an odd number, we often ended up having to play other children, who were mostly better at tennis than we were, or the instructor, who was definitely better at tennis than we were. I whined about this injustice to my father one evening, to which he responded, "You should always play against people who are better than you. That's the only way you'll get better."

I built my dissertation committee by following my dad's advice, and have never regretted it. Each one of them built their career through an unswerving devotion to tackling hard questions about politics. Each was not only distinctly better at doing this than I was, but was willing to teach, push, and encourage me to do better.

Gabe Lenz, who embodies the following dictum from Richard Feynman, taught me "how not to fool yourself, because you are the easiest person to fool." Gabe poured time into making me a competent and above all ethical researcher: supervising my first forays into Stata, shepherding my master's thesis when I was stuck between comparative and American politics, asking "could something else be driving your results?" at every lab meeting. As a supervisor, he is a warm, kind, and unrelenting skeptic, which is one of the best things a young researcher could hope for in their graduate training. One of the most memorable moments I had as a graduate student came at his suggestion. After gently identifying that my abstract writing was not as strong as it could be, he and Laura Stoker suggested I send them a new abstract of my dissertation—every day. I sent them my best efforts each morning, and by early evening would receive comments. After two weeks of the exercise, I pled exhaustion and fled—but his concern (and perhaps to a lesser extent, the training) stuck, and I've never had a conference proposal turned down since.

Laura Stoker was the first Berkeley political science faculty member I ever met, one of the first to recruit me to "the dark side" (American politics), and the first person to get on the phone with me when things got tough. I first met Laura before I had even accepted my offer to UC Berkeley; I took BART to go to the admissions visit weekend, and sat next to a woman with curly hair answering emails on her phone. I pulled out my admissions brochure, and noticed that she was suddenly looking over. I looked up, and she said, "Are you an admit to the political science department?" From that moment on, Laura made it her business to make sure I was heading in the right direction, both literally and figuratively. She whupped my butt through her classes, often writing two pages of sharp comments for every page I wrote, and never accepted the slightest hint of mediocrity. At my first practice job talk, which was 37 minutes long, Laura stayed to give me an hour and 45 minutes of comments. She is, simply, more comprehensively brilliant than anyone else I've ever met. The good parts of my research and teaching all bear her fingerprints.

Sarah Anzia began her tenure on my committee as my outside member and ended as one of my closest collaborators. Though Sarah had sworn an oath not to do further work on gender and politics, in the end her commitment to answering the hard questions—and her remarkable kindness—made her willing to work with me, first as a supervisor, and then as a coauthor. Sarah is undoubtedly the most rigorous, diligent, and careful person I've met in academia, and it has been an incredible opportunity to observe her work firsthand. This, combined with our mutual "inability to undertake straightforward projects," made for a fantastic learning environment, one that I hope will continue long after our paper together finds a home.

Where Sarah brought care and sustained diligence, Amy Lerman brought the lightning. While she and I did not meet until much later in graduate school, I left every interaction with a thousand new ideas, perhaps two of which are good enough that she and I will eventually get around to testing. Despite the intellectual firepower Amy contributes to the committee, some of the biggest contributions she made came in terms of in tangible sponsorship: nominating me for invitation-only conferences, grant-writing for research assistantships, and underwriting costly airfare to Oxford for an interview that became a post-doc. Her amazing sense of humor, warmth, and generosity helped sustain me during some of the dark times that occur for everyone during the slog of graduate school.

In contrast to Amy, Ruth Berins Collier is the longest-standing member of my committee, and thus perhaps the least sure as to why she is still on it, though there has never been any question in my mind that she should be. She and David Collier originally recruited me to Berkeley as a Latin Americanist; unsurprisingly, I leapt at the chance. Ruth and David have produced a long line of exceptional scholars, and although our paths diverged early on, Ruth has continued to (bemusedly) school me in how to be a better thinker and writer, and I am grateful that she has chosen to take the path less traveled with me. [Ruth, the split infinitive in the last sentence is for you.] Ruth's feedback typically consists of a dreadfully hard question, asked in a gentle tone that gives one the impression the question is answerable. It may be, but certainly not in the time in which one has stuttered out an answer. Such questions are, nonetheless, the only ones worth answering.

My committee has been helped by many other advocates, cheerleaders, and Good Samaritans. Eric Schickler, my master's thesis supervisor, gave me invaluable advice during my first semester: "it's always good to keep [research designs] simple, because the results will complicate themselves." He also made an excellent chair, colleague, collaborator, advisor, and all-around mensch. Eric, I'm glad I got to be your boss, even if just in Lea's eyes. Leo Arriola was the first to recruit me to Berkeley, and patiently suffered through the period of time where I had not yet learned R. Sean Gailmard was the first to recruit me to American politics, and I will never forget that he was the first to offer to podcast classes for me when my partner's broken back prevented me from attending class. Alison Post gave me my first research assistantship, and endless encouragement even as I transitioned from comparative to American politics. Rob Van Houweling sat through and commented on an innumerable series of practice job talks. Terri Bimes provided warmth and encouragement and real financial support whenever possible. Jack Citrin was a brilliant instructor of political psychology and entrusted me with running the IGS poll during the most important election in recent history. Jas Sekhon never let me forget how smart I was, even when I was feeling exceptionally stupid in his classes. Joel Middleton, listening to me complain about a failed replication, gave me the best advice I received in graduate school: "it turns out doing ethical research is way harder than doing unethical research." I owe thanks to Wendy Brown, Aila Matanock, Michaela Mattes, and Jennifer Bussell for encouraging me (and the other women in the department) to stay the course. Jason Wittenberg shared his love of physics and sense of humor. Kinch Hoekstra, last but certainly not least, allowed an ignorant Americanist and lapsed classicist to audit his class and in so doing rediscover what she loved about studying politics.

Beyond the department, my coauthors made sure I had the breadth and depth of expertise in gender and politics necessary to make real contributions to the field. Shauna Shames was my first APSA discussant, first coauthor, and first to celebrate my successes and encourage me through my failures. Dawn Teele got me to EGEN, read drafts, took phone calls from me late at night and early in the morning, and made calls on my behalf when I was on the market even though I wasn't her student. Mirya Holman lent wigs and quotation inspiration and job market materials and connected me with the greater GPP community. The women of EGEN and GPP, in turn (too many to name, though I am grateful to them all), shared data, commented on papers, and cheered for me. More

importantly, they taught me how to be the kind of academic I had dreamt of but begun to believe was impossible.

Just as the dissertation does not cease with the Ph.D. granted, it also does not begin in graduate school. Earlier mentors convinced me to pursue my education, and more broadly, to pursue anything difficult that I might love, "because any fool knows that to work hard at something you want to accomplish is the only way to be happy" (Eugene O'Neill). Gabe Del Real taught me to think, and how to love myself even when I am wrong, which is often. Jim Lever gave me my first research job, first publication, and first lessons on how joyful research can be. Daryl Press convinced me not to leave the classroom because it was full of dudes (his words) and, more importantly, that the classroom would be better with me in it. Jim Doig wrote me rec letter after rec letter, including an incredibly charming one to a friend who was the chair of a top research department, which said—as they were only hiring in IR—that I was a "quick study." Paul Christesen taught me to write, and how to thrive on the difficult. He has only himself to blame if I have been difficult ever since.

I also benefited from what I believe, immodestly, is the best department of political science graduate students in the country. I made the decision to come to Berkeley because of them, and they were worth every drop of sweat, tears, and blood along the way. My cohort—Gabrielle Elul, Sean Freeder, Jake Grumbach, Adam Lichtenheld, Nathan Pippenger, Paul Thissen, and Sherry Zaks—are all smarter than I, and more importantly, warm and funny and willing to read endless drafts of the master's thesis and prospectus and conference papers and dissertation. I would not have made it through graduate school without them. Extra thanks go to Sean, my consortium buddy, coauthor, and dissertation co-parent, who has shaped my thinking and writing for the better. Thanks are also due to my "cohort-plus:" Doug Ahler, Natalie Ahn, Richard Ashcroft, Caroline Brandt, Laurel Eckhouse, Ryan Hubert, Rachel Strohm, Fiona Shen-Bayh, Rochelle Terman, and Guadalupe Tuñon contributed much love, useful commentary, job market advice, cocktails, and other ingredients of a successful Ph.D. Jae Yeon Kim was a good student, excellent GSI, and better friend. My Ph.D. students, including Melissa Carlson, Julia Christesen, Casey Ste. Claire, Charlotte Hill, Laura Jakli, Jiali Li, Elizabeth Mitchell, David Nield, and Jacqueline Vokoun Duran brightened both my personal and professional life, and give me hope that the next generation of political scientists will be even better.

My dissertation would not exist at all without my research assistants and research subjects. Marissa Lei Aclan, Grayson Dimick, Rikio Inouye, Julia Konstantinovsky, Jesse Luís Lopez, Caitlin Olson, Ernesto Rojas, and Alan Yan (now coauthor) filled in endless gender codings, tracked down obscure election results, and made me smile every week I got to work with them. Following Mirya Holman's advice, I only work with people smarter than I, and these students and former students were the ones who fit the bill. There would be very little data left to analyze without their contributions. Likewise, I owe the many anonymous Turkers, SSI, and YouGov respondents thanks for taking my surveys—especially the Turkers, who often bore the brunt of debugging pilot studies. Finally, I owe Emerge California and Emerge America, both staff and alumnae, a huge debt of gratitude for sharing so much sensitive and awe-inspiring data about what it takes to be a woman in American politics in the twenty-first century. Andrea Dew Steele, Kimberly Ellis, and Nazneen Rydhan-Foster went above and beyond by enabling me to observe their programs and people at work, and I look forward to paying that debt by getting more women in office.

Finally, there could be no dissertation without a village of people (who can't understand how it could possibly take six years) standing behind it and propping it up. My family, Peter, Susan, Erica, and Dan, listened to me kvetch, patiently problem-solved, duct-taped what could be duct-taped, and "walked off" what could not. Mi segunda familia, Miguel, Annel, Adrian, Kristelle, y Ale: son la mejor familia que una mujer podría pedir. My six sisters from Athenian, Chelsea, Beth, Sarah, Monica, Sarah, and Claire, are the best kind of nasty women: brilliant, funny, loving, and willing to

drop anything to proofread. My Dartmouth buddies, Hillary, Arran, Chris, Allen, Tom, Alejandro, Rissie, Bean, Elisse, Allie, Heather, Diane, Lauren, Xandra, Jenna, and Vanessa. Friends old and new—Alex, Ray, Gary, Emily, Chris, and Kerry—brought so much joy into my life and picked me up every time I was down. Above all, I owe my husband Miguel thanks. He has read draft upon draft, listened to complaint upon complaint; donated frequent flyer miles to conference travel and nightstands to library books; made dinners, breakfasts, cups of tea, arepas, brownies; mailed grant applications, filed paperwork, purchased scanners and pens; he has gone to bed alone to the sound of keystrokes in another room; he has flown halfway around the world expecting to be on vacation only to hear that a term paper, a conference paper, a job market paper is not quite done; and he has, for all that, never called me anything worse than a turd. Miguel, this dissertation is for you.

Chapter 1

Tinder Decides: Mate Desirability Influences Votes

Democracy is premised on voters' ability to identify qualified candidates for office. However, extensive evidence suggests that candidate appearance has a non-trivial impact on voter decision-making. Voters' evaluations of candidates' photos for perceived competence, attractiveness, dominance, and sex typicality predict real election outcomes. Social scientists often argue that the brain's tendency to take cognitive shortcuts explains this phenomenon, but this still fails to explain why the brain takes one shortcut over another. This article provides experimental and observational evidence that, when visual cues are provided, the candidates whom voters believe to be most qualified are those perceived to be most desirable as romantic partners. I find evidence of this "Tinder mentality" even in reasonably high-information real elections where voters see candidates' other qualifications. Voters select female candidates on their physical attractiveness and apparent nurturing ability, and male candidates on their attractiveness and ability to provide for their families. When aggregated, these tendencies regularly affect election outcomes and may have attendant pernicious consequences for descriptive representation of women, as well as for democratic accountability in locales that hold direct elections.

How do we judge whether candidates are qualified to hold office? Worryingly, research suggests that visually appealing politicians fare better with voters than their qualifications alone would predict (Ahler, Citrin, Dougal, & Lenz, 2016; Banducci et al., 2008; Lawson et al., 2010; Todorov et al., 2005). Other evidence suggests that voters find it especially challenging to assess female candidates' qualifications objectively: even though women tend to be more qualified and more effective in office than men (Anzia & Berry, 2011; Fulton, 2012; Milyo & Schosberg, 2000), voters are more likely to inquire about women's qualifications than men's and to penalize women when their qualifications are in doubt (Ditonto, 2016; Ditonto, Hamilton, & Redlawsk, 2014).

Evaluating candidates' qualifications is hard, which incentivizes the use of heuristics. It may therefore seem unsurprising that we rely on snap judgments of appearance instead. Nevertheless, these empirical findings provide an incomplete picture of the psychological mechanisms behind this heuristic. Looking at faces may be easier than analyzing complex political information, but scholars do not yet understand what makes someone "look" qualified—let alone whether such a heuristic might have different implications for men's and women's candidacies. Moreover, evaluating candidate qualifications appears to be a difficult task whether voters have a great deal of information or insufficient information about the candidates.

When faced with a problem that is difficult to solve, humans often inadvertently substitute an easier one (Tversky & Kahneman, 1974). I argue that when presented with the difficult and unfamiliar problem of evaluating candidate qualifications we instead inadvertently ask a question familiar because humans have had to answer it for millennia. Who do we think would be a good partner? Who would we trust with our kids? Who would we want to see every day?

I present evidence from both survey experiments and analyses of real elections in Oregon that perceived mate desirability—a more familiar heuristic—predicts voting behavior. I find that when shown a photo, voters exhibit a "Tinder mentality," substituting an assessment of the individual's appeal as a long-term partner for a more holistic evaluation of the candidate's qualifications in both surveys and in real elections. In keeping with other research on mate selection, this means that voters evaluate female candidates primarily on their physical attractiveness and secondarily on their perceived ability to nurture, while voters evaluate male candidates primarily on their perceived ability to provide for and protect others and secondarily on their physical attractiveness (Buss & Schmitt, 1993; Darwin, 1888; Eagly & Wood, 2013).

This behavior has troubling consequences for descriptive representation and democratic accountability. Candidates running for office who do not fit a socially prescribed mold will face a harder road to office, i.e., discrimination. For instance, voters could select a man who is politically inexperienced, but who appears to be a good provider and protector, over a woman who is more politically experienced but lacks sex appeal. Moreover, increasing amounts of available information, as is case in the U.S. in the twenty-first century, may not increase the likelihood that voters make better decisions about which candidate to vote for. If a person-centered evaluation task, rather than a low information setting, is sufficient to trigger use of heuristics, candidate-centered elections in many polities may be influenced by such cognitive shortcuts.

THEORY

We base our decisions on superficial information like candidates' appearance because modern direct democratic elections pose a formidable test of citizen competence. In the United States, thousands of offices are now elective, from president down to mayor and tree warden. Scholars of elections dating back to at least Berelson et al. (1954) suggest that under these circumstances most individuals fall short of fully informed, economically rational voting behavior. Instead, our behavior suggests that we are

cognitive misers attempting to maximize the utility of the limited information we do have while avoiding the time-consuming search needed for a fully informed vote (e.g., Conover & Feldman, 1989; Redlawsk, 2004). In many low-salience races, particularly at the state and local levels, sleuthing out information about candidates is the province of only the most sophisticated voters.[1] Moreover, many of these candidates are new to politics and thus have no record. Given this reality, the scholarship on heuristics examines whether and to what extent voters can approximate full information given a limited information environment.

As Tversky and Kahneman (1974) point out, "these heuristics are quite useful, but sometimes they lead to severe and systematic errors" (p. 1124). Heuristics operate by providing simplifying principles; these principles are revealed when framing the same decision problem in different ways leads to different (and predictable) outcomes (Tversky & Kahneman, 1981, p. 453). For instance, a shape that is slightly blurred at the edges will usually be perceived as farther away than one that is crisply outlined; an easy question, "how blurry is the object?" has thus been substituted for the more involved "exactly how far away is the object?" (Tversky & Kahneman, 1974). This process, called attribute substitution, saves time and cognitive effort—without such simplifications, we would be unable to navigate through a busy intersection—but has the potential to produce systematic distortions.

Voters may be even more likely to substitute assessments of other traits for female candidates' qualifications because women are stereotyped as less qualified. Though a major review of the American literature on gender and politics concluded that "discrimination has fallen out of favor as an explanation for women's absence from electoral politics. The public's attitudes toward women in politics have evolved" (Lawless, 2015, p. 352), other work suggests that voters may routinely overlook female candidates' qualifications. Psychological studies find that women are stereotyped as warm, while men are stereotyped as competent (Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002). Studies of real elections find that female candidates are typically better-qualified than male candidates (Fulton, 2012; Milyo & Schosberg, 2000), and secure more benefits for their constituents once in office (Anzia & Berry, 2011). In interviews with political elites, Dittmar (2015) finds evidence of a broad conviction that women must prove their credentials while men's are assumed. Experimental studies show that voters doubt women's qualifications and penalize women with dubious qualifications more harshly (Ditonto, 2016; Ditonto et al., 2014). Perhaps most concerningly, providing voters with explicit information about women's qualifications attenuates but does not eradicate bias against female candidates in a significant portion of voters (Mo, 2015). In other words, if voters fail to perceive or focus on a woman's qualifications, they may instead be evaluating other aspects of her person or platform, including appearance.

If the nature of modern direct elections makes it challenging for voters to assess numerous candidates' qualifications, and gender stereotypes exacerbate this problem for voters evaluating women, what criteria might be substituted instead? Significant evidence suggests that candidate appearance has a non-trivial impact on voters' behavior. Voters' evaluations of candidates' photos for perceived competence (Todorov, Mandisodza, Goren, & Hall, 2005; Olivola & Todorov, 2010; Lawson, Lenz, Baker, & Myers, 2010; Lenz & Lawson, 2011), attractiveness (Little, Burriss, Jones, & Roberts, 2007; Banducci et al., 2008; Little, Jones, & DeBruine, 2011), dominance (Little et al., 2007), and sex typicality (Carpinella & Johnson, 2013b, 2013a; Hehman, Carpinella, Johnson, Leitner, &

[1] In an era of party polarization, the simplicity and high signal-to-noise ratio conveyed by candidate partisanship and endorsements make for straightforward voting heuristics. Accordingly, political scientists tend to look upon these heuristics more favorably (e.g., Arceneaux & Kolodny, 2009) than they do more complex heuristics like retrospective voting (e.g., Achen & Bartels, 2016; Healy & Malhotra, 2013). Nonetheless, in many—perhaps most—of the electoral races in which Americans are eligible to vote, heuristics based on partisanship are irrelevant or serve little purpose.

Freeman, 2014; Carpinella, Hehman, Freeman, & Johnson, 2015) all predict both experimental and real election outcomes.[2]

Researchers frequently ascribe these behaviors to a failure of the brain's System 1/System 2 processing. The intuitive System 1, upon seeing a candidate, offers an automatic, valenced response (e.g., "This candidate is good-looking") that the lazy but rational System 2 fails to detect or correct, influencing subsequent judgments of the candidate (Kahneman, 2011; see Mo, 2015, p. 357 and Todorov et al., 2005, p. 1624). Nevertheless, the appearance literature lacks an explanation for why so many conceptually distinct measures should all meaningfully predict voting behavior. It seems implausible that such findings are compatible because different traits share common physical features: babyfacedness appears to predict competence (Poutvaara, Jordahl, & Berggren, 2009), testosterone-driven sexual dimorphism predicts dominance (Little et al., 2007), and facial symmetry, attractiveness (Little et al., 2011). Yet many of these features are typically incompatible (e.g., babyfacedness and high testosterone expression). Likewise, Spezio et al. (2012) find that when candidates' faces are hidden in photos, respondent evaluations of the non-facial cues in photos still predict election outcomes. One possibility is that each of these traits represents an aspect (of variable importance) of a more complex assessment of socially prescribed partner desirability.

ARGUMENT

> *"Powerful men are sexy, sexy women are powerful, and these propositions are not at all the same." – Kathleen Jamieson (1995, p. 151)*

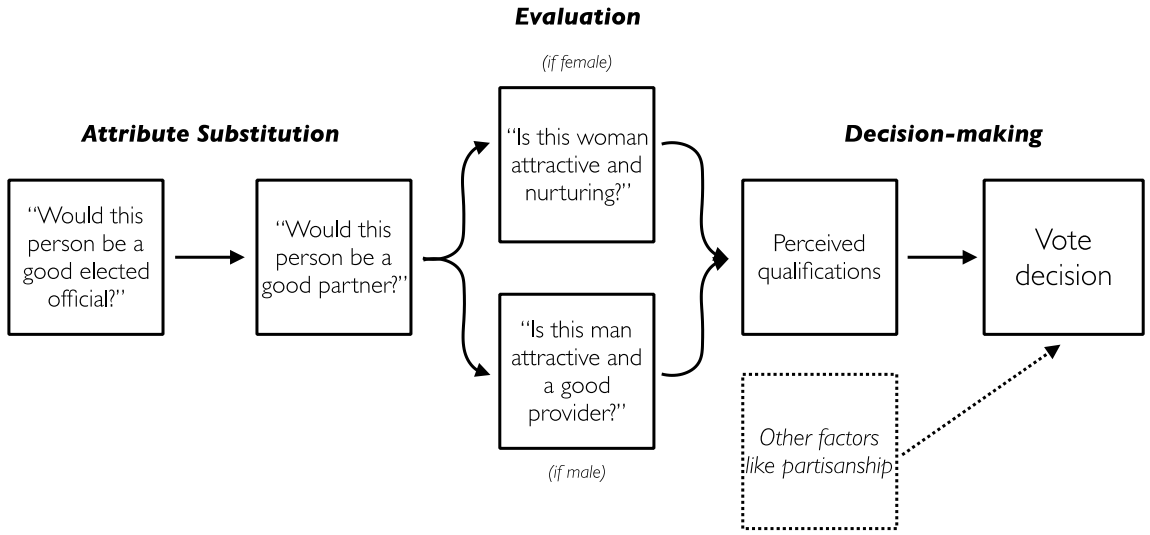
I argue that voters employ a three-step process to evaluate candidates. First, faced with the hard problem of assessing candidate qualifications, voters will engage in attribute substitution, inadvertently assessing instead a more familiar question: how appealing a candidate seems as a long-term partner. Second, during the assessment itself, voters will judge female and male candidates on different criteria. Third, this information will be turned back into what the voter believes is an assessment of the candidate's qualifications, which is then combined with other non-mate criteria (e.g., partisanship) to make a vote decision. Figure 1.1 outlines the proposed psychological process.

This type of task is not, in fact, unfamiliar to the human brain: to vote for candidates, we must decide whether we want to make some reasonably long-term, albeit impersonal, commitment to a stranger. Whether one believes that the criteria on which we evaluate potential for long-term partnership reflects evolutionary strategy, culturally instilled preference, or some combination of the two, the reality is that we spend years of our lives evaluating mate potential—not just to find our own partners, but on behalf of friends and relatives as well—and mere days voting. The familiarity of the former task and unfamiliarity of the latter could not be more stark. In other words, we "know" how to evaluate mate desirability. We do not "know" how to assess a candidate's fitness for office. In the face of uncertainty, our choices will err towards candidates who we believe make appealing partners. This generates the first hypothesis:

H1: voters will judge candidates' faces against the traits of an ideal mate.

[2] Though some research raises questions about whether this effect reflects voter preferences and not strategic candidate entry (Atkinson, Enos, & Hill, 2009), more recent experimental work suggests that photos have effects in real-world elections (Ahler, Citrin, Dougal, & Lenz, 2016), and the findings hold across a number of countries and levels of elections (Lawson, Lenz, Baker, & Myers, 2010).

Figure 1.1. The Evaluation Process



Attribute substitution allows our lazy System 2 to reduce a holistic evaluation of a candidate's qualifications to a quick assessment of mate desirability carried out by System 1. In Kahneman and Tversky's nomenclature, an assessment of mate desirability might be termed a representativeness heuristic:[3] "does this person resemble an ideal partner?" Other psychologists call the same phenomenon judgment against a prototype (e.g., Johnson, Murphy, Zewdie, & Reichard, 2008). Regardless of nomenclature, the expectation is that people compare to an ideal, rather than on their own personal preferences.

