

GENDER AND POLITICAL REPRESENTATION

A Dissertation

by

KOSTANCA DHIMA

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Michelle Taylor-Robinson
Co-Chair of Committee,	Matt Golder
Committee Members,	William Roberts Clark Jessica Gottlieb Diana Z. O'Brien
Head of Department,	William Roberts Clark

May 2022

Major Subject: Political Science

Copyright 2022 Kostanca Dhima

ABSTRACT

In my dissertation, I present three papers that evaluate the causes of women’s political underrepresentation through a supply, demand, and institutions framework. First, I focus on elite demand. Drawing on theories of gender bias, group attachment, and partisan identity, I conduct a field experiment in Canada to examine whether political elites exhibit gender discrimination when responding to political aspirants. The results indicate that legislators are more responsive to female aspirants and more likely to provide them with helpful advice when they ask about how to get involved in politics. This pro-women bias, which exists at all levels of government, is stronger among female legislators and those associated with left-leaning parties. Next, I focus on mass demand. Drawing on theories of gender bias, gender stereotypes, and role congruity, I conduct a choice experiment in South Korea to examine how candidate sex and gender expression shape voter preferences. I find that voters, on average, prefer female candidates. Despite this pro-woman bias, however, voters don’t think that women will win the election. These results suggest that we shouldn’t necessarily infer voter behavior simply from voter preferences. When it comes to how voters evaluate candidates who deviate from gender norms, I find that voters tend to prefer candidates who run counter to gender stereotypes: they prefer women candidates who present a “tough” approach to politics and men candidates who present a “compassionate” approach. The third paper takes a more aggregate-level approach and looks at how supply-side and demand-side factors interact to affect women’s representation while controlling for institutional context. Existing empirical studies treat supply-side and demand-side factors separately and ignore the inherent interaction at the theoretical core of the supply and demand framework. However, women’s descriptive representation should only be high when supply and demand are both sufficiently high. I test the implications of my theory using a new global dataset on women’s representation from 1990 to 2018. The results are consistent with my theory and are substantively important because

they indicate the conditions under which we can expect supply-side and demand-side factors to actually translate into greater female political representation.

DEDICATION

To my family. Faleminderit për gjithëçka. Ju dua me gjithë zemër.

ACKNOWLEDGMENTS

I feel incredibly blessed and lucky to have had the opportunity to learn from and get to know many amazing people throughout my graduate studies. I'm not sure I'll ever find the words to express my gratitude. I can only commit to trying my best to share what I've learned from them with others.

First and foremost, I am indebted to my co-chairs. Matt Golder has been the most dedicated advisor. The trajectory of my graduate career and life forever changed when I started to work with Matt. I am most grateful. Thank you for your countless hours spent teaching me how to think and write, for always pushing to bring the best out of me, and for believing in me whenever I've needed it most. I am in awe of and inspired by your generosity and dedication. I cannot thank you enough. A special shout out to my friend, Christopher Boylan, for encouraging me to work with Matt.

Likewise, Michelle Taylor-Robinson has been an exemplary advisor who's always been very supportive of me and my work. Thank you for always keeping me grounded, for carefully reading my work and providing me with timely and detailed feedback, and for always being so open and encouraging. Thanks, too, for being one of the nicest people I know and for elevating everyone around you to be their best selves.

I'm also most thankful to William R. Clark. I came to A&M largely because of Bill, and having had the opportunity to learn from him has been one of my most valuable experiences. Taking the writing course with Bill, in particular, has been one of the highlights of my graduate studies. I am in awe of his wisdom, depth and breadth of knowledge well beyond political science, and his ability to always make your work better. Thank you, Bill, for your precision and eloquence.

I'm grateful to Diana O'Brien for her help and support throughout these years. She has always read my work carefully, and asked smart and thoughtful questions that have encouraged me to think more deeply and carefully about my research. My work has improved tremendously from

your example, Diana. Thank you!

I also thank Jessica Gottlieb for serving on my dissertation committee. Thanks, Jessica, for your sharpness and kindness, and for always asking thoughtful and difficult questions. I always feel smarter and better off after having talked with you. I only hope that our paths continue to cross.

Sona Golder has also been instrumental to my success in graduate school. She is an inspiration to me and I thank her for all the time she has invested in me. She's read my work carefully, given me invaluable guidance on all aspects of my life, shown up at all my panels at conferences, and has always lifted me up. Thank you, Sona, for your ceaseless support, keen insights, and good humor.

I'm grateful to Sarah Fulton for welcoming me with open arms at A&M. Besides being incredibly supportive, Sarah has also created research opportunities for me, inviting me to collaborate on several papers. I think fondly of the time we've spent together.

I also owe a great deal to other faculty members at A&M – Zé Cheibub, Scott Cook, David Fortunato, Maria Escobar-Lemmon, Matthew Fuhrmann, Nehemia Geva, Paul Kellstedt, Mike Koch, Alex Pacek, Erik Peterson, Ahmer Tarar, Joe Ura, and Guy D. Whitten – whose courses and feedback shaped my intellectual curiosity in the discipline.

I've also been fortunate to work with amazing co-authors who've been valuable mentors and friends, including Ray Block, Charles Crabtree, Yaoyao Dai, Jennifer Piscopo, Laura Stephenson, and Karine Van der Straeten. My dissertation work has also benefited from the feedback I've received from Ben Ferland, Florian Foos, and Jonathan Homola.

I am also thankful for my A&M friends: Nathalie Mendez for always centering me, Samantha Zuhlke for your wise advice, John Niehaus for always lending an ear, Alejandro Medina for your sense of fairness often accompanied with coffee, Flavio Souza for being such a gracious host, Brenna Armstrong for the laughs, Weiwei Yin for tutoring us all in game theory, Thiago da Silva

Nascimento for your cheerfulness, Ben Peterson for our philosophical talks, Janica Magat for being my methods study buddy, Hwalmin Jin for your honesty and humor, Max Allamong for sharing your R skills, Thiago Talzzia for your joyous laugh, Andrea Junqueira for your sincerity, as well as Morgan Winkeler, Yewon Kwon, Yohan Park, Austin McCrea, Miyeon Song, Yoo Sun Jung, Austin Mitchell, Ali Kagalwala, Lauren Bondarenko, Andrew Philips, and Seung-Ho An, among others.

I'm grateful to my CP Group fam who've been there for me every step of the way: Charles Crabtree for always sharing your wealth, Yaoyao Dai for your good taste in all things, Molly Ariotti for the retail therapy adventures, and Boyoon Lee for our immediate bond. Thank you!

I am also indebted to my Knox College community: my BBK friends and the professors who inspired me to pursue graduate school, including Robert Seibert, Karen Kampwirth, Stephen Bailey, Mary Crawford, Lane Sunderland, Sue Hulett, Magali Roy-Féquièrè, and John Haslem. Thanks to Mësuese Thomaida, Mrs. Edwards, Mrs. Eblen, and Mike Neberz for guiding me before Knox.

Thank you to all my students who've motivated and inspired me to keep going during the pandemic and the job market season. Being your teacher is a privilege.

Above all else, I dedicate this dissertation to my family, without whom none of this would have been possible. To my parents, Sokrat and Zhuljeta, who are simply the best of the best. I am forever blessed to have them as my parents. Thank you for all your sacrifice, for always being there for me and holding me up. To Kristjan, zemra e motrës, thank you for being a light in my life and for always having my back. To Fiolla, thank you for your fearlessness and for always encouraging me. To Ricardo, thank you for being my biggest cheerleader. You've expanded my world and it's meant everything. To my grandparents – Mihal and Martha Dhima, and Irakli and Kleoniqe Çani – thank you for your love, perseverance, and foresight. We are here because of you. Ju dua shumë!

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professors Matt Golder (co-advisor) of the Department of Political Science at Pennsylvania State University, Michelle Taylor-Robinson (co-advisor), William R. Clark, and Diana Z. O'Brien of the Department of Political Science, and Professor Jessica Gottlieb of the Bush School of Government and Public Service.

The data presented in Chapter Two were provided by the Represent Civic Information API. Chapter Two was published in *Politics & Gender* in 2022. The data presented in Chapter Four were provided by the Inter-Parliamentary Union, the World Bank, the World Values Survey, the European Values Survey, and the Quota Adoption and Reform Over Time (QAROT).

All other work conducted for the dissertation was completed by the student independently.

Funding Sources

Graduate study was supported by a fellowship from from the College of Liberal Arts at Texas A&M University. The research presented in Chapter Three was supported by the Tokyo Foundation for Policy Research.

NOMENCLATURE

ACIE	Average Component Interaction Effects
AMCE	Average Marginal Component Effect
EVS	European Values Survey
NDP	New Democratic Party
OLS	Ordinary Least Squares
QAROT	Quota Adoption and Reform Over Time
WVS	World Values Survey

TABLE OF CONTENTS

	Page
ABSTRACT	ii
DEDICATION	iv
ACKNOWLEDGMENTS	v
CONTRIBUTORS AND FUNDING SOURCES	viii
NOMENCLATURE	ix
TABLE OF CONTENTS	x
LIST OF FIGURES	xiii
LIST OF TABLES.....	xv
1. INTRODUCTION.....	1
2. DO ELITES DISCRIMINATE AGAINST FEMALE POLITICAL ASPIRANTS? EVIDENCE FROM A FIELD EXPERIMENT ¹	5
2.1 Should Gender Affect Elite Responsiveness?.....	9
2.2 Experimental Research Design	16
2.3 Results and Discussion.....	20
2.4 Conclusion.....	28
3. DO VOTERS DISCRIMINATE AGAINST FEMALE CANDIDATES? EVIDENCE FROM AN EXPERIMENT IN SOUTH KOREA.....	32
3.1 Voter Preferences for Women Candidates	35
3.1.1 Effect of Candidate Sex on Candidate Evaluations	35
3.1.2 Effect of Candidate Gender Expression on Candidate Evaluations	36
3.2 Voter Expectations for Women Winning	39
3.3 Experimental Research Design	40
3.3.1 Sample	44

¹Reprinted with permission from “Do Elites Discriminate Against Female Political Aspirants? Evidence from a Field Experiment” by Dhima, Kostanca, 2022. *Politics & Gender*, 18(1), 126-157, 2022 by Cambridge University Press.

3.4	Results and Discussion.....	45
3.4.1	Effect of Candidate Sex on Being the <i>Preferred Candidate</i>	45
3.4.2	Effect of Candidate Gender Expression on Being the <i>Preferred Candidate</i> ...	49
3.4.3	Effect of Candidate Sex on Being the <i>Winning Candidate</i>	51
3.4.4	Effect of Candidate Gender Expression on Being the <i>Winning Candidate</i>	53
3.5	Conclusion.....	54
4.	REEXAMINING THE SUPPLY AND DEMAND FRAMEWORK FOR EXPLAINING WOMEN’S DESCRIPTIVE REPRESENTATION.....	56
4.1	Theory: The Supply × Demand Framework	57
4.2	Empirical Analysis	60
4.3	Conclusion.....	72
5.	SUMMARY AND CONCLUSIONS	74
	REFERENCES	80
	APPENDIX A.	104
A.1	What about Quebec?	104
A.2	Sample of Legislators	109
A.3	Staff Responses.....	111
A.4	Including Responses After the 2 Week Cutoff.....	113
A.5	Analyzing the Content of the Email Responses	117
A.5.1	‘Micro’-mentorship Approach	117
A.5.2	Quality of Response Approach	118
A.5.3	Length of Response Approach.....	119
A.5.4	Results	119
	APPENDIX B.	123
B.1	Experimental Design in Korean	123
B.2	Descriptive Information About the Survey & Sample.....	124
B.2.1	Survey	124
B.2.2	Sample	125
B.3	Full Regression Results – <i>Preferred Candidate</i>	126
B.3.1	Regression Results	126
B.3.2	AMCE of Candidate Attributes	127
B.3.3	Marginal Means of Candidate Attributes.....	129
B.4	Full Regression Results – <i>Winning Candidate</i>	130
B.4.1	Regression Results	130
B.4.2	AMCE of Candidate Attributes	131
B.4.3	Marginal Means of Candidate Attributes.....	133

B.5	Mixed-Sex Pairings	134
B.5.1	AMCE of Candidate Attributes – Mixed-Sex Pairings	134
B.5.2	Marginal Means of Candidate Attributes – Mixed-Sex Pairings	137
B.6	Effects of Candidate Attributes by Respondent Sex	139
APPENDIX C.	142
C.1	Sample and Descriptive Statistics	142
C.1.1	Descriptive Statistics	142
C.1.2	Sample	143
C.1.2.1	Model (1): <i>Supply</i> (Labor Market) and <i>Demand</i> (Political Leaders).....	144
C.1.2.2	Model (2): <i>Supply</i> (Labor Market) and <i>Demand</i> (Jobs Scarce)	144
C.1.2.3	Model (3): <i>Supply</i> (Tertiary Education) and <i>Demand</i> (Political Leaders).....	145
C.1.2.4	Model (4): <i>Supply</i> (Tertiary Education) and <i>Demand</i> (Jobs Scarce).....	146
C.1.3	Visualizing Supply and Demand.....	147
C.2	Different Operationalizations for Supply and Demand.....	149
C.2.1	Model (2): <i>Supply</i> (<i>Labor Market</i>) and <i>Demand</i> (<i>Jobs Scarce</i>).....	150
C.2.2	Model (3): <i>Supply</i> (<i>Tertiary Education</i>) and <i>Demand</i> (<i>Political Leaders</i>).....	152
C.2.3	Model (4): <i>Supply</i> (<i>Tertiary Education</i>) and <i>Demand</i> (<i>Jobs Scarce</i>)	154
C.3	Controlling for electoral system	156
C.4	Relaxing the Linear Interaction Assumption	158
C.4.1	Linear Interaction Diagnostic Plot.....	159
C.4.2	Binning Estimator	162
C.4.3	Kernel Estimator.....	166

LIST OF FIGURES

FIGURE	Page
2.1 Email Sent to Legislators	18
2.2 Cumulative Response Rates by Level of Government	22
3.1 The Relationship between Candidate Sex and Gender Expression with Respect to Candidate Support	38
3.2 Experimental Design	44
3.3 Effect of Candidate Sex on Being Preferred	46
3.4 Effect of Candidate Sex on Being Preferred by Respondent Sex	48
3.5 Effects of Candidate Attributes on Being Preferred Conditional on Candidate Sex ...	50
3.6 Effect of Candidate Sex on Being Winner	51
3.7 Effect of Candidate Sex on Being Winner by Respondent Sex	52
3.8 Effects of Candidate Attributes on Being Winner Conditional on Candidate Sex	54
4.1 The Interaction between Supply and Demand with Respect to Women’s Descriptive Representation	60
4.2 Visualizing the Conditional Relationship between <i>Supply</i> , <i>Demand</i> , and <i>Women’s</i> <i>Representation</i>	66
4.3 The Conditional Effects of <i>Supply</i> and <i>Demand</i> on Women’s Descriptive Re- presentation	69
5.1 Visualizing the Conditional Impact of Supply, Demand, and Institutions on Women’s Descriptive Representation	78
B.1 Experimental Design in Korean	123
B.2 Effects of Candidate Attributes on Being Preferred – AMCEs	128
B.3 Effects of Candidate Attributes on Being Preferred – MMs	129

B.4	Effects of Candidate Attributes on Being Winner – AMCEs	132
B.5	Effects of Candidate Attributes on Being Winner – MMs	133
B.6	Effects of Candidate Attributes on Being Preferred – AMCEs for Mixed-Sex Pairings	135
B.7	Effects of Candidate Attributes on Being Winner – AMCEs for Mixed-Sex Pairings .	136
B.8	Effects of Candidate Attributes on Being Preferred – MMs for Mixed-Sex Pairings ..	137
B.9	Effects of Candidate Attributes on Being Winner – MMs for Mixed-Sex Pairings	138
B.10	Effects of Candidate Sex on Being Preferred by Respondent Sex	140
B.11	Effects of Candidate Sex on Being Winner by Respondent Sex	141
C.1	Scatter Plots of <i>Supply</i> and <i>Demand</i>	148
C.2	The Conditional Effects of <i>Supply</i> and <i>Demand</i> on Women’s Descriptive Representation – Model (2)	151
C.3	The Conditional Effects of <i>Supply</i> and <i>Demand</i> on Women’s Descriptive Representation – Model (3)	153
C.4	The Conditional Effects of <i>Supply</i> and <i>Demand</i> on Women’s Descriptive Representation – Model (4)	155
C.5	Linear Interaction Diagnostic Plots	161
C.6	GAM Plots of the Relationship between Women’s Representation, Supply, and Demand	162
C.7	Conditional Marginal Effects from Binning Estimator	165
C.8	Conditional Marginal Effects from Kernel Estimator	167

LIST OF TABLES

TABLE	Page
2.1 Demographic Balance Across Treatment Groups	20
2.2 Response Rates by Treatment Name and Legislator Gender	23
2.3 Response Rates by the Treatment Name, Level of Office, and Legislator Gender	26
2.4 Response Rates by the Treatment Name, Party Ideology, and Legislator Gender	29
3.1 Attributes for Candidate Profiles in Conjoint Experiment	41
4.1 Women’s Legislative Representation, 1990-2018	65
A.1 Response Rates by Treatment Name and Legislator Gender: Only Quebec	105
A.2 Response Rates by Treatment Name and Legislator Gender: Excluding Quebec	106
A.3 Response Rates by the Treatment Name, Level of Office, and Legislator Gender: Excluding Quebec	107
A.4 Response Rates by Treatment Name and Legislator Gender: Excluding Staff Re- sponses	112
A.5 Response Rates by Treatment Name and Legislator Gender: Including Later Re- sponses	113
A.6 Response Rates by the Treatment Name, Level of Office, and Legislator Gender: Including Later Responses	115
A.7 Response Rates by the Treatment Name, Party Ideology, and Legislator Gender: Including Later Responses	116
A.8 Quality of Response by Treatment Name	120
A.9 Quality of Response by Treatment Name and Legislator Gender	122
B.1 Demographics of Survey Sample	125
B.2 Effects of Candidate Attributes on Being Preferred	126

B.3	Effects of Candidate Attributes on Being Winner	130
C.1	Descriptive Statistics, 1990-2018	143
C.2	Determinants of Women’s Legislative Representation	157

1. INTRODUCTION

Relative to their share of the voting population, women remain systematically and disproportionately under-represented in politics around the world. While women make up half of the population in every country, today they comprise, on average, just 26% of the representatives in lower house legislatures globally (Inter-Parliamentary Union 2022).