The literature examining ideal mate criteria is contentious, to say the least. Darwin (1888) originally developed the theory of sexual selection to explain speciation, and subsequent evolutionary biologists and psychologists expanded the theory, arguing that the relative costliness of reproduction for each sex determines mate preferences: men select female partners primarily on their physical attractiveness and secondarily on their apparent nurturing ability, while women select men about equally on their ability to provide for and protect their families and on their physical attractiveness (Buss, 1989; Buss & Barnes, 1986; Buss & Schmitt, 1993). In contrast, some sociologists and psychologists have argued that these criteria are culturally determined and symptomatic of men's structural advantages over women (Eagly & Wood, 2013; Wood & Eagly, 2012; Zentner & Eagly, 2015): if women were the ones holding power, they would prioritize attractiveness and nurturing skills in men, while men would seek powerful women (Zentner & Eagly, 2015). Indeed, some studies find that a society's gender equality correlates with lessened sex differentiation in mate preferences (Zentner & Mitura, 2012). Immense differences in theoretical origin and implications notwithstanding, both sides tend to agree that physical attractiveness and ability to care for offspring, which I term

[3] "Tversky and Kahneman conjectured that observers expect the statistics of a sample to closely resemble (or 'represent') the corresponding population parameters, even when the sample is small. This 'representation hypothesis' soon led to the idea of a 'representativeness heuristic,' according to which some probability judgments (the likelihood that X is a Y) are mediated by assessments of resemblance (the degree to which X 'looks like' a Y)" (Kahneman & Frederick, 2002, p. 879).

motherliness/fatherliness, are the most predictable criteria on which we select long-term romantic partners. This generates the second hypothesis:

> **H1A: voters will judge female candidates' faces primarily on attractiveness and secondarily on perceived motherliness, and male candidates' faces primarily on perceived fatherliness and secondarily on attractiveness.**

Moreover, unlike other theories of appearance cues, mate selection theory predicts an interaction between candidate age and candidate sex. An important subtext of arguments made by biologists about mate desirability is that physical attractiveness functions as a proxy for fertility. For women, fertility declines rapidly, while for men, fertility declines very slightly over the course of the human lifespan (Velde, R, & Pearson, 2002).[4] If mate selection is at work, we should expect that respondents' perceptions of women's qualifications to decline more rapidly as women age than perceptions of men's qualifications do as men age.

> **H1B: voters will judge older female candidates as much less qualified than younger female candidates, while male candidates will face only a slight age penalty.**

Finally, prior research demonstrates that appearance cues affect behavior both in the ultra-low-information setting of survey experiments and in the slightly higher information context of real elections. In particular, Todorov et. al (2005) demonstrate that ratings of facial competence, rather than cues like facial attractiveness or facial dominance, best predict election outcomes. Accordingly, any novel theory of appearance cues should explain voting in both low-information (experimental) and high-information (election) contexts, including controlling for competence ratings, to be considered a meaningful contribution.

> **H2: voters' assessments of mate desirability should predict votes in both low-information (surveys) and high-information (elections) settings.**

EMPIRICAL STRATEGY

My proposition makes claims about both underlying psychological processes and observable voting behavior. To address both, I focus on a real-world case, the 2000-2014 Oregon state legislative elections, for which we can readily determine the information voters are likely to have about candidates.

Because Oregon is an entirely vote-by-mail state in which all voters receive a state-issued voting pamphlet, this case substantially improves the study of candidate evaluation in three ways. First, voters are likely to receive these cues (e.g., candidate photos), so inferences do not depend on voters being highly informed or receiving information through the media. Second, voting-by-mail means that we can measure some of the information available to voters; this makes a better case for selection on observables than is true in other contexts. In races with almost no information, voters might be likely to seek out additional information, while voters in high-salience races might be exposed to significant additional information via the media. Either makes omitted variable bias more likely than in the Oregon context. Third, to the extent that voters are better informed about candidates' policy

[4] It is worth noting that the study of human women's fertility is undergoing rapid change: while older studies seemed to indicate that women's fertility declined precipitously even as early as age 30, more recent research has questioned that account for reasons both historical and methodological. Rather than specify an age of infertility for women to be treated as a quasi-discontinuity, I have opted for a simple sex*age interaction term.

preferences than predicted, the effects of any given heuristic—like appearance cues—should be attenuated, rather than exacerbated.

In turn, this makes three contributions to existing work on descriptive representation of women. First, using real candidates' photos and occupations as experimental stimuli increases external validity over experiments that use more limited realizations of these variables (e.g., factorial designs) or artificial vignettes about candidates. Second, it brings new data to bear: existing work has often been confined to Congressional and gubernatorial races due to the difficulty of collecting data in state and local elections, but the salience of partisanship in such races makes it hard to tell whether voters assess male and female candidates differently (cf. Hayes 2011). Third, state and local races are critical to understanding the pipeline through which women emerge as candidates for higher office, as women are more likely than men to start their political careers in local office (Carroll and Sanbonmatsu, 2013). If voters select for a certain type of female candidate, or impede others, that is worth knowing.

When voters sit down to decide, they see a mix of standardized and optional information. Figure 1.2, below, shows an example of a voting pamphlet. All candidates are listed with at least one partisan affiliation, an occupation and occupational history, a photo, their education, and their prior government experience. Candidates can also choose to include a statement or other personal information. In general, voters have more comprehensive information about candidates than most survey experiments testing heuristics or stereotyping provide. The standardized inclusion of information on partisanship and candidates' qualifications suggest that reliance on visual cues and stereotypes should be *less* than in environments where this information is not provided.

To create my sample, I scrape the 2000-2014 voting pamphlets to collect the standardized data: candidates' names, partisan affiliation, type of race, photos, occupation, education, and prior government experience. Data on election returns were scraped from the Oregon Secretary of State's website. Candidates were dropped for races that were not contested by both a Democrat and a Republican and for which one or both candidates did not submit photos. This created a pool of 816 candidates, out of which 789 had unique photos. 228 of the candidates are women, and 561 are men.[5]

I coded education and prior government experience using simple least-to-most scales, with graduate degrees and seat incumbency serving as the top points of their respective scales; details of the coding schemes are available in the SM. I rated candidates' photos and occupations on multiple traits using large samples of survey respondents. The aggregated ratings give each candidate a mean score for each trait. As I use four surveys and experiments to test **H1**-**H2**, I describe the procedures study-by-study in the results section.[6]

MATE DESIRABILITY PREDICTS VOTES

Design and Procedures

In Study 1, I assess whether mate desirability predicts vote preferences (**H1**). 3,245 Mechanical Turk (MTurk) respondents were recruited and randomly assigned to rate 30 candidate photos out of a subset (n=529) on one of three measures. One-third rate candidates on the dependent variable, vote choice ("how likely would you be to vote for this individual?"). One-third rate candidates on a single-question measure of the independent variable, mate desirability ("how appealing do you think others would find this person as a long-term romantic partner?"). One-third rate candidates on perceived competence ("how competent do you think this individual is?"), the question best shown to predict voting behavior using appearance cues. All three questions use a seven-point scale (e.g., "extremely

[5] Additional descriptive data is available in the Supplemental Materials (SM), available at: http://bit.ly/2vOgorp.
[6] Full details and results for each are available in the SM.

Figure 1.2. Example of an Oregon Voting Pamphlet

State Senator, 10th District

Jackie Pierce

Democrat (DEM)

Occupation: Retired Social Worker

Occupational Background: Psychiatric Social Worker, Oregon State Hospital 1989-2006; Fairview Training Center 1987-1989; Navy and Air Force Family Services Centers 1983-1986

Educational Background: B.A. Social Work California State University, MSW University of Southern California.

Prior Governmental Experience: Precinct Committee Person, Polk County

Personal: Husband Al Pierce

Community Service: Volunteer for American Red Cross, World Beat Festival, Friends of Bush Gardens, Member of Marion, Polk, Yamhill Counties Labor Council.

As an Oregonian I believe in respecting every individual. *(As your State Senator, I will take initiative in fighting for you)*. We must all work together *(with integrity)* to ensure the well being of all of Oregon.

JACKIE PIERCE
A Senator that will work for all Oregonians

As your State Senator my priorities will be:

Lay a solid foundation to help local small businesses succeed and hire workers
- Invest in technology and innovation to create green jobs
- Speed up construction jobs to rebuild our aging roads and schools

Protect vital services like education, senior care, and Head Start
- Extend unemployment benefits to out-of work Oregonians
- Provide better access to education, health care and jobs to our returning troops
- Make sure seniors and disabled Oregonians can remain in their own homes

Protect the natural places that make Oregon special
- Protect Oregon's waterways, parks beaches and watersheds
- Eliminate toxic chemicals from drinking water and our food supply

Hold Wall Street banks and credit card companies accountable when they break the law
- Give the Attorney General the power to go after health insurance companies when they break the law
- Provide incentives to reduce nickel-and-dime fees like ATM fees and credit card interest rates

(Human Dignity, Freedom and Choice)
- Reproductive Choices
- Choices for Death and Dying issues
- Freedom to choose who to marry

(This information furnished by Jackie Pierce.)

The above information has not been verified for accuracy by the State of Oregon.

State Senator, 10th District

Jackie Winters

Republican (REP)
Independent (IND)

Occupation: Small Business Owner; Jackie's Ribs, State Senator

Occupational Background: Public Agency Administrator

Educational Background: Jefferson HS; Portland, Oregon State System of Higher Education Continuing Education

Prior Governmental Experience: State Senator, Member; Joint Ways and Means Committee, Human Resources Subcommittee, Emergency Board, Quality Education Model Review Committee, State of Oregon Ombudsman, Assistant to Governor Atiyeh, Oregon State Executive Service, U.S. Air Force Academy Board of Visitors

A TRUE LEADER:

"Jackie Winters is committed to improving our economy and supporting Oregon businesses, families, and children. One of Jackie's many strengths is bringing people together to solve challenging issues facing our community and state. Please join us in voting for Jackie Winters." Dick Withnell

Oregon Building Trades Council
Oregon Business Association
National Federation of Independent Business/Oregon
Oregon Farm Bureau Federation
Oregon AFSCME Council 75
Oregon State Fire Fighters Council

BIPARTISAN SOLUTIONS IN THE FIGHT AGAINST METH

"Jackie Winters is a strong ally in our fight against the meth epidemic, and is committed to making our neighborhoods and families safer." Walter M. Beglau; Marion County District Attorney

Jason Myers; Marion County Sheriff
Oregon State Sheriffs' Association
Oregon State Police Officers' Association
Oregon Police Chiefs for Safer Communities

QUALITY CARE FOR OUR SENIORS AND CHILDREN

Senator Winters is a tireless advocate for our children, seniors and people with disabilities. She has always done everything possible to take care of the most vulnerable members of our society.

Wes Ediger; Retired Educator, Salem/Keizer School District
Oregon Health Care Association
Oregon Nurses Association
Oregon State Council for Retired Citizens
Citizens' Alliance for Responsible Education

SERVING AS YOUR SENATOR IS AN HONOR

"I learned from Governors Tom McCall and Victor Atiyeh that Oregon's greatest asset are its people and that our role as public servants is to help our constituents solve problems and represent them with honor and dignity. Thank you for the opportunity to serve you in the State Senate. I ask for your vote."

Jackie Winters

www.jackiewinters.com

(This information furnished by Friends of Jackie Winters.)

The above information has not been verified for accuracy by the State of Oregon.

unappealing" to "extremely appealing"). Every photo received approximately 57 unique respondents' ratings for each trait (vote choice, partner appeal, competence), which were aggregated into a mean rating for each candidate. For this and subsequent studies, the SM describes the procedures and findings in detail.

To analyze the results, I regress ratings of willingness to vote for a candidate on ratings of their appeal as a long-term partner. I add ratings of competence from the same study to a second regression model to see if the results hold even after accounting for the most likely alternative explanation.

Results

I test **H1** by regressing respondents' vote choices on ratings of candidates' mate desirability. Figure 1.3 shows evidence that each candidate's photographic partner appeal strongly predicts respondents' vote choice in surveys ($p<.001$). In Figure 1.3, each point represents a single candidate. Table 1.1 shows the same data for individual candidates broken out by candidate sex. Partner appeal predicts vote choice for both male (B=.34) and female (B=.29) candidates even after I control for facial competence, the strongest alternative explanation provided by existing literature on appearance cues as a predictor of vote preference. However, in a multivariate regression, the coefficient for competence for both sexes of candidates (B=.48) is significantly larger than that for mate appeal (B=.35), $\chi^2(1, N = 551) = 10.24$, $p<.01$. As my theory argues that partner appeal *influences* perceptions of competence, we should not be surprised to see that the coefficients for both decline in a joint regression, nor to see instability in the coefficients (which occurs when two variables are substantially correlated). Nonetheless, mate desirability continues to meaningfully predict variation in respondent vote choice.

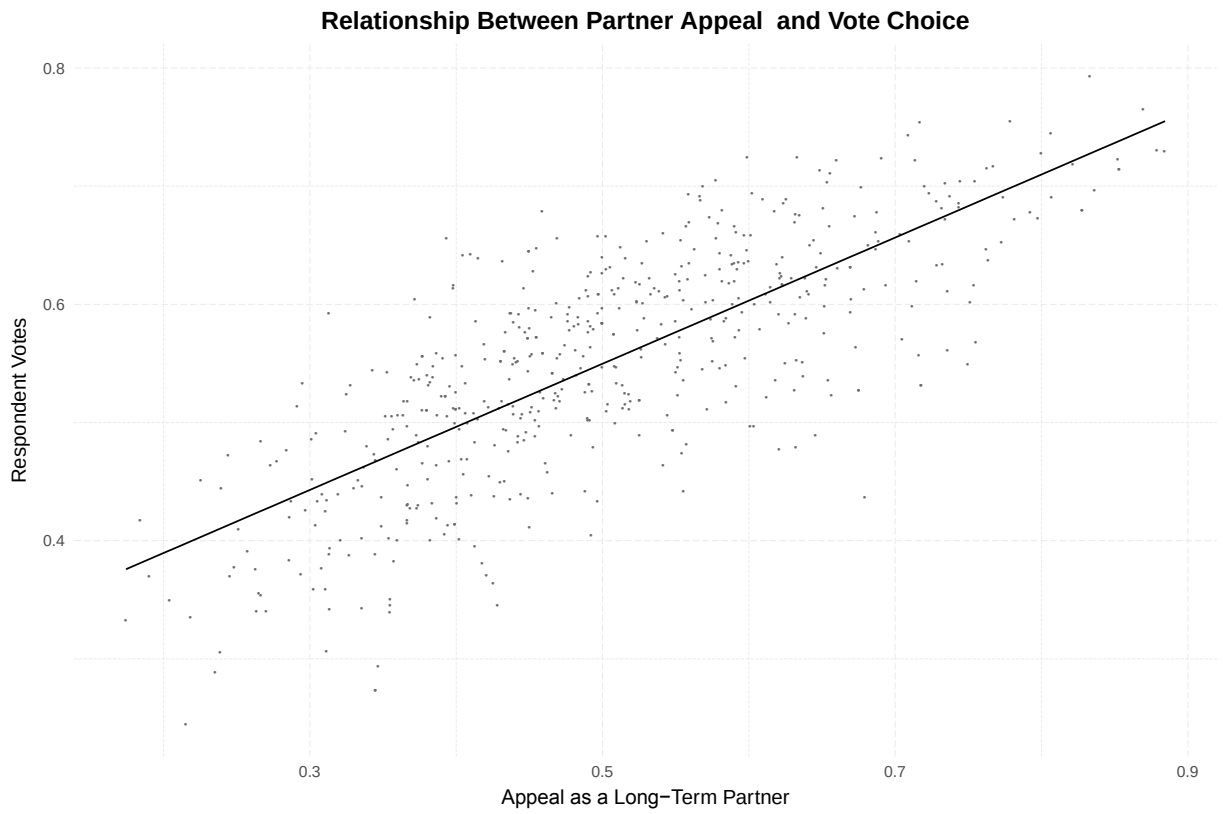
Table 1.1.

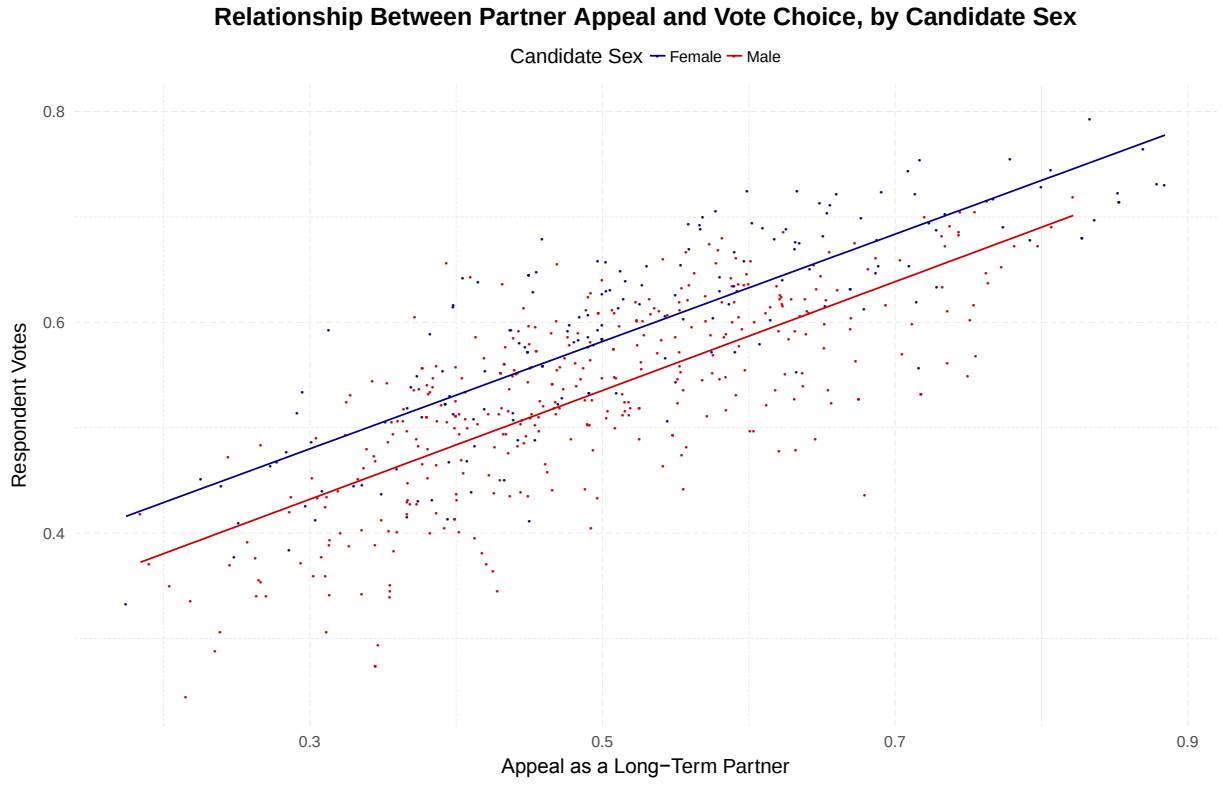
Partner Appeal vs. Competence

	Dependent variable:					
	Willingness to Vote for Candidate					
	Female Candidates	Male Candidates	Female Candidates	Male Candidates	Female Candidates	Male Candidates
Appeal as a Partner	.509***	.516***			.288***	.343***
	(.026)	(.025)			(.029)	(.022)
Perceived Competence			.806***	.735***	.486***	.491***
			(.040)	(.035)	(.045)	(.031)
Constant	.327***	.277***	.059*	.050*	.120***	.041*
	(.015)	(.012)	(.027)	(.023)	(.022)	(.018)
Observations	172	381	172	381	172	381
R^2	.688	.539	.705	.541	.815	.720
Adjusted R^2	.686	.538	.704	.540	.813	.718
Residual Std. Error	.054 (df = 170)	.062 (df = 379)	.052 (df = 170)	.062 (df = 379)	.042 (df = 169)	.048 (df = 378)
F Statistic	374.425*** (df = 1; 170)	443.360*** (df = 1; 379)	407.059*** (df = 1; 170)	447.434*** (df = 1; 379)	371.816*** (df = 2; 169)	485.523*** (df = 2; 378)

Note: Coefficients are scaled 0-1. *p<0.05; **p<0.01; ***p<0.001

Figure 1.3.


Relationship Between Partner Appeal and Vote Choice
Respondent Votes
0.8
0.6
0.4
0.3
0.5
0.7
0.9
Appeal as a Long−Term Partner


Relationship Between Partner Appeal and Vote Choice, by Candidate Sex
Candidate Sex Female Male
Respondent Votes
0.8
0.6
0.4
0.3
0.5
0.7
0.9
Appeal as a Long−Term Partner

MATE DESIRABILITY PREDICTS TRAIT RATINGS

Design and Procedures

Studies 2 and 3 create photo ratings for each candidate by asking survey respondents to evaluate a set of faces on a trait or traits. In Study 2, 7,036 MTurk survey respondents rated the 789 unique photos on attractiveness, competence, dominance, and gender typicality (which I refer to as femininity throughout), plus motherliness (for women) or fatherliness (for men). Each respondent rated approximately 25 unique, randomly selected candidate photos on a single, randomly assigned trait (e.g., competence). Every photo received at least fifty unique respondents' ratings for each trait (competence, attractiveness, dominance, gender typicality, motherliness/fatherliness); these ratings were then aggregated into a mean rating for each candidate (e.g., mean attractiveness). Competence, attractiveness, and gender typicality were measured on seven-point scales (e.g., very incompetent to very competent), while dominance and motherliness/fatherliness were measured on five-point scales (e.g., not at all dominant to very dominant).

In Study 3, I replicated these ratings for a subset of photos (n=187) on a sample of 4,551 registered voters recruited through Survey Sampling International. Respondents again rated candidates' faces for competence, attractiveness, dominance, gender typicality, and motherliness/fatherliness, using the same question wordings and scales. Each respondent rated four candidate photos (two male, two female) on each trait; the four photos were drawn randomly for each trait. All respondents rated photos on competence and gender typicality; half of respondents rated sets of photos on dominance, attractiveness, and motherliness/fatherliness (hereafter referred to as parentliness for brevity).[7] Traits rated by all respondents (e.g., competence) received approximately 94 unique ratings each, while traits rated by half of respondents (e.g., attractiveness) received approximately 47 unique ratings each. As with Study 1, ratings were aggregated into a mean trait rating score for each candidate. All significant results from Study 1 replicate in Study 2 (for details, see SM).

The trait ratings from Study 2 and 3 form the independent variables of this analysis. I assess which traits (attractiveness, dominance, gender typicality, and parentliness) predict vote choice (willingness to vote for this person, collected in Study 1) using a multivariate OLS regression.

Results

A mate selection argument suggests that voters will evaluate women primarily on attractiveness and secondarily on apparent nurturing qualities, and men about equally on attractiveness and ability to provide. Put another way, both sexes but especially women should benefit from high scores on attractiveness, while both but especially men should benefit from high scores on parentliness. In contrast, if existing research is correct that competence is what voters search for, and competence is "constructed from facial cues of attractiveness, masculinity, and confidence" (Todorov, 2017, p. 127), we should see that high scores on attractiveness and dominance, and low scores on femininity, benefit candidates of both sexes.

I find strong evidence, shown in Table 1.2, that mate selection predicts which traits voters will evaluate candidates on, and that male and female candidates are evaluated on different traits as hypothesized in **H1A**. Attractiveness (B=.74) and motherliness (B=.178) significantly predict willingness to vote for female candidates. For male candidates, attractiveness (B=.48) and fatherliness

[7] I had photos rated on other traits, including the Bem Sex Role Inventory, to address two sets of alternative explanations: first, that voters might engage in gender stereotyping rather than prototype judgments, and second, to rule out alternative explanations based on survey wording. I describe the results in the SM.

Table 1.2.

Facial Traits as Qualifications for Office by Candidate Sex

	Dependent variable:	
	Willingness to Vote for Candidate	
	Female Candidates	Male Candidates
Attractiveness	.740***	.481***
	(.055)	(.031)
Dominance	-.076	-.151***
	(.066)	(.036)
Femininity	-.061	.019
	(.036)	(.034)
Parentliness	.178**	.274***
	(.059)	(.030)
Constant	.182***	.198***
	(.042)	(.027)
Observations	133	331
R^2	.661	.528
Adjusted R^2	.650	.522
Residual Std. Error	.055 (df = 128)	.063 (df = 326)
Note:	Coefficients are scaled 0-1. *p<0.05; **p<0.01; ***p<0.001	

(B=.27) predict willingness to vote. The differences in coefficients between attractiveness and parentliness are as predicted by a theory of mate selection: voters select women primarily on attractiveness, while both attractiveness and parentliness play large roles in voters' decision-making about men. Dominance also has a significant negative marginal effect (B=-.15) for men. Femininity fails to correlate with willingness to vote for candidates of either sex; these latter two findings run strongly counter to the predictions of a model suggesting that voters look for masculine, dominant, "leaderly" behavior from their candidates.

Mate selection theory also suggests that age should play an important role in voters' assessments of candidates. As predicted in **H1B**, the positive, significant effect of *Male*Age* across all three assessments suggests that women face a steeper penalty in evaluations as they age than men do, though age has a negative effect on evaluations for both men and women. Figure 1.4 and Table 1.3 show the results are similar whether the trait evaluated is willingness to vote or perceived competence. Women face the strongest penalty for each additional year when respondents are rating partner appeal (B=-.011), as expected, but there is clear evidence that perceptions of women's competence and willingness to vote for women decline as women get older. This occurs even though candidates are likely to have significantly more government experience as they age (p<.01, see SM for details). In contrast, though men never receive the desirability "boost" that young, attractive women receive, respondents' attitudes towards men stay static as they age. For each year, perceptions of men's

Figure 1.4.

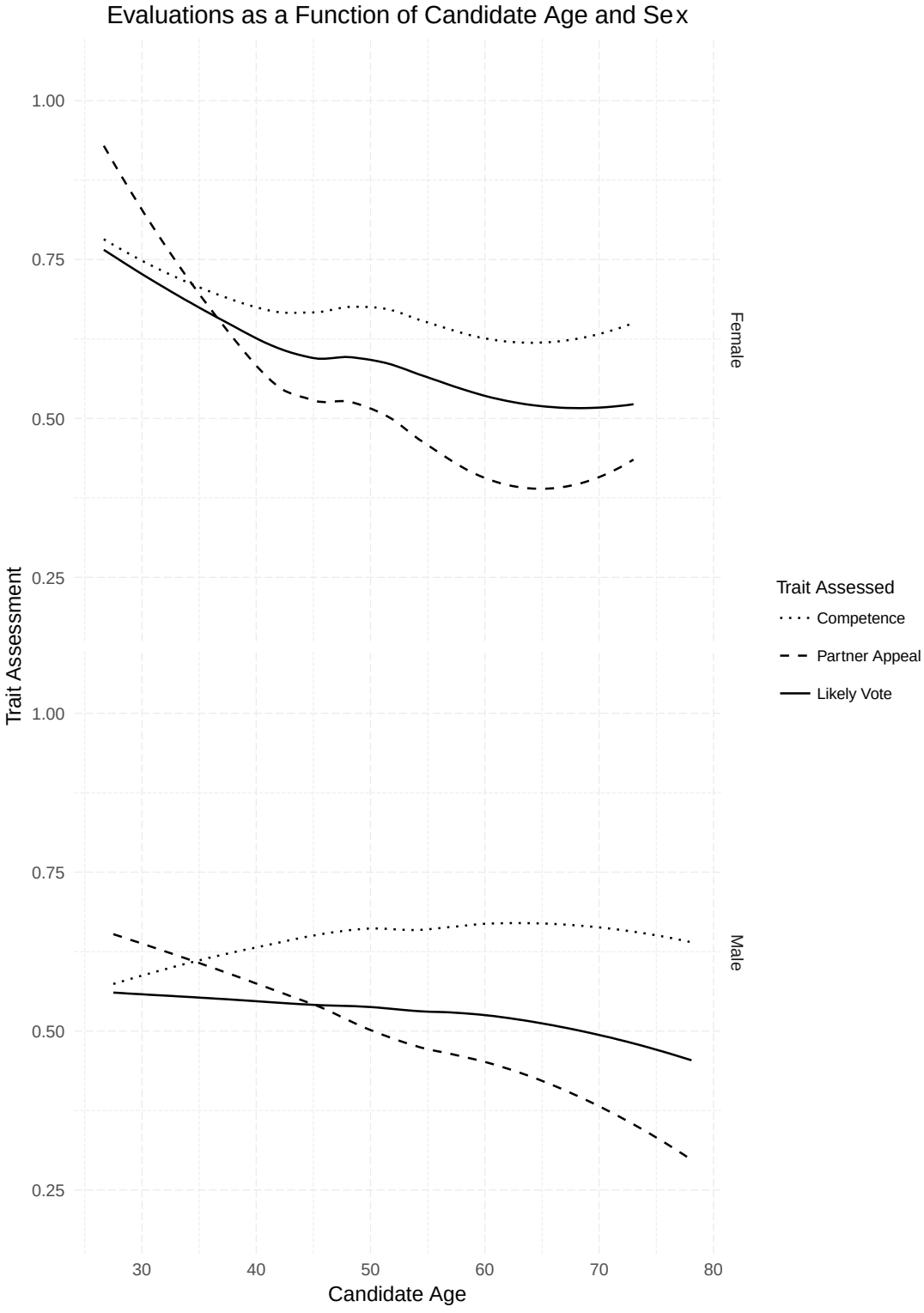

Evaluations as a Function of Candidate Age and Sex
Trait Assessment
1.00
0.75
0.50
0.25
Female
Male
30
40
50
60
70
80
Candidate Age
Trait Assessed
Competence
Partner Appeal
Likely Vote

Table 1.3.

Candidate Evaluations by Age and Sex

	Dependent variable:		
	Partner Appeal	**Perceived Competence**	**Likely Vote**
Age	-.011***	-.003***	-.005***
	(.001)	(.001)	(.001)
Male	-.227***	-.240***	-.242***
	(.056)	(.045)	(.043)
Male*Age	.005***	.005***	.004***
	(.001)	(.001)	(.001)
Constant	1.050***	.811***	.851***
	(.047)	(.037)	(.036)
Observations	548	548	548
R^2	.326	.052	.191
Adjusted R^2	.322	.047	.187
Residual Std. Error (df = 544)	.114	.092	.087
Note:	Age is in years. *p<0.05; **p<0.01; ***p<0.001		

competence increase slightly (B=.002), effectively remaining static over the lifetime. Willingness to vote for older men is barely net negative (B=-.001) with each year.

MATE DESIRABILITY PREDICTS VOTING BEHAVIOR

Design and Procedures

I conduct one additional study to gather ratings of candidates' occupations, which are the only outstanding item of information from the voting pamphlets remaining, before proceeding to an analysis of voting behavior. Study 4 uses similar procedures to Studies 1-3 to create ratings of candidates' listed occupations rather than photos. Due to the complexity of occupation as a signal, I created three measures of occupation. The first two, occupational class (a three-point scale) and political feeder profession (a binary variable) I coded myself, again detailed in the SM. These are coarse measures, so I also asked respondents to assess whether someone holding a given occupation would make an effective legislator if they had no other political experience (a five-point scale, from "not at all effective" to "very effective"). 906 MTurk respondents each rated 10 of the 99 unique occupations derived from the data on this measure, and the ratings were aggregated into a mean score using the same procedure described above for the photo ratings. [8] Each occupation's state legislator qualifications were rated by 91 respondents on average.

[8] Study 4 also contained an experiment that compared ratings of occupations for legislators against ratings of occupations for prospective romantic partners or dates. The design and results are described in the SM.

I use this data assess whether citizens vote in real elections for candidates they regard as desirable mates. I use the accumulated data from Studies 1-4 in a multivariate fixed-effects regression to assess real voting behavior. Two-party vote share, scraped from the election returns, is the dependent variable. I use the aggregated ratings of the information from the voting pamphlets—photo, prior government experience, occupation, and education—collected in Studies 1-4 as independent variables. I use ratings of partner appeal as my main explanatory variable.[9] For occupation, I present models both with hand-coded measures (class and feeder profession) and ratings of occupational qualification to be a state legislator as an alternative measure. Fixed effects for party-district, year, and office (state senate or legislature) are included. As with the first study, I then re-run the models after including perceived competence as a control to assess whether the partner appeal findings are robust to its inclusion.

Results

Finally, I evaluate whether mate desirability predicts actual election outcomes. In Table 1.4A, I show fixed-effects regression models with partner appeal alone. In Table 1.4B, I show the same regressions, but controlling for perceived competence as well. In all models, it is important not to consider the coefficients generated as treatment effects: moving from a 0 to 1 on a seven-point scale of (for instance) perceived competence is not the same as counting out pills in a medical trial. The coefficients represent real-life associations between the variables of interest, not experimental manipulations.

Table 1.4A suggests that judgments of candidates' appearance on mate desirability, as well as candidates' previous government experience, meaningfully predict real votes for both male and female candidates regardless of which model is used. Candidates garner around 15% more vote share when they are rated as the most appealing partners, compared to those rated as the least appealing partners. A candidate with one standard deviation (SD=.20) more partner appeal than an opponent at the mean would receive about 3.2% more vote share than their opponent. This may sound small, but 29 (7%) of the 407 races in the dataset were decided by a smaller margin. Candidates who are sitting incumbents garner around 20% more vote share than those with no prior government experience. Occupation and education do not predict votes in any of the regressions. This supports the argument that judgments of mate desirability play a role in real voting behavior.

Table 1.4B tells a more complex story. Estimates of the effect of partner appeal on vote share for female candidates are significant and consistent with estimates from Table 1.4A, even with competence included. However, for male candidates, both competence and partner appeal seem to be significant predictors. The coefficients for partner appeal and competence are not significantly different for male candidates except in the first model, which uses the appearance cues only. For male but not female candidates, competence seems to play a meaningful role in voters' considerations as well. Per the results in the previous section, attractiveness plays a much larger role in voters' evaluations of female candidates for both partner appeal and competence; accordingly, there may be little other variation for "perceived competence" to explain.

[9] Some candidates use the same picture in more than one election cycle. All rated pictures have the aggregate ratings imputed for each use of that photo: e.g., a candidate who uses the same photo in 2004 and 2008 will have the same competence, attractiveness, etc. ratings for both years. This means that in analyses where real election vote share is the dependent variable, some candidates will appear more than once, with the same values of the independent variable but different values of the dependent variable.

Table 1.4A.

Relationship of Cues to Vote Share

	Sex of Candidate					
	Subset					
	Women	Men	Women	Men	Women	Men
Desirable Partner Based on Photo	.172**	.115*	.164**	.147***	.159**	.149***
	(.062)	(.049)	(.051)	(.040)	(.052)	(.041)
Prior Government Experience			.202***	.202***	.211***	.208***
			(.049)	(.023)	(.048)	(.023)
Professional Class			-.043	.050		
			(.086)	(.038)		
Feeder Profession			.008	.007		
			(.029)	(.016)		
Beneficial Profession					.098	.097
					(.136)	(.052)
Education			-.046	-.026	-.072	-.027
			(.051)	(.021)	(.045)	(.019)
Constant	.514***	.330***	.380***	.370***	.366***	.353***
	(.066)	(.040)	(.047)	(.046)	(.051)	(.050)
Party-District, Office, and Year Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	172	381	168	380	167	380
R^2	.856	.700	.906	.793	.908	.793
Adjusted R^2	.727	.577	.814	.704	.818	.704
Residual Std. Error	.070 (df = 90)	.081 (df = 270)	.059 (df = 84)	.068 (df = 265)	.058 (df = 84)	.067 (df = 266)

Note: Coefficients are scaled 0-1. Standard errors are robust (HC1) SEs. *p<0.05; **p<0.01; ***p<0.001

Table 1.4B.

Relationship of Cues to Vote Share

	Sex of Candidate					
	Subset					
	Women	Men	Women	Men	Women	Men
Desirable Partner Based on Photo	.169*	.068	.164**	.114**	.159**	.121**
	(.067)	(.051)	(.052)	(.042)	(.053)	(.043)
Perceived Competence Based on Photo	.101	.152**	.015	.107*	.009	.098*
	(.087)	(.056)	(.078)	(.047)	(.074)	(.047)
Prior Government Experience			.199***	.198***	.209***	.203***
			(.054)	(.023)	(.052)	(.022)
Professional Class			-.044	.054		
			(.088)	(.038)		
Feeder Profession			.007	.004		
			(.029)	(.016)		
Beneficial Profession					.096	.081
					(.138)	(.053)
Education			-.047	-.030	-.073	-.028
			(.052)	(.021)	(.046)	(.019)
Constant	.447***	.271***	.373***	.330***	.363***	.316***
	(.088)	(.047)	(.060)	(.051)	(.059)	(.054)
Party-District, Office, and Year Fixed Effects?	Yes	Yes	Yes	Yes	Yes	Yes
Observations	172	380	168	379	167	379
R^2	.860	.711	.906	.799	.908	.798
Adjusted R^2	.731	.592	.812	.711	.816	.711
Residual Std. Error	.070 (df = 89)	.079 (df = 268)	.059 (df = 83)	.067 (df = 263)	.058 (df = 83)	.067 (df = 264)

Note: Coefficients are scaled 0-1. Standard errors are robust (HC1) SEs. *p<0.05; **p<0.01; ***p<0.001

DISCUSSION

Broadly, mate desirability seems to predict both citizens' survey responses and real voting behavior. We should not be surprised. Even the most thoughtful and careful souls among us may quail at the thought of opening a two-hundred-page voting pamphlet (as mine was in California in November 2016). Voters overwhelmed with high information may be just as likely to look for clues—a competent face, a gendered first name—to get the chore of voting done as voters who have little other information in front of them. "In the great blooming, buzzing confusion of the outer world we pick out what our culture has already defined for us" (Lippmann, 1922, p. 81). However, the data presented are subject to several caveats and limitations.

Causal identification (i.e., does increasing mate desirability cause increasing willingness to vote for the individual) is a perennial challenge when studying sensitive topics like gender or appearance. In the present case, social desirability bias limits the possibilities for survey experiments: respondents tend to alter their responses to "would you be willing to vote for this candidate?" after being asked to rate the same person on attractiveness so as not to seem superficial. Accordingly, I use between-subjects designs intended to minimize this bias. Social desirability bias is also a chronic problem for studies of gender and other stereotypes; when observed in the context of an experiment or survey, voters worry they may seem sexist and revise their responses accordingly. In both cases, there is a tendency to find a null effect when there may be a true pattern of behavior. Nonetheless, future work would be improved by identifying experimental ways to test the theory.

Analyses of survey data are always subject to generalizability concerns. In many of the studies, I use convenience samples from Mechanical Turk, which we know to be an unrepresentative population already. However, I expect these results to be fairly universal; prior work finds that Indian and Brazilian raters can predict U.S. and Canadian election outcomes (Lenz & Lawson, 2011). Study 2, which used a sample of registered voters, recreated many of the findings presented (see SM), which suggests that the convenience sample may not be unduly usual compared to a similar online (but more representative) sample. Moving forward, machine learning using these (or similar) ratings as a training set might enable improved measurement of such variables. Until then, the current case represents a large influx of new data on descriptive representation and voting behavior that would be difficult to construct otherwise.

Underlying both the survey and election data is an additional problem of candidate selection and strategy. Even if the estimated effects are accurate, in whatever sense we might mean that, it is hard to assess whether a given result (e.g., for candidate age and attractiveness for women) reflects selection issues. For instance, if younger women who run for office know they face a hurdle to being seen as qualified, more capable candidates in this category may put intensive effort into looking more appealing—an argument for a common cause (candidate preparation) for both mate desirability and voting behavior. Similarly, all the survey data reflects whatever idiosyncrasies of candidate selection exist in the real Oregon data. If Oregon has more skydivers running for office than is representative of the general Oregon population, all the aggregate ratings from surveys are also skydiver heavy.

However, using the real data in surveys and experiments comes with a positive trade-off, which is that it allows immediate testing of the generalizability of one's findings. Unclear or limited generalizability of findings is a chronic concern of survey and experimental work, which the present empirical strategy addresses by collecting the data needed to make the comparison as part of the analysis. The use of both survey and experimental data and election returns attempts to address this concern.

BROADER IMPLICATIONS

Inadvertent reliance on mate desirability criteria in lieu of more thorough assessments of candidates' qualifications has real-world political implications. Voters may select candidates perceived to be desirable mates over those with greater qualifications, especially when the candidates are older women. For instance, voters could select a man who is less politically experienced, but who appears to be a good provider and protector, over a woman who is more politically experienced but lacking maternal or physical appeal. The results suggest that reliance on direct elections in which voters select individuals, especially with few other relevant cues, maximizes voters' unconscious propensity to err in favor of mate desirability.

What does our tendency to follow cognitive shortcuts mean for democracy? Broadly, this tendency mounts a challenge to getting good representation. Canonically, we break representation into four components: descriptive, substantive, symbolic, and formalistic representation (Pitkin, 1967). The behavior described has potential implications for each form of representation. First, an unconscious preference for certain types of people (e.g., nurturing women) affects descriptive representation; some types of people will have little chance of acquiring a representative that "looks like them," and in turn individuals running for office who do not fit a desirable pattern will face a harder road to office, i.e., discrimination. Second, theory and data both suggest that lack of descriptive representation can affect substantive representation—for instance, female representatives spend more time on bills that affect women's health (Swers, 1998, 2002)—and symbolic representation: a recent study, for instance, finds that citizens trust government less and are more likely to see a decision on sexual harassment as illegitimate if it is made by an all-male committee than a mixed-sex committee (Clayton, O'Brien, & Piscopo, n.d.). To the extent that unconscious preferences shift citizens' decisions away from qualified legislators in favor of attractive legislators, they may receive less capable substantive representation, which in turn may make them feel that government does not work for them, even if we assume there is no failure of descriptive representation. Finally, the flip side of the unconscious preferences' effect on descriptive representation is that they can also affect democratic accountability, the cornerstone of formalistic representation: legislators who do a bad job, but can make themselves look appealing to voters, may escape the consequences of bad performance where a less appealing-looking representative cannot.

Critically, this behavior creates the most severe problems in a democracy with many candidate-centered elections, such as the United States—and these problems may occur whether voters have lots of information about candidates, as in the Oregon state legislative elections, or very little information about candidates, as they do in survey experiments. The twenty-first century finds us at a unique moment where we are more overwhelmed with potential sources of information about candidates than ever before at the same time that we are less likely to know candidates personally than ever before. Each of these elections is an opportunity for us to default to shortcuts over scrupulousness. In contrast, a democracy in which voters select parties, not officials, and for fewer offices, requires less information and may trigger less pattern-seeking (e.g., based on visual appearance), though these benefits may not be costless.

Looking beyond politics, these findings suggest a need to reexamine the role of psychologically reductive strategies more broadly, given the wide range of decision-making tasks in which a lazy System 2 might abdicate its responsibility. Social and evolutionary psychologists debate whether gender socialization or evolutionary strategies explain mate preferences (Conroy-Beam & Buss, 2016; Eagly & Wood, 2013; Schmitt, 2014; Zentner & Eagly, 2015); this research agenda is even more critical if mate preferences influence non-relationship outcomes. Mate selection criteria appear to affect political choices, opening the possibility that they also influence other evaluative tasks like hiring and

salary decisions. Moreover, while I find that low-information environments with visual cues may exacerbate inadvertent reliance on this "Tinder mentality," this behavior likely extends to other environments. The broader literatures on motherhood and evaluations of women in the workforce or in politics (see e.g., Deason, Greenlee, & Langner, 2015; Eagly & Karau, 2002; Hochschild & Machung, 2012) suggest that these are far from the only circumstances under which we substitute evaluations of partner desirability for leadership assessments of women.

The most critical area for future investigation is to determine what conditions or interventions may circumvent the attribute substitution process. Moving democracies away from direct elections of candidates to party lists to circumvent implicit bias is a tall order; it may be more fruitful to identify interventions amenable to experimental testing and pursue those changes in voting policy. For instance, removing candidate photos from voting pamphlets might lessen the likelihood that voters with little other information about the candidates will rely on appearance as a visual cue. Envisioning how to address differences in perceptions of men's and women's qualifications based on their professional experience remains especially problematic. The apparent relevance of occupational history to real qualifications to hold office, as well as Mo's (2015) finding that providing explicit information about women's qualifications does not eradicate bias, mean that more work must be done to discover effective interventions.

Chapter 2

Does Gender Stereotyping Affect Women at the Ballot Box? Evidence from Local Elections in California, 1995-2016[1]

Research demonstrates that many voters use gender stereotypes to evaluate candidates, but does that stereotyping affect women's electoral success? In this paper, we try to make headway in answering that question by combining a novel empirical strategy with local election data from California. Our empirical strategy relies on two key findings from the existing literature: first, that individuals are more likely to rely on stereotypes when they have less information about the candidates, and second, that the average voter in elections held concurrently with national elections has less information about local candidates than the average voter in off-cycle elections. We propose that we can therefore estimate the electoral effect of increased gender stereotyping by examining the difference in women's win rates in higher-information (off-cycle) and lower-information (on-cycle) elections—and how that difference varies by constituency and the office sought. Our results show that the effect of increased stereotyping is more negative for female candidates in mayoral races than in city council races, and also that the effect of greater stereotyping is more negative for women running in conservative cities than in more liberal cities. Thus, we conclude that there probably isn't a single, one-size-fits-all answer to the question of how gender stereotyping affects female candidates, but rather that the direction and magnitude of the effect varies across contexts.

[1] This chapter is derived from a paper of the same name coauthored with Sarah F. Anzia.

The vast literature on women in American politics has devoted considerable attention to two questions: Do voters use gender stereotypes in evaluating candidates? And does gender stereotyping hurt (or possibly even help) female candidates at the ballot box? The answer to the first is clearly "yes": some voters tend to infer that female candidates are more liberal than men and also more compassionate and collaborative, more competent on certain issues like education, and less competent on other issues such as foreign policy (Alexander and Andersen, 1993; Eagly and Karau, 2002; Huddy and Terkildsen, 1993a and 1993b; Kahn, 1996; Sanbonmatsu, 2002). On the second question, however, research has not produced a consensus (Dolan, 2010).

If there is a dominant viewpoint in the women in politics literature, it is that once women decide to run for office, they fare as well as men—and that gender stereotyping by voters makes little to no difference to women's electoral fortunes (e.g., Brooks, 2013; Burrell, 1994; Duerst-Lahti, 1998; Hayes and Lawless, 2016; Newman, 1994; Seltzer, Newman, and Leighton, 1997). However, if the average female candidate is of higher quality than the average male candidate, then men and women receiving equal average vote share is evidence of the *presence* of gender bias—not its absence (Anzia and Berry, 2011; Ditonto, 2016; Fulton, 2012; Milyo and Schosberg, 2000; Mo, 2015). Other work proposes that the effect of gender stereotyping depends on context, including the issues salient in the election (Lawless, 2004), the type of office being sought (Huddy & Terkildsen, 1993b), characteristics of the voters (Bauer, 2015), and the availability and salience of party cues (Hayes, 2011; Hayes and Lawless, 2016; Lawless, 2015). As it stands, the literature does not offer a clear answer as to whether—or under what conditions—gender stereotyping harms (or helps) female candidates.

One reason for the debate may be the significant methodological challenges inherent in estimating the effect of gender stereotyping on women's electoral success. While some studies use observational data to study this question, most political science research on gender stereotyping uses survey experiments or lab experiments, for good reason: it is more straightforward to isolate theoretical effects of interest with controlled experiments. Critics of studies that use election returns to study gender bias point to omitted variable bias problems, such as the difficulty of controlling for candidate quality (e.g., Fulton, 2012). And when the researcher only has data on how people in a certain jurisdiction voted (as is the case with aggregate election returns), it is very difficult to tease out the effects of psychological processes like gender stereotyping.

That said, there are also disadvantages to the experimental approach—and advantages to using real election data. A common critique of experimental research on gender stereotyping is that unrepresentative samples and heavy-handed experimental designs can exaggerate the effects of stereotypes on voting (see, e.g., Brooks, 2013). Even the most careful experiments may not replicate how real voters make decisions in real elections. And given that our question of interest is whether and how gender stereotyping affects female candidates in real elections, there are advantages to using data on how people actually vote—even with the inferential challenges involved.

In this paper, we try to make headway in evaluating the effect of gender stereotyping on women's electoral success by combining a novel empirical strategy with local election data from California. Our dataset, built from data in the California Elections Data Archive (CEDA), tracks the results of 15,239 local races featuring 62,926 candidates over a 22-year period, 1995 to 2016. To overcome one of the major challenges involved in studying gender stereotyping using data on election returns, we rely on two key findings from the existing literature: First, individuals are more likely to rely on stereotypes when they have less information about the issues and candidates (e.g., Lupia, 1994; Matson and Fine, 2006). Second, the average voter in on-cycle elections (those held concurrently with national elections) has less information about local candidates and issues than the average voter in off-cycle elections (Oliver and Ha, 2007). As a theoretical baseline, if gender stereotyping has no effect on the electoral success of female candidates, a woman running in a lower-information environment (on-cycle elections) should be just as likely to win as she would in a higher-information environment (off-cycle

elections). By contrast, if gender stereotyping negatively (or positively) affects female candidates, the woman should fare worse (or better) in the lower-information environment than in the higher-information environment.

What, then, should our expectations be? Rather than propose a single, one-size-fits-all answer, we survey the existing literature to develop hypotheses about how the effects of increased gender stereotyping will depend on the context of the election. In our empirical analysis, we test these hypotheses by examining what happens to women's win rates in the shift from off-cycle to on-cycle elections—our proxy for increased gender stereotyping—and how the effect of that shift varies by type of race and constituency.

Our empirical results reveal interesting patterns. First, we find that the effect of shifting from a high- to a low-information electoral environment—which we assume results in more gender stereotyping—is usually *more negative* for female candidates in mayoral races than in city council races. Second, contrary to our expectation that the effect of greater gender stereotyping would be more positive for women in school board races than in city council races (because of the issues involved), we find no difference in the effect of more stereotyping across the two types of races. Finally, in mayoral races, the effect of shifting to a low-information environment is generally more negative for women running in conservative cities than in more liberal cities. While there are some inferential concerns that arise from our approach and data, our findings suggest that there isn't a single answer to the question of how gender stereotyping affects female candidates, but rather that the direction and magnitude of the effect is conditional.

LITERATURE

Extensive research has established that some voters do employ gender stereotypes when evaluating female candidates. As with all stereotypes, we can expect their use to vary with the amount of information voters have about candidates (Ditonto, Hamilton, and Redlawsk, 2014; Lau and Redlawsk, 2001; Lupia, 1994) and voters' political sophistication (Bauer, 2015; Zaller, 1992). There is also little dispute regarding the content of these stereotypes, which tend to fall into three categories. Voters tend to perceive female candidates as more liberal than male candidates (Sanbonmatsu, 2002)—a "beliefs" stereotype. They often view female candidates as more compassionate and communitarian (Bem, 1981; Eagly and Karau, 2002) and more honest (Brooks, 2013) than male candidates—"trait" stereotypes. Female candidates are also assumed to be more competent than male candidates on so-called women's issues such as education and healthcare (an "issue competency" or "domain" stereotype), and weaker than men on issues like foreign affairs (Huddy and Terkildsen, 1993a; Sanbonmatsu, 2002).

The question, then, is whether these stereotypes affect women's ability to win elective office. In one of the most comprehensive studies to date, Brooks (2013) concludes that the answer is no. Brooks carries out a series of well-designed experiments using hypothetical male and female candidates, asking respondents to rate the candidates on favorability and likely effectiveness in office. She finds that while voters show signs of using some classic gender stereotypes, it does not look as though stereotyping systematically affects respondents' evaluations of men and women's favorability and likely effectiveness. The conclusion, then, is that the use of stereotypes does not appear to have either a negative or a positive net effect on female candidates' electoral chances—and that the source of women's underrepresentation in public office must lie elsewhere (such as lower political ambition, e.g., Fox and Lawless, 2010).