Previous research shows that the near universal under-representation of women in decision-making institutions not only undermines political equality (Mansbridge 1999), but it also yields inferior public policy outcomes (Bratton and Ray 2002; Schwindt-Bayer and Mishler 2005; Kittilson 2008; Schwindt-Bayer, Malecki and Crisp 2010; Mendelberg, Karpowitza and Goedert 2014; Piscopo and Franceschet. 2014) and depresses both participation in the public sphere (Wolbrecht and Campbell 2007; Desposato and Norrander 2009; Barnes and Burchard 2012; Dassonneville and McAllister 2018) and trust in political institutions (Mansbridge 1999; Clayton, Piscopo and O'Brien 2019). Therefore, it is important that we try to understand the factors that lead to the political under-representation of women.

In my dissertation, I present three papers that evaluate the causes of women's political under-representation. These papers all fit into an overarching supply, demand, and institutions framework. Broadly-speaking, *supply-side factors* determine the size of the pool of women with the resources and ambition to run for office. In contrast, *demand-side factors* shape the preferences that actors have for women in politics. The demand for women in politics depends on the preferences of both elites and masses. The impact of these supply-side and demand-side factors on women's political representation varies across different institutional contexts. *Institutional factors* such as regime type and electoral rules define the context or environment in which supply and demand influence women's representation. Each of the papers studies women's political under-representation by examining different components of this supply, demand, and institutions framework.

The first two papers focus on demand for women’s political representation. While much of the existing literature focuses on mass or voter demand, my first paper, “Do Elites Discriminate Against Female Political Aspirants? Evidence from a Field Experiment”, focuses on *elite* demand. Elite demand is more difficult to study in part because it’s not easy to get party elites into a lab to participate in experiments, or to get them to use randomized campaign tactics in the field. In my paper, I use one of the first gender-focused politics field experiments ever conducted outside of the United States to examine whether elites exhibit gender discrimination when responding to political aspirants. One explanation for why women are under-represented in politics is that they do not seek a career in politics at the same rate as men (Lawless and Fox 2010; Fulton et al. 2006; Pruyser and Blais 2019). Importantly, existing research suggests that women are much more likely to run for office if they receive encouragement from political elites (Fox and Lawless 2004; Niven 2006; Karpowitz, Monson and Preece 2017). In this chapter, I use a pre-registered experimental design that draws on theories of gender discrimination, group attachment, and partisan identity to examine whether elected officials at the municipal, provincial, and federal levels in Canada exhibit gender bias when responding to individuals who are seeking advice about a career in politics. Overall, I find no evidence of elite gender discrimination against female political aspirants. Indeed, I find that legislators at all levels of Canadian politics, but especially female legislators and those associated with left-leaning parties, are significantly *more* responsive to female aspirants than male ones. I also find that female aspirants receive more detailed and helpful advice than their male counterparts. This research is important because it suggests that political elites are open to increasing female political representation, at least in this very early and informal stage of the political recruitment process, and the overall message should be seen as encouraging for female political aspirants.

My second paper, “Do Voters Discriminate Against Female Candidates? Evidence From an Experiment in South Korea,” focuses on *mass* demand. I use a pre-registered conjoint survey experiment that draws on theories of gender bias, gender stereotypes, and role congruity to exam-

ine how voters in South Korea respond to political candidates who deviate from gendered norms. Existing experimental research in this area tends to focus on candidate sex and hence whether individuals are willing to vote for male or female candidates. There is mixed evidence as to whether voters are unwilling to support female candidates (Schwarz and Coppock 2020), and one potential explanation for these mixed results is that scholars have paid too little attention to how candidates *express* their gender. This is problematic because a preference for a male candidate might not necessarily be a preference for a man but rather a preference for a *masculine* candidate. Similarly, a preference for a female candidate might not be a preference for a woman but a preference for *femininity*. In my conjoint survey experiment, I use a fully-crossed factorial design that allows me to examine the interaction between a candidate's sex and their expressed level of masculinity/femininity across different dimensions. Overall, I find evidence of a gender bias in favor of women political candidates. Indeed, I find that voters, on average, prefer female candidates rather than male candidates. This overall pro-woman bias is driven by women voters; male voters actually exhibit a smaller pro-man bias. While voters indicate an overall preference for female as opposed to male candidates, they also indicate that they think that the male candidate has a better chance of winning the election. These findings suggest that it is important to carefully interpret the results from these types of experiments as preferences and expectations about who will win interact to determine final voting behavior. We should not necessarily infer voter behavior simply from voter preferences. When it comes to how voters evaluate candidates who deviate from gender norms, I find that voters tend to evaluate women and men candidates with the same attributes similarly, except for approach to politics. Voters tend to prefer candidates who run counter to gender stereotypes: they prefer women candidates who present a "tough" approach to politics and men candidates who present a "compassionate" approach. These findings have important implications for both female and male candidates and party elites who recruit them.

In my third paper, "Reexamining the Supply and Demand Framework for Explaining Women's Descriptive Representation," I take a more aggregate-level approach and look at how *supply-side*

and demand-side factors interact to affect cross-national variation in women's descriptive representation around the world. The empirical analysis controls for institutional context. Theoretically, scholars have recognized that women's representation can be explained within a supply and demand framework (Norris and Lovenduski 1993; Inglehart and Norris 2003; Matland 2005; Paxton, Kunovich and Hughes 2007). Unfortunately, previous empirical tests of the supply and demand framework have produced mixed results. In this chapter, I argue that a potential explanation for these mixed results has to do with the gap between theory and empirics. Specifically, all existing empirical studies treat supply-side and demand-side factors separately and ignore the inherent interaction between supply and demand (Paxton 1997; Kenworthy and Malami 1999; Studlar and McAllister 2002; Paxton and Kunovich 2003; Tremblay 2007; Swiss 2009; Stockemer 2011; Rosen 2011; Ruedin 2012). Theoretically, we should only expect to see high levels of women's representation when *both* supply and demand are sufficiently high. Using a new global dataset of women's legislative representation from 1990 to 2018, I conduct the first test of the conditional implications at the heart of the supply and demand framework. The results are consistent with my theory and are substantively important because they indicate the conditions under which we can expect supply-side factors (such as increased female labor force participation) and demand-side factors (such as less traditional gender attitudes with respect to political office holders) to actually translate into greater female political representation.

As I've demonstrated above, my dissertation research can be situated in a supply, demand, and institutions framework. Two of the papers focus on the independent components of this framework and one of the papers focuses on how these components combine to drive aggregate patterns in women's representation. I hope that the findings of my dissertation research help to increase women's political representation by encouraging women early in the political recruitment process (Chapter Two), widening the profile of favored gender expressions among candidates (Chapter Three), and honing our understanding of the optimal conditions under which supply-side factors and demand-side factors are likely to increase women's political representation (Chapter Four).

2. DO ELITES DISCRIMINATE AGAINST FEMALE POLITICAL ASPIRANTS? EVIDENCE FROM A FIELD EXPERIMENT*

Do elites exhibit gender bias when responding to political aspirants? Relative to their share of the voting population, women are significantly underrepresented in politics. The political recruitment process — the process of going from an eligible citizen to an elected representative — occurs in three distinct stages (Norris and Lovenduski 1993; Krook and Norris 2014). First, eligible citizens must select themselves as potential candidates. Second, those who aspire to political office must be selected as candidates by political parties. Finally, candidates must be selected as representatives by the voters. Gender distortions can occur at each of these stages, and studies have generally shown that women face higher barriers to representation than similarly situated men (Norris and Lovenduski 1993; Fox and Lawless 2004, 2010; Fulton et al. 2006; Fulton 2012; Anzia and Berry 2011). Most research has focused on the last stage of the recruitment process, where voters select their representatives. Here, though, I focus on a much earlier stage by examining whether elites discriminate against female political aspirants. Specifically, I examine whether elites are equally responsive to female and male political aspirants who signal an interest in a political career.

Most studies of the political recruitment process rely on observational data, which makes it difficult to isolate the effect of gender discrimination because of potential problems with omitted variable bias, selection bias, and post-treatment bias. Even when we observe differences between men and women, it's hard to know whether these differences are really due to the gender of the candidate or to some other unobserved or difficult to measure factor such as quality (Fulton 2012; Fulton and Dhima 2020), experience (McDermott 2005) or attractiveness (Lenz and Lawson 2011; Ahler et al. 2017). The issue of post-treatment bias, which has received relatively little attention

*Reprinted with permission from “Do Elites Discriminate Against Female Political Aspirants? Evidence from a Field Experiment” by Dhima, Kostanca, 2022. *Politics & Gender*, 18(1), 126-157, 2022 by Cambridge University Press.

in the literature, is particularly problematic (Montgomery, Nyhan and Torres 2018; King and Zeng 2006). An individual's gender is often determined early in an individual's life and affects most things they do or experience. Many of the things that scholars like to control for to isolate the effect of gender in their analyses, such as education and experience, are likely influenced by an individual's gender. This makes it extremely difficult to control for all of the potential confounders in an observational study of gender discrimination without inducing post-treatment bias (Crabtree 2019). Worryingly, post-treatment bias can be in any direction (Montgomery, Nyhan and Torres 2018). The methodological difficulties faced by observational studies in accurately identifying gender discrimination can have enormous potential costs. For example, some observational studies may erroneously indicate the presence of gender discrimination, leading policy makers and other actors to waste significant time and money attempting to remedy a problem that does not in fact exist. Similarly, other observational studies may incorrectly indicate that gender discrimination is absent, inducing policy makers and other actors to overlook the very real negative consequences of discrimination felt by particular gender groups.

To avoid the potential methodological problems associated with observational studies, I employ an experimental research design, specifically an audit (or correspondence) experiment, to examine elite gender discrimination. Audit experiments are especially useful for investigating sensitive topics, like gender discrimination, because they allow researchers to avoid both selection bias concerns that arise when the people who are likely to discriminate opt out of studies and social desirability concerns that arise when people have incentives to downplay their discriminatory behavior to avoid perceived social and legal sanctions.¹ With an audit experiment, the researcher varies some characteristic of individuals, keeping everything else the same, and then sends these individuals, or messages from these individuals, into the field to see whether the randomized char-

¹While they are relatively new to political science, audit experiments have a long history dating back to the 1940s and 1950s (Gaddis 2018a). Large-scale audit studies were first conducted in the United States and the United Kingdom in the late 1960s following the introduction of legislation that sought to make various kinds of racial discrimination illegal.

acteristic affects some outcome of interest (Bertrand and Duflo 2017). Unlike survey experiments, which can only get at reported *attitudes* toward discrimination in a hypothetical scenario, audit experiments provide a real-world *behavioral* measure of discrimination. In my audit experiment, I send an email message from a political aspirant inquiring about a career in politics to elected representatives at different levels of government in Canada.² By varying the sex of the political aspirant, I can examine whether political elites respond to men and women at the same rates. Replying (versus not replying) to an email message like this is recognized as important because responses can be considered “a type of ‘micro’-mentorship where even a small act of encouragement can teach an aspirant about the profession and provide cues about whether he or she will be welcome” (Kalla, Rosenbluth and Teele 2018, 338).

My study is one of the few gender and politics audit experiments to be conducted outside the United States and the first to address one of the stages of the political recruitment process.³ To my knowledge, there has been only one audit experiment that looks at elite gender discrimination in the political recruitment process. In contrast to previous observational studies of political elite behavior, the audit study conducted by Kalla, Rosenbluth and Teele (2018) in the United States finds no evidence of any gender discrimination by legislators against female political aspirants. A common concern with experiments is that they are often not replicated and, as a result, it’s hard to know whether their findings are generalizable to other contexts (Clark and Golder 2015). Indeed, much of the discussion regarding the drawbacks of experiments is framed around concerns with external validity.⁴ One way to address concerns with external validity and evaluate the robust-

²Email audit experiments are increasingly common (Crabtree 2018) and have recently been employed to study things like racial discrimination (Bertrand and Mullainathan 2004; Brushman and Bonacci 2004; Butler and Broockman 2011; Hogan and Berry 2011; Gell-Redman et al. 2018), sexual discrimination (Neumark, Bank and Nort 1996; Ahmed, Anderson and Hammarstedt 2013), age discrimination (Ahmed, Anderson and Hammarstedt 2012; Baert et al. 2016), and gender discrimination (Neumark, Bank and Nort 1996; Milkman, Akinola and Chugh 2015; Kalla, Rosenbluth and Teele 2018).

³While several audit experiments have focused on gender in the economic sphere, very few have examined the gendered behavior of political actors (Kalla, Rosenbluth and Teele 2018). Most audit studies dealing with public legislators have instead focused on race (Butler and Broockman 2011; Broockman 2013; Butler 2014; Gell-Redman et al. 2018).

⁴Internal validity refers to our ability to determine whether a treatment effect – the difference between the out-

ness of results is to replicate an experiment in a different context (Krupnikov and Levine 2014). Replication is especially pertinent when the findings of an experiment, such as the one conducted by Kalla, Rosenbluth and Teele (2018), run counter to expectations in the literature. My audit experiment builds on and extends Kalla, Rosenbluth, and Teele’s analysis to the Canadian context.

There are at least three reasons why Canada is a good context for evaluating elite gender discrimination against female political aspirants. First, Canada is useful from a practical perspective when it comes to extending the Kalla, Rosenbluth and Teele (2018) study. While Canada is a parliamentary democracy and the United States is a presidential one, both countries employ a single-member district plurality electoral system. And, with the exception of Quebec, which has a large French-speaking population, both countries are predominantly English speaking. This means that I can use the same experimental treatment, thereby addressing potential concerns about minimal replication in research and the generalizability of experimental studies. Moreover, Canada has a qualitatively similar level of women’s legislative representation to that found in the United States. Second, prior research on Canadian politics suggests that the persistent underrepresentation of women has less to do with voter reluctance to elect female candidates (Black and Erickson 2003; Young 2006; Goodyear-Grant 2010; Bashevkin 2011) and more to do with discrimination in the earlier stages of the political recruitment process (Erickson 1991; Thomas and Bodet 2013). This suggests that a focus on elite gender discrimination is particularly pertinent in the Canadian case. Finally, the inclusion of women in politics has become a salient issue over the last few years in Canada, and there has been growing pressure to increase diversity and make political representation more inclusive. During the 2015 federal elections, for example, the Liberal Party made a commitment to gender-equal cabinets. Upon coming to power, the Liberal Party fulfilled some of its promises by forming Canada’s first gender-balanced government at the federal level, further

come when exposed to the treatment and the outcome when not exposed to the treatment – reflects a *causal* relationship (Shadish, Cook and Campbell 2002). In contrast, external validity refers to our ability to determine whether the magnitude and significance of a treatment effect differs across people and settings, or more generally “context” (Morton and Williams 2008; McDermott 2011).

propelling issues of gender equality onto the national agenda.

Based on responses from 1,774 legislators in Canada, I find no evidence of elite gender discrimination against female political aspirants. Indeed, I find consistent evidence that Canadian legislators are significantly more responsive to female political aspirants than male ones and more likely to provide them with helpful advice. This pro-women bias in a very early and informal phase of the political recruitment process exists at all levels of government in Canada and tends to be stronger among female legislators and those associated with left-leaning parties. These results obviously don't imply that female political aspirants have not historically faced elite discrimination in this phase of the political recruitment process or that they don't face higher hurdles than men in later and more formal stages of the recruitment process. They do suggest, however, that contemporary political elites in Canada may be open to increasing female political representation. This should be treated as positive news. When combined with the absence of elite gender discrimination at the same stage of the political recruitment process in the United States (Kalla, Rosenbluth and Teele 2018), the results from my experiment should serve as welcome encouragement for women to put themselves forward as potential candidates and pursue their political ambitions.