This conclusion aligns with a different segment of the literature: one that compares the vote shares of male and female candidates in real elections to determine whether voters penalize the

women. Most of these studies find that the average vote share of female candidates is no different from the average vote share of male candidates (Burrell, 1994; Brooks, 2013; Duerst-Lahti, 1998; Newman, 1994; Seltzer, Newman, and Leighton, 1997). In large part because of these studies, the conventional wisdom in the women in politics literature is that voters are *not* biased against women (e.g., Burrell, 1994; Darcy et al., 1994; Fox, 2006)—and that, by implication, gender stereotyping by voters is not having an impact on women's electability.

There is also research that challenges that conclusion. Fulton (2012) and Milyo and Schosberg (2000) find that once variation in candidate quality is accounted for, female candidates actually *do* receive fewer votes than male candidates. Moreover, extensive work in the field of psychology suggests that female leaders are penalized for exhibiting the masculine traits that are associated with strong leadership (Eagly and Karau, 2002; Koenig, Eagly, Mitchell, and Ristikari, 2011; Rudman and Phelan, 2008), which potentially puts female candidates at a disadvantage. Some more recent work also suggests that voters assume that men are more qualified than women—and that they penalize women because of it (Ditonto, 2016; Ditonto et al., 2014; Mo, 2015). With such divergent findings in the literature, it is difficult to draw clear conclusions about the effects of gender stereotyping on women's electability.

From a theoretical standpoint, it is not even clear that there *should* be a single, one-size-fits-all answer to the question of how gender stereotyping affects women's electoral success. A finding of no effect of gender stereotyping on women's vote share makes theoretical sense: most group stereotypes contain a mix of positive and negative content (Fiske et al., 2002), and if gender stereotyping leads some voters to penalize women and other voters to support women, the result may be no visible effect in the aggregate. But findings of negative or even positive effects are also theoretically plausible, and the direction and magnitude of the effects may depend on the context of the election.

This theme—the conditionality of gender stereotyping effects—has not been a central one in the women in politics literature, but there are some examples of work along these lines. Focusing on the issues salient in elections, for example, Lawless (2004) finds that during wartime, voters prefer more dominant leaders, which disadvantages female candidates (see also Herrnson, Lay, and Stokes, 2003; Kahn, 1996). The findings of Huddy and Terkildsen (1993b) also suggest that stereotypes of communitarianism benefit women running for legislative office and possibly hurt women running for executive office. And a growing line of research finds that Democrats (both voters and elites) tend to be more supportive of women than Republicans (King and Matland, 2003; Kitchens and Swers, 2016), suggesting that the effect of gender stereotyping might vary by the partisan or ideological composition of the jurisdiction.

Together, these findings point in the direction of *varying* effects of gender stereotyping—some positive, some negative, and some neutral—rather than to a single, unconditional answer about how voter stereotyping affects female candidates at the polls. But hypotheses about the conditions that shape gender stereotyping effects remain underdeveloped. Moreover, for the hypotheses that have been proposed in the literature, there have been very few empirical tests—and, to our knowledge, no empirical tests of multiple hypotheses using the same data.

Empirical Strategy

Given the methodological challenges involved, how should one go about evaluating the effects of gender stereotyping on women's electoral success? Both the experimental approach and analyses of election returns have their strengths and weaknesses.

As we suggested earlier, there are good reasons why scholarship focused on stereotyping mainly relies on surveys and experiments. Stereotyping, after all, is a psychological phenomenon. It is not something one can readily measure by observing individuals' behavior. Even if we *could* observe a behavior that we understand to represent bias, individuals being observed are often reluctant to demonstrate behaviors viewed as socially undesirable. It makes sense, then, that researchers primarily study the use and effects of stereotypes through experiments: the researcher manipulates the conditions presented to respondents and then assesses how the varying information (about gender, for example) affects respondents' evaluations of a given candidate. Election returns, by comparison, provide little if any information about the psychological processes of voters. Well-designed studies comparing the vote shares of male and female candidates can uncover the extent of overall "bias" against or in favor of women (even if scholars dispute the degree to which omitted variables might explain such findings), but they cannot readily speak to the psychological processes underpinning voters' behavior.

At the same time, if our interest is in understanding the *effects* of those psychological processes on women's chances of being elected, there are considerable advantages to focusing on the stage at which citizens cast real votes. Election returns allow us to observe voters' actual choices. By contrast, in experimental work, respondents are usually presented with hypothetical candidates, and sometimes just one candidate at a time—not two or more competing candidates from which they must choose. Therefore, while there are advantages to using experiments to assess the prevalence and content of stereotypes, if we want to evaluate the effects of stereotype usage in real elections—where it counts—there are benefits to using data on election returns, as imperfect as they may be.

We are certainly not the first to argue the advantages of incorporating observational data into the assessment of the impacts of gender stereotypes (Dolan, 2010). However, studies that use election data almost all focus on congressional and gubernatorial elections,[2] and they find that any effects of gender cues are overwhelmed by other cues—namely, political party (Hayes, 2011). This has led scholars to conclude that in today's polarized politics, gender cues matter little for women's electability (Hayes, 2011; Hayes, Lawless, and Baitinger, 2014).

Nonetheless, in most elections throughout the United States, voters cannot easily rely on party affiliation as a heuristic to decide between candidates. In state and national primary races, for example, voters are usually asked to choose from candidates of the same party. In the states that have adopted "top two" primaries, moreover, partisanship may not even be a useful cue in general elections. And in elections for the nation's nearly 90,000 local governments, candidates' party affiliations usually do not appear on the ballot—and therefore are not readily available to be used as a cue by voters. Candidate gender, by contrast, is usually apparent to voters via candidates' first names. Therefore, it is important to study the effects of candidate gender in nonpartisan elections: because the absence of party cues may increase the use of gender stereotypes, and also because research suggests that women are more likely than men to start their political careers in local races (Carroll and Sanbonmatsu, 2013).

The challenge discussed above remains, however: how can one infer something about a psychological process using aggregate election returns? We propose that one way of doing this is to study voters' choices (in the aggregate) under conditions in which individuals are more or less likely to use stereotypes. Research shows that when voters know a great deal about the candidates, they are less likely to rely on heuristics generally (Popkin, 1991) and gender stereotypes in particular (McDermott, 1997; McDermott, 2005). Moreover, experimental research shows that voters with a low level of attention to politics are more likely to rely on gender stereotypes (Bauer, 2015).

Critically, the amount of information voters have about local candidates and issues varies with the timing of the elections. Voter turnout is much lower in off-cycle elections than in on-cycle elections

[2] A notable exception is Crowder-Meyer et al. (2015).

(e.g., Hajnal, Lewis, and Louch, 2002). Not only does low turnout in off-cycle elections create opportunities for organized groups to have greater influence (Anzia, 2014; Berry and Gersen, 2010), but it also means that the average voter knows more about local candidates and issues than the average voter in on-cycle elections (Oliver and Ha, 2007). And that is what we should expect. When local elections are held on the same day as national elections, many people only vote in local races because they are already at the polling place to vote in national races. In off-cycle local elections, by contrast, most of the people who participate do so *because* they have interest in the local races. It makes sense, then, that the typical voter in on-cycle local elections has less information about local candidates and issues.

By combining these two empirical findings—about stereotype use and about election timing—we can analyze differences in voters' choices in off-cycle and on-cycle local elections as a way of evaluating the effects of gender stereotypes. If a female candidate receives a lower vote share when the election is held on-cycle rather than off-cycle, that suggests that greater use of gender stereotyping has a negative effect on the candidate. If she receives higher vote share in the on-cycle election, that suggests that greater gender stereotyping helped her. In sum, we can use election timing as a proxy for average voter information about the candidates—and thus their tendency to use gender stereotypes—as a way of evaluating the effects of stereotyping on females' electoral success.

Two sets of caveats are worth emphasizing at the outset. First, by making comparisons between women and men's win rates in off-cycle and on-cycle elections, we are attempting to draw conclusions about the effects of *more* gender stereotyping on women's electoral success. Because we do not know how much gender stereotyping occurs in the baseline condition (the higher information environments), we cannot say anything about the *base rate* of gender stereotyping. However, we can rely on the literature's findings to assume that voters will be more likely to use stereotypes in the on-cycle condition than in the off-cycle condition, and then compare any gap in women and men's electoral fortunes in the two conditions across contexts.

Second, as we discussed earlier, omitted variable bias is a central concern with any observational data analysis. Throughout our own analysis, we must be cognizant of several threats to causal inference. One is the possibility that some other difference between on- and off-cycle elections—whether within or across cities—explains variation in women's electoral success. For example, off-cycle electorates tend to be older (Anzia, forthcoming; Kogan et al., forthcoming), are less diverse racially and ethnically (Hajnal and Trounstine, 2005), and can feature voters with different preferences than those in on-cycle elections (Anzia, 2014; Berry and Gersen, 2010, 2011; Kogan et al., forthcoming; Meredith, 2009). To the extent that these differences in the electorate alter the vote share and win rates of female versus male candidates, they would bias our results (although the direction of any bias is theoretically unclear). In addition, candidates' decisions to enter races probably depend on their expectations of winning, which we think should attenuate any effects we estimate. For example, if more women decline to run in on-cycle city council races than off-cycle city council races because they anticipate that the on-cycle electorates will be unfavorable to them, then that should reduce any difference in women's success rates in the two types of elections. In what follows, our empirical results are preliminary, but we expect to address these inferential concerns more thoroughly in future work.

Data

To carry out the analysis, we downloaded data on local elections from CEDA, which tracks the results of most local elections throughout California, including counties, municipal governments, and school districts from 1995 onward. Given that we are interested in leveraging variation in election timing, we focus on the municipal governments and school districts, because their elections are held at different

times—some concurrent with national elections in November of even-numbered years, and some off-cycle. Our dataset includes all regular and special elections for city council, mayor, and school board held between 1995 and 2016. In total, this includes 62,926 candidates running in 15,239 races.

Table 2.1 presents the numbers of races and candidates in our dataset, broken down by the office sought. In total, there are 6,037 city council races featuring 27,448 candidates. 5,720 of the races are regular races for full terms, and 317 are for shortened terms. Of the city council races, 2,057 were in November of even-numbered years during presidential elections, 1,693 were in November of even-numbered years during midterm elections, and the remaining 2,287 were held at other times. Most of the city council races were contested: 5,438 of them featured more candidates running than there were seats up for election.

Not all cities in California have separately elected mayors, so there are fewer mayoral races in the dataset. As we show in Table 2.1, there are 1,190 unique mayoral races (in 176 municipal governments) and 3,175 mayoral candidates. All but 44 of the races for mayor were for regular, full terms, and again, most were competitive: 929 featured two or more candidates, and 261 were uncontested. About 70% of the mayoral races were held concurrently with national elections: 417 were held concurrently with presidential elections, and 421 were concurrent with midterm elections.

There are almost twice as many school districts in California as there are municipal governments, and so there are far more school board elections than city council elections: 8,012 unique races featuring 32,303 candidates. Again, most of the school board races were for full terms (7,432 of them), and nearly all of them were contested (7,816). 34 percent of the school board races in our dataset were held on-cycle concurrent with presidential elections, 30% were concurrent with midterm elections, and the remaining 36% were held off-cycle.

The CEDA dataset includes rich information about the elections and candidates, including the election date; number of seats up for election; candidate names; the number of votes received by each candidate; an indicator for whether the candidate won, lost or advanced to a run-off; and whether the candidate was an incumbent or a non-incumbent. The CEDA data do not have an indicator for candidate sex or gender, however, so we used a two-step process to code it: First, we used the genderizeR package in R, which uses the first name of each candidate and U. S. Census data to generate a probability that the person is female.[3] Using the probabilities created by the R package, we code a candidate as female if 80% or more of the people in the U.S. with that name are female, and we code a candidate as male if 20% or fewer of the people with the name are female. This allows us to code candidate gender for 96% of the city observations. For all mayoral, council, and school board candidates that are not categorized using this rule, we had research assistants code them as male or female on the basis of the candidate's first name and, if possible, the ballot designation (e.g., "businesswoman"). In all, we were able to code candidate gender for 27,133 city council candidates, 3,147 mayoral candidates, and 31,898 school board candidates.

As we show at the bottom of Table 2.1, the share of candidates who are women, and the share of candidates who are incumbents, varies depending on the office sought. School board elections have the highest proportion of female candidates: 42%. By comparison, 27% of the candidates for city council are women, and the number is even lower for mayoral races, where only 21% of the candidates are women. School board races also feature the highest share of incumbent candidates: 34%, compared to 26% in both city council and mayoral races.

[3] Kamil Wais (2016). genderizeR: Gender Prediction Based on First Names. R package version 2. 0. 0. https: //CRAN. R-project. org/package=genderizeR.

Hypotheses

With this dataset, we can devise tests of three main hypotheses—all of which are hypotheses about how the net effect of gender stereotyping will depend on the context of the election. First, we draw on research about a particular trait stereotype: that women are less well suited for executive office than the legislature. If some voters do use such a stereotype—and if that stereotype isn't offset by other voters who think that women are *better* suited for executive than legislative office, then the effect of increased gender stereotyping on women's electoral success should be *more negative* in executive races than in legislative races. We can test this using our municipal election data by comparing women in mayoral and city council elections. Specifically, our first hypothesis is:

> ***H1: The effect of increased gender stereotyping on women's electoral success will be more negative in mayoral races than in city council races.***

Our strategy for testing this hypothesis is to examine the effect of being a woman (as opposed to a man) on a candidate's electoral success under different conditions. As a baseline, in city council races, we can estimate the effect of being a woman on candidates' electoral success—and, more importantly, how that effect differs (if at all) in off-cycle and on-cycle elections. If the effect of being female is *lower* in on-cycle elections than in off-cycle elections, that would imply that greater stereotyping by voters leads to penalties for female candidates. A positive change would suggest the opposite: that as voters have less information about candidates and rely more on stereotypes, female candidates actually do better.

That, then, is the baseline. But H1 implies a comparison. As a next step, we can estimate the same effects for female candidates in mayoral elections: How does being a woman affect electoral success in off-cycle elections, and how (if at all) does that female effect change in the shift to on-cycle elections? The test of H1, then, is whether the change in the female effect when moving from an off-cycle to an on-cycle election is *lower* or *more negative* in mayoral races than in city council races.

Our second hypothesis is based on an issue competency stereotype: that women are more competent than men in policy areas like education and health and less competent in areas such as crime, the economy, and foreign affairs. If some voters apply these issue competency stereotypes, then we should expect the effect of greater stereotyping to be more favorable to women in school board elections—where education is the only issue—than the effect in a different kind of legislative race: city council elections, where law enforcement and economic development are central issues. Therefore, Hypothesis 2 is as follows:

> ***H2: The effect of increased gender stereotyping on women's electoral success will be more positive in school board races than in city council races.***

To test H2, we employ the same empirical set-up as for H1, this time making the comparisons between city council and school board races rather than city council and mayoral races. Specifically, we estimate the change in the effect of being a woman in city council races when moving from an off-cycle to an on-cycle context, then estimate the change in women's success when moving from off-cycle to on-cycle school board races, and finally test whether the latter effect is greater, or more positive, than the former.

Our third hypothesis is rooted in a well-documented beliefs stereotype: voters are inclined to think female candidates are more liberal than male candidates. We propose that the effect of greater gender stereotyping on women's electoral success in this case should depend on whether a given electorate views liberal candidates favorably. Thus, we hypothesize that the effect of a lower-

information electoral environment will be more negative in contexts where being liberal is perceived unfavorably than in contexts where being liberal is perceived favorably. Specifically:

H3: The effect of increased gender stereotyping on women's electoral success will be more negative in conservative cities than in more liberal cities.

To test H3, we combine our municipal election data with city-level presidential vote share data from the 2012 election: the two-party presidential vote that went to Mitt Romney, accessible through the California Secretary of State's office.

Empirical Analysis

As we discussed earlier, the main advantage of using real election data is that we can observe how voters made decisions (in the aggregate) about real candidates in real elections. Throughout our analysis, then, our dependent variable is a binary indicator, *Win*, which equals 1 if he or she won the election and 0 if he or she lost.[4]

Our goal is to assess how any gap in male and female candidates' win rates changes in the shift from off-cycle (high information) to on-cycle (low information) elections—and how those changes vary across contexts (for example, city council and mayoral races). Naturally, however, there are challenges inherent in using real election data to draw conclusions about the effects of more gender stereotyping. As a first step in our analysis, then, we consider some of those challenges and describe how we attempt to address them.

For starters, our empirical strategy for testing the effects of gender stereotyping is almost certainly better suited for non-incumbents than for incumbents. Recall that our approach is premised on the finding that voters have less information about the candidates and issues in on-cycle local elections. For incumbents, it is not clear that this holds—at least to the same degree as for non-incumbents: incumbents have already been in public office and are probably recognizable to many voters, even in on-cycle elections. Also, incumbency is a strong cue to voters (e.g., Darcy and Choike, 1986)—one that candidates often signal through their ballot designation—and the presence of this competing cue probably decreases voters' reliance on gender stereotypes. We therefore limit our analysis to non-incumbent candidates.

That said, incumbents do win at much higher rates, even in these local elections, and so the electoral success of non-incumbent candidates should depend on whether they are competing with incumbents. We create a variable called *Incumbent Ratio*, which equals the number of incumbents running in the race divided by the number of seats up for election in the race (where number of seats is usually greater than one in at-large city council races). We find that this variable is indeed a strong predictor of whether or not a non-incumbent candidate wins. It is also correlated with our main independent variables of interest: The average incumbent-to-seat ratio is larger in on-cycle elections than off-cycle elections, for all three types of races. Moreover, average *Incumbent Ratio* is significantly smaller in races that feature non-incumbent female candidates compared to races in which the non-incumbents are exclusively male.[5] Because *Incumbent Ratio* is correlated with both win rates *and* the main independent variables, we need to account for the presence of incumbents in our analysis.

[4] If a candidate in a given race advanced to a runoff, we exclude that observation.

[5] Specifically, average *Incumbent Ratio* is 0.65 for council races with only male non-incumbents, but it is 0.48 for council races with female non-incumbents. For mayoral elections, the average is 0.73 for races with only male non-incumbents, and it is 0.38 for races with female non-incumbents. For school board races, the numbers are 0.77 and 0.51, respectively. This suggests that female non-incumbents are less likely than male non-incumbents to enter races when there are incumbents running.

The overall competitiveness of elections is another potential source of bias. We define the variable *Competition* as the number of candidates running in the race divided by the number of seats up for election in the race. Naturally, greater competition should lower the probability that any given candidate will win. But for mayoral and council races, *Competition* is also correlated with election timing: on average, on-cycle races feature fewer candidates per seat than off-cycle races. *Competition* is also positively correlated with the presence of a female candidate in the race—a finding that aligns with work on congressional primaries by Lawless and Pearson (2008): for all three types of office, races that feature at least one female candidate tend to have larger numbers of candidates running per seat.[6] Also, the magnitude of these cross-gender differences in competitiveness also vary by the office type (and are largest in mayoral races). Thus, in order to make comparisons between male and female win rates in on-cycle and off-cycle elections, and to assess how they vary across contexts, it is important to account for the portion of the variation attributable to competitiveness.

Another set of potential challenges relates to variation in the nature of on-cycle elections. Our empirical strategy takes advantage of the fact that on-cycle local elections involve many voters who are mainly (or only) interested in state and national races—and who vote in the local races because they occur on the same day on the same ballot. But the relative size of this part of the electorate—and perhaps even their propensity to vote for female candidates in local races—may well depend on the particulars of the state and national elections happening concurrently.

Some of this variation stands to be analytically useful. In particular, it is well established that far more people turn out for presidential general elections than for midterm elections—and those turnout differences also show up in local races (Hajnal, Lewis, and Louch 2002). We should therefore expect the proportion of low-information voters to be highest in local elections concurrent with presidential elections—and somewhat lower in local elections concurrent with midterms (although still higher than off-cycle elections). We can use this pattern as a validity check for our approach: to the extent we find effects of more gender stereotyping on female candidates' win rates, we should find that they are greater in local elections concurrent with presidential elections than in local elections concurrent with midterm elections.

The effects of other kinds of variation in on-cycle elections are more difficult to predict, but it is worth considering some possible ways this might affect our analysis. For example, national public mood varies over time (e.g., Mayhew, 2001), and some election cycles are more favorable to one major party over the other (e.g., Abramowitz, 2012; Tufte, 1975). Perhaps on-cycle local elections taking place during years favoring Democrats will tend to favor female candidates, or perhaps there are swings in public mood that are more or less favorable to women (e.g., Koch and Thomsen, 2017). We do not see much reason for why this would affect female legislative and executive candidates differently, but it may be a reason why the overall effects of more gender stereotyping would vary across election cycles or years.

Another possibility is that prominent female candidates running for national or state offices might affect voters' propensity to support female candidates in local elections held on the same day or around the same time. Our data span 22 years, and in four of those years (2007-8 and 2015-16), Hillary Clinton was running for president. 2008 also featured a female vice presidential nominee (Sarah Palin), and other years featured women running for California governor (Meg Whitman) and U.S. Senate (Carly Fiorina, Kamala Harris, and Loretta Sanchez). There are at least three ways in which this might affect voters' evaluations of female candidates in local races. First, observing women at the top

[6] For city council elections, in races with at least one female, there is an average of 2.6 candidates running per seat, compared to 2.07 candidates per seat in races without a female running. For mayoral elections, the average is 3.42 candidates in races with a female and 2.11 candidates in races without a female. In school board races, the gap is smaller but still significant: 2.06 compared to 1.96.

of the ballot may prime or activate voters' gender stereotypes (Bargh, Chen, and Burrows, 1996; Bauer 2015; Kunda and Spencer, 2003), leading to exaggerated gender stereotyping effects in local races when women are running for higher offices. Second, voters motivated to turn out to support women at the top of the ballot may be more pro-woman (Dolan, 2008; Herrnson, Lay, and Stokes, 2003; Sanbonmatsu, 2002; Seltzer et al., 1997), and perhaps their positive affect toward women carries over into local races. Third, there may be "coattail" effects for female candidates—but this is unexplored in the literature: the literature on coattails focuses on partisanship (e.g., Meredith, 2013) rather than gender. While we do not have clear predictions about how women running for national and state offices affect women running in local races, we at least want to consider the possibility that women running at the top of the ticket affect women running in local races—in ways that might vary by the office sought.

With all of these considerations in mind, we begin our tests of H1 with simple comparisons of men and women's win rates in city council and mayoral elections. In the top half of Table 2.2, we start by calculating the average win rates of male and female non-incumbent city council candidates in off-cycle elections—meaning elections not concurrent with any statewide or national election.[7] There, we can see that on average, the men win 26.9% of the time, and the women win 30.3% of the time—which implies a female advantage of 3.4 percentage points (see the rightmost column labeled "Female Advantage / (Disadvantage)"). The next row presents average win rates for male and female city council candidates in local elections held concurrently with midterms: the win rate for men grows by 2.1 percentage points to 29%, and the win rate for women grows by 3.6 percentage points to 33.9%. Thus, as we can see in the rightmost column, the female advantage in city council elections *grows* in the shift from off-cycle to midterm elections—by 1.5 percentage points. Finally, we calculate the average win rates for male and female city council candidates during presidential elections: 29.1% and 38%, respectively. In presidential elections, therefore, female city council candidates have an advantage of 8.8 percentage points—an increase of 5.5 percentage points from the off-cycle condition. Therefore, as voter turnout increases, and average voter information about local candidates decreases, the female advantage in California city council races grows.

Next, we turn to non-incumbent candidates for mayor. The average win rates for male and female mayoral candidates in off-cycle elections are very similar: 16.7% and 17.8%, respectively. Also, both men and women win at higher rates during midterm elections than in off-cycle elections, although the increase in male win rates is slightly larger than that of women. This trend is even more pronounced for the shift from off-cycle to presidential elections: male win rates grow by 10.1 percentage points to 26.8%, whereas female win rates grow by only 5.8 percentage points to 23.6%. In terms of the direction of the change, then, this pattern is different than it was for city council races. An influx of lower-information voters in mayoral elections *decreases* any female advantage by 4.3 percentage points (to the point where being a woman is actually a slight disadvantage during presidential elections), whereas for council races, it *increases* it by 5.4 percentage points.

These are just the raw win rates, however, and we have already discussed how the presence of incumbents and competitiveness vary by election timing, candidate gender, and office type. In the bottom half of Table 2.2, therefore, we instead analyze the residuals from an OLS regression of *Win* on *Incumbent Ratio* and *Competition*. (The residuals can be interpreted as the variation in win rates not attributable to incumbents running or competitiveness.) Again, we see that in city council races, men's win rates increase in the shift to midterm and especially presidential elections, but that women's win rates increase even more. Thus, the female advantage grows in the shift from off-cycle to midterm elections, and grows further still in the shift to presidential elections. For mayoral elections, we see the opposite pattern. Instead, the female *disadvantage* grows in magnitude in the shift from off-cycle to

[7] We exclude the small number of local elections held concurrently with a statewide primary or other statewide elections.

midterm elections, and grows further still in the shift to presidential elections. Specifically, the female advantage in city council elections grows by 4.3 percentage points in the shift from off-cycle to presidential elections. For mayoral elections, there is a female disadvantage, and that disadvantage grows by 3.5 percentage points in the shift from off-cycle to presidential elections.

This, then, is preliminary evidence in support of H1: the effect of greater gender stereotyping on female candidates' win rates appears to be more negative for mayoral candidates than for city council candidates. The averages in Table 2.2 also serve as a validity check for our overall empirical strategy, because just as we expect, the effects are more muted for midterm elections (where turnout is somewhat lower) than for presidential elections (where turnout is highest). In what follows, then, we focus our analysis on the comparisons between off-cycle local elections and local elections concurrent with presidential elections; we present the results for local elections concurrent with midterms in the online appendix.

As a second step toward evaluating H1, we explore whether these general relationships hold across time, or whether they are sensitive to the cyclical nature of state and national elections or the presence of prominent women running for state and national office. In Table 2.3, we once again analyze the residuals from the model that partials out the effects of incumbents running and competitiveness, but this time we break our data into four-year increments and calculate separate averages for each time period. (Because we have 22 years of data, the final period—2015-2016—only has two years.) As in Table 2.2, the top row for each time period shows the average residuals for male and female candidates in off-cycle city council elections, as well as the difference between the female and male averages ("Female Advantage"). In the second row, we show the same figures but for council elections concurrent with presidential elections. The third row shows the effect of a change from off-cycle to presidential elections (the difference between the first two rows), such that the number in bold in the third row represents what we are referring to as the effect of more gender stereotyping in city council elections. The next three rows present the same statistics but for mayoral elections. Finally, in this table, the rightmost column labeled "Mayor/Council Difference" shows the effect of more stereotyping in mayoral elections minus the effect of more stereotyping in city council elections.

Focusing on the Mayor/Council Difference—which is the most relevant statistic for evaluating H1—we see that the difference is negative for all time periods except one: 2015-2016. And for most of the time periods, the effect is similar in magnitude: the effect of more gender stereotyping is roughly 6 to 10 percentage points lower (more negative) in mayoral elections than in city council elections. Two time periods stand out, however. The first is 2007-2010, when the effect of more gender stereotyping is much larger for mayoral elections, and in the negative direction: almost 32 percentage points. The second is 2015-2016, when suddenly the shift to on-cycle elections works more in favor of women running for mayor than women running for city council.

We did not expect the patterns in 2007-10 and 2015-16 to deviate from the overall trend, but it may be that this reveals something interesting about heterogeneity in the effects. These are the two time periods in which Hillary Clinton was running for president: in 2007-08 she was vying for the Democratic nomination, and in 2015-2016 she was the Democratic nominee. One possibility is that Clinton's campaigns increased the salience of gender and made voters more likely to use gender as a cue in evaluating female candidates for local offices—*and* in a way that affected female candidates for executive office differently than female candidates for legislative office (since Clinton was running for president, not Congress). But if the deviations are a Clinton effect, why does the direction of that effect change from 2007-08 to 2015-16? We cannot know for sure, but one possibility is that the nature of the campaigns in 2007-08 and 2015-16 were dramatically different: In 2007-08, Clinton was competing with Barack Obama, did not emphasize gender in her campaign, and did not ultimately become the nominee. In 2015-16, by contrast, Clinton did emphasize gender, and she was competing against Donald Trump in a general election campaign that was highly charged on both gender and

race dimensions. Given the stark differences in the nature of these campaigns, it seems reasonable that they would shape local candidates' electoral prospects in different ways. Alternatively, the 2015-16 averages might reflect a pro-woman turnout effect in the November 2016 election: women motivated to turn out to support Hillary Clinton might have been more favorable to female mayoral candidates as well.

This is all speculation, of course. And it is important to point out that any given time period features relatively few non-incumbent women running for mayor. 2015-16 in particular only spans two years, and in that two-year period, only 7 female non-incumbents ran for mayor in off-cycle elections—and every one of them lost. Meanwhile, in the November 2016 election, 11 female non-incumbents ran for mayor, and their overall win rate—31%—was higher than in past mayoral elections concurrent with presidential elections. It may be, then, that 2015-2016 was just a peculiar cycle and not indicative of effects of the highly gender-charged presidential campaign. At a minimum, though, we need to be cognizant that the effects we are estimating appear to be different in 2015-2016, and in moving forward, we will want to carry out our tests both with and without the 2015-16 election cycle. As a next step, we test H1 with the following linear probability model:[8]

$$\begin{aligned} Win_{ijk} = {} & \beta_0 + \beta_1 Female_{ijk} + \beta_2 Oncycle_{jk} + \beta_3 (Female_{ijk} \times Oncycle_{jk}) + \beta_4 Mayor_{jk} \\ & + \beta_5 (Female_{ijk} \times Mayor_{jk}) + \beta_6 (Oncycle_{jk} \times Mayor_{jk}) \\ & + \beta_7 (Female_{ijk} \times Mayor_{jk} \times Oncycle_{jk}) + \beta_8 Competition_{jk} \\ & + \beta_9 Incumbent\ Ratio_{jk} + \alpha_k + \varepsilon_{ijk} \end{aligned}$$

Subscript *i* denotes the candidate, *j* denotes the race, and *k* denotes the jurisdiction. *Female*, *Incumbent Ratio*, and *Competition* are described above, and *Mayor* equals 1 if race *j* in city *k* is a mayoral election and 0 if it is a city council election. *On-cycle* equals 1 if the local race was concurrent with a presidential general election and 0 if it was not concurrent with any statewide or national election. We run models with and without city fixed effects (α_k), but when they are included, they partial out the effects of any time-constant city characteristics that are correlated with city election timing, the type of candidates who run for office, competitiveness, and the like. These models put the focus on within-city variation—and shed light on differences in the experiences of female candidates running for city council and mayor within the same city.[9] We cluster the standard errors by city.

Because our model includes all interactions of *Female*, *On-cycle*, and *Mayor*, it is worth discussing how to interpret coefficients β_1 through β_7 as well as what to look for in the test of H1. Most straightforwardly, β_1 can be interpreted as the effect of being a woman (as opposed to a man) on the likelihood of winning in off-cycle city council races (i.e., where *On-cycle* and *Mayor* equal 0). β_2, then, shows us the average difference in win rates for male candidates running in on-cycle races versus off-cycle races, and β_3 can be interpreted as the change in the effect of being a woman when moving from an off-cycle to an on-cycle election. Thus, β_3 is what we are referring to as the effect of more stereotyping on women's win rates in city council races: When electorates are composed of voters with less information on average, does the effect of being female on win rates increase, decrease, or stay the same? We can then estimate the same quantities for mayoral races: $\beta_1 + \beta_5$ is the effect of being a woman in off-cycle mayoral races, $\beta_1 + \beta_3 + \beta_5 + \beta_7$ is the effect of being a woman in an on-cycle mayoral race, and therefore the change in the effect of being a female when moving from an off-cycle to an on-cycle race is $\beta_3 + \beta_7$. That means that the *difference* between the effect of more

[8] We use a linear probability model rather than logistic regression because it allows us to estimate marginal effects without calculating probabilities.

[9] This allows us to go beyond Palmer and Simon's (2010) description of "women-friendly" districts to see how women fare at different times within a particular jurisdiction.

stereotyping in city council races and mayoral races is simply β_7—the coefficient on the triple interaction term. Specifically, if $\beta_7 < 0$, that would support H1. If $\beta_7 \geq 0$, that would be evidence against H1.

We start by estimating the model without city fixed effects, including all years of data, 1995-2016. The coefficient estimates and standard errors are presented in column 1 of Table 2.4. The coefficient on *Female* is positive, suggesting that in the average off-cycle city council race, being a female candidate is associated with a slightly higher probability of winning—about 3 percentage points higher. That female advantage grows in the shift to on-cycle city council elections: the coefficient on the interaction of *Female* and *On-cycle* is 0.042 and statistically significant. β_5 is negative but not statistically distinguishable from zero, suggesting that the effect of being female is not significantly different in off-cycle city council and mayoral races. That said, when we add together β_1 and β_5, as we do at the bottom of column 1, we find that in off-cycle mayoral races, being a woman has no significant effect on the likelihood of winning. Furthermore, when we examine the effect of being a woman in on-cycle mayoral races, also shown at the bottom of column 1, we estimate a coefficient of -0.042—although not statistically significant (p=0.187). Thus, the advantage for female council candidates grows in the shift to on-cycle elections, but there is no female advantage to speak of in mayoral elections—and if anything, the shift to on-cycle elections works to the disadvantage of female mayoral candidates.

Ultimately, the test of H1 is the coefficient on the triple interaction term, and in column 1, that coefficient is negative but not significant. However, we have just shown that 2015-16 was unusual in featuring a rare advantage of female mayoral candidates in on-cycle elections, so in column 2 of Table 2.4, we exclude those two years of data from the analysis. There, with 2015-16 excluded, the coefficient on the triple interaction term is -0.116 and statistically significant. What we find more generally is that there is still a female advantage in city council elections—one that grows with a shift to on-cycle election timing. For female mayoral candidates in off-cycle elections, there is no advantage, and in on-cycle mayoral elections, women are at a statistically significant disadvantage: all else equal, women are 7.2 percentage points less likely to win. And focusing on the negative coefficient on the triple interaction term, we conclude that the effect of the low-information environment—and more stereotyping—on women's probability of winning is about 11.6 percentage points smaller for the average mayoral race than the average city council race. Our findings here are therefore supportive of H1.

In column 3 of Table 2.4, we add city fixed effects, and our findings are substantively the same. Most importantly for testing H1, the estimate of β_7 is still negative and significant, suggesting that the effect of more gender stereotyping is more negative for women in mayoral races than in council races. In column 4, in addition to dropping 2015-16, we drop 2007-08 (because Table 2.3 showed that the effects were larger during Hillary Clinton's first presidential campaign). As expected, the estimate of β_7 is smaller in magnitude than in column 2: it is -0.091, and only significant at the 10% level in a one-tailed test (p=0.09). Even in this model, however, we find that the effect of being a woman in on-cycle mayoral elections is negative and significant: holding constant race competitiveness and incumbents running, female candidates are 7.2 percentage points less likely to win than men in on-cycle mayoral races.

Finally, we consider whether it makes sense to limit the analysis to competitive races with both male and female candidates running. One could argue that voters cannot easily rely on gender stereotypes to decide among candidates if all candidates are of the same gender,[10] so perhaps we should

[10] That said, we think it possible that candidates of the same sex could provide more or fewer feminine cues, such as through ballot designations (e.g., "attorney/mother" vs. "attorney".)

focus on races in which voters could plausibly use gender as a cue. However, a closer look at the distribution of mixed-sex races across office types and election timing gives us pause. In particular, we find that for city council races, a larger share of the on-cycle races are mixed-sex races than the off-cycle races: specifically, 65% of the off-cycle council races are mixed-sex, whereas 69% of the on-cycle races are. By contrast, in mayoral races, there is no significant difference (36% of off-cycle races are mixed-sex, while 35% of on-cycle races are). It is possible that these patterns point to strategic entry by female candidates in *anticipation* of the differential effect of more gender stereotyping: Perhaps more women enter on-cycle council races than off-cycle council races because they expect that the larger, less-informed electorate will tend to favor them. And perhaps potential female candidates also know that that is not the case for mayoral races. By itself, that would be a pattern supportive of H1. But it also suggests that by limiting our model to mixed-sex races, we reduce our ability to detect a difference in the effect of more gender stereotyping across contexts.

Even so, for the sake of evaluating the robustness of our results, we estimate the same model as in column 2, this time limiting the sample to mixed-sex races that have more candidates running than open seats. As we show in column 5, we lose about a quarter of the observations when we do this. We still find evidence of a female advantage in off-cycle council elections, and we still find a positive effect for women in the shift to on-cycle council elections. As before, we also find that the effect of being a woman diminishes the probability of winning in on-cycle mayoral elections. Our test of H1, however, is based on the coefficient on the triple interaction term. That coefficient is -0.106, but here it is only significant at the 10% level in a one-tailed test.

To summarize our findings so far, we do find evidence supportive of H1, although the magnitude and precision of our main estimates depend on whether we limit the model to competitive races featuring both male and female candidates, and whether we include 2015-16 and 2007-08 in the estimation (both of which deviate from the overall trends). We suspect that strategic candidate entry is contributing to the attenuation of the main coefficient of interest in column 5, and the distribution of mixed-sex races across office types and election timing is consistent with that. We also speculate that Hillary Clinton running for president may have changed the salience of candidate gender—for executive offices in particular—during the two periods in which she was running for president. However, with our current data, we cannot test for these factors, so we consider these results preliminary and tentative.

Next, we turn to H2: a test of whether the effect of increased stereotyping is larger (or more positive) in school board races than in city council races. Our approach is the same as before, except that now we use the sample of city council and school board races. Table 2.5 shows the average residuals from a regression of *Win* on *Incumbent Ratio* and *Competitive*, as we did with the city council and mayoral race comparisons. As before, we find that there is a slight advantage for female candidates in off-cycle city council races and that the advantage grows in the shift to on-cycle elections. Interestingly, we find that the female advantage in off-cycle school board elections (6.1 percentage points) is larger than that of city council elections (2.9 percentage points). We also find that that female advantage in school board elections grows to 9.2 percentage points in the shift to on-cycle elections. However, our hypothesis here is that the shift to a lower-information environment will help female school board candidates *more* than female city council candidates, and the evidence in Table 2.5 does not support that. Instead, we find the female advantage in city council elections and school board elections grows by roughly the same amount when elections are shifted to on-cycle—by 3.1 to 3.9 percentage points. Moreover, in the bottom half of the table, we show that the quantity most relevant for H2—the School Board / City Council Difference—is negative in some years, positive in others, and close to zero in the period 1999 to 2002.

It is no surprise, then, that when we estimate the linear probability model of *Win* (using the same specification as before except replacing *Mayor* with *School Board*), we do not find that the estimates

support H2. See Table 2.6. Just as in Table 2.5, our coefficient estimates show that both female city council and school board candidates have an advantage over men in off-cycle elections—and that that advantage grows with more gender stereotyping. If anything, the female advantage in off-cycle elections appears to be larger for school board candidates than for city council candidates: the coefficient on *Female* × *School Board* is 0.033, although it is not significant. However, we do not find that the shift from off-cycle to on-cycle elections leads to a significantly *greater* advantage for female school board candidates than it does for female city council candidates: the coefficient on the triple interaction term is indistinguishable from zero. Thus, contrary to H2, the results here do not suggest that the effect of greater stereotyping is more positive for school board races than for city council races. Rather, the shift to on-cycle election timing lifts female electoral prospects roughly equally in both types of races.[11]

Finally, we turn to our third hypothesis, which is that due to beliefs stereotypes that female candidates are more liberal, the effect of more gender stereotyping on women's electoral success will be more negative in conservative constituencies than in liberal constituencies. We focus here on city-level races because we have a measure of conservatism—Republican two-party vote share from the 2012 presidential election—at the city level. (We do not have presidential vote share at the school district level.) We test H3 for each type of race separately—city council races and mayoral races—to explore whether the potential effects of this stereotype vary across contexts.

We begin once again with an analysis of the residuals from OLS regressions of *Win* on *Incumbent Ratio* and *Competitive* (this time running separate models for city council and mayoral candidates). Figure 2.1 presents two lowess plots: the dashed line is a lowess plot of the residuals for male city council candidates in off-cycle elections and city Republican presidential vote share. The solid line is the same but for male city council candidates in on-cycle elections. The histogram in grey depicts the frequency of city-level Republican presidential vote share across the male non-incumbent city council candidates in the sample. Then, in Figure 2.2, we present the same lowess plots but for female city council candidates. These plots and the plots to follow exclude the data from 2015-16.

We are focused here on the difference between the dashed and solid lines (the shift from off-cycle to on-cycle elections) and whether the gap between the two lines changes in a systematic way as we move from more liberal to more conservative cities. For city council candidates, shown in Figures 2.1 and 2.2, we do not see any clear patterns. For male city council candidates (Figure 2.1), the shift to on-cycle elections brings a slight disadvantage in the most liberal cities, a slight advantage for cities between 20% and 40% Republican vote share, a slight disadvantage for cities between 40% and 60% Republican vote share, and then an advantage in the most conservative cities (where the data are sparse). For female city council candidates (Figure 2.2), the plots show that women have a consistent advantage in on-cycle elections over off-cycle elections, regardless of how conservative the city is. Notably, Figure 2.2 does not show that the size of the gap between the dashed and solid lines narrows with increasing Republican presidential vote share.

In Figures 2.3 and 2.4, we create the same graphs for male and female mayoral candidates, and there we do see clear patterns. Specifically, Figure 2.3 presents lowess plots of the residuals for male mayoral candidates in off-cycle (dashed line) and on-cycle (solid line) elections. We can see that for the most liberal cities, the solid line is well below the dashed line—suggesting that men's win rates in liberal cities tend to be lower in on-cycle elections than in off-cycle elections. But as city Republican vote share increases, the gap between the lines shrinks. And at about 20% Republican presidential vote share, male mayoral candidates' on-cycle disadvantage turns into an advantage: on-cycle election

[11] Our conclusions are substantively the same when we drop 2015-16, when we also drop 2007-08, when we include county fixed effects, and when we limit our estimation to competitive, mixed-sex races. See the online appendix for those results.

timing becomes *good* for male candidates. That advantage expands as cities become even more Republican (with the exception of the most conservative cities, where there is very little data). Thus, what starts out as relative disadvantage for male mayoral candidates in on-cycle elections in liberal cities turns to positive and widening advantage in on-cycle elections in more conservative cities.

Figure 2.4 presents the same lowess plots but for female mayoral candidates.[12] The first feature of the plot worth noting is the dearth of female mayoral candidates in more conservative cities: for example, 27% of the non-incumbent mayoral candidates in the most liberal cities (less than 10% Republican presidential vote) are female, but only 16% of the non-incumbent mayoral candidates in more conservative cities (greater than 50% Republican vote share) are female. Equally striking is the pattern we see in the cities where most of the data are: cities between 20% and 50% Republican vote share. For the more liberal cities in this interval (e.g., 20% Republican vote share), shifting from an off-cycle election to an on-cycle election gives a slight boost to female mayoral candidates. That boost shrinks, however, in the shift to more conservative cities—and past about 25% Republican presidential vote share, on-cycle election timing actually lowers female candidates' win rates. This pattern does not hold for the most liberal cities—such as those with less than 10% Republican vote share—but the data there are thin. It may also be that in the most liberal cities (such as Berkeley) the liberal cue provided by female candidates is not a particularly meaningful one for voters. Beyond the most liberal cities, however, the patterns in Figures 2.3 and 2.4 are supportive of H3: for mayoral elections, it looks as though the shift to on-cycle elections helps female candidates in liberal cities and increasingly hurts female candidates as one moves to more conservative cities.

We test this with linear probability models in Table 2.7. For these models, we center Republican presidential vote share around its average in the data—0.375—so that the coefficients on *Female* and *Female* × *On-cycle* represent the effect of being female in cities with average Republican presidential vote share. In column 1, we start with a model of city council candidates. The coefficients on *Female* and *Female* × *On-cycle* show that in a city with average Republican vote share, female candidates fare somewhat better than male candidates in off-cycle council elections, and also that the effect of being female more than doubles when council elections are shifted to on-cycle. What, then, of cities that are more Republican than the average city? The coefficient on *Female* × *Republican vote* is negative but statistically insignificant, indicating that there is no clear difference in female council candidates' electoral fortunes in more Republican cities holding off-cycle elections. Nor is the effect of the shift to on-cycle elections significantly different from the effect for more liberal cities: the coefficient on the triple interaction term is negative but not significant. Therefore, as in Figures 2.1 and 2.2, for city council races, we do not see clear evidence that the effect of a shift to a lower-information environment is more negative in Republican cities than in Democratic cities. The same is true when we add city fixed effects, drop 2015-16, and limit the estimation to candidates running in competitive, mixed-sex races (not shown). There is an advantage for women in city council races—one that grows in the shift to lower-information environments—but that advantage does not change significantly depending on the conservatism of the city.

In the remaining columns of Table 2.7, we estimate models for mayoral candidates, and our conclusions are quite different. In column 2, we present the results with all years of data, but we have already shown that the dynamics in mayoral elections were different in 2015-16, so in column 3, we drop those two years. As we expect, the results are clearer when we drop 2015-16, so we focus our discussion on that column of estimates. To start, we find that in cities with off-cycle elections and average Republican presidential vote share, there is no effect of being a woman on the likelihood of

[12] Here, for presentation purposes, we drop one race in which a female mayoral candidate (in an off-cycle election) won in a very conservative city. We do this for presentation purposes only—this particular candidate has a very large residual in the direction we would expect.

winning. However, the effect of being a woman in on-cycle city council races in these cities is *negative*—and significant at the 5% level (see the bottom of column 3). On average, the effect of being a female candidate in a mayoral race is a 7.2 percentage point reduction in the likelihood of winning. Moreover, the change in the effect of being a woman resulting from the shift from off-cycle to on-cycle elections is negative and significant. Thus, even in cities with average partisanship, the shift to a lower-information environment generates penalties for female non-incumbents running for mayor.

What about women's win rates in more conservative cities? At the bottom of Table 2.7, we combine coefficients to show the effect of being female in more Republican cities—specifically, cities with one standard deviation more Republican presidential vote share (16 percentage points more, or cities with about 53% Republican vote share). There, we can see that for these more conservative cities with off-cycle elections, there is no significant advantage or disadvantage of being a female candidate (although the coefficient is positive). In on-cycle elections, the female effect turns negative (-0.12) and is statistically significant. However, we are especially interested the effect of shifting from an off-cycle to an on-cycle mayoral election in these more conservative cities, and that effect is shown by *Female* × *On-cycle* + *Female* × *On-cycle* × *Republican vote*. At the bottom of the table, we estimate that the effect of the shift to a lower-information environment in these cities and races is negative and statistically significant, suggesting that more gender stereotyping hurts female candidates running for mayor in more conservative cities. Furthermore, the test of H3 is based on the coefficient on the triple interaction term, and that coefficient is large and negative. This means that the shift from an off-cycle to an on-cycle mayoral election has a more negative effect on women's win rates in more conservative cities than in cities with average Republican vote share.

In column 4, we drop an outlier: a female non-incumbent mayoral candidate who won in an off-cycle election in a conservative city (with 59% vote share for Romney in 2012). When we drop this outlier, the coefficients of interest are slightly smaller, but they point in the same direction: Even in the cities with average Republican vote share, there is a disadvantage for female candidates in local elections concurrent with presidential elections. Also, that on-cycle disadvantage becomes larger in more conservative cities.

In column 5, we include city fixed effects (and still drop the outlier), and we find the same patterns. Moreover, in column 6, where we limit the analysis to competitive, mixed-sex races—which by our logic should attenuate the effects—we still find a significant, negative difference between the effect of more gender stereotyping on female win rates in more and less conservative cities. Thus, on average, with the exception of the 2015-16 cycle, more gender stereotyping in mayoral races *lowers* women's likelihood of winning, and that disadvantage tends to be larger in more conservative cities.

Discussion

This empirical analysis is preliminary, and there is more to be done before we can make broad assessments about the conditions under which gender stereotyping helps or hurts (or has no effect on) women's success at the ballot box.

First, in future empirical work, we will attempt to evaluate whether variation in candidate quality by candidate sex is affecting our estimates. One of the main critiques of research that uses electoral returns to demonstrate the presence or absence of gender bias is that it is difficult to account for the effects of candidate quality (e.g., Anzia and Berry, 2011; Fulton, 2012). All else equal, we should expect a higher quality candidate to win greater vote share. Thus, potential differences between the average quality of male and female candidates in our analysis may explain some of our estimated effects of being a woman on win rates. To address this concern, we have started collecting survey data on how citizens evaluate candidates' ballot designations, and we intend to use those evaluations as a

measure of quality for many of the candidates in our dataset. By including this measure in our models, we will assess whether variation in candidate quality explains some of the difference in men and women's win rates, and we can also test whether the returns to candidate quality are different for male and female candidates.

Second, we plan to evaluate whether patterns of strategic candidate entry are reducing our ability to detect effects of greater gender stereotyping by voters. As we discussed earlier, the number of women running for office appears to depend on the office sought and on the timing of the election. Moreover, the patterns we observe in the data are consistent with an account in which women are more likely to run in contexts more favorable to them. If women are more likely to run in on-cycle than off-cycle city council elections in anticipation of the effects of stereotyping, that would reduce our ability to detect a difference between the effects in city council and mayoral races. Our measure of candidate quality will help us evaluate this possibility to some extent: for example, we can compare the quality of female candidates in mayoral and council elections and in on-cycle and off-cycle elections. For the time being, we note that the patterns we observe in our electoral data are consistent with this account, and that it is reasonable to expect that our estimates here are smaller because of strategic candidate entry.

Third, our future work will evaluate whether other differences in on-cycle and off-cycle electorates—not just differences in the prevalence of stereotyping—might account for differences in women's electoral success. A growing body of work shows that there are indeed differences in the electorates of on- and off-cycle elections. One of us, for example, has provided evidence that the low turnout of off-cycle elections increases the electoral influence of organized groups (Anzia, 2014). Importantly, though, her argument is that the overall effect of increased group influence in off-cycle elections should depend on the nature of group activity and competition in a polity—and we do not see any *a priori* reason for why this effect should disproportionately favor or disfavor female candidates. That said, she does show evidence that teacher unions are advantaged in the average off-cycle school board election, and that municipal employee unions benefit from off-cycle city elections. Again, it is not clear whether this pattern would affect the win rates of female candidates, but in future work, we will examine patterns of candidate endorsements by these groups to evaluate whether they are more or less likely to support women.