2.1 Should Gender Affect Elite Responsiveness?

On the whole, research suggests that gender distortions are greater earlier in the political recruitment process than later. While studies of voter discrimination find that stereotypes are still pervasive in voters' evaluation of candidates (Bauer 2015), there's only mixed evidence as to whether and when voters are unwilling to support female candidates (Schwarz and Coppock 2020). In fact, some analyses find that voters are *more* willing to vote for female candidates than male ones (Black and Erickson 2003; Lawless and Pearson 2008; Golder et al. 2017).⁵ Research suggests that parties are less enthusiastic than voters when it comes to selecting female candidates. For ex-

⁵Some scholars caution that equal success rates of female and male candidates among voters do not necessarily indicate a gender-neutral electoral environment (Lawless and Pearson 2008; Anzia and Berry 2011; Fox and Lawless 2004; Fulton 2012; Fulton and Dhima 2020; Mo 2015). For example, it may be the case that female candidates only "do as well as men" because they have stronger valence characteristics, such as higher qualifications, thereby suggesting that the electoral environment is still biased against women.

ample, evidence indicates that party leaders have historically been more likely to recruit men than women to run for political office (Niven 1998, 2006; Sanbonmatsu 2006; Fox and Lawless 2010; Schwindt-Bayer 2011) and that parties, when they do select female candidates, often nominate them in districts where they're less likely to win (Erickson 1991; Sanbonmatsu 2002; Thomas and Bodet 2013) or place them lower on party lists (Luhiste 2015). Gender distortions are arguably even stronger at the self-selection stage, with numerous studies showing that women are much less likely to put themselves forward as potential candidates than similarly situated men (Lawless and Fox 2010; Fulton et al. 2006; Pruyssers and Blais 2019) because of gender role socialization (Clark, Hadley and Darcy 1989), family obligations (Fulton et al. 2006), perceptions of qualifications (Fox and Lawless 2011), lack of party support and recruitment (Fox and Lawless 2004, 2010, 2011; Fulton et al. 2006; Niven 1998; Sanbonmatsu 2006), and election aversion (Kanthak and Woon 2015).

As the early stages of the recruitment process appear to be more distortionary, women's underrepresentation is often considered a supply-side issue rather than a demand-side one (Fox and Lawless 2004; Htun 2016).⁶ More specifically, it's frequently assumed that the low level of female representation is primarily caused by the lack of women running for political office rather than low voter demand (Htun 2016, 90-91). But it's important to remember that demand also comes from political elites. Elite demand matters not only because elites have the power to change formal institutions, such as electoral rules, and thus make the political opportunity structure more permissive to female political inclusion, but also because their informal messages and behavior can have a significant impact on encouraging female political aspirants to put themselves forward as potential candidates. Research suggests, for example, that political aspirants are twice as likely to think about running if they're encouraged by political elites (Fox and Lawless 2004) and that women

⁶The level of women's legislative representation in a country is determined by both demand-side and supply-side factors (Inglehart and Norris 2003; Paxton, Kunovich and Hughes 2007). While *supply-side factors* shape the size of the pool of women with the experience and willingness to compete for political office, *demand-side factors* have to do with the preferences that individuals have for female representatives.

are more likely to run and get elected if elites promote messages of female inclusion (Karpowitz, Monson and Preece 2017). Encouragement from political elites is especially important for female candidates as “women are simply unlikely to run in the face of elite discouragement” (Niven 2006, 485).

Can we expect political elites to be equally responsive to female aspirants when they seek advice on how to start a political career? By looking at political aspirants who contact legislators for advice, I focus on an *informal* phase of the recruitment process to examine possible elite gender bias. To date, relatively little is known about how informal institutions shape candidate emergence. Most studies of gender bias tend to focus on the impact of formal institutions such as quotas (Jones 1998; Frechette, Maniquet and Morelli 2008; Krook 2009), electoral systems (Kittilson and Schwindt-Bayer 2010), district magnitude (Shugart 1994; Taagepera 1994; Schmidt 2009), and ballot structures (Jones and Navia 1999; Schmidt 2009; Thames and Williams 2010; Luhiste 2015). The scholarship that exists on informal institutions tends to address the challenges that female representatives face once they’re in the legislature (Kathlene 1994; Hawkesworth 2003; Heath, Schwindt-Bayer and Taylor-Robinson 2005; Escobar-Lemmon and Taylor-Robinson 2009). With a few exceptions (Bjarnegard 2013; Bjarnegard and Kenny 2015), this scholarship on informal institutions is not mirrored to the same extent in the literature that addresses political recruitment.

There are several reasons why we might expect political elites to exhibit gender bias when responding to individuals who are thinking about a career in politics. As noted earlier, existing research suggests that there’s gender discrimination in the political recruitment process and numerous studies find that elites have historically been less likely to recruit female candidates than male ones (Fox and Oxley 2003; Sanbonmatsu 2006; Fox and Lawless 2010). As an example, Niven (1998) finds that the majority of women holding local office across four U.S. states report having been discouraged from running for office by party leaders. Even when female and male candidates report receiving similar levels of encouragement from political elites, there seems to be some bias when it comes to the districts in which they’re selected to run. For example, fe-

male candidates for state house and senate races in Florida in 2000 and 2002 report having been discouraged from running in favorable districts and instead encouraged to run in unfavorable districts, while men received the opposite messages (Niven 2006). Similarly, there's evidence that female candidates for federal office in Canada in 2008 and 2011 were more likely than men to be nominated in non-competitive districts (Erickson 1991; Thomas and Bodet 2013).⁷ One potential reason why political elites discriminate against female candidates has to do with how implicit and explicit gender stereotypes influence who they deem appropriate for political office. When envisioning a strong legislative candidate, studies have found that party leaders tend to describe someone with stereotypically masculine traits (Niven 1998). If political elites believe that female candidates are not suited to holding political office, then they're less likely to encourage them to run for office. This reasoning leads to the *Gender Bias Hypothesis*.

Gender Bias Hypothesis: Political elites will be less responsive to female political aspirants than to male political aspirants.

There are reasons to believe that the level of gender bias may vary depending on the gender of the political elite because of in-group and out-group bias. According to social identity theory (Tajfel and Turner 1979), individuals have a natural tendency to categorize people into groups based on shared identity traits such as gender, race, and religion. These shared group identities create a sense of connection and belonging, which can lead to a more favorable evaluation and treatment of 'in-group' as opposed to 'out-group' members, even in the absence of any conscious pro-ingroup bias (McPherson, Smith-Lovin and Cook 2001). According to social identity theory, therefore, female political aspirants should experience more discrimination from political elites who don't share their gender (men) than those who do (women). Consistent with the claim that people favor members of their in-group, Niven (1998) finds that male party chairs in the United States prefer candidates who resemble themselves on a range of traits, including gender, occupation, and personality characteristics. Since women are not free of these biases, female political

⁷Though Medeiros, Forest and Erl (2019) find that this was not the case for the federal elections in 2015.

elites should favor female political aspirants. While Tremblay and Pelletier (2001) find that female party presidents in Canada don't prefer candidates with stereotypically feminine traits, Cheng and Tavits (2011), as well as (Medeiros, Forest and Erl 2019), find that they are more likely to nominate women candidates in their constituency. This reasoning suggests a gender affinity story in which male elites will be more responsive to men and female elites will be more responsive to women. Note that this gender affinity story is consistent with the predictions from the *Gender Bias Hypothesis*. The fact that contemporary political elites are primarily made up of men means that we should observe a gender bias on average against female political aspirants.

Gender Affinity Hypothesis: Political elites will be more responsive to political aspirants who share their gender. In other words, male political elites will be less responsive to female political aspirants and female political elites will be more responsive to female political aspirants.

The gender affinity story suggests that we'll see less discrimination against female political aspirants when the numeric or descriptive representation of female elites is high. This is because female political aspirants will enjoy more positive in-group bias and less negative out-group bias when women make up a larger percentage of the political elite. A common claim in the literature is that fewer women hold political office as we move up the levels of government (Blais and Gidengil 1991; Baxter and Wright 2000; Palmer and Simon 2001, 2010), leading some scholars to talk of a "glass ceiling" when it comes to women's representation (Baxter and Wright 2000; Ferree and Purkayastha 2000; Cotter et al. 2001; Folke and Rickne 2016). Political hierarchies, such as those that often exist between different levels of government, are commonly associated with increased discrimination against marginalized groups such as women, with the most powerful offices typically restricted to men and other high status elites. Putnam (1976, 33) refers to this as the "law of increasing disproportion." Bashevkin (1993, 92) finds evidence of this hierarchical impact on women, which she summarizes as "the higher, the fewer," within Canadian parties. She later suggests that a similar result holds across the different levels of government in Canada

(Bashevkin 2009, 4). The reasoning here is captured in the *Levels of Government Hypothesis*.

Levels of Government Hypothesis: Political elites will be less responsive to female political aspirants than to male political aspirants at all levels of the government. However, this negative effect will grow as we move from the local to the national level.

The extent to which this general theoretical hypothesis applies in the Canadian case is somewhat open to question, though. One reason is that there's debate about the degree to which the different levels of government — municipal, provincial, and federal — represent a clear political hierarchy.⁸ Constitutionally, the federal and provincial levels enjoy equal status, with municipalities occupying a subordinate status. However, some have argued that provinces, despite their formal and constitutional equality, are also subordinate in practice to the federal level, or at least perceived to be so, partly because of their limited monetary resources (Dyck 1998, 225). There's also debate as to whether the descriptive representation of women declines as we move from the municipal to the federal level. Tolley (2011, 585) finds that women experienced a “municipal advantage” in about 60% of jurisdictions in 2009 and that the level of women legislators at the federal level was lower than at both the provincial and municipal levels from 2004 to 2009. However, these differences were substantively small and not necessarily reflective of earlier time periods. It remains an open empirical question, therefore, whether the degree to which political elites respond differently to female and male political aspirants will vary across the different levels of government in Canada.

How much gender bias political elites exhibit against female political aspirants should also depend on their partisan affiliation. Political elites are nested within political parties, and there is compelling evidence that parties differ in their ideological and behavioral commitment to gender egalitarianism. Past studies have shown, for example, that parties on the left of the ideological spectrum are more responsive to gender-related demands than parties on the right (Caul 1999; Kittilson 2006; Salmond 2006; O'Brien 2018). The commitment of left-wing parties to issues of

⁸Canada has 3,573 municipalities, 10 provinces, and 3 territories. For the purpose of this paper, I treat the territories as having quasi-provincial status and use the word “province” to refer to both provinces and territories.

gender equality is discernible in their gender-egalitarian policies (Beckwith 2000; Young 2000), their greater incorporation of women within leadership structures (Caul 2001), and their initiatives to increase the presence of women in politics by recruiting more female candidates. Since legislators select into parties and egalitarian attitudes on gender issues are associated with left-wing party membership (Tremblay and Pelletier 2000), it's likely that political elites from left-leaning parties will be more responsive to female political aspirants than political elites from right-leaning parties. This reasoning is captured in the *Left-Wing Partisan Hypothesis*.

Left-Wing Partisan Hypothesis: Political elites from left-leaning parties will be more responsive to female political aspirants than political elites from right-leaning parties.

Although most research assumes that political elites will be biased against women who are thinking about a political career, there are also reasons to expect that they'll be equally responsive to male and female political aspirants. First, some scholars argue that when it comes to their reelection and political careers (Fiorina 1989; Grose 2011), interactions with constituents, like personal communications with them, are as important, if not more important, for political elites than legislative behavior (Fenno 1978). Since elites are likely to be vote-maximizing agents and voters can sanction them based on their interactions, they have an incentive to be responsive to all of their constituents irrespective of any potential biases they might otherwise have. Evidence for this comes from a recent audit experiment by Loewen and MacKenzie (2019) showing that 202 randomly selected Canadian legislators at the federal and provincial levels were equally responsive to requests for assistance from men and women. Second, if political elites believe that voter demand for female candidates is similar to voter demand for male candidates, as some existing research indicates (Black and Erickson 2003; Lawless and Pearson 2008; Schwarz and Coppock 2020; Golder et al. 2017), then strategic incentives will again encourage political elites to be equally responsive to female and male political aspirants. This reasoning is captured in the *Equal Response Hypothesis*.

Equal Response Hypothesis: Political elites will be equally responsive to female and male political aspirants.

2.2 Experimental Research Design

Identifying gender bias in the responsiveness of political elites is difficult with observational data because of potential problems with omitted variable bias, selection bias, and post-treatment bias. I avoid these methodological problems by conducting the first gender and politics audit experiment on an aspect of the political recruitment process outside the United States. In the audit experiment, I send an email message from a political aspirant inquiring about a career in politics to legislators at the municipal, provincial, and federal levels in Canada. By randomizing the sex of the political aspirant, it's possible to determine whether political elites respond at equal rates to women and men. Whether political elites are willing to reply to an email from a political aspirant seeking advice on how to start a career in politics is important as responses serve as a visible signal of inclusion, indicating whether the female and male aspirants are welcome in the political profession. This type of “micro-mentorship” is often considered especially important for female aspirants as women are significantly more likely to put themselves forward as candidates if they're encouraged and actively recruited to run for office (Carroll and Sanbonmatsu 2013; Fox and Lawless 2004, 2014). Audit experiments, like the one conducted here, are well suited to investigating sensitive topics, such as elite gender discrimination, as they allow researchers to directly evaluate actual behavior, as opposed to attitudes or *reported* behavior, while mitigating selection bias and social desirability concerns.⁹

⁹The benefits of audit experiments for measuring discrimination have long been recognized by academics and governments alike (Gaddis 2018*b*). For example, the Race Relations Board, created by the British Parliament in the 1960s, was an early adopter of audit experiments to measure levels of racial discrimination (Daniel 1968; Smith 2015). Similarly, the U.S. Department of Housing and Urban Development has conducted multiple audit studies over the years looking at discrimination in the housing market (Johnson, Porter and Mateljan 1971; Quillian et al. 2017). Although audit experiments provide an effective behavioral measure of discrimination, they require that scholars engage in deception and eschew standard informed consent procedures. Deception is necessary as participants are virtually guaranteed to behave differently if they know that they're taking part in, say, an experiment on gender discrimination. Similarly, informed consent is not feasible given that the mere knowledge of taking part in a study, even a “cover” experiment, is likely to change participant behavior and thereby invalidate causal inferences. In my own study, I followed best practices as they relate to the ethical implementation of audit experiments. First, I obtained institutional review board (IRB) approval and preregistered my analysis at Evidence in Governance and Politics (EGAP). Second, I respect participant confidentiality by reporting only aggregate results and no specific responses. Third, I tried to minimize the time burden on participants by keeping the question in my email message short and straightforward to answer. Some evidence that I was successful on this last point comes from the fact that the median response to my

The email message I sent to legislators in Canada is shown in Figure 2.1.¹⁰ The email contained a request to learn about how the legislator entered politics and to give advice on how to start a career in politics. The email was sent from a hypothetical university student and each legislator received just one email. Having a university student, as opposed to a high school or middle school student, ask for advice about how to start a political career allows for a more credible inquiry since university students are more likely to have thought seriously about their career choices and taken steps in pursuit of their career objectives.¹¹ The only difference in the email sent to each legislator was whether the email was sent from (and signed by) an email account with a female or male first name. In effect, the randomized experimental treatment is the gendered name of the putative student. In line with the broader literature on audit studies (Bertrand and Mullainathan 2004), I use multiple female and male names to avoid the possibility that differences in elite responsiveness might be driven by a particular name effect as opposed to the gender of the political aspirant.¹²

The names I use were the same as those used by Kalla, Rosenbluth and Teele (2018) in their study of political elite responsiveness in the United States. The thirteen first names for women were Amanda, Ashley, Brittany, Emily, Hannah, Jessica, Kayla, Lauren, Megan, Rachel, Samantha, Sarah, Stephanie. The thirteen first names for men were Andrew, Brandon, Christopher, Daniel, David, James, John, Joshua, Matthew, Michael, Nicholas, Ryan, and Tyler. The 26 surnames are

Allen, Anderson, Brown, Clark, Davis, Hall, Harris, Jackson, Johnson, Jones, King, Lee, Lewis, email message was just 44 words long. Fourth, the experiment doesn't place an undue burden on vulnerable groups and poses minimal risk to the participants and wider community.

¹⁰The email message was always sent in English to replicate as closely as possible the empirical strategy employed by Kalla, Rosenbluth and Teele (2018). This has implications for Quebec with its large French-speaking population. As I address in more detail in Appendix A.1, while the response rate was lower for Quebec than the overall response rate, the *pattern* of responses was almost identical. Importantly, my upcoming results are robust to the inclusion or exclusion of the legislators in Quebec.

¹¹As I noted earlier, my study replicates and extends a recent audit experiment conducted in the United States by Kalla, Rosenbluth and Teele (2018). One minor change was needed in the wording of the email message to make it suitable for the Canadian context. Specifically, the email message mentions a "second-year university student" rather than a "college sophomore." This change was necessary because Canadians refer to students by their year and because "university" in Canada refers to a four-year degree granting institution, whereas "college" usually refers to a community college or technical school.

¹²The use of multiple names also reduces the likelihood that legislators in the same office or building would become aware of the experimental intervention by observing emails that came from the same student.

Figure 2.1: Email Sent to Legislators

From: [Treatment: Student Sex]
To: [Legislator's email]
Subject: Help on a class project?

Dear [LEGISLATOR],

My name is [MALE/FEMALE] and I am a second-year university student. I'm interviewing politicians for a class project to learn about how they entered their field and what advice they might have for students interested in politics. As someone who really cares about my community, one day I hope to be a politician. What advice would you give me?

Sincerely,
[MALE/FEMALE]

Martin, Miller, Moore, Robinson, Smith, Taylor, Thomas, Thompson, Walker, White, Williams, Wilson, and Young.¹³ I considered adding last names that would signal the race/ethnicity of the student, for example, a French- or South Asian- sounding last name. However, I ultimately refrained from doing this as the relatively small number of Canadian legislators limits my statistical power and makes factorial experimental designs that jointly manipulate the gender and race of the student less practical. What this means, though, is that my audit experiment is limited to testing whether political elites discriminate against Anglo-Canadian college-educated women. I randomly combined the first and last names to create 26 unique names. Finally, I generated Gmail accounts for each hypothetical student that took the following form: `firstname.lastnameXXXX@gmail.com`, where XXXX represents four random digits.