Along the same lines, off-cycle electorates tend to be older (Anzia, forthcoming; Kogan et al., forthcoming) and less racially and ethnically diverse (Hajnal and Trounstine, 2005) than on-cycle electorates. To the extent that younger and more diverse electorates are more inclined to elect female candidates, it might explain some of the positive effects of the shift to on-cycle election timing for women running in city council and school board races. But we are hard-pressed to come up with an explanation for how this could account for the opposite pattern for women running in mayoral races—especially in more conservative constituencies. Thus, while we certainly recognize that off-cycle and on-cycle electorates differ in ways other than the amount of stereotyping by voters, we have little reason to think that those differences can explain our findings.

In sum, we propose that our empirical strategy offers a credible way of testing the effects of gender stereotyping on women's electoral success. The literature features a great deal of work on gender stereotyping, but to the extent we want to know about the effects of stereotyping in real elections, there are advantages to analyzing real election data. The challenge up to this point has been the difficulty of studying psychological processes like stereotyping using aggregate data on how people voted. We have developed a strategy for doing that. In addition, we have proposed that the effects of greater stereotyping on women's win rates should vary by context. Our evidence supports this expectation and suggests that there isn't a one-size-fits-all answer to the question of how women are affected by gender stereotyping by voters. Instead, the answer is conditional: the effect of greater

stereotyping is more negative for women running in executive races than in legislative races, especially in more conservative constituencies.

Supplemental Materials

Table 2.1. Numbers of Races and Candidates

	All Offices	City Council	Mayor	School Board
Number of races	15,239	6,037	1,190	8,012
Full-term	14,298	5,720	1,146	7,432
Concurrent with presidential election	5,209	2,057	417	2,735
Concurrent with midterm election	4,482	1,693	421	2,368
Contested	14,183	5,438	929	7,816
With gender coded	15,503	6,306	1,187	8,010
Number of candidates	62,926	27,448	3,175	32,303
Incumbents	19,168	7,247	836	11,085
With gender coded	62,178	27,133	3,147	31,898
Female	21,272	7,205	664	13,403
Female non-incumbents	14,073	5,122	505	8,446

Table 2.2. The Effect of More Gender Stereotyping, City Council and Mayor

City Council, Win Rates

	Male	*Female*	*Female Advantage / (Disadvantage)*
Off-cycle	0.269	0.303	0.034
Midterm	0.290	0.339	0.050
Change, off-cycle to midterm	0.021	0.036	**0.015**
Presidential	0.291	0.380	0.088
Change, off-cycle to presidential	0.023	0.077	**0.054**

Mayor, Win Rates

	Male	*Female*	*Female Advantage / (Disadvantage)*
Off-cycle	0.167	0.178	0.011
Midterm	0.245	0.234	(0.011)
Change, off-cycle to midterm	0.077	0.056	**(0.021)**
Presidential	0.268	0.236	(0.032)
Change, off-cycle to presidential	0.101	0.058	**(0.043)**

City Council, Residuals

	Male	*Female*	*Female Advantage / (Disadvantage)*
Off-cycle	-0.018	0.013	0.031
Midterm	-0.008	0.036	0.044
Change, off-cycle to midterm	0.010	0.023	**0.013**
Presidential	-0.008	0.065	0.073
Change, off-cycle to presidential	0.009	0.051	**0.043**

Mayor, Residuals

	Male	*Female*	*Female Advantage / (Disadvantage)*
Off-cycle	0.042	0.032	(0.010)
Midterm	-0.011	-0.032	(0.021)
Change, off-cycle to midterm	-0.052	-0.063	**(0.003)**
Presidential	-0.010	-0.053	(0.043)
Change, off-cycle to presidential	-0.052	-0.085	**(0.035)**

Table 2.3. City Council and Mayoral Elections by Time Period

		Male	*Female*	*Female Advantage*	*Mayor/Council Difference*
1995-1998	**City Council**				**(0.073)**
	Off-cycle	-0.027	0.028	0.055	
	Presidential	-0.013	0.056	0.069	
	Change, off-cycle to presidential	0.014	0.027	**0.014**	
	Mayor				
	Off-cycle	-0.063	0.003	0.066	
	Presidential	-0.071	-0.065	0.007	
	Change, off-cycle to presidential	-0.008	-0.068	**(0.059)**	
1999-2002	**City Council**				**(0.065)**
	Off-cycle	-0.001	0.018	0.019	
	Presidential	-0.021	0.051	0.073	
	Change, off-cycle to presidential	-0.021	0.033	**0.054**	
	Mayor				
	Off-cycle	0.089	0.045	(0.043)	
	Presidential	0.0002	-0.054	(0.054)	
	Change, off-cycle to presidential	-0.088	-0.099	**(0.011)**	
2003-2006	**City Council**				**(0.057)**
	Off-cycle	-0.037	0.037	0.074	
	Presidential	-0.012	0.033	0.045	
	Change, off-cycle to presidential	0.025	-0.004	**(0.029)**	
	Mayor				
	Off-cycle	0.015	0.061	0.046	
	Presidential	-0.013	-0.053	(0.040)	
	Change, off-cycle to presidential	-0.028	-0.114	**(0.085)**	
2007-2010	**City Council**				**(0.317)**
	Off-cycle	-0.022	-0.026	(0.004)	
	Presidential	-0.023	0.074	0.097	
	Change, off-cycle to presidential	-0.001	0.100	**0.101**	
	Mayor				
	Off-cycle	0.010	0.161	0.151	
	Presidential	-0.007	-0.072	(0.066)	
	Change, off-cycle to presidential	-0.017	-0.233	**(0.216)**	
2011-2014	**City Council**				**(0.104)**
	Off-cycle	-0.014	0.005	0.019	
	Presidential	0.012	0.079	0.067	
	Change, off-cycle to presidential	0.025	0.074	**0.049**	
	Mayor				
	Off-cycle	0.112	0.005	(0.107)	
	Presidential	0.029	-0.133	(0.162)	
	Change, off-cycle to presidential	-0.083	-0.138	**(0.055)**	
2015-2016	**City Council**				**0.257**
	Off-cycle	0.007	0.049	0.042	
	Presidential	0.005	0.085	0.080	
	Change, off-cycle to presidential	-0.002	0.036	**0.038**	
	Mayor				
	Off-cycle	0.078	-0.161	(0.239)	
	Presidential	-0.011	0.044	0.055	
	Change, off-cycle to presidential	-0.089	0.205	**0.294**	

Table 2.4. Effect of an Increase in Gender Stereotyping in Mayoral versus City Council Elections

	(1)	(2)	(3)	(4)	(5)
Female	0.03	0.029	0.027	0.036	0.03
	(0.016)	(0.016)	(0.017)	(0.017)	(0.017)
On-cycle	0.007	0.007	0.003	0.01	0.001
	(0.009)	(0.009)	(0.016)	(0.010)	(0.010)
Female × On-cycle	0.042	0.041	0.048	0.027	0.046
	(0.021)	(0.022)	(0.023)	(0.024)	(0.023)
Mayor	0.074	0.07	0.04	0.073	0.111
	(0.027)	(0.030)	(0.021)	(0.028)	(0.043)
Female × Mayor	-0.042	-0.025	-0.02	-0.045	-0.07
	(0.045)	(0.048)	(0.047)	(0.050)	(0.061)
On-cycle × Mayor	-0.074	-0.068	-0.039	-0.075	-0.096
	(0.031)	(0.035)	(0.026)	(0.034)	(0.051)
Female × On-cycle × Mayor	-0.072	-0.116	-0.118	-0.091	-0.106
	(0.058)	(0.062)	(0.062)	(0.068)	(0.079)
Competition	-0.066	-0.065	-0.057	-0.064	-0.056
	(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Incumbent Ratio	-0.213	-0.218	-0.215	-0.214	-0.18
	(0.008)	(0.009)	(0.010)	(0.009)	(0.009)
Constant	0.588	0.584		0.58	0.525
	(0.020)	(0.020)		(0.020)	(0.017)
Model	All years	1995-2014	1995-2014, city FE	1995-2006, 2009-14	1995-2014, competitive mixed-sex
R-squared	0.1	0.1	0.14	0.1	0.08
Observations	13,551	11,703	11,703	9,928	9,037
Female + (Female × On-cycle)	0.072	0.070	0.075	0.063	0.076
	(0.013)	(0.015)	(0.015)	(0.016)	(0.016)
Female + (Female × Mayor)	-0.012	0.004	0.008	-0.009	-0.039
	(0.046)	(0.049)	(0.049)	(0.051)	(0.062)
Female + (Female × On-cycle) +	-0.042	-0.072	-0.062	-0.072	-0.099
(Female × Mayor) + (Female × On-Cycle × Mayor)	(0.031)	(0.035)	(0.036)	(0.043)	(0.046)
(Female × On-cycle) +	-0.030	-0.075	-0.070	-0.063	-0.060
(Female × On-cycle × Mayor)	(0.056)	(0.062)	(0.062)	(0.067)	(0.078)

Notes: Standard errors clustered by city in parentheses. Dependent variable = 1 if candidate won and 0 if he/she lost.

Table 2.5. City Council and School Board Elections

		Male	*Female*	*Female Advantage*	*School Board / City Council Difference*
All Years	**City Council**				**(0.008)**
	Off-cycle	-0.024	0.005	0.029	
	Presidential	-0.028	0.040	0.068	
	Change, off-cycle to presidential	-0.004	0.035	**0.039**	
	School Board				
	Off-cycle	-0.005	0.056	0.061	
	Presidential	-0.023	0.070	0.092	
	Change, off-cycle to presidential	-0.018	0.013	**0.031**	
1995-1998					0.042
1999-2002					(0.009)
2003-2006					0.145
2007-2010					(0.102)
2011-2014					(0.042)
2015-2016					(0.129)

Table 2.6. Effect of an Increase in Gender Stereotyping in School Board versus City Council Elections

Female	0.028
	(0.013)
On-cycle	0.00005
	(0.012)
Female × On-cycle	0.041
	(0.022)
School board	0.025
	(0.010)
Female × School board	0.033
	(0.022)
On-cycle × School board	-0.018
	(0.013)
Female × On-cycle × School board	-0.009
	(0.027)
Competition	-0.09
	(0.006)
Incumbent Ratio	-0.214
	(0.009)
Constant	0.664
	(0.025)
R-squared	0.1
Observations	25,606
Female + (Female × On-cycle)	0.069
	(0.016)
Female + (Female × School)	0.062
	(0.012)
Female + (Female × On-cycle) +	0.093
(Female × School) + (Female × On-cycle × School)	(0.018)
(Female × On-cycle) +	0.031
(Female × On-cycle × School)	(0.021)

Notes: Standard errors clustered by county in parentheses.

Figure 2.1. Residuals, Male City Council Candidates

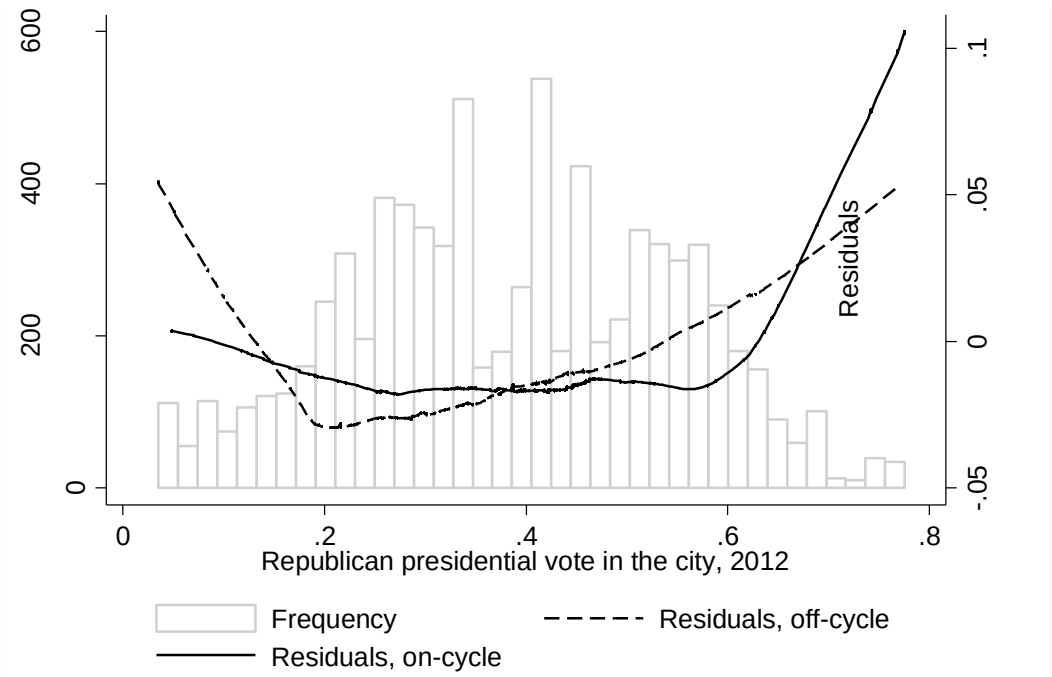


Figure 2.2: Residuals, Female City Council Candidates

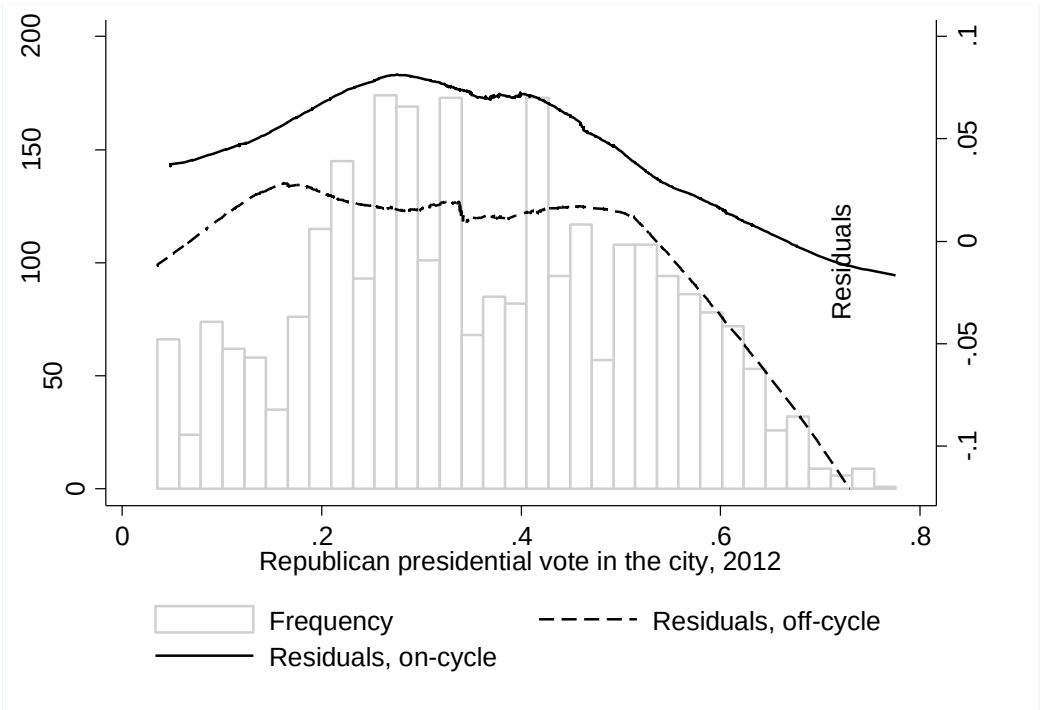


Figure 2.3. Residuals, Male Mayoral Candidates

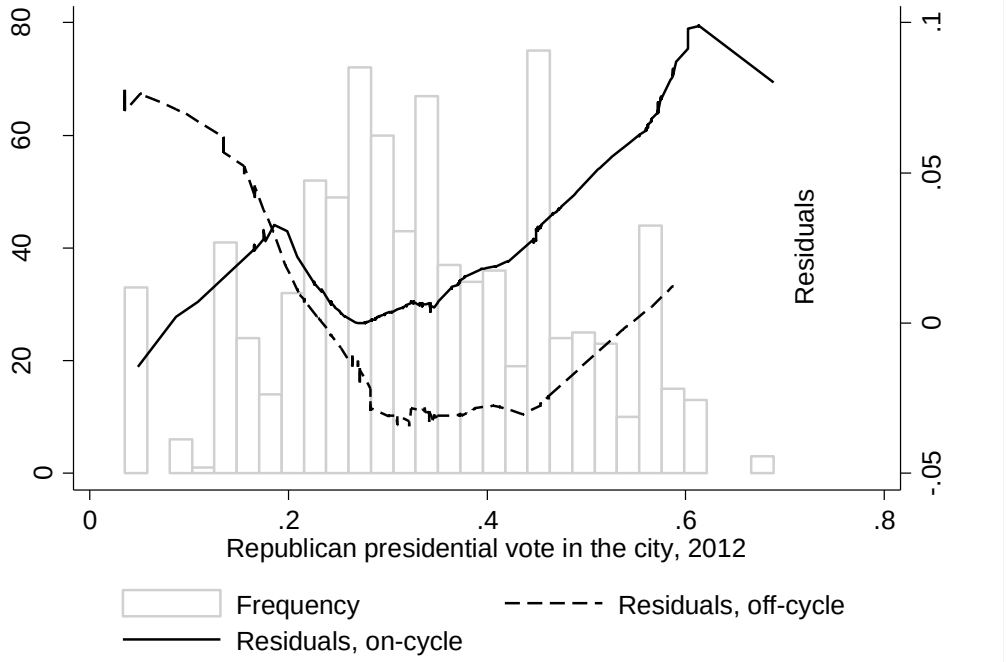


Figure 2.4. Residuals, Female Mayoral Candidates

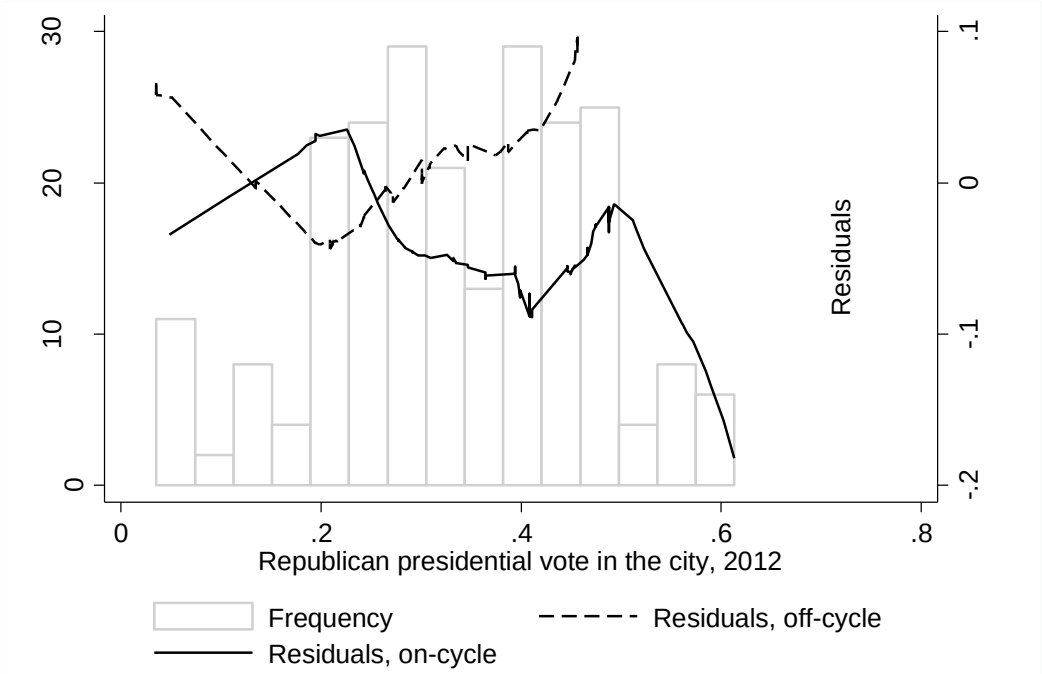


Table 2.7. Effect of an Increase in Gender Stereotyping, by City Republican Presidential Vote

	City Council			*Mayor*		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.028	0.041	0.064	0.041	0.042	0.065
	(0.016)	(0.062)	(0.068)	(0.069)	(0.082)	(0.091)
On-cycle	0.001	0.026	0.037	0.037	-0.011	0.03
	(0.009)	(0.022)	(0.024)	(0.024)	(0.044)	(0.042)
Female × On-cycle	0.046	-0.087	-0.136	-0.113	-0.105	-0.148
	(0.021)	(0.071)	(0.078)	(0.078)	(0.093)	(0.103)
Republican vote	0.024	-0.303	-0.322	-0.323	-7.344	-0.209
	(0.051)	(0.102)	(0.113)	(0.113)	(1.074)	(0.200)
Female × Republican	-0.047	0.477	0.555	0.394	0.371	0.517
	(0.109)	(0.371)	(0.411)	(0.400)	(0.459)	(0.510)
On-cycle × Republican	-0.004	0.415	0.51	0.511	0.143	0.538
	(0.062)	(0.142)	(0.164)	(0.164)	(0.438)	(0.251)
Female × On-cycle × Republican	-0.066	-0.526	-0.859	-0.698	-0.765	-0.956
	(0.142)	(0.471)	(0.502)	(0.494)	(0.570)	(0.621)
Competition	-0.076	-0.039	-0.038	-0.038	-0.043	-0.028
	(0.005	(0.006)	(0.006)	(0.005)	(0.008)	(0.005)
Incumbent Ratio	-0.217	-0.202	-0.213	-0.215	-0.222	-0.146
	(0.009)	(0.017)	(0.018)	(0.018)	(0.024)	(0.018)
Constant	0.624	0.476	0.464	0.465		0.379
	(0.017)	(0.034)	(0.035)	(0.035)		(0.047)
Model	All years	All years	1995-2014	1995-2014, no outlier	1995-2014, city FE, no outlier	1995-2014, competitive mixed-sex, no outlier
R-squared	0.1	0.14	0.15	0.15	0.22	0.12
Observations	12,290	1,261	1,084	1,083	1,083	596
Female effect, on-cycle, average partisanship	0.074	-0.046	-0.072	-0.072	-0.063	-0.083
	(0.013)	(0.033)	(0.036)	(0.036)	(0.043)	(0.046)
Female effect, off-cycle, Republican cities	0.020	0.117	0.153	0.104	0.101	0.148
	(0.026)	(0.114)	(0.126)	(0.125)	(0.148)	(0.165)
Female effect, on-cycle, Republican cities	0.056	-0.054	-0.120	-0.120	-0.126	-0.153
	(0.016)	(0.052)	(0.054)	(0.054)	(0.064)	(0.072)
Change in female effect, Republican cities	0.036	-0.171	-0.273	-0.225	-0.227	-0.301
	(0.031)	(0.126)	(0.138)	(0.137)	(0.162)	(0.181)

Notes: Standard errors clustered by city in parentheses.

Chapter 3

Wearing the Pants? The Role of Gendered Leadership Styles in Candidate Evaluations

Must female politicians adopt a masculine style to be taken seriously as leaders? Social psychological literature suggests that women face a "double bind" in leadership evaluations: when women exhibit masculine behavior, others perceive them to be "bossy" or "unlikable," but when they exhibit feminine behavior, they are not seen as leadership material. In two survey experiments of registered voters in September and October 2016 (n=1800 each), I find that on average, voters seemed to prefer both male and female candidates when they were described with feminine leadership styles. However, voter partisanship moderates this relationship: Democrats preferred a feminine style, while Republicans tended to have no preference or prefer masculinity. In contrast to earlier work implying there is a "right way" (or no right way) to present oneself as a woman in politics, the findings suggest that the right (gendered) leadership style to emphasize depends more on the partisanship of the candidate's base than it does on the candidate's sex, for both male and female politicians.

What explains the persistent gap in female representation in the U.S.? At the end of 2016, women made up only 19% of the United States Congress and 25% of state legislatures; 80 other nations have a higher percentage of female national legislators (Center for American Women and Politics [CAWP], n.d.). For many, the failure of Hillary Clinton's campaign for the presidency also prompted renewed concern about potential voter discrimination.

Over the last forty years, scholars of gender and women in politics have attempted to answer this question by exploring a variety of potential causes, including voter discrimination, incumbency effects that inherently favor men, biased media coverage, and lower political ambition among women. Many studies scrutinizing the gap between men and women's political attainment have achieved little consensus on how and when stereotypes play a role: results from empirical studies are often mixed or outright contradictory. In many regards, the currently predominant explanation rejects the notion of a biased public: "discrimination has fallen out of favor as an explanation for women's absence from electoral politics" (Lawless, 2015, p. 4).

Nonetheless, the possibility exists that voters do have preferences about political candidates' personalities and leadership styles that may interact with the candidate's sex, sometimes problematically for female candidates. On the one hand, there is reason to believe that candidates should try to present themselves as more masculine: a wealth of psychological literature suggests that when individuals evaluate a candidate's leadership potential, especially in a low-information setting, they are still using a fundamentally masculine set of criteria (Koenig, Eagly, Mitchell, & Ristikari, 2011), and some evidence seems to suggest that masculine women will fare better (Holman, Merolla, & Zechmeister, 2016; Huddy & Terkildsen, 1993a; Lawless, 2004). On the other, research also suggests that women will be penalized for non-normative, i.e. masculine, behavior even though this behavior is needed to be seen as a leader, while men suffer no such "backlash effect" (Rudman & Phelan, 2008), and Herrnson, Lay, and Stokes (2003) suggest that running "as a woman" is in fact beneficial for female candidates' chances of success.

A second possibility is that voter preferences about candidates' leadership styles may interact not just with candidate sex, but with voter partisanship. Lakoff (2010), for instance, suggests that conservatives seek a "Stern Father" figure from their government, while liberals prefer a "Nurturant Parent" model. Empirically, Winter (2010) finds a similar result: popular images of Republicans are gendered masculine, and Democrats gendered feminine. Analyses of aggregate voter preferences might therefore fail to find a result of discrimination, when in fact Democrats and Republicans (or liberals and conservatives) hold distinct preferences that cancel each other out.

In two survey experiments of registered voters, conducted in September and October 2016, I find that partisans do have distinct preferences over candidates' leadership style. Told that one of nine national-level candidates had a distinctive leadership style, Democrats felt more favorable toward the same candidate when that style was feminine, while Republicans typically had no preference or a slight preference for masculinity. The aggregate, Democratic-leaning sample preferred the feminine style across both male and female candidates. In contrast to earlier work implying there is a "right way" (or no right way) to present oneself as a woman in politics, the findings suggest that the right leadership style to emphasize depends more on the partisanship of the candidate's base than it does on the candidate's sex, for both male and female politicians.

THEORY

All societies segregate gender roles based on sex (Bem, 1981). Stereotyping, then, should be understood not as an inherently negative process, but a process of categorization and organization that creates gender roles. Psychology breaks these roles into two components: descriptive norms,

which are "consensual expectations about what members of a group actually do," and prescriptive norms, which are "consensual expectations about what a group of people ought to do or ideally would do" (Eagly and Karau 2002, p. 574).

Many psychology researchers studying gender use one of two (overlapping) frameworks when discussing the contents of those norms: agency (masculine) vs. communion (feminine), and competence (masculine) vs. warmth (feminine). The agency-communion framework is the most commonly used framework, and suggests respondents describe women with communal attributes – concerns with the welfare of others – such as affection, compassion, and sensitivity, while they identify men with agentic attributes, such as assertiveness, leadership, and self-confidence (see e.g. Eagly and Karau 2002).

In contrast, the competence-warmth framework, also known as the Stereotype Content Model, asks respondents how they would rate various groups (e.g., the elderly, Asians, teachers, women) on competence and warmth, and finds that men are perceived to be more competent and women, more warm (Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002). Typically, high-status groups are perceived as more competent (e.g., the wealthy, Jews). Psychologists studying gender less commonly use the latter framework, as it is broader but less nuanced than the agency-communion framework described above. However, it has been picked up by some political science researchers due to the known effects of competence in particular on the vote (see discussion in Carpinella and Johnson 2013a; Carpinella and Johnson 2013b). Both frameworks tend to emphasize the possible penalties that women face for role incongruity, i.e., non-normative or masculine behavior, even though this behavior is needed to be seen as a leader (Eagly & Karau, 2002; Rudman & Phelan, 2008; Tyler & McCullough, 2009), though a recent study by Bauer (2016) suggests this penalty may not apply in politics. This suggests the following hypothesis:

Hypothesis 1A: voters will prefer gender-normative candidates to non-normative candidates (role congruity theory).

In contrast, political scientists have spent less time thinking about whether female candidates are penalized for deviating from norms (Bauer, 2016 is an exception); instead, they focus on how voters translate these norms into political stereotypes of male and female candidates. This focus has produced two different sets of content-oriented stereotypes: belief stereotypes, which suggest that female candidates are more liberal than male candidates (Huddy & Terkildsen, 1993a; Kahn, 1994; Koch, 2002), and trait stereotypes, which suggest that female candidates are more competent on issues like health care and education because women are felt to be more compassionate, sensitive, and so on, i.e., communal (Huddy and Terkildsen 1993a; Alexander and Andersen 1993). In so doing, political scientists have married the communal and warmth (feminine) aspects of the psychological theories above to their expected effects on perceptions of candidate policy stances and competencies, with some success in repeated experiments.

Nevertheless, this literature has been less successful at empirically tying these stereotypes to the vote, often obtaining contradictory results. Some researchers find that women perform better when they emphasize their feminine characteristics (consonant with role congruity theory); for instance, Herrnson, Lay, and Stokes (2003) argue that women are more successful when they run on women's areas of issue ownership and target women's groups. Many other studies, however, find that highlighting one's femininity is detrimental to one's electoral prospects (e.g., Holman, Merolla, & Zechmeister, 2016): Kahn notes that "women candidates should continue their present strategy of stressing 'male' traits in their campaign appeals" (Kahn, 1996, p. 136)—stating that at least journalists are biased, even if voters may not be. In a similar vein, Huddy and Terkildsen (1993a) argue that female candidates should emphasize their masculine traits, i.e., act gender non-normative, in order to succeed with voters. The latter findings echo a recent meta-analysis by Koenig et al.

(2011), which examined three different paradigms of gender and stereotypes—"think male, think manager," "agency-communion," and "masculinity-femininity"—in 69 different studies and found much larger correlations between masculinity and leadership than femininity and leadership in all three paradigms. This suggests an alternative hypothesis:

> **Hypothesis 1B: voters will prefer masculine candidates to feminine candidates (leadership-masculine theory).**

Finally, research such as that by Lakoff (2010) and Winter (2010), described in the introduction, suggests that voter partisanship predicts preferences over leadership styles. Similar findings have been confirmed across a variety of candidate cues. For instance, Carpinella and Johnson (2013a, 2013b) also find that Democratic and Republican voters react differently to gendered facial cues: facial femininity increased perceptions of competence by Democrats relative to Republicans, even when no information about candidate partisanship was provided. Similarly, Hayes (2005) and others have found that the traits typically associated with Republicans (e.g., "strong") and Democrats (e.g., "compassionate") are also highly gendered. The complex interrelationship between partisanship and gender at a minimum suggests that heterogeneous treatment effects are likely. This generates the following, final hypothesis:

> **Hypothesis 2: Democratic voters will prefer feminine candidates; Republican voters will prefer masculine candidates.**

DATA AND DESIGN

I test these hypotheses in two survey experiments, conducted through YouGov in September and October 2016. The main treatment is a description of the candidate as having a particular leadership style, which is randomly assigned to be either typically masculine or typically feminine. In Study 1, Hillary Clinton was the only candidate described; in Study 2, nine national-level candidates— Hillary Clinton, Elizabeth Warren, Jill Stein, Barack Obama, Joe Biden, Bernie Sanders, Tim Kaine, Mike Pence, and Donald Trump—were varied. Analyses are all simple t-tests for significant differences (e.g., between experimental conditions, or between respondent parties).

Samples for both studies were recruited through YouGov to be a sample of 1800 registered California voters (3600 total) as part of a series of collaborative Field Poll-Institute for Governmental Studies surveys leading up to the November election. Respondents were recruited to be representative of the overall California population of registered voters on partisanship, gender, age, ethnicity, and education. Sample characteristics are set out in Table 3.1 below.

Respondents took an approximately 15-minute online survey that focused heavily on California politics and policy, particularly related to the November ballot initiatives. The questions for this study were interspersed with these other questions throughout. In Study 1, respondents received one of the questions in Figure 3.1.

The treatments were designed to evoke either a communitarian (feminine) leadership style, or agentic (masculine) leadership style. In Study 1, a separate set of respondents were asked whether they thought this leadership style was more common among men or among women to confirm that the treatment did indeed evoke perceptions of gender-typical leadership styles. 600 respondents saw the feminine condition, 600 saw the masculine condition, and 600 received the content check (more common among men or women) question.

Study 2 added eight additional candidates and changed the wording of the leadership style question slightly to evaluate whether the results from the first study might be due to, for instance, a particularly positive or negative treatment wording for one of the conditions. Like in Study 1,

respondents saw only one treatment condition (one candidate, described in either masculine or feminine terms, as set out in Figure 3.2) during another 15-minute survey focused on California politics and policy.

Study 2 also added a direct manipulation check several questions later, which asked the respondent whether they thought the same candidate (Elizabeth Warren in this example) was more likely to make a tough decision on her own (an agentic style) or after listening to others (a communitarian style). Finally, Study 2 also had a content check question to ensure that the variation on the survey wording questions again manipulated perceptions of gender-typical leadership styles. Approximately 180 people evaluated each candidate (90 in each style condition) and answered the manipulation check about each candidate, and all 1800 respondents answered the content check several questions after that.

Table 3.1. Representativeness of Sample Characteristics

	Present Sample	**2016 Field Poll (weighted)[1]**	**2016 CCES (unweighted)[2]**	**2016 ACS[3]**	**2016 ACS[4]**
Region	*California*	*California*	*US*	*California*	*US*
Population	*Reg. voters*	*Reg. voters*	*Reg. voters*	*All adults*	*All adults*
% Women	53.63%	53%	52.95%	50.3%	50.8%
% Democrats	46.75%	48%	40.40%	--	--
% White	56.61%	56%	72.76%	61.3%	76.0%
Education (25+)					
Less than HS	1.75%	5%	2.10%	17.9%	13.0%
HS/equivalency	15.88%	15%	23.21%	20.6%	27.5%
Some college/AA	39.59%	33%	36.04%	29.5%	29.1%
Bachelor's	26.79%	21%	24.50%	20.1%	18.8%
Grad./prof. degree	16.00%	21%	14.14%	11.9%	11.5%
Age (18+)					
18-24	7.80%	10%	6.86%	13.4%	12.7%
25-44	29.08%	30%	34.12%	36.8%	34.3%
45-64	40.55%	38%	39.55%	32.9%	34.1%
65+	22.55%	22%	19.46%	16.9%	18.9%

[1] Obtained at https://web.archive.org/web/20170131002627/http://www.field.com/fieldpollonline/subscribers/Rls2545.pdf.
[2] Calculated by author, data obtained from http://www.people-press.org/2016/09/13/2016-party-identification-detailed-tables/.
[3] Obtained at https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S0501&prodType=table.
[4] Obtained at https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?pid=ACS_16_5YR_S0501&prodType=table.

Figure 3.1. Study 1 Treatment Wording

26a. Hillary Clinton is said to have a distinctive leadership style. She believes that listening to and working with others is more important than taking a stand and sticking with it.

Do you think that sort of leadership style is what America needs right now?

1 ○ Strongly agree
2 ○ Somewhat agree
3 ○ Neither agree nor disagree
4 ○ Somewhat disagree
5 ○ Strongly disagree

26b. Hillary Clinton is said to have a distinctive leadership style. She believes that taking a stand and sticking with it is more important than listening to and working with others.

Do you think that sort of leadership style is what America needs right now?

1 ○ Strongly agree
2 ○ Somewhat agree
3 ○ Neither agree nor disagree
4 ○ Somewhat disagree
5 ○ Strongly disagree

Figure 3.2. Study 2 Treatment Wording

22a(3). Democratic Senator Elizabeth Warren is said to have a distinctive leadership style. She believes that being compassionate and working with others is more important than being assertive and aggressively pursuing goals.

Do you think that sort of leadership style is what America needs right now?

1 ○ Strongly agree
2 ○ Somewhat agree
3 ○ Neither agree nor disagree
4 ○ Somewhat disagree
5 ○ Strongly disagree

22a(4). Democratic Senator Elizabeth Warren is said to have a distinctive leadership style. She believes that being assertive and aggressively pursuing goals is more important than being compassionate and working with others.

Do you think that sort of leadership style is what America needs right now?

1 ○ Strongly agree
2 ○ Somewhat agree
3 ○ Neither agree nor disagree
4 ○ Somewhat disagree
5 ○ Strongly disagree

RESULTS

Per Figure 3.3, both studies find that, in the lead-up to the November 2016 election, voters preferred candidates when they were described with a feminine leadership style (B=.117, p<.001), regardless of the candidate's sex. As such, we reject **Hypothesis 1B**, which predicted that voters would on average prefer masculine candidates (see Figure 3.6. in the Supplemental Materials (SM) for breakdowns for individual candidates). This effect may be specific to the context of the 2016 election, which featured both the first-ever female candidate in the general presidential election and a strong wave of anti-traditional candidate sentiment. As such, voters (and particularly registered voters in California) may have been particularly likely to feel that a feminine leadership style represented a symbolic good—the imminent shattering of a longstanding glass ceiling—or a refreshingly new (and therefore anti-traditional) way of conducting politics.

Nevertheless, the evidence for **Hypothesis 1A** (role congruity theory), shown in Figure 3.4, is weaker than the rejection of the leadership-masculine hypothesis might suggest. While respondents did prefer female candidates (B=.129, p<.001) when they were described as having a feminine leadership style (confirming to prescriptive gender stereotypes), respondents also preferred male candidates who were described as feminine (B=.098, p<.001), which seems to fit neither the predictions of role-congruity theory or leadership-masculine theory.

The results for Hypothesis 2 explain the odd results for Hypotheses 1A and 1B: per Figure 3.5, Democrats preferred candidates with a feminine style regardless of the partisanship of the candidate (B=.179, p<.001). However, Republicans did not favor masculine styles as predicted (B=.006, p=.83), regardless of whether the candidates were their co-partisans (see Table 3.3. in the SM for a breakdown by both respondent and candidate partisanship). Independents tended to fall between Democrats and Republicans, exhibiting a moderate preference for feminine conditions (B=.095, p<.001). This heterogeneity means that a sample of progressive-leaning voters (California registered voters) shows an aggregate preference for feminine leadership.

The effects estimated are substantively meaningful, given that respondents are likely to have reasonably solid preexisting attitudes toward these well-known national candidates, especially at the height (September and October) of a presidential election year. The Cohen's *d* for the overall preferences is .325, indicating a small-to-moderate size effect.

In both studies, the content check was successful: respondents perceived the feminine leadership styles to be more typical of women, and masculine styles to be more typical among men. The size of the difference was a little less than one point on the five-point scale (B=.825, p<.001 for September, and B=.961, p<.001 for October). The effects are slightly larger for female respondents than male respondents; this may be evidence of some slight social desirability bias wherein male respondents are reluctant to suggest that women are not typically assertive or independent leaders, though the difference in means is still significant after controlling for gender.

Finally, the manipulation check in Study 2 found that, for eight of the nine candidates, respondents who saw a feminine condition were more likely to believe that the candidate would make tough decisions by listening to others, confirming that they saw the candidates as more communion-oriented than agentic (overall t-test results show masculine condition as significantly more agentic, B=.330, p<.001). However, the manipulation check failed for one candidate, Donald Trump. Given the timing of Study 2, this may be unsurprising: Trump was in the news for an audiotape in which he made lewd comments about women, and several news outlets had recently featured articles suggesting that he was not prone to taking advice from his campaign staff. Considering this, it seems plausible to imagine that many voters resisted the feminine condition

Figure 3.3.

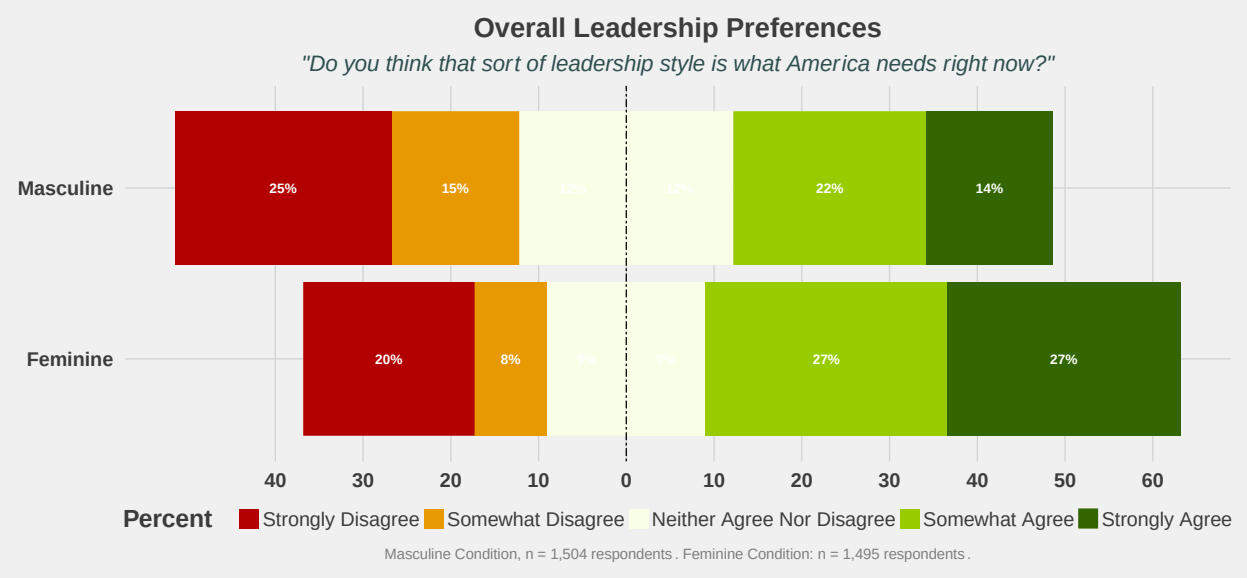


Figure 3.4.

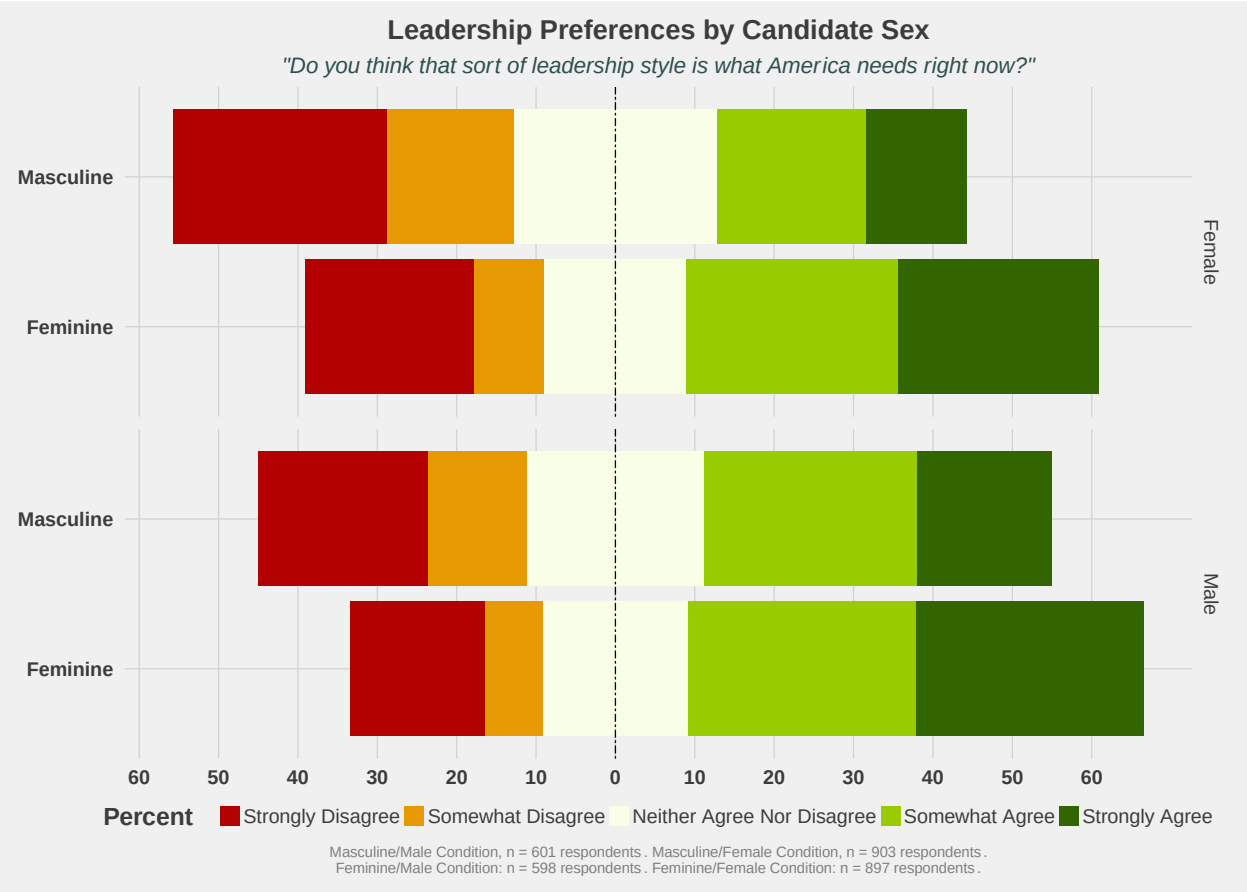
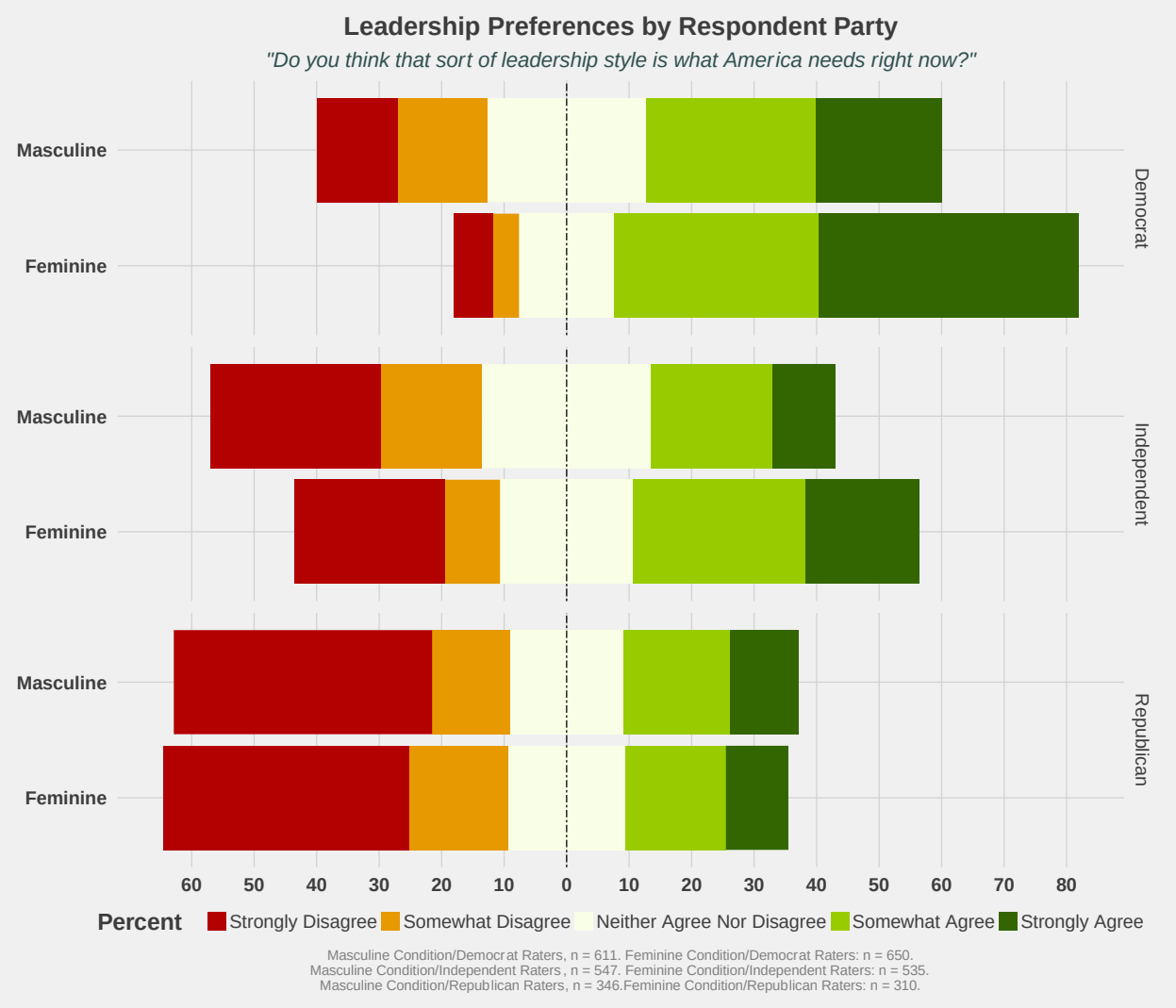


Figure 3.5.



Trump treatment. The overall findings hold whether respondents who viewed the feminine style Trump condition are included or excluded.