Given my interest in political elite bias, the population of interest is the universe of legislators in Canada. The names and contact information for Canadian legislators come from the Represent

¹³The first and last names were the most popular names in the United States in the 1990s based on information from the U.S. Census and Social Security Administration, and should therefore be common among second-year university students when the audit experiment was conducted in January 2018. One approach would have been to substitute these names with the most popular first and last names in Canada in the 1990s. However, Statistics Canada, the equivalent of the U.S. Census Bureau, doesn't collect data on the popularity of baby names. I checked websites that had information on the popularity of baby names in Canada and the most common first and last names were very similar to those used in the Kalla, Rosenbluth and Teele (2018) analysis. Thus, to keep the experimental treatment as similar as possible across the two studies, I chose to leave the first and last names unchanged. The one exception is that I excluded Hispanic last names as Hispanics are not a salient visible minority in Canada.

Civic Information API. The original sample consisted of 1,936 legislators.¹⁴ However, I dropped several legislators for two reasons. First, I was forced to drop those legislators for whom an email address was not provided. Second, I dropped duplicate legislators. These were legislators who held multiple elected positions, such as city councillor and regional councillor. This left me with 1,779 unique legislators across all of the levels of government in Canada. Specifically, there are 854 municipal legislators (28.6% women), 591 provincial legislators (31.6% women), and 334 federal legislators (26.7% women). Of the 1,779 legislators in the final sample, five could not be reached because of an invalid email address. In line with common practice, I exclude these observations from the upcoming analyses (Butler and Broockman 2011). This means that the results reported in the paper refer to the 1,774 legislators who actually received an email.¹⁵

The emails were sent on January 20 and January 21, 2018, with legislators randomly assigned to receive their message on one of these days. To better test whether male and female legislators respond at different rates, I block randomized the email messages on the gender of the legislator (Moore and Schnakenberg 2012). This means that I first divided the legislators into two groups — male and female — and then I randomly assigned the treatment within these two groups.¹⁶ The benefit of block randomization is that we can ensure that roughly equal numbers of male and female legislators are assigned to each experimental treatment (Gerber and Green 2012). The information contained in Table 2.1 confirms that the randomization procedure was successful and that the two experimental treatment groups (male or female sender) are balanced demographically. Because

¹⁴Although this doesn't represent the full number of legislators in Canada, the API notes that it's "the most comprehensive source in Canada for elected officials and electoral districts." More descriptive information about the federal, provincial, and municipal legislators in the API sample can be found in Appendix A.2

¹⁵In a recent audit study in Canada, Loewen and MacKenzie (2019) exclude legislators from Prince Edward Island on the grounds that the provincial legislators there are likely to know most of their constituents and may therefore be suspicious of an email from someone they don't know. However, the pattern of responses from the legislators on Prince Edward Island is similar to that found elsewhere and a careful read of the responses doesn't indicate anything problematic. As a result, I include the responses from Prince Edward Island in my upcoming analyses. I note, though, that my inferences are robust to excluding these responses.

¹⁶To identify whether a legislator is female or male, I looked up each legislator online. Most organization websites (councils, assemblies, parliament) include a profile for each of the legislators that provides a photo and/or biography from which it's possible to determine the sex of the legislator. I also consulted newspaper articles, Facebook accounts, Twitter accounts, and other sources for the more difficult cases.

Table 2.1: Demographic Balance Across Treatment Groups

	Male Name	Female Name	<i>p</i> -value of Difference
Female Legislator	0.29 (0.46)	0.29 (0.46)	0.99
Male Legislator	0.71 (0.46)	0.71 (0.46)	0.99
Municipal Legislator	0.47 (0.50)	0.49 (0.50)	0.62
Provincial Legislator	0.33 (0.47)	0.33 (0.47)	0.97
Federal Legislator	0.19 (0.40)	0.18 (0.39)	0.55
Left Party Ideology	0.10 (0.30)	0.11 (0.32)	0.48
Center Party Ideology	0.16 (0.37)	0.16 (0.37)	0.84
Right Party Ideology	0.16 (0.36)	0.14 (0.35)	0.32
% Bounced Email	0.00 (0.05)	0.00 (0.06)	0.65
N	890	889	

Note: Table 2.1 indicates the means for different demographic variables across the two treatment groups; standard deviations are shown in parentheses. The column *p-value of Difference* refers to the *p*-value from a difference-in-means test across the two treatment groups.

I'm interested in the responsiveness of political elites, my outcome measure, *Email Response*, is coded 1 if a response came from an email account associated with the legislator within two weeks, and 0 otherwise; I don't count auto-responses as replies.¹⁷

2.3 Results and Discussion

Before evaluating the specific hypotheses, I discuss the response rate in general. Almost half of the legislators in Canada responded to the email they received. Specifically, 864 (49%) of the 1,774 emails that were successfully sent received a response. This response rate was significantly higher than the 26% response rate in the United States (Kalla, Rosenbluth and Teele 2018). The

¹⁷As with all audit experiments of this type, there's no guarantee that the legislator is the person to receive and respond to the student's email message. Technically, therefore, the unit of analysis is the email address of the legislator and not the legislator. In many cases, it's possible to identify whether a staff member has sent the email response rather than the legislator. My inferences are robust to excluding those responses that are identified as coming from a staff member. A more detailed discussion of this issue can be found in A.3.

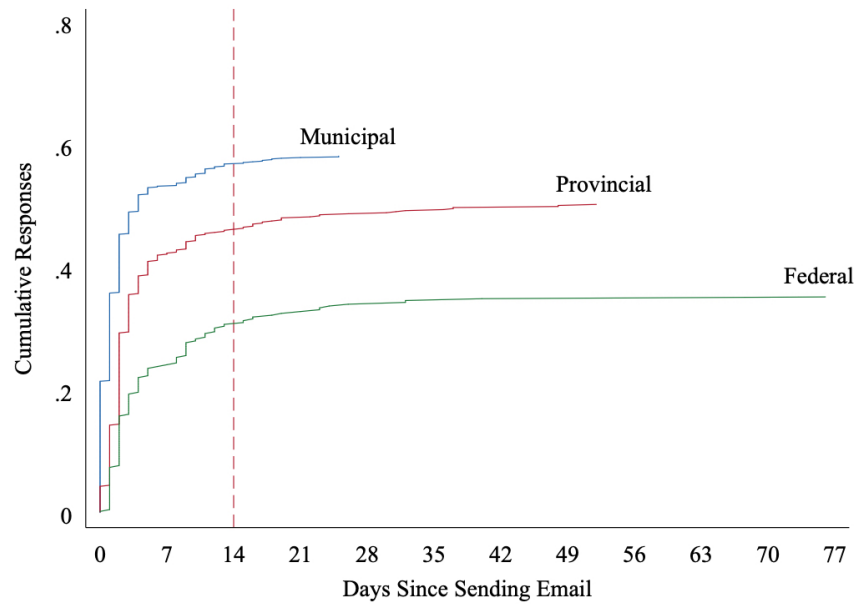
response rates decrease as we move from the local to the national level — the response rate was 57% for legislators at the municipal level, 46% for legislators at the provincial level, and 31% for legislators at the federal level. The upcoming reported response rates are based on responses that were received within two weeks of the original email message being sent. In Figure 2.2, I show the cumulative response rates across the various levels of government over time. As Figure 2.2 indicates, the temporal pattern of responses was very similar across the three levels of government, and almost all of the legislators who responded did so within two weeks of receiving the email from the hypothetical student. This is consistent with previous audit studies dealing with political elites, in which almost all responses were received within an initial two week window (Costa 2017).¹⁸ Interestingly, there's no substantive difference in the overall response rates for female (48%) and male (49%) legislators.

Table 2.2 provides information about response rates by treatment name and legislator gender. The first row shows how the overall response rate of the legislators varies depending on whether the hypothetical student sending the email message is female or male. The second and third rows show the response rates broken down by legislator gender. 95% confidence intervals are shown in square brackets. Recall that the *Gender Bias Hypothesis* predicts that legislators will be more responsive to male political aspirants than female ones and that the *Equal Response Hypothesis* predicts that the gender of the political aspirants will have no effect on the response rates. The results from the audit experiment falsify both hypotheses. Canadian legislators respond to female students (52%) at higher rates than male students (45%). This pro-women bias of 7 percentage points is statistically significant ($p = 0.01$) and can be attributed solely to the gendered name manipulation in the experiment.¹⁹ When Kalla, Rosenbluth and Teele (2018) conducted the same experiment on

¹⁸The two week window is consistent with the research design in my pre-registration plan. As I demonstrate in Appendix A.4, the inferences from my upcoming analysis are robust to including the responses that came after the two week cutoff.

¹⁹The reported p -values come from a linear probability model (LPM) where the binary dependent variable is *Email Response*. In the LPM, I include strata fixed effects, where the strata are defined in terms of legislator gender. I also employ robust standard errors to deal with potential heteroskedasticity in the LPM and cluster these errors on the email account to take account of the fact that observations using the same email account may not be completely independent.

Figure 2.2: Cumulative Response Rates by Level of Government



Note: Figure 2.2 shows the cumulative response rate across the different levels of government. The horizontal axis indicates the number of days since the original email message was sent; the last response was received 76 days after the original email was sent. The vertical axis indicates the cumulative proportion of responses received. The vertical dashed line at 14 days indicates the two-week cutoff for my upcoming analyses.

legislators in the United States, they found no evidence of gender discrimination by political elites against female aspirants. This result, which ran counter to expectations in the literature, raised concerns about external validity and made one wonder whether similar results would be found in other contexts. The results from my audit experiment in Canada also reveal no evidence of gender discrimination against female political aspirants. Indeed, unlike the study in the United States, I find evidence of a pro-women bias.

But does the gender of the legislator matter? According to the *Gender Affinity Hypothesis*, male legislators are expected to be more responsive to male political aspirants and female legisla-

There are 26 email accounts used in this study — one for each female and male name. One criticism of cluster-robust standard errors is that they're asymptotic to the number of clusters and I only have 26 email accounts (Wooldridge 2003, 135). My results are robust to not clustering, as well as using a difference-in-proportion test that employs a cluster-robust bootstrap procedure (Cameron and Trivedi 2010, 420-1). While I follow current practices in the experimental literature in using an LPM because it's easy to interpret and provides an unbiased estimate of the average treatment effect (Lin 2013; Judkins 2016), my inferences are also robust to using a simple difference-in-proportions test or estimating a logit model.

Table 2.2: Response Rates by Treatment Name and Legislator Gender

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
All Legislators	45% [42%, 48%] <i>N</i> = 888	52% [49%, 56%] <i>N</i> = 886	7 [2, 12]	0.01
Female Legislators	42% [36%, 49%] <i>N</i> = 259	54% [48%, 60%] <i>N</i> = 258	11 [4, 19]	0.01
Male Legislators	46% [42%, 50%] <i>N</i> = 629	52% [48%, 56%] <i>N</i> = 628	6 [-0.3, 12]	0.06

Note: The first two columns of Table 2.2 show the response rates to the email messages sent from male and female students for different sets of legislators. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-women bias. The 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

tors are expected to be more responsive to female ones. The fact that we already know that women are more likely on average to receive a reply than men is, on its face, an argument against the gender affinity story. Since there are more men in elected positions in Canada, the gender affinity story predicts that we should have observed a higher average response rate for male political aspirants. Nonetheless, we can examine the *Gender Affinity Hypothesis* more directly by looking at the response rates of female and male legislators separately. The core finding is that both female and male legislators exhibit a substantively large and statistically significant pro-women bias in their response rates. The pro-women bias exhibited by female legislators (11 percentage points) is, after accounting for rounding, twice as large as that exhibited by male legislators (6 percentage points).²⁰

²⁰These results differ slightly to those found in the United States. Kalla, Rosenbluth and Teele (2018) find that female legislators respond to men (27%) and women (27%) at similar rates, but that male legislators exhibit a small

There's evidence that this pro-women bias in response rates also translates into a pro-women bias in terms of meaningful responses and micro-mentorship more generally. To evaluate this requires looking at the actual content of the email responses. There are several ways to do this.²¹ One approach is to examine whether the responses contained substantive advice or not (Kalla, Rosenbluth and Teele 2018).²² For example, email responses in which the legislator suggested that the student should (i) run for student government, (ii) learn about the issues, (iii) attend local party or political meetings, (iv) learn to be extroverted, or (v) always put their values first were coded as providing substantive advice. About 26% of the responses were coded as giving substantive advice. Of the responses that didn't specifically offer substantive advice, many encouraged the student to call or set up a meeting. Overall, there's a statistically significant 3 percentage points pro-women bias among all legislators when it comes to receiving substantive advice. A second approach employs a new measure of elite responsiveness for audit studies developed by Costa (2020). This measure distinguishes between responses that are meaningful and those that only satisfy some minimum required effort on the part of the legislator. Using this "quality of response" measure, I again find a statistically significant pro-women bias (6 percentage points) among Canadian legislators. Finally, we might suspect that longer responses are more substantively meaningful than shorter ones. On this metric, I find that the responses to female political aspirants are significantly longer and contain more characters than those to male political aspirants.

While there's an overall pro-women bias in terms of responsiveness among legislators in Canada, it's still possible that there's a glass ceiling effect where female political aspirants do less well as

pro-women bias (27% versus 24%). There's no significant difference in the rates at which female and male legislators respond to female political aspirants. In neither study do male or female legislators ever exhibit gender discrimination against female political aspirants.

²¹Space constraints limit my ability to fully discuss the content of the email responses in the main text. However, a more in depth discussion can be found in Appendix A.5. I find an overall pro-women bias across nine different metrics; this bias is statistically significant in eight of these cases.

²²Evaluating the content of the email responses can lead to post-treatment bias as the responses are a consequence of the treatment (Montgomery, Nyhan and Torres 2018). To avoid this possibility, I redefine the outcome measure for my analysis of the content so that it's not conditional on having received a response (Coppock 2019). Practically-speaking, this means also coding non-responses as not providing substantive advice.

we move from the local to the national level. Table 2.3 provides information about response rates by treatment name, level of office, and legislator gender. There are three main sections that each relate to legislators at either the municipal, provincial, or federal levels. Within each section, the first row shows how the overall response rate at the specified level of government varies depending on whether the hypothetical student sending the email message is female or male. The second and third rows in each section show the response rates broken down by whether the legislator is female or male. Again, 95% confidence intervals are shown in square brackets. Contrary to the *Levels of Office Hypothesis*, there's a pro-women bias at all three levels of government. Significantly, the magnitude of this bias is fairly consistent across the different levels. Specifically, the pro-women bias is 7 percentage points at the municipal level, 8 percentage points at the provincial level, and 6 percentage points at the federal level. This overall pro-women bias is only statistically significant at conventional levels at the municipal level. However, this may well be due to the fact that the sample size shrinks markedly as we move from the municipal to the federal level. While a pro-women bias is exhibited by both female and male legislators at each level of government, the magnitude of the pro-women bias is typically larger for female legislators. At the municipal level, the pro-women bias exhibited by female legislators (12 percentage points) is 2.5 times larger than that exhibited by male legislators (5 percentage points). At the provincial level, the pro-women bias exhibited by female legislators (15 percentage points) is three times larger than that exhibited by male legislators (5 percentage points). There's no substantive difference in the magnitude of the pro-women bias across female and male legislators at the federal level. Indeed, it's only at the federal level that the pro-women bias exhibited by female legislators is not statistically significant.

Are there partisan effects? According to the *Left-Wing Partisan Hypothesis*, legislators from left-leaning parties will be more responsive to female political aspirants than legislators from right-leaning ones. In what follows, I focus on the three largest political parties that exist at the federal level in Canada: the New Democratic Party (NDP), the Liberal Party, and the Conservative Party. On a left-right ideological scale, the NDP is on the left, the Liberal Party is center-left, and the

Table 2.3: Response Rates by the Treatment Name, Level of Office, and Legislator Gender

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
<i>Municipal:</i>				
All Legislators	54% [49%, 59%] <i>N</i> = 421	61% [56%, 65%] <i>N</i> = 429	7 [2, 11]	0.01
Female Legislators	48% [39%, 57%] <i>N</i> = 123	60% [51%, 69%] <i>N</i> = 119	12 [1, 22]	0.03
Male Legislators	56% [51%, 62%] <i>N</i> = 298	61% [56%, 66%] <i>N</i> = 310	5 [-0.5, 10]	0.08
<i>Provincial:</i>				
All Legislators	42% [37%, 48%] <i>N</i> = 295	51% [45%, 56%] <i>N</i> = 295	8 [-2, 18]	0.12
Female Legislators	41% [31%, 51%] <i>N</i> = 93	56% [46%, 66%] <i>N</i> = 93	15 [-1, 31]	0.07
Male Legislators	43% [36%, 50%] <i>N</i> = 202	48% [41%, 55%] <i>N</i> = 202	5 [-6, 16]	0.35
<i>Federal:</i>				
All Legislators	28% [21%, 35%] <i>N</i> = 172	34% [27%, 41%] <i>N</i> = 162	6 [-4, 16]	0.21
Female Legislators	30% [16%, 45%] <i>N</i> = 43	35% [20%, 49%] <i>N</i> = 46	5 [-13, 22]	0.60
Male Legislators	27% [19%, 35%] <i>N</i> = 129	34% [25%, 42%] <i>N</i> = 116	6 [-4, 17]	0.20

Note: The first two columns of Table 2.3 show the response rates to the email messages sent from male and female students for different sets of legislators separated by level of office. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-women bias. The 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

Conservative Party is on the right. Demographically, Canadian parties differ in their gender composition. In line with the literature (Caul 1999, 2001; Kittilson 2006; Salmond 2006; Medeiros, Forest and Erl 2019), the proportion of female legislators in my sample is highest in the left-leaning NDP (101 men and 88 women) and lowest in the right-leaning Conservative Party (215 men and 48 women); the proportion of female legislators in the center-left Liberal Party is in between (200 men and 85 women). Ideologically, the three parties differ in their support for enhancing political access for women. The NDP is a social democratic party that has been at the forefront of enhancing

the electoral representation of women (Matland and Studlar 1996; Young and Cross 2003; Cross 2004; Pruysers and Cross 2016). Furthermore, the NDP and the Liberal Party have both taken more concrete steps than the Conservative Party to increase diversity and in particular increase the representation of women (Young and Cross 2003; Cross 2004; Pruysers and Cross 2016). At the provincial level, there's considerable variation in party systems, both in terms of the identities of the parties competing but also in the extent to which these parties are organizationally integrated with the parties competing at the national level (Thorlakson 2009). It's possible, however, to classify many of the provincial parties into broad ideological categories associated with the left, the center, and the right.²³ Elections at the municipal level tend to be non-partisan; as a result, my upcoming discussion focuses on the effect of partisanship at only the federal and provincial levels.

Table 2.4 provides information about response rates by treatment name, party ideological type, and legislator gender. There are three main sections in the table depending on whether the legislator is associated with a left, center, or right party. Within each section, the first row shows how the overall response rate in the specified party varies depending on whether the hypothetical student sending the email message is female or male. The second and third rows in each section show the response rates broken down by whether the legislator is female or male. As before, 95% confidence intervals are shown in square brackets. The legislators from all three party types exhibit an overall pro-women bias. However, in line with the *Left-wing Partisan Hypothesis*, the legislators from the left-wing parties respond at higher rates (57%) to female political aspirants than legislators from the right-wing parties (48%). Moreover, the pro-women bias exhibited by legislators from the left-wing parties (11 percentage points) is almost two times larger than that exhibited by legislators

²³Parties that are coded as left-wing include the Alberta New Democratic Party, the New Democratic Party, the New Democratic Party of British Columbia, the New Democratic Party of Manitoba, the New Democratic Party of Ontario and the Nova Scotia New Democratic Party. Parties that are coded as centrist include the Alberta Liberal Party, the Liberal Party, the Manitoba Liberal Party, the Nova Scotia Liberal Party, the Ontario Liberal Party, the Prince Edward Island Liberal Party, and the Alberta Party. Parties that are coded as right-wing include the Conservative Party, the Prince Edward Island Progressive Conservative Party, the Progressive Conservative Association of Nova Scotia, the Progressive Conservative Association of Alberta, the Progressive Conservative Party of Manitoba, the Progressive Conservative Party of Ontario, the Saskatchewan Party and the United Conservative Party. Provincial legislators from other parties are omitted from the upcoming analysis.

from the right-wing parties (6 percentage points). On the whole, both female and male legislators from each of the party types respond to female political aspirants at a higher rate than male political aspirants. The only time when this isn't the case comes when we look at female legislators from the right-wing parties; these particular legislators demonstrate a pro-men bias (−7 percentage points).

2.4 Conclusion

Women remain significantly underrepresented in politics in virtually every country in the world. The political recruitment process consists of three stages: (i) self-selection, (ii) party selection, and (iii) voter selection. While there's evidence that gender discrimination against women exists at all three of these stages, the earlier stages appear to be more problematic for women's representation (Fox and Lawless 2004, 2010; Thomas and Bodet 2013). In this chapter, I've examined whether there's elite gender discrimination in an informal phase of the recruitment process where political aspirants are seeking encouragement or mentorship from elites to start a political career. This is a particularly important phase of the recruitment process as studies have repeatedly shown that elite encouragement, especially for women, plays an influential role in getting political aspirants to put themselves forward as potential candidates (Fox and Lawless 2004; Niven 2006; Karpowitz, Monson and Preece 2017).

To identify if there's gender discrimination against female political aspirants, I employ an email audit experiment. Audit experiments are particularly well suited to investigating sensitive topics such as gender discrimination as they mitigate concerns that researchers might have with social desirability and selection biases and help overcome methodological problems with omitted variable and post-treatment biases that affect studies that rely on observational data. Audit studies also have the advantage that they provide us with a behavioral, as opposed to a reported behavioral or attitudinal, measure of discrimination. For those interested in increasing women's political representation, especially in Canada, the results from my experiment are promising. Overall, I find no evidence that Canadian legislators discriminate against female political aspirants who contact

Table 2.4: Response Rates by the Treatment Name, Party Ideology, and Legislator Gender

	Male Sender	Female Sender	Difference	<i>p</i> -value of Difference
<i>Left Party:</i>				
All Legislators	46% [35%, 56%] <i>N</i> = 90	57% [47%, 67%] <i>N</i> = 99	11 [-5, 27]	0.16
Female Legislators	44% [27%, 60%] <i>N</i> = 39	55% [41%, 70%] <i>N</i> = 49	11 [-13, 36]	0.34
Male Legislators	47% [33%, 61%] <i>N</i> = 51	58% [44%, 72%] <i>N</i> = 50	11 [-9, 31]	0.27
<i>Center Party:</i>				
All Legislators	35% [27%, 43%] <i>N</i> = 141	44% [36%, 53%] <i>N</i> = 144	9 [-2, 20]	0.12
Female Legislators	30% [15%, 45%] <i>N</i> = 40	51% [36%, 66%] <i>N</i> = 45	21 [6, 37]	0.01
Male Legislators	38% [28%, 47%] <i>N</i> = 101	41% [32%, 51%] <i>N</i> = 99	3 [-12, 19]	0.62
<i>Right Party:</i>				
All Legislators	42% [34%, 50%] <i>N</i> = 138	48% [39%, 57%] <i>N</i> = 124	6 [-9, 20]	0.41
Female Legislators	61% [39%, 82%] <i>N</i> = 23	54% [33%, 76%] <i>N</i> = 24	-7 [-32, 18]	0.59
Male Legislators	38% [29%, 47%] <i>N</i> = 115	47% [37%, 57%] <i>N</i> = 100	9 [-6, 24]	0.24

Note: The first two columns of Table 2.4 show the response rates to the email messages sent from male and female students for different sets of legislators separated by party ideology. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-women bias. The 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

them.²⁴ Indeed, legislators in Canada appear to be both significantly more responsive and helpful to female political aspirants than male ones. This pro-women bias, which exists at all levels of Canadian government, is stronger among female legislators and those associated with left-leaning parties.

One of the goals of this research was to examine the generalizability of the results from a similar

²⁴As mentioned in footnote 13, my audit experiment technically only speaks to gender discrimination as it relates to university-educated Anglo-Canadian female political aspirants.

audit experiment conducted by Kalla, Rosenbluth and Teele (2018) in the United States. That study found that there was no significant difference in the response rates of legislators to female and male political aspirants. One of the common concerns raised with experimental research has to do with external validity. To what extent do the results from one experiment generalize to other contexts? This concern is particularly pertinent when the findings of an experiment, like the one conducted by Kalla, Rosenbluth and Teele (2018), run counter to expectations in the literature. As my audit experiment indicates, the result that female political aspirants don't experience elite gender discrimination in this informal phase of the political recruitment process does generalize beyond the United States, at least to Canada. I encourage scholars to further evaluate the external validity of these findings by examining whether they generalize to additional cases.

While the responsiveness of legislators to female political aspirants who express an interest in politics doesn't necessarily imply that parties are actually going to nominate them as candidates,²⁵ the results of these audit experiments in the United States and Canada should be encouraging for women who are thinking about a career in politics. Furthermore, these results suggest that to the extent that gender discrimination against women does exist in the early stages of the political representation process, it doesn't occur in this informal phase — when political aspirants are seeking advice on how to start a political career — but at some other point in the process of going from a citizen to a legislator. On this point, it's worth noting that the email requests for advice in these audit experiments come from “self-starters” who had already self-identified as political aspirants and decided to reach out for help (Kalla, Rosenbluth and Teele 2018). It's possible that a gender bias exists even earlier in the recruitment process when men and women are thinking about whether a career in politics is for them. To the extent that women are concerned that they won't receive support if they put themselves forward, the results from these audit experiments should be particularly reassuring.

²⁵It's important to remember, for example, that elected legislators are not typically directly involved in the selection of party candidates (Carty and Eagles 2005; Sayers 1999; Cross 2002, 2006, 2016; Cross and Pruyers 2019; Pruyers and Cross 2016).

Ultimately, if women's political representation is to increase, it's important to identify exactly where gender discrimination occurs and why. Audit experiments are an important part of the methodological toolkit for studying gender discrimination as they can help us understand exactly where and how gender distortions are occurring.

3. DO VOTERS DISCRIMINATE AGAINST FEMALE CANDIDATES? EVIDENCE FROM AN EXPERIMENT IN SOUTH KOREA

Women remain significantly under-represented in politics. While women make up half of the population in every country, they comprise, on average, just 26% of the representatives in lower house legislatures globally (Inter-Parliamentary Union 2022). One potential cause of women's political under-representation is low mass demand for women representatives. Much of the research on gender discrimination has focused on mass demand and there is mixed evidence as to whether and when voters are unwilling to support female candidates (Schwarz and Coppock 2020). In fact, some analyses find that voters are *more* willing to vote for female candidates than male ones (Black and Erickson 2003; Lawless and Pearson 2008; Golder et al. 2017).

The vast majority of these studies, though, have focused on the advanced industrialized democracies in Western Europe and North America (e.g., USA: Lawless and Pearson (2008); Holman, Merolla and Zechmeister (2016); Teele, Kalla and Rosenbluth (2018); Doherty, Dowling and Miller (2019); Costa (2021); Kirkland and Coppock (Forthcoming); UK: Eggers, Vivyan and Wagner (2018); Germany: Senninger and Bischof (Forthcoming); Denmark: Dahl and Nystrup (2020)). Much less is known about how voters evaluate women candidates elsewhere, which raises questions about generalizability of findings to other countries and regions, and especially those with more conservative gender role attitudes.

It is also unclear whether the results of these past studies of voter bias reveal a preference for the candidate's sex or the candidate's gender expression – related to their expressed level of femininity/masculinity. Existing experimental research tends to focus on candidate sex and hence whether individuals are willing to vote for male or female candidates. In contrast, little attention has been paid to how candidates express their gender and the impact of gender norms. This is problematic because a preference for a male candidate might not necessarily be a preference for

men but a preference for *masculinity*. Likewise, a preference for a female candidate might not be a preference for women but a preference for *femininity*.

To examine the interaction between a candidate's sex and their expressed level of masculinity/femininity across different dimensions, I employ a conjoint survey experiment¹ with a fully-crossed factorial design in South Korea. My study is one of the few candidate choice experiments designed to study gender in Asia (Kage, Rosenbluth and Tanaka 2018; Ono and Yamada 2020; Horiuchi, Smith and Yamamoto 2020). Gender role attitudes tend to be fairly conservative in Asia (Ono and Yamada 2020), and in recent years, gender relations have become increasingly salient in the political sphere. This is the case in South Korea, for example, where recent data from the World Values Survey indicates more than half of the respondents in South Korea either agreed or strongly agreed with the claim that “men make better political leaders, and should be elected rather than women” (WVS 2020). That this large of a percentage of respondents report an explicit preference for men as political leaders suggests that low mass demand for women representatives may be a reason as to why South Korean women are under-represented in politics. Women currently comprise just 18% of the seats in the National Assembly (Inter-Parliamentary Union 2022).

The conjoint design has several advantages to examining voter gender bias. With a candidate choice conjoint experiment, the researcher asks respondents to review the profiles of two hypothetical candidates that are randomly generated from the set of attributes and then to choose between them. First, given that the attribute values are randomized, the design allows me to identify the effect of each candidate attribute on the probability of being selected as the preferred candidate. Put differently, the randomization provides new leverage to disentangle the effects of correlated attributes such as sex and political experience or sex and policy expertise. The conjoint design thus

¹Conjoint experiments are increasingly common (Bansak et al. 2019) and have recently been employed to study voting (Franchino and Zucchini 2014; Aguilar, Cunow and Desposato 2015; Carnes and Lupu 2016; Kirkland and Coppock Forthcoming; Horiuchi, Smith and Yamamoto 2020; Teele, Kalla and Rosenbluth 2018; Clayton et al. 2020), immigration attitudes (Hainmueller, Hangartner and Yamamoto 2015; Bansak, Hainmueller and Hangartner. 2016; Wright, Levy and Citrin 2016; Clayton, Ferwerda and Horiuchi 2021), and policy preferences (Ballard-Rosa, Martin and Scheve 2017; Bansak, Bechtel and Margalit 2021).

enables me to disentangle between the effect of candidate sex and masculine/feminine traits. This way, I can examine how a candidate's gender stereotype conforming/non-conforming attributes affects their electability beyond simply their sex. Moreover, given that I vary all the attributes and measure their effects on the same scale, the design allows me to examine the attributes' relative importance. For instance, I can compare the effect of candidate sex with that of having a particular issue specialization. Second, I can examine possible interactions in the effect of the candidate attributes (e.g., does a particular issue specializations matter more for male or female candidates?). This allows me to test the conditions under which some attributes matter more or less. Finally, the design allows me to consider interactions between respondent and candidate characteristics (e.g., who is more likely to prefer the female candidate?).

Based on responses from 2,258 South Korean voters, I find no evidence of voter discrimination against women candidates. Indeed, all else equal, I find the average effect of being a woman (versus a man) is associated with an approximately 2.6 percentage point increase (S.E. = 0.65) in being selected as the preferred candidate. This pro-woman bias is driven by women voters. When it comes to winning the election, though, all voters think that the male candidate has a better chance of winning. Being a woman (versus a man) is associated with an approximately 4.5 percentage point decrease (S.E. = 0.65) in the probability of being selected as the winning candidate. These findings suggest that we should be careful when interpreting results from these types of experiments as preferences and expectations about who will win interact to determine voting behavior. In other words, we should not necessarily infer voter behavior simply from voter preferences.

When it comes to how voters evaluate candidates who deviate from gender norms, I find that overall, voters tend to evaluate women and men candidates with the same attributes similarly, except for personality trait. When it comes to personality traits, voters tend to prefer candidates who run counter to gender stereotypes: they prefer women candidates who present a "tough" approach to politics and men candidates who present a "compassionate" approach. These findings

have important implications for both female and male candidates and party elites who recruit them.

3.1 Voter Preferences for Women Candidates

One potential cause of women's political under-representation is voter bias against women candidates. In vote choice experiments, scholars have sought to evaluate voter bias by examining whether individuals are willing to vote for male or female candidates (e.g. Kage, Rosenbluth and Tanaka (2018); Teele, Kalla and Rosenbluth (2018); Ono and Yamada (2020); Clayton et al. (2020)). In other words, is the average effect of being a female candidate (relative to a male candidate) on electoral support positive, negative, or null? Meta studies find that there is only mixed evidence as to whether and when voters are unwilling to support female candidates (Schwarz and Coppock 2020). In their meta-analysis of 67 survey experiments, Schwarz and Coppock (2020) find that there is an overall positive effect – the average effect of being a woman (relative to a man) is a gain of approximately 2 percentage points. Yet, the vast majority of these studies focus on advanced industrialized democracies in Western Europe and North America (Lawless and Pearson 2008; Holman, Merolla and Zechmeister 2016; Teele, Kalla and Rosenbluth 2018; Doherty, Dowling and Miller 2019; Costa 2021; Kirkland and Coppock Forthcoming; Eggers, Vivyan and Wagner 2018; Senninger and Bischof Forthcoming; Dahl and Nyrup 2020), which raises questions about generalizability of findings to other countries and regions, and especially to those with more traditional gender role attitudes. Indeed, Schwarz and Coppock (2020) report heterogeneous effects across contexts.

3.1.1 Effect of Candidate Sex on Candidate Evaluations

There are several reasons why we might expect voters – and especially those in countries with more traditional gender role attitudes – to exhibit gender bias when evaluating female political candidates. Research suggests that voters discriminate on the basis of sex and that voters are indeed biased against female political candidates (Bermeo and Bhatia 2017; Blackman and Jackson 2021; Ono and Yamada 2020). One potential reason why voters prefer male candidates is that they tend to

consider them as more appropriate for political office. The composition of most decision-making institutions in the world has historically been dominated by men, after all. In South Korea, women comprise just 18% of the seats in the National Assembly, despite making up half of the population, ranking 123 out of 193 countries in women's descriptive representation (Inter-Parliamentary Union 2022). Voters might then prefer male candidates because they expect elected officials to be men. This leads to the *Female Bias Hypothesis*.

Female Bias Hypothesis: On average, female candidates will receive a less favorable evaluation than male candidates.

There are reasons to believe that the level of bias against female candidates may vary depending on the sex of the voter because of in-group and out-group bias. Some voters prefer, when given the option – all else equal – to vote for a candidate of the same sex or the same race or some other shared characteristic – that is, engage in ‘affinity voting’ (Converse et al. 1961; Besley and Coate 1997; Dolan 1998; Sanbonmatsu 2002). A voter might feel that a candidate who shares her background is more likely to have faced similar experiences, to hold similar views, and to be more aligned on unforeseen issues that could arise during the upcoming electoral term (Mansbridge 1999). Consistent with the claim that people favor members of their in-group, Kim and Kweon (2022), for example, find that group threat is associated with negative attitudes toward legislative gender quotas among young men in South Korea. This reasoning suggests a gender affinity story in which male voters will prefer male candidates and female voters will prefer female candidates.

Affinity Voting Hypothesis: On average, candidates will receive a more favorable evaluation from voters who share their gender. In other words, female candidates are going to be evaluated more favorably by female voters and male candidates are going to be evaluated more favorably by male voters.

3.1.2 Effect of Candidate Gender Expression on Candidate Evaluations

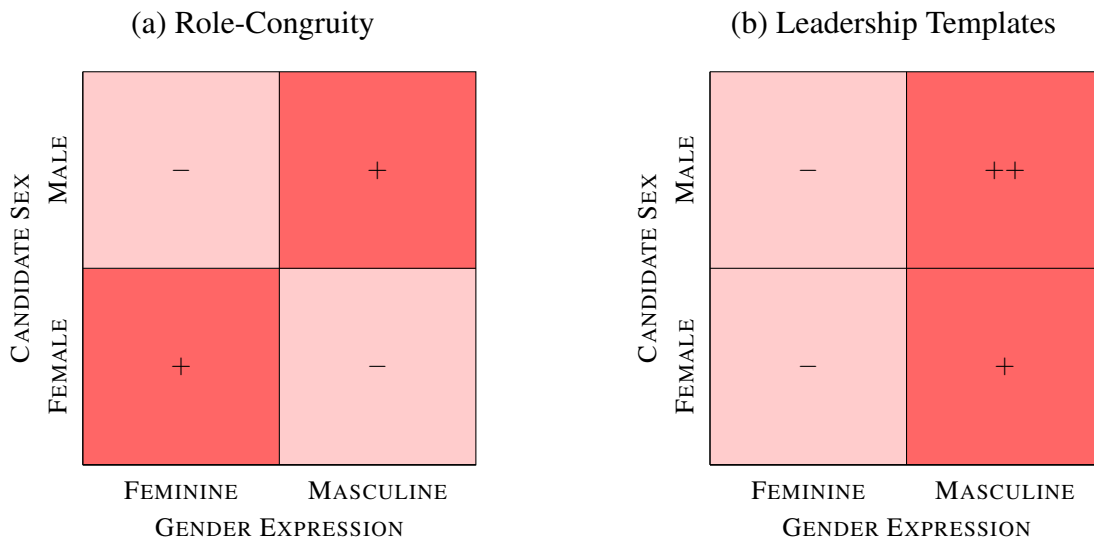
Apart from this aggregate bias against female candidates, voters may also discriminate against female political candidates in more implicit ways. Studies of voter discrimination, for example,

find that stereotypes are still pervasive in voters' evaluation of candidates (Bauer 2015). With few exceptions (Ono and Yamada (2020); Teele, Kalla and Rosenbluth (2018); Clayton et al. (2020); Saha and Weeks (2020)), most choice experiments tend to focus on the effect of candidate sex (e.g. Franchino and Zucchini (2014); Carnes and Lupu (2016); Aguilar, Cunow and Desposato (2015); Kirkland and Coppock (Forthcoming); Goggin, Henderson and Theodoridis G. (2020); Horiuchi, Smith and Yamamoto (2020)). It is therefore unclear whether the results from these studies reveal a preference for candidate sex or candidate gender expression. This is problematic because a preference for a male candidate might not necessarily be a preference for men but a preference for *masculinity*. Likewise, a preference for a female candidate might not be a preference for women but a preference for *femininity*.

The effect of candidate gender expression on candidate evaluation depends on the stereotypes that voters have about the kinds of attributes – or gender roles – that women and men are expected to embody (Eagly and Karau 2002). For example, women tend to be associated with more communal characteristics, like being more compassionate, caring, sensitive. In contrast, men tend to be associated with agentic characteristics including being more tough, aggressive, self-confident, and decisive (Alexander and Andersen 1993; Huddy and Terkildsen 1993; Lawless 2004). Role congruity theory proposes that people are more positively evaluated when their characteristics are recognized as aligning with that group's typical social roles (Eagly and Karau 2002). The theoretical relationship between candidate sex and gender expression is shown in Figure 3.1(a).² According to the role congruity theory, we should only expect a positive evaluation of political candidates when they conform to gender roles. In other words, women and men will be more positively evaluated when their gender expression (of gender roles) aligns with their sex. This reasoning leads to the *Role Congruity Hypothesis*.

²I have dichotomized candidate sex and gender expression in Figure 3.1 purely for presentational purposes. In reality, neither gender nor gender expression are dichotomous.

Figure 3.1: The Relationship between Candidate Sex and Gender Expression with Respect to Candidate Support



Role Congruity Hypothesis: On average, candidates who align with gender role stereotypes will receive a more favorable evaluation than candidates who diverge from gender role stereotypes.

Gender role stereotypes are thought to spill over into the workplace and, in turn, shape whether women and men are viewed as a good ‘fit’ for a job (Gutek and Morasch 1982; Eagly and Karau 2002; Koenig et al. 2011). When we consider what it takes to be a political leader, good leadership is often described in masculine terms, including being self-confident, decisive, objective, aggressive, competitive, etc (Huddy and Terkildsen 1993; Eagly 2007). Since stereotypically masculine traits are thought to be necessary to do the job and to be an effective leader, it is likely that voters value candidates with masculine gender expressions irrespective of the candidate being male or female. In other words, since the job is defined in masculine terms, those who embody a masculine gender expression will be evaluated more favorably than those who do not (Brescoll, Dawson

and Uhlmann 2010; Koenig et al. 2011). This reasoning also suggests that since leadership and masculine stereotypes are more congruent for men, embodying a masculine gender expression will matter more for them than for female candidates. The positive effect of embodying masculine attributes should be larger for men because when men express those masculine traits, they are also behaving in line with gender role expectations. Women, on the other hand, are in a double bind because when female candidates express these desired masculine traits, they are diverging from gender norm expectations. As Eagly and Karau (2002, 575) explain, “thinking about female leaders, people would combine their largely divergent expectations about leaders and women, whereas in thinking about male leaders, people would combine highly redundant expectations”, producing a disadvantage for women. The reasoning here is illustrated in Figure 3.1(b) and is captured in the *Leadership Template Hypothesis*.

Leadership Template Hypothesis: On average, candidates with masculine gender expressions will receive a more favorable evaluation than candidates with feminine gender expressions. The magnitude of this positive effect will be stronger among male candidates.

3.2 Voter Expectations for Women Winning

While understanding voter preferences is important to studying voter demand for women representatives, they do not necessarily indicate voter behavior. Unfortunately, we have growing evidence that we don't really know how attitudes translate into behavior or the extent that they do across contexts (Fazio and Roskos-Ewoldsen 2005; Gross and Niman 1975; Schuman and Johnson 1976; Jerolmack and Khan 2014). In other words, we can't simply look at preferences and assume that they directly feed into behavior. Even supposing that we find no gender differences or even a pro-woman bias in voters' reported preferences, it does not necessarily mean that voters are equally or more likely to vote for a woman candidate. Voters engage in strategic voting, where they take into account whether the candidate is likely to win. If they think that their more preferred candidate does not have a reasonable chance of winning, they could vote in favor of a less preferred

candidate who has a more realistic chance of winning. As such, I also evaluate how candidate sex and gender expression shape voters' perceptions of the candidate's electoral prospects.

Impressions about women's electoral prospects could be shaped by women's current limited presence in politics. That men currently dominate most decision-making institutions could have created the impression among voters that women are not viable candidates or representatives. Voters could thus use women's current low levels of representation as an indication of their likelihood to win the election. Women's under-representation could also indicate to some voters that the system or other voters are discriminatory against women, and so, even if they, themselves, prefer a female candidate, they might still think she does not have a realistic chance to win the election. This line of reasoning leads to the *Winning Hypothesis*.

Winning Hypothesis: On average, female candidate will receive a less favorable evaluation of their electoral prospects than male candidates.

3.3 Experimental Research Design

To assess the effect of candidate sex and gender expression on voter evaluations, I employ a choice-based conjoint design. Following a short introduction explaining the exercise, respondents were asked to review the profiles of two hypothetical candidates that were randomly generated from the set of attributes. They were then instructed to act as if they were a voter in their district and choose between the profiles of two hypothetical political candidates.

The profiles of the hypothetical candidates for the National Assembly are randomly generated from the set of eight attributes. Table 3.1 contains the full list of the attributes and values.³ Each of the attributes can take on multiple values and they were randomly assigned for each profile. The

³There are two reasons why I do not include a candidate's party label in the candidate profiles. First, since voters tend to vote along party lines (Dolan 2014; Bauer 2015; Holman, Merolla and Zechmeister 2016), including party affiliation would limit my ability to estimate the effects of the other candidate attributes included in the profiles. Second, from a practical perspective, excluding party label minimizes the possibility of creating implausible candidate profiles.

Table 3.1: Attributes for Candidate Profiles in Conjoint Experiment

Attributes	Values
<i>Sex</i>	Female Male
<i>Personality Trait</i>	Compassionate approach to politics Tough approach to politics
<i>Issue Specialization</i>	Social welfare Healthcare National security Foreign policy
<i>Ideological Placement</i>	Left-leaning Right-leaning
<i>Number of Years in Politics</i>	None 3 years 8 years
<i>Age</i>	35 years old 45 years old 65 years old
<i>Number of Children</i>	No children 1 child 3 children
<i>Marital Status</i>	Single Married Divorced

Note: Table 3.1 shows the attributes and attribute values that are used to generate the candidate profiles for the conjoint experiment.

attributes are also varied independently for each of the two candidates, allowing me to simultaneously measure (and compare) the independent effect of each value (Hainmueller, Hopkins and Yamamoto 2014). The attributes and values were chosen in line with the existing literature on gender stereotypes and “gendered” pathways to political office. To evaluate the effect of candidate sex on voter evaluations, I indicate whether the candidate is male or female. To evaluate the effect

of candidate gender expression on evaluations, I also include information about the candidate's *personality trait*, *issue specialization*, *ideological placement*, *number of years in politics*, *age*, *number of children*, and *marital status*. In what follows, I provide detail about each attribute in turn.

Informed by the literature of gender stereotypes (Alexander and Andersen 1993; Huddy and Terkildsen 1993; Kahn 1994, 1996; Lawless 2004), I randomize *personality trait*: whether the candidate has a compassionate or tough approach to politics. Studies find that voters make assumptions about a candidate's traits based on the candidate's sex (Alexander and Andersen 1993; Huddy and Terkildsen 1993; Kahn 1994). Women tend to be viewed as compassionate, sensitive, honest, collaborative, and caring; men as aggressive, tough, and decisive (Alexander and Andersen 1993; Huddy and Terkildsen 1993; Kahn 1996; Lawless 2004). Thus, the "compassionate approach to politics" attribute represents a stereotypically feminine trait and the "tough approach to politics" attribute represents a stereotypically masculine trait.⁴

Gender stereotypes are thought to spill over into issue area specializations (Alexander and Andersen 1993; Huddy and Terkildsen 1993; Koch 1999; Sanbonmatsu and Dolan 2009; Holman, Merolla and Zechmeister 2011, 2016). Women tend to be evaluated more favorably and perceived as more qualified to deal with stereotypically feminine policy areas, including healthcare, social welfare, childcare, poverty, and education. Men, on the other hand, are viewed more favorably in stereotypically masculine policy areas, like military, national defense, and foreign policy. Therefore, I randomize *issue specialization*: whether the candidate specializes in social welfare, healthcare, national security, or foreign policy. The "social welfare" and "healthcare" policy attributes represent a stereotypically feminine policy specialization, and the "national security" and "foreign policy" attributes represent a stereotypically masculine specialization.

Previous research finds that voters also draw inferences about the candidate's ideological po-

⁴Research, though, finds that women politicians do not tend to be viewed as having feminine traits (Schneider and Bos 2014).

sition based on the candidate's sex (McDermott 1997; Koch 2000, 2002). Women candidates and leaders tend to be perceived as more left leaning than their male counterparts (Koch 2000; O'Brien 2019). In line with this research, I randomize *ideological placement*: whether the candidate is left-leaning or right-leaning. The "left-leaning" ideological placement attribute thus represents a stereotypically feminine trait and the "right-leaning" ideological placement attribute represents a stereotypically masculine trait.

Informed by the literature on pathways to political office (Lawless and Fox 2005; Carroll and Sanbonmatsu 2013; Teele, Kalla and Rosenbluth 2018), I also randomize the candidate's *prior political experience* (as the number of years in politics – none, 3 years, or 8 years), *age* (35, 45, or 65 years old), *number of children* (no children, 1 child, or 3 children), and *marital status* (single, married, or divorced). Gender-role expectation also apply to one's family structure and could influence how voters evaluate candidates. I expect that voters will have a preference for a candidate with traditional family roles (married with children). This preference, however, poses a double bind for women and not for men because those women candidates who are married with children may be perceived as neglecting their familiar responsibilities and those who are single and without children may be perceived as violating societal gender-role expectations – namely not being married and having children – to pursue their political ambitions (Teale, Kalla and Rosenbluth 2018; Clayton et al. 2020). While it is unclear how female candidates will be evaluated based on their family status, in line with the *Role Congruency Hypothesis*, I expect that male candidates who are married with kids will be evaluated most favorably.

Below the candidate profiles, respondents were asked to evaluate the candidate as displayed in Figure 3.2 (the original design written in Korean is presented in the Appendix B.1). The first question asks respondents to report a preference for one of the candidates as a member of the National Assembly. The second question asks them to report which candidate they think has a better chance of winning the election. Finally, participants are asked to rate each of the candidates

Figure 3.2: Experimental Design

	Candidate 1	Candidate 2
Sex	Male	Female
Ideological Placement	Left-leaning	Right-leaning
Issue Specialization	National security	Healthcare
Personality Trait	Compassionate approach to politics	Tough approach to politics
Number of Years in Politics	3 years	8 years
Age	35	45
Marital Status	Single	Married
Number of Children	0	1

Based on the information above, if you had to choose between them, which candidate would you prefer as a member of Parliament?

Candidate 1
 Candidate 2

Based on the information above, which candidate do you think has a better chance of winning the election?

Candidate 1
 Candidate 2

Note: Figure 3.2 shows an example of one set of candidate profiles that was presented to a respondent in the conjoint experiment.

on a 10-point scale.⁵ Each respondent evaluates six comparisons between pairs of candidates, each displayed on a new screen.

3.3.1 Sample

My data come from a survey of Korean citizens 18 years and older administered online through Lucid Fulcrum Exchange platform⁶ in May 2020. The sample consists of 2,258 respondents.⁷

⁵Abramson, Kocak and Magazinnik (N.d.) have recently criticized candidate choice experiments for confusing vote choice and strength of preference. While I do not control for intensity of preference in my upcoming analysis, I intend to do so in future iterations of this paper.

⁶Lucid an online recruiting source that aggregates survey respondents from many respondent providers. Research examining Lucid samples suggests that they can be considered as representative samples (Coppock 2019).

⁷A more detailed description of the demographics of the sample is provided in Appendix B.2.

Because each respondent evaluated six pairs of candidates, there are a total of 27,096 evaluated profiles, or 13,548 pairings.

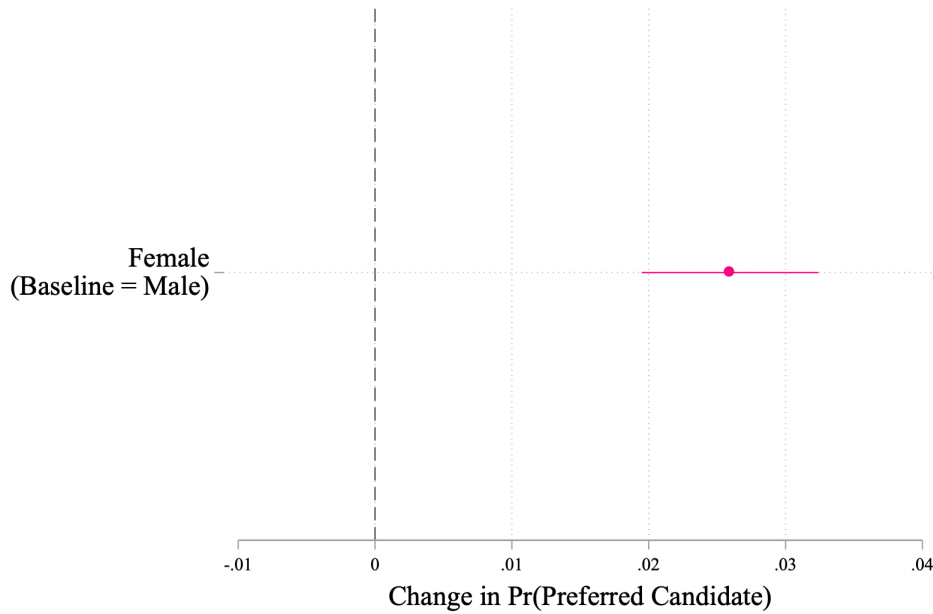
3.4 Results and Discussion

3.4.1 Effect of Candidate Sex on Being the *Preferred Candidate*

First, I evaluate the effect of candidate sex on being chosen as the preferred candidate. *Preferred Candidate* is coded 1 if the respondent preferred that candidate profile, and 0 otherwise. To test the causal effect of candidate sex on being the preferred candidate, I follow the statistical approach developed in Hainmueller, Hopkins and Yamamoto (2014) and estimate the *average marginal component effect (AMCE)*, which represents the marginal effect of a specific attribute over the joint distribution of all other attributes. Simply put, the AMCE of *Candidate Sex – Female* is the marginal effect of changing candidate sex from male to female, all else constant. This is estimated by regressing an indicator for whether the respondent chooses a given candidate on the various candidate characteristics listed above. I am able to estimate this simple linear regression because each attribute was randomly assigned independently of all other attributes. I use cluster-robust standard errors at the respondent level to take account of the fact that observations from the same respondent may not be completely independent.

Since I am interested in evaluating whether voters have a lower regard for female politicians, I focus on the average effect of being female (as opposed to male) on being chosen as the preferred candidate. The estimates are based on the benchmark regression model described above where the *Preferred Candidate* variable is regressed on sets of indicator variables for each level of each candidate attribute (omitting the reference categories). The full regression model and results for the other attributes are presented in Appendix B.3. Figure 3.3 displays the key result of the effect of *Candidate Sex* on being the *Preferred Candidate*. The dot indicates the point estimate for the AMCE of being a female candidate as opposed to male on the probability of being selected as

Figure 3.3: Effect of Candidate Sex on Being Preferred



Note: Figure 3.3 shows average marginal component effect (AMCE) of presenting a hypothetical candidate as female as opposed to male on the probability that respondents choose the profile as their preferred candidate. Estimates are based on the benchmark OLS model with clustered standard errors detailed in Appendix B.3; bars represent 95% confidence intervals.

the preferred candidate for the National Assembly.⁸ A positive (negative) AMCE indicates that the attribute makes a candidate more attractive (less attractive). The horizontal bars indicate 95% confidence intervals.

Recall that the *Female Bias Hypothesis* predicts that on average, female candidates will receive a less favorable evaluation than male candidates. Despite the low number of women in Korean politics and contrary to my expectations, I find that respondents do not exhibit any bias against female political candidates in their preferences. In fact, the average effect of being female (versus male) is associated with an approximately 2.6 percentage point increase (S.E. = 0.65) in being the preferred candidate.⁹ This boost for female candidates may even be underestimated since the randomly gen-

⁸For *Preferred Candidate*, the AMCE of female versus male is calculated by the proportion of all realized female candidate profile that were chosen, calculating the proportion of all male profiles that were chosen, and taking the difference between the two (Bansak et al. 2019).

⁹Following the advice of Leeper, Hobolt and Tilley (2020), I also report unadjusted marginal means results in Appendix B.3. My inferences remain robust – being a female candidate increases profile favorability and being a male

erated profiles include both mixed-sex (one male and one female candidate) and same-sex (two male or two female candidates) candidate pairings. Among the 13,548 evaluated candidate pairings, 6,743 (49.77%) are mixed-sex pairings and 6,805 (50.23%) are same-sex pairings. Indeed, when I exclude these same-sex pairings and evaluate only mixed-sex candidate pairings, I find that the average boost for female (versus male) candidates is 5.2 percentage points (S.E. = 1.3).¹⁰

But does this average effect vary by respondent gender? According to the *Gender Affinity Hypothesis*, female candidates are expected to be evaluated more favorably by female voters and male candidates are expected to be evaluated more favorably by male voters. As suggested by Leeper, Hobolt and Tilley (2020), I estimate differences in marginal means to evaluate these heterogeneous effects by respondent sex.¹¹ Figure 3.4 shows the effect of respondent sex on *Preferred Candidate* across female and male respondents.¹² The top figure shows the conditional marginal mean by respondent sex, and the bottom figure shows the differences in conditional marginal means. Conditional marginal means are interpreted as probabilities: for example, a marginal mean of 0.5 indicates that respondents select profiles with that feature level with probability 0.5. In line with the *Gender Affinity Hypothesis*, female respondents prefer female candidates and male respondents prefer male candidates. While less than 49% of male respondents would pick a female candidate as their preferred candidate, holding everything else equal, nearly 54% of female respondents would pick a female candidate. And, while around 46% of female respondents would pick a male candidate, holding everything else equal, nearly 51% of male respondents would pick a male candidate. These effects are matched in size by the difference in preferences between female and male respondents shown in Figure 3.4(b). Female and male voters indeed have different preferences when it comes to candidate sex.

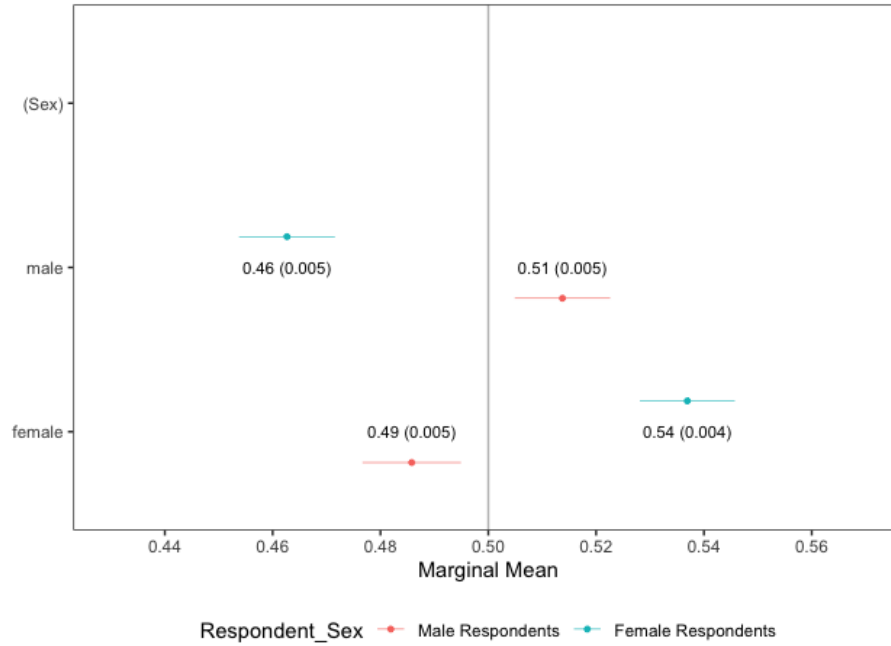
candidate decreases profile favorability.

¹⁰A more detailed discussion of this analysis can be found in Appendix B.5.

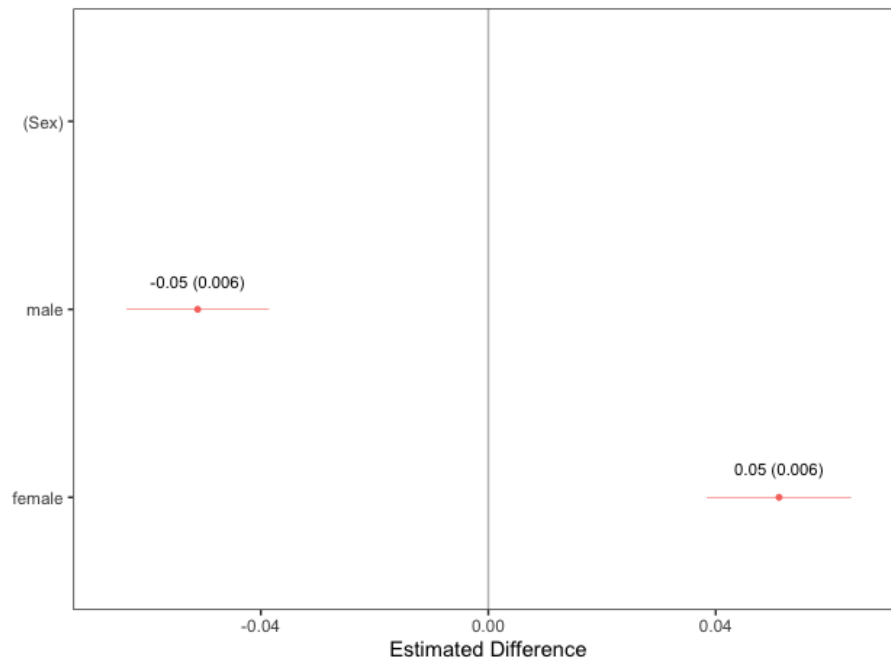
¹¹Leeper, Hobolt and Tilley (2020) recommend to estimate subgroup difference using conditional marginal means and differences between conditional marginal means instead of conditional AMCEs. They note that presenting the differences in the AMCE for subgroups analysis would yield to a misleading representation of the patterns of preferences because AMCE differences may be sensitive to the reference category chosen in the analysis.

¹²The results for the other attributes are presented in Appendix B.6.

Figure 3.4: Effect of Candidate Sex on Being Preferred by Respondent Sex



a) Marginal Means by Respondent Sex



b) Difference in Marginal Means across Respondent Sex

Note: Figure 3.4 shows the effect of candidate's sex on *Preferred Candidate* across respondent sex. Top plot shows marginal means and bottom plot shows the differences in marginal means. These plots show estimates of the effects of the randomly assigned candidates' sex on the probability of being chosen as the preferred candidate (for the sake of simplicity, the rest of the attributes are not displayed) by respondent sex. Estimates are based on the regression estimators with clustered standard errors; bars represent 95% confidence intervals.

3.4.2 Effect of Candidate Gender Expression on Being the *Preferred Candidate*

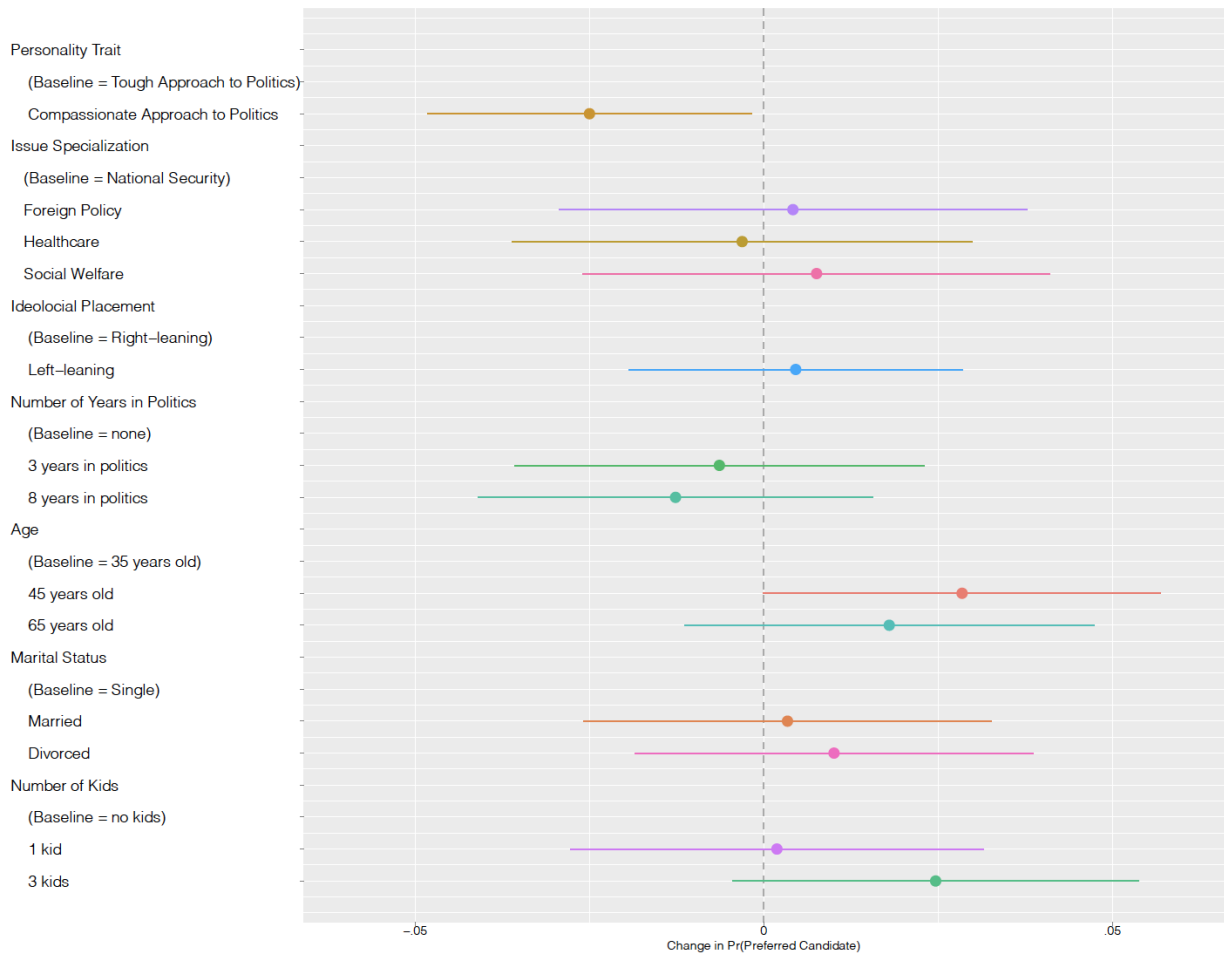
To evaluate the effect of candidate gender expression on being chosen as the preferred candidate, I interact candidate sex with the other candidate attributes and estimate average component interaction effects (ACIE). Figure 3.5 presents the plots of average component interaction effect estimators (with 95% confidence intervals) when the candidate is female. The ACIE estimates here represent the percentage point differences in the AMCEs of attributes between a male candidate and a female candidate. Each value indicates the extent to which our respondents reward or punish female candidates with a certain attribute. If the estimate of the attribute is positive (negative/zero), having that attribute (relative to the base category) is associated with an increase (decrease/no change) in the probability that the respondent picked that female candidate as the preferred candidate. Recall that the *Role Congruency Hypothesis* predicts that candidates who align with gender role stereotypes will receive a more favorable evaluation. Since the ACIE estimates reported in Figure 3.5 are for female candidates, the role congruity story predicts that the feminine attributes will have a positive effect. The leadership template story, on the other hand, predicts that masculine attributes will have a positive effect.

When it comes to *personality trait*, contrary to the gender congruity story, I find that women are actually punished when they conform to gender stereotypes in their approach to politics. Women candidates with a compassionate approach (versus a tough approach) to politics receive on average a 2.5 percentage points (S.E. = 0.12) decrease in being preferred. Note that this is also contrary to our leadership template story since male candidates with a compassionate approach to politics are rewarded.¹³

When it comes to gender congruity in term of *policy specialization* and *ideological placement*, I find that women candidates are neither rewarded nor punished for conforming to gender-based expectations. First, although national security is stereotypically masculine, I find no statistically

¹³Note that if we were to graph the ACIE estimates for male candidates, the estimates would be all symmetric to the estimates for female candidates shown in Figure 3.5.

Figure 3.5: Effects of Candidate Attributes on Being Preferred Conditional on Candidate Sex

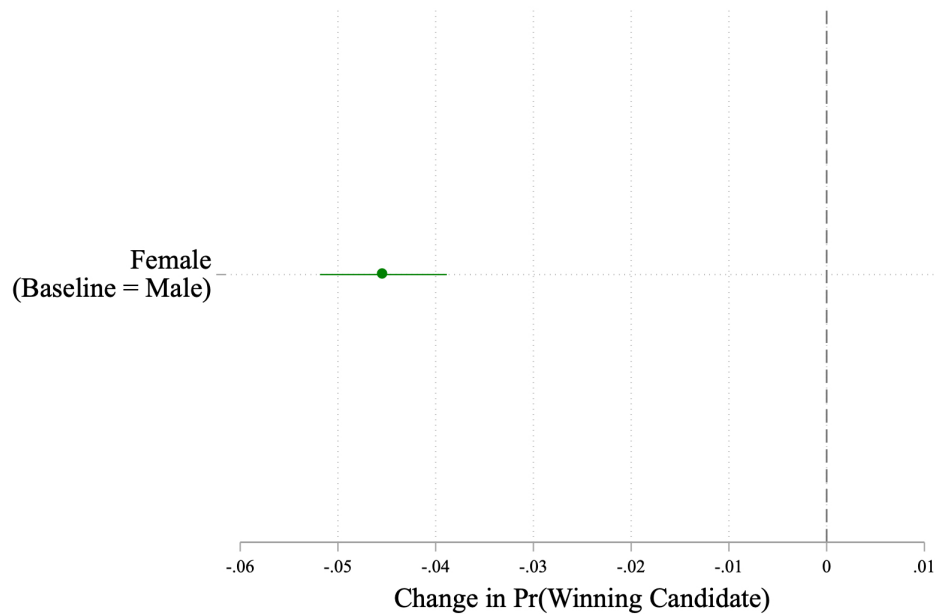


Note: Figure 3.5 presents the average component interaction effect (ACIE) estimators when the candidate is female. Bars represent 95% confidence intervals. It shows the difference between male and female candidates in their estimated average effects of the randomly assigned candidate attributes on the probability of being preferred by voters.

significant difference between being a national security expert and healthcare or social welfare expert. Second, while female candidates tend to be viewed as more liberal and progressive than their male counterparts (Koch 2000; McDermott 1997), I find no statistically significant difference between being right-leaning or left-leaning. These findings suggest that a candidate's policy specialization and ideological position do not systematically affect how voters evaluate female and male candidates.

In terms of having a more traditional family structure, I find that female candidates are not

Figure 3.6: Effect of Candidate Sex on Being Winner



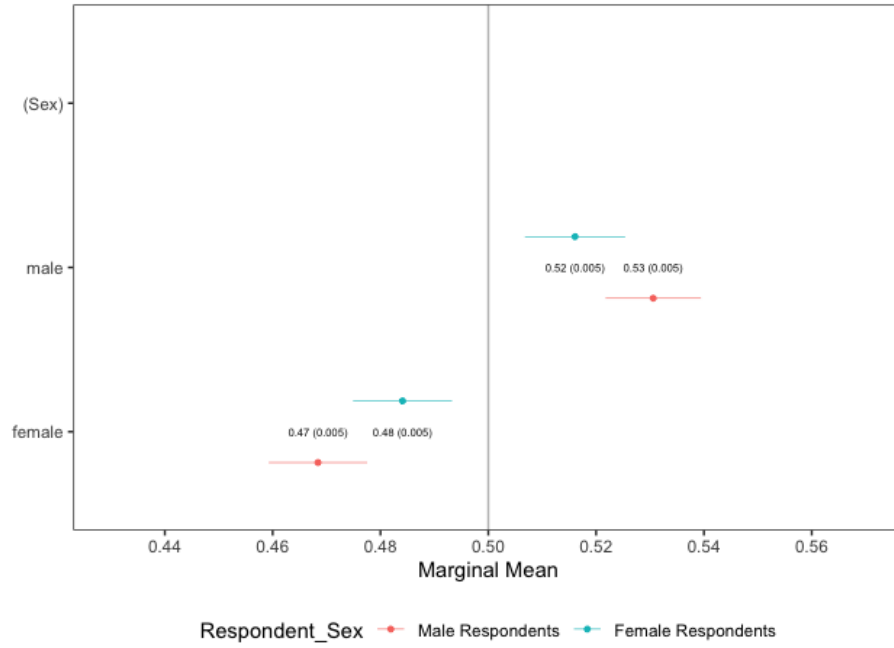
Note: Figure 3.6 shows average marginal component effect (AMCE) of presenting a hypothetical candidate as female as opposed to male on the probability that respondents choose the profile as their winning candidate. Estimates are based on the benchmark OLS model with clustered standard errors detailed in Appendix B.4; bars represent 95% confidence intervals.

rewarded when they are married (versus being single) and when they have children (versus being childless). When it comes to prior political experience, women with more political experience do not receive an additional bump in being preferred.

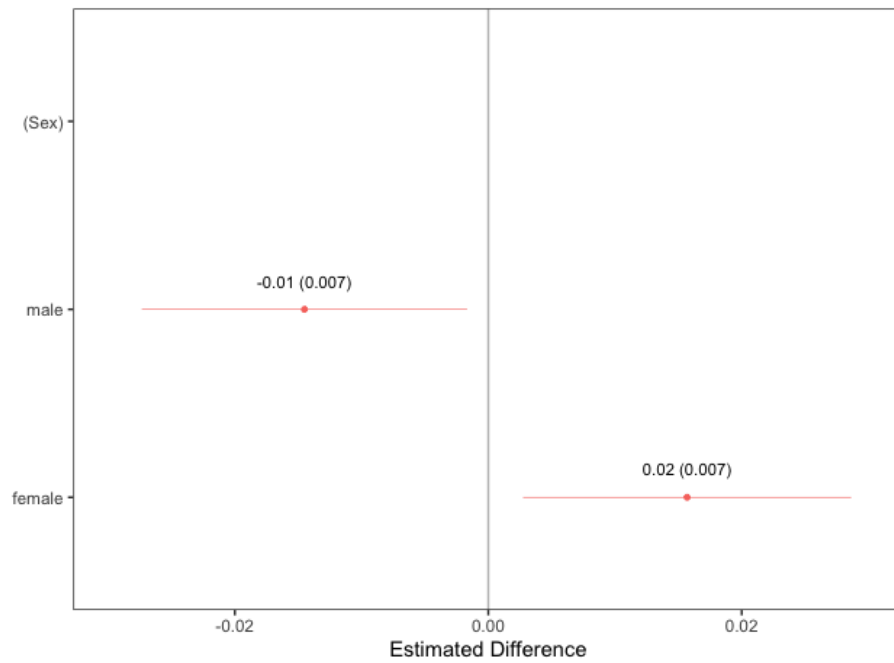
3.4.3 Effect of Candidate Sex on Being the *Winning Candidate*

Since we can't simply look at preferences and assume that they directly feed into behavior, I also evaluate the effect of being a woman candidate on voters' perceptions on her capacity to win. Recall that the *Winning Hypothesis* predicts that female candidate will receive a less favorable evaluation of their electoral prospects than male candidates. Figure 3.6 displays the AMCE for candidate sex on being chosen as the *Winning Candidate* for all respondents. The full regression model and results for the other attributes are presented in Appendix B.4. In line with the *Winning Hypothesis*, I find that despite the fact that the effect of candidate sex was associated with an in-

Figure 3.7: Effect of Candidate Sex on Being Winner by Respondent Sex



a) Marginal Means by Respondent Sex



b) Difference in Marginal Means across Respondent Sex

Note: Figure 3.7 shows the effect of candidate's sex on *Winning Candidate* across respondent sex. Top plot shows marginal means and bottom plot shows the differences in marginal means. These plots show estimates of the effects of the randomly assigned candidates' sex on the probability of being chosen as the winning candidate (for the sake of simplicity, the rest of the attributes are not displayed) by respondent sex. Estimates are based on the regression estimators with clustered standard errors; bars represent 95% confidence intervals.

crease in *Preferred Candidate*, when it comes to her prospects of winning the election, the average effect of being a woman (versus a man) is associated with an approximately 4.5 percentage point decrease (S.E. = 0.65).

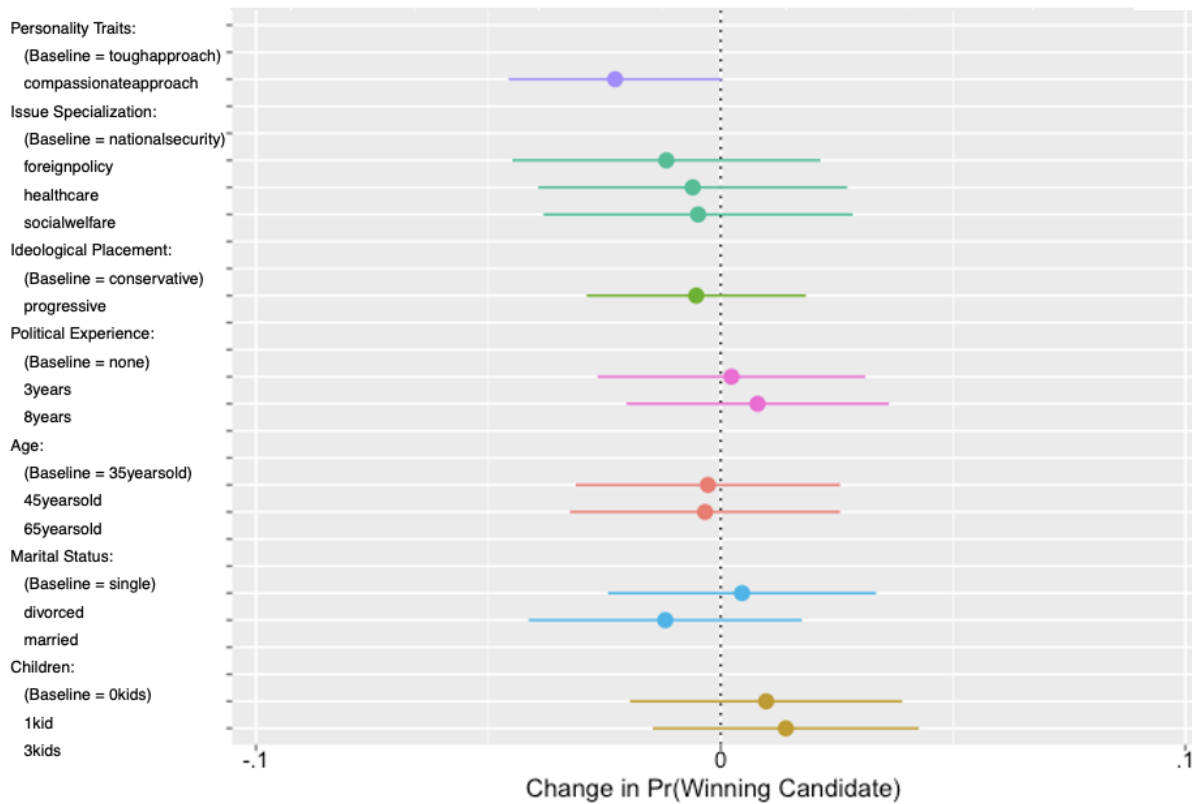
To evaluate whether female and male respondents have similar perceptions about her electoral prospects, I estimate differences in marginal means. Figure 3.7 shows the effect of respondent sex on *Winning Candidate* across female and male respondents. The top figure shows the conditional marginal mean by respondent sex, and the bottom figure shows the differences in conditional marginal means. Both female and male respondents think that the male candidate has a better chance of winning than the female candidate does.

3.4.4 Effect of Candidate Gender Expression on Being the *Winning Candidate*

To evaluate the effect of candidate gender expression on being chosen as the winning candidate, I interact candidate sex with the other candidate attributes and estimate average component interaction effects (ACIE). Figure 3.8 presents the plots of average component interaction effect estimators (with 95% confidence intervals) when the candidate is female. Recall that the ACIE estimates represent the percentage point differences in the AMCEs of attributes between a male candidate and a female candidate. If the estimate of the attribute is positive (negative/zero), having that attribute (relative to the base category) is associated with an increase (decrease/no change) in the probability that the respondent picked that female candidate as the winning candidate. Overall, there are no effects – it does not appear that conforming with gender roles makes a difference to voters' perceptions about her capacity to win the election.

In sum, the results from my candidate-choice experiment do not reveal either outright gender bias against women candidates or implicit bias against those who deviate from gender-role societal expectations. Except for the candidate's approach to politics, voters tend to evaluate women and men candidates with the same attributes similarly.

Figure 3.8: Effects of Candidate Attributes on Being Winner Conditional on Candidate Sex



Note: Figure 3.8 shows estimates of the effects of candidate sex (female) on the probability of winning the election by respondent attributes. Estimates are based on the benchmark OLS model with clustered standard errors detailed in Appendix B.4; bars represent 95% confidence intervals.

3.5 Conclusion

One explanation as to why women remain significantly under-represented in politics is low mass demand for women representatives. In this chapter, I evaluate whether (and how) South Korean voters discriminate against women candidates. To identify if voters discriminate against women political candidates, I employ a conjoint experiment that allows me to distinguish between the effect of candidate sex and the effect of candidate gender expression on voter evaluations. Overall, I find no evidence that South Koreans discriminate against women political candidates. Indeed, all else equal, I find that the average effect of being a woman (versus a man) is associated

with an approximately 2.6 percentage point increase (S.E. = 0.65) in being selected as the preferred candidate. This pro-woman bias is driven by women voters. Yet, despite indicating a preference for the female (as opposed to the male) candidate, voters do not think she has a better chance of winning the election. The average effect of being a female candidate on winning the election is negative and statistically significant even among women voters. These findings suggest that it is important that we are careful when interpreting the results from these types of experiments because preferences and expectations about who will win interact to determine voting behavior. We should not necessarily infer voter behavior simply from voter preferences. When it comes to how voters evaluate candidates who deviate from gender norms, I find that overall, voters tend to evaluate women and men candidates with the same attributes similarly. However, they do prefer candidates who run counter to gender stereotypes when it comes to personality traits. They prefer women candidates who have a tough approach to politics and men candidates who have a compassionate approach. These findings have important implications for both female and male candidates and the party elites who recruit them.

4. REEXAMINING THE SUPPLY AND DEMAND FRAMEWORK FOR EXPLAINING WOMEN'S DESCRIPTIVE REPRESENTATION

Historically, many scholars have sought to explain variation in women's descriptive representation in terms of demand-side and supply-side factors (Norris and Lovenduski 1993; Inglehart and Norris 2003; Matland 2005; Paxton, Kunovich and Hughes 2007). Broadly speaking, *supply-side factors* determine the size of the pool of women with the resources and ambition to run for office. When women have the resources and ambition to effectively compete for office, the 'supply' of qualified female candidates is high. In contrast, *demand-side factors* shape the preferences that individuals have for women representatives. When people want female representatives, the 'demand' for women's descriptive representation is high.

Despite widespread acceptance of the theoretical supply and demand framework, empirical results from cross-national studies have been somewhat mixed regarding the effect of supply-side and demand-side factors on women's representation. While some studies find that an increase in the supply of qualified female candidates is associated with significantly higher levels of women's descriptive representation (Moore and Shackman 1996; Paxton and Kunovich 2003), others find no such association (Norris and Lovenduski 1995; Paxton 1997; Kenworthy and Malami 1999; Kunovich and Paxton 2005). Perhaps to a lesser extent, mixed results also hold with respect to the impact of increased demand for female legislators on the level of women's representation. In this particular regard, much seems to depend on exactly how demand is operationalized (Paxton and Kunovich 2003).

In this chapter, I reexamine the utility of the supply and demand framework and argue that one potential explanation for these mixed results has to do with the gap between theory and empirics when it comes to evaluating the implications of the supply and demand framework. Specifically, all existing cross-national empirical studies treat supply-side and demand-side factors separately

and ignore the inherent theoretical interaction between supply and demand (Paxton 1997; Kenworthy and Malami 1999; Studlar and McAllister 2002; Paxton and Kunovich 2003; Tremblay 2007; Swiss 2009; Stockemer 2011; Rosen 2011; Ruedin 2012). Importantly, an increase in the supply of qualified women candidates should have only a limited effect on women's descriptive representation when voters exhibit little desire for female legislators. And similarly, an increase in the demand for female legislators should have little effect on women's representation, at least in the short run, when the supply of qualified women candidates is low. In effect, we should only expect to see high levels of women's descriptive representation when *both* supply and demand are sufficiently high. To appropriately evaluate this core aspect of the supply and demand framework requires utilizing an interactive, rather than an additive, research design (Brambor, Clark and Golder 2006)

My empirical analysis, which controls for key aspects of a country's institutional structure, provides the first cross-national test of the conditional implications at the heart of the supply and demand framework. Drawing on a new global dataset of women's descriptive representation from 1990 to 2018, I find, consistent with my expectations, that supply-side and demand-side factors interact to determine the level of women's descriptive representation. Specifically, I find that increased demand leads to higher levels of women's representation only when supply is sufficiently high and that increased supply leads to higher levels of women's representation only when demand is sufficiently high. These results are substantively important as they indicate the conditions under which we can expect supply-side factors such as increased female labor force participation and demand-side factors such as less traditional gender attitudes among the masses with respect to the holders of political office to actually translate into greater female political representation.

4.1 Theory: The Supply \times Demand Framework

Historically, many scholars have sought to explain cross-national variation in women's descriptive representation in terms of a supply and demand framework (Inglehart and Norris 2003;

Matland 2005; Paxton, Kunovich and Hughes 2007). ‘Demand’ refers to the preferences that individuals have for female legislators and is high when people want female legislators and are willing to vote for them (Black and Erickson 2003; Golder et al. 2017; Schwarz and Coppock 2020; Sevi, Arel-Bundock and Blais 2019; Dhima et al. 2022). In many respects, demand is determined by the extent to which voters hold traditional or progressive attitudes regarding gender roles, especially as they relate to the political sphere. As such, the level of demand for women’s descriptive representation in a country is likely shaped by both economic and cultural factors. Cultural modernization theory, for example, argues that human development leads to a shift away from traditional ‘survival values’ that focus on physical and economic security to more progressive ‘self-expression values’ that focus on individual autonomy and promote gender equality (Inglehart and Welzel 2005). Cultural factors related to religion are likely to be particularly important in determining the demand for women legislators. This is because religious denominations have varying attitudes towards ‘appropriate’ gender roles (Kloppenborg and Hanegraaff 1995; Franzmann 2000; Inglehart and Norris 2003) and thereby influence the opportunities for women’s emancipation and involvement in politics (Mayer and Smith 1985). Here too, though, economic factors play a role. For example, secularization theory suggests that economic development leads to a decline in both religiosity and socially conservative attitudes (Gaskins, Golder and Siegel 2013*a,b*; Dhima and Golder 2021).

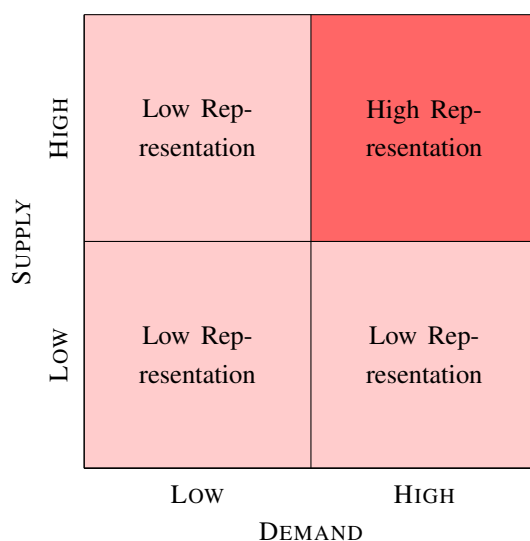
‘Supply’ refers to the size of the pool of potential female candidates with the resources and ambition to compete effectively for political office (Paxton, Kunovich and Hughes 2007; Fulton et al. 2006; Fulton 2012; Fulton and Dhima 2020). The supply of female candidates is strongly influenced by the level of women’s labor force participation and educational attainment (Matland 1998; Kenworthy and Malami