DISCUSSION

If there is a "right way" to present oneself as a woman in politics, it is by playing up the a feminine leadership style to a Democratic base, and perhaps a neutral style for a Republican base. While the literature on gender in politics has long known that partisanship plays a bigger role in candidate evaluations than the sex of the candidates, the importance of partisanship does not mean the sex of the candidate is irrelevant. If women tend on average to have a more feminine style—and the content check suggests they may—they are likely be evaluated more positively by Democratic voters than by Republican voters. While the current study does not measure the extent to which evaluations of leadership style influence voters' final vote decisions, this finding is suggestive in light of the fact that female Republican candidates have struggled to make electoral inroads the way female Democratic candidates have. Though many factors surely influence this gap, it may be worth

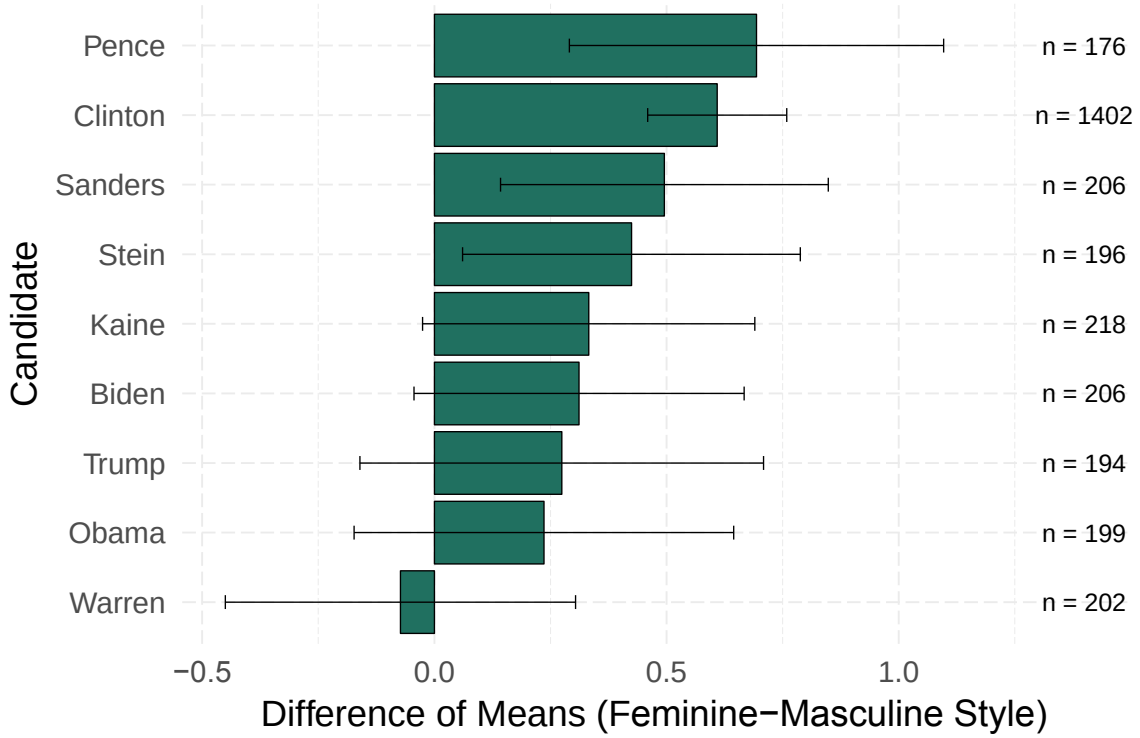
exploring in more detail the extent to which Republican women feel "cross-pressured" to present with a neutral or masculine leadership style. Such a study might also find that Democratic women (and men) feel pressure to exaggerate a feminine style to appear consistent with their base's preferences.

The study also suggests possibilities for future research that uses campaign advertising (mailers, TV ads, etc.) to manipulate perceptions of candidates' personas and by extension to influence real voter behavior, not just attitudes on a survey. While in this study we cannot say for sure that preferences for certain types of leadership styles *affect* vote choice, future research might fruitfully explore (for instance, via a field experiment) whether persistent framing of a candidate's leadership style as masculine or feminine could affect voting behavior.

Finally, it may be worth attempting interventions with female candidates or would-be candidates to influence their political ambition. Research by Bernhard, Shames, & Teele (n.d.) suggests that graduates of female candidate training programs believe they need to adopt more masculine traits to be successful during campaigns than they need to be successful once in office, and a broad literature on women's political ambition suggests that many women feel they are not a good "fit" for politics in terms of their personality. If some of these decreases in ambition stem from women's perceptions that voters will penalize them for a more feminine leadership style, sharing information about positive evaluations of candidates with a feminine leadership style may increase ambition. For instance, in a recent experiment, Holman and Schneider (2016) find that women's political ambitions are influenced by their perception of the gender gap as stemming from demand-side (e.g., voter) discrimination rather than supply-side (e.g., women's low ambition) factors. While studies tend to study either demand-side or supply-side factors in the gender gap, more work is needed to understand the mutually reinforcing nature of the two, as well as how that reinforcement may work differently for Democratic and Republican women, given the massive influence of partisanship in American elections.

Supplemental Materials

Figure 3.6. Leadership Style Evaluations: Results by Individual Candidate



For no candidate did raters express a significantly positive preference for the masculine leadership style (95% confidence intervals are shown above). In four of the nine cases, including for Hillary Clinton (which had the largest sample size, as it was evaluated in both Study 1 and Study 2), respondents expressed a significant ($p<.05$) preference for the feminine leadership style, and eight of the nine candidates received more positive evaluations in the feminine leadership style condition.

Table 3.2. Partisan Preferences When Evaluating In-Party versus Out-Party

Respondent Party	Candidate Co-Partisan?	Difference in Means	Mean (Masculine)	Mean (Feminine)	T-Statistic	P-Value
Democrat	No	-0.570	2.228	2.798	-2.545	0.012
Democrat	Yes	-0.751	3.429	4.180	-11.457	0.000
Independent	(Rep.)	-0.676	2.671	3.347	-2.720	0.008
Independent	(Dem.)	-0.351	2.690	3.041	-3.961	0.000
Republican	No	-0.002	3.921	3.923	-0.008	0.994
Republican	Yes	0.057	2.256	2.199	0.514	0.607

Democratic respondents favored the feminine leadership style in both in- and out-party candidates (for the purposes of this analysis, Jill Stein is treated as a co-partisan for Democrats). Independents did prefer the feminine style in both Republican and Democratic candidates as well. Republicans exhibited no preference between feminine and masculine leadership style, either with co-partisans or out-party candidates.

Bibliography

Abramowitz, A. (2012). Forecasting in a polarized era: The time for change model and the 2012 presidential election. *PS: Political Science & Politics 45*(4), 618-19.

Ahler, D. J., Citrin, J., Dougal, M. C., & Lenz, G. S. (2016). Face value? Experimental evidence that candidate appearance influences electoral choice. *Political Behavior*, *39*(1), 77–102.

Alexander, D., & Andersen, K. (1993). Gender as a factor in the attribution of leadership traits. *Political Research Quarterly*, *46*(3), 527. https://doi.org/10.2307/448946

Anzia, S. F. Forthcoming. When does a group of citizens influence policy? Evidence from senior citizen participation in city politics. *Journal of Politics*.

Anzia, S. F. (2014). *Timing and turnout: How off-cycle elections favor organized groups*. Chicago: University of Chicago Press.

Anzia, S. F., & Berry, C. R. (2011). The Jackie (and Jill) Robinson effect: Why do congresswomen outperform congressmen? *American Journal of Political Science*, *55*(3), 478–493.

Atkinson, M. D., Enos, R. D., & Hill, S. J. (2009). Candidate faces and election outcomes: Is the face-vote correlation caused by candidate selection. *Quarterly Journal of Political Science*, *4*(3), 229–249.

Banducci, S. A., Karp, J. A., Thrasher, M., & Rallings, C. (2008). Ballot photographs as cues in low-information elections. *Political Psychology*, *29*(6), 903–917.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, *71(2)*, 230.

Bauer, N. M. (2016). The effects of counterstereotypic gender strategies on candidate evaluations. *Political Psychology*, *38*(2), 279–295. https://doi.org/10.1111/pops.12351

Bauer, N. M. (2015). Who stereotypes female candidates? Identifying individual differences in feminine stereotype reliance. *Politics, Groups, and Identities 3*(1), 94–110.

Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review 88*(4), 354–364. http://dx.doi.org/10.1037/0033-295X.88.4.354

Bernhard, R., Shames, S. L., & Teele, D. L. (n.d.). To emerge? Breadwinning and income in women's decisions to run. Presented at the Midwest Political Science Association Annual Meeting, Chicago, April 6, 2017.

Berry, C. R., & Gersen, J. (2011). The implications of election timing for public policy. *Quarterly Journal of Political Science 6*(2), 103-35.

Berry, C. R., & Gersen, J. (2010). The timing of elections. *The University of Chicago Law Review 77*(Winter), 37-64.

Brooks, D. J. (2013). *He runs, she runs: Why gender stereotypes do not harm women candidates.* Princeton University Press.

Brooks, D. J. (2011). Testing the double standard for candidate emotionality: Voter reactions to the tears and anger of male and female politicians. *The Journal of Politics*, *73*(02), 597–615. https://doi.org/10.1017/S0022381611000053

Burrell, B. C. (1994). *A woman's place is in the house: campaigning for congress in the feminist era*. Ann Arbor, MI: University of Michigan Press.

Buss, D. M. (1989). Sex differences in human mate preferences: Evolutionary hypotheses tested in 37 cultures. *Behavioral and Brain Sciences*, *12*(1), 1–14.

Buss, D. M., & Barnes, M. (1986). Preferences in human mate selection. *Journal of Personality and Social Psychology*, *50*(3), 559–570. https://doi.org/10.1037/0022-3514.50.3.559

Buss, D. M., & Schmitt, D. P. (1993). Sexual strategies theory: An evolutionary perspective on human mating. *Psychological Review*, *100*(2), 204–232. https://doi.org/10.1037/0033-295X.100.2.204

Carpinella, C. M., Hehman, E., Freeman, J. B., & Johnson, K. L. (2015). The gendered face of partisan politics: Consequences of facial sex typicality for vote choice. *Political Communication*, *33*(1), 21–38. https://doi.org/10.1080/10584609.2014.958260

Carpinella, C. M., & Johnson, K. L. (2013a). Appearance-based politics: Sex-typed facial cues communicate political party affiliation. *Journal of Experimental Social Psychology*, *49*(1), 156–160. https://doi.org/10.1016/j.jesp.2012.08.009

Carpinella, C. M., & Johnson, K. L. (2013b). Politics of the face: The role of sex-typicality in trait assessments of politicians. *Social Cognition*, *31*(6), 770–779.

Carroll, S. J., & Sanbonmatsu, K. (2013). *More women can run: Gender and pathways to the state legislatures.* New York: Oxford University Press.

Center for American Women and Politics (CAWP). (2016). *Current numbers of women officeholders.* Eagleton Institute of Politics. Retrieved from http://www.cawp.rutgers.edu/current-numbers

Clayton, A., O'Brien, D., & Piscopo, J. M. (n.d.). All male panels? Representation and democratic legitimacy. Presented at the Midwest Political Science Association Annual Meeting, Chicago, April 6, 2017.

Conover, P. J., & Feldman, S. (1989). Candidate perception in an ambiguous world: Campaigns, cues, and inference processes. *American Journal of Political Science*, *33*(4), 912–940. https://doi.org/10.2307/2111115

Conroy-Beam, D., & Buss, D. M. (2016). Do mate preferences influence actual mating decisions? Evidence from computer simulations and three studies of mated couples. *Journal of Personality and Social Psychology*, *111*(1), 53–66. https://doi.org/10.1037/pspi0000054

Crowder-Meyer, M., Gadarian, S. K., & Trounstine, J. (2015). Electoral institutions, gender stereotypes, and women's local representation. *Politics, Groups, and Identities 3*(2), 318–334.

Darcy, R., & Choike, J. R. (1986). A formal analysis of legislative turnover: Women candidates and legislative representation. *American Journal of Political Science 30*(1), 237-55.

Darcy, R., Welch, S., & Clark, J. (1994). *Women, elections, and representation* (2nd ed.). Lincoln: University of Nebraska Press.

Darwin, C. (1888). *The descent of man, and selection in relation to sex* (Vol. 1). Murray. Retrieved from https://books.google.com/books?hl=en&lr=&id=NaPu24dY4iAC&oi=fnd&pg=PA5&dq=darwin+sexual+selection&ots=AFFhzZi3zS&sig=j6z7YXZW_k8OSHjadW_ZxUUDYLY

Deason, G., Greenlee, J. S., & Langner, C. A. (2015). Mothers on the campaign trail: Implications of politicized motherhood for women in politics. *Politics, Groups, and Identities*, *3*(1), 133–148. https://doi.org/10.1080/21565503.2014.992792

Ditonto, T. M. (2016). A high bar or a double standard? Gender, competence, and information in political campaigns. *Political Behavior 39*(2), 301-325. https://doi.org/10.1007/s11109-016-9357-5

Ditonto, T. M., Hamilton, A. J., & Redlawsk, D. P. (2014). Gender stereotypes, information search, and voting behavior in political campaigns. *Political Behavior*, *36*(2), 335–358.

Dittmar, K. (2015). *Navigating gendered terrain: Stereotypes and strategy in political campaigns*. Temple University Press.

Dolan, K. (2010). The impact of gender stereotyped evaluations on support for women candidates. *Political Behavior 32*(1), 69–88.

Duerst-Lahti, G. (1998). *The bottleneck: Women becoming candidates*. In S. Thomas & C. Wilcox (Eds.), *Women and Elective Office* (pp. 15-25). New York: Oxford University Press.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, *109*(3), 573–598. https://doi.org/10.1037//0033-295X.109.3.573

Eagly, A. H., & Wood, W. (2013). The nature–nurture debates: 25 years of challenges in understanding the psychology of gender. *Perspectives on Psychological Science*, *8*(3), 340–357.

Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences*, *11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology 82*(6), 878–902. https://doi.org/http://dx.doi.org/10.1037/0022-3514.82.6.878

Fox, R. L. (2006). Congressional elections: Where are we on the road to gender parity? In S. J. Carroll & R. L. Fox (Eds.), *Gender and elections* (pp. 97–116). New York, NY: Cambridge University Press.

Fulton, S. A. (2012). Running backwards and in high heels: The gendered quality gap and incumbent electoral success. *Political Research Quarterly 65*(2), 303–314.

Hajnal, Z. L., Lewis, P. G., & Louch, H. (2002). *Municipal elections in California: Turnout, timing, and competition*. Public Policy Institute of California.

Hajnal, Z., & Trounstine, J. (2005). Where turnout matters: The consequences of uneven turnout in city politics. *Journal of Politics 67*(2), 515-35.

Hayes, D. (2011). When Gender and Party Collide: Stereotyping in Candidate Trait Attribution. *Politics & Gender 7*(2), 133–65.

Hayes, D. (2005). Candidate qualities through a partisan lens: A theory of trait ownership. *American Journal of Political Science*, *49*(4), 908. https://doi.org/10.2307/3647705

Hayes, D., & Lawless, J. L. (2016). *Women on the run: Gender, media, and political campaigns in a polarized era*. New York: Cambridge University Press.

Hayes, D., Lawless, J. L., & Baitinger, G. (2014). Who cares what they wear? Media, gender, and the influence of candidate appearance. *Social Science Quarterly 95(*5), 1194–1212.

Hehman, E., Carpinella, C. M., Johnson, K. L., Leitner, J. B., & Freeman, J. B. (2014). Early processing of gendered facial cues predicts the electoral success of female politicians. *Social Psychological and Personality Science*, *5*(7), 815–824. https://doi.org/10.1177/1948550614534701

Herrnson, P. S., Lay, J. C., & Stokes, A. K. (2003). Women running "as women": Candidate gender, campaign issues, and voter-targeting strategies. *The Journal of Politics*, *65*(1), 244–255.

Hochschild, A., & Machung, A. (2012). *The second shift: Working families and the revolution at home*. Penguin. Retrieved from https://books.google.com/books?hl=en&lr=&id=St_6kWcPJS8C&oi=fnd&pg=PT25&dq=arlie+hochschild&ots=8F5Xo-yYkg&sig=wDSoWAqhR0uErgtXVhR8_7zqbwg

Holman, M. R., Merolla, J. L., & Zechmeister, E. J. (2016). Terrorist threat, male stereotypes, and candidate evaluations. *Political Research Quarterly*, *69*(1), 134–147. https://doi.org/10.1177/1065912915624018

Holman, M. R., & Schneider, M. C. (2016). Gender, race, and political ambition: how intersectionality and frames influence interest in political office. *Politics, Groups, and Identities*, *6*(2), 264–280. https://doi.org/10.1080/21565503.2016.1208105

Huddy, L., & Terkildsen, N. (1993a). Gender stereotypes and the perception of male and female candidates. *American Journal of Political Science*, *37*(1), 119. https://doi.org/10.2307/2111526

Huddy, L., & Terkildsen, N. (1993b). The consequences of gender stereotypes for women candidates at different levels and types of office. *Political Research Quarterly*, *46*(3), 503. https://doi.org/10.2307/448945

Jamieson, K. H. (1995). *Beyond the double bind: Women and leadership*. Oxford University Press on Demand. Retrieved from https://books.google.com/books?hl=en&lr=&id=67DB9krBq2oC&oi=fnd&pg=PA3&dq=jamieson+double+bind&ots=D49CtflLXf&sig=4trcE9azzIF5WuA_ilq5uYbNJg0

Johnson, S. K., Murphy, S. E., Zewdie, S., & Reichard, R. J. (2008). The strong, sensitive type: Effects of gender stereotypes and leadership prototypes on the evaluation of male and female leaders. *Organizational Behavior and Human Decision Processes*, *106*(1), 39–60. https://doi.org/10.1016/j.obhdp.2007.12.002

Kahn, K. F. (1996). *The political consequences of being a woman: How stereotypes influence the conduct and consequences of political campaigns*. Columbia University Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=wb01N-2jEf4C&oi=fnd&pg=PR9&dq=kahn+woman+media&ots=nAbJhmjlx7&sig=Myl0tiPFvfPZJM0y11nxj7zw2w0

Kahn, K. F. (1994). The distorted mirror: Press coverage of women candidates for statewide office. *The Journal of Politics*, *56*(1), 154–173. https://doi.org/10.2307/2132350

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan. Retrieved from http://books.google.com/books?hl=en&lr=&id=SHvzzuCnuv8C&oi=fnd&pg=PP2&dq=kahneman+thinking+fast+slow&ots=NRplQD3oJC&sig=PoPK4LmzjBDpCXzWD195SkVTlsw

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and Biases* (pp. 49–81). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511808098.004

King, D. C., & Matland, R. E. (2003). Sex and the grand old party: An experimental investigation of the effect of candidate sex on support for a republican candidate. *American Politics Research 31*(6), 595–612.

Kitchens, K. E., & Swers, M. L. (2016). Why aren't there more republican women in congress? gender, partisanship, and fundraising support in the 2010 and 2012 elections. *Politics & Gender 12*(4), 648–76.

Koch, J. W. (2002). Gender stereotypes and citizens' impressions of house candidates' ideological orientations. *American Journal of Political Science*, *46*(2), 453. https://doi.org/10.2307/3088388

Koch, J., & Thomsen, D. M. (2017). Gender equality mood across states and over time. *State Politics & Policy Quarterly 17*(4), 351-60.

Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin 137*(4), 616–642. https://doi.org/10.1037/a0023557

Kogan, V., Lavertu, S., & Peskowitz, Z. Forthcoming. Election timing, electorate composition, and policy outcomes: Evidence from school districts. *American Journal of Political Science.*

Kunda, Z., & Spencer, S. J. (2003). When do stereotypes come to mind and when do they color judgment? A goal-based theoretical framework for stereotype activation and application. *Psychological Bulletin 129*(4), 522.

Lakoff, G. (2010). *Moral politics: How liberals and conservatives think*. University of Chicago Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=MTKR3o2PICMC&oi=fnd&pg=PR7&dq=lakoff+moral+politics&ots=SAHaEJ-u34&sig=MhRawj1ji4aO8r-b-gcUAUnN3OI

Lau, R. R., & Redlawsk, D. P. (2001). An experimental study of information search, memory, and decision making during a political campaign. In Kuklinski, J. H. (Ed.), *Citizens and Politics: Perspectives from Political Psychology* (pp. 136–59). Cambridge: Cambridge University Press.

Lawless, J. L. (2015). Female candidates and legislators. *Annual Review of Political Science*, *18*(1), 349–366. https://doi.org/10.1146/annurev-polisci-020614-094613

Lawless, J. L. (2004). Women, war, and winning elections: Gender stereotyping in the post-September 11th era. *Political Research Quarterly 57*(3), 479–490.

Lawless, J. L., & Pearson, K. (2008). The primary reason for women's underrepresentation? Reevaluating the conventional wisdom. *The Journal of Politics 70*(1), 67–82.

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

Lenz, G. S., & Lawson, C. (2011). Looking the part: Television leads less informed citizens to vote based on candidates' appearance. *American Journal of Political Science*, *55*(3), 574–589. https://doi.org/10.1111/j.1540-5907.2011.00511.x

Lippmann, W. (1922). *Public opinion*. Transaction Publishers. Retrieved from https://books.google.com/books?id=YhXLOVc6BsoC&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior*, *28*(1), 18–27. https://doi.org/10.1016/j.evolhumbehav.2006.09.002

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: Evolutionary based research. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *366*(1571), 1638–1659.

Lupia, A. (1994). Shortcuts versus encyclopedias: Information and voting behavior in California insurance reform elections. *American Political Science Review 88*(1), 63–76.

Matson, M., & Fine, T. S. (2006). Gender, ethnicity, and ballot information: Ballot cues in low-information elections. *State Politics & Policy Quarterly*, *6*(1), 49–72.

Mayhew, D. R. (2001). *Divided we govern: Party control, lawmaking, and investigations, 1946-1990*. Yale University Press.

McDermott, M. L. (2005). Candidate occupations and voter information shortcuts. *Journal of Politics 67*(1), 201–219.

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), 270-83.

Meredith, M. (2013). Exploiting friends-and-neighbors to estimate coattail effects. *American Political Science Review 107*(4), 742–765.

Meredith, M. (2009). The strategic timing of direct democracy. *Economics and Politics 21*(1), 159-177.

Milyo, J., & Schosberg, S. (2000). Gender bias and selection bias in house elections. *Public Choice 105*(1–2), 41–59.

Mo, C. H. (2015). The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters. *Political Behavior*, *37*(2), 357–395. https://doi.org/10.1007/s11109-014-9274-4

Newman, J. (1994). *Perception and reality: A study comparing the success of men and women candidates*. Washington, DC: National Women's Political Caucus.

Oliver, J. E., & Ha, S. E. (2007). Vote choice in suburban elections. *American Political Science Review 101*(3), 393-408.

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, *34*(2), 83–110. https://doi.org/10.1007/s10919-009-0082-1

Palmer, B., & Simon, D. (2010). *Breaking the political glass ceiling: Women and congressional elections* (2nd edition). New York: Routledge.

Pitkin, H. F. (1967). *The concept of representation*. Univ of California Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=AgUVWLswTNEC&oi=fnd&pg=PA1&dq=pitkin+representation&ots=15xOkI0lq7&sig=C9xKqvLtc3QAu9f33GrzZjWAsNo

Popkin, S. L. (1991). *The reasoning voter: Communication and persuasion in presidential campaigns*. Chicago: University of Chicago Pres.

Poutvaara, P., Jordahl, H., & Berggren, N. (2009). Faces of politicians: Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology*, *45*(5), 1132–1135.

Redlawsk, D. P. (2004). What voters do: Information search during election campaigns. *Political Psychology*, *25*(4), 595–610.

Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, *28*, 61–79. https://doi.org/10.1016/j.riob.2008.04.003

Sanbonmatsu, K. (2002). Gender stereotypes and vote choice. *American Journal of Political Science 46*(1), 20-34.

Schmitt, D. P. (2014). On the proper functions of human mate preference adaptations: Comment on Eastwick, Luchies, Finkel, and Hunt (2014). *Psychological Bulletin*, *140*(3), 666–672. https://doi.org/10.1037/a0036225

Seltzer, R. A., Newman, J., & Leighton, M. V. (1997). *Sex as a political variable*. Boulder, CO: Lynne Reinner.

Spezio, M. L., Loesch, L., Gosselin, F., Mattes, K., & Alvarez, R. M. (2012). Thin-slice decisions do not need faces to be predictive of election outcomes. *Political Psychology*, *33*(3), 331–341. https://doi.org/10.1111/j.1467-9221.2012.00897.x

Swers, M. L. (2002). *The difference women make: The policy impact of women in Congress*. University of Chicago Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=gGnsoel8ZFkC&oi=fnd&pg=PR9&dq=swers+female+candidate&ots=nlINOBQ62U&sig=jBe2YOvXItAFrUPzde9TpVqXadc

Swers, M. L. (1998). Are women more likely to vote for women's issue bills than their male colleagues? *Legislative Studies Quarterly*, *23*(3), 435–448.

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton University Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=mH8rDgAAQBAJ&oi=fnd&pg=PP1&dq=todorov+face+value&ots=yWnQuc5VcV&sig=Vht-jk7enIn8mpRjycHPk7vinIo

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

Tufte, E. R. (1975). Determinants of the outcomes of midterm congressional elections. *American Political Science Review 69*(3), 812-826.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

Tyler, J. M., & McCullough, J. D. (2009). Violating prescriptive stereotypes on job resumes: A self-presentational perspective. *Management Communication Quarterly*, *23*(2), 272–287. https://doi.org/10.1177/0893318909341412

Velde, E. R., & Pearson, P. L. (2002). The variability of female reproductive ageing. *Human Reproduction Update*, *8*(2), 141–154. https://doi.org/10.1093/humupd/8.2.141

Winter, N. J. G. (2010). Masculine republicans and feminine democrats: Gender and Americans' explicit and implicit images of the political parties. *Political Behavior*, *32*(4), 587–618. https://doi.org/10.1007/s11109-010-9131-z

Wood, W., & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. In J. M. Olson & M. P. Zanna (Eds.), *Advances in Experimental Social Psychology* (Vol. 46, pp. 55–123). Elsevier. https://doi.org/10.1016/B978-0-12-394281-4.00002-7

Zaller, J. (1992). *The nature and origins of mass opinion*. New York: Cambridge University Press.

Zentner, M., & Eagly, A. H. (2015). A sociocultural framework for understanding partner preferences of women and men: Integration of concepts and evidence. *European Review of Social Psychology*, *26*(1), 328–373. https://doi.org/10.1080/10463283.2015.1111599

Zentner, M., & Mitura, K. (2012). Stepping out of the caveman's shadow: Nations' gender gap predicts degree of sex differentiation in mate preferences. *Psychological Science*, *23*(10), 1176–1185. https://doi.org/10.1177/0956797612441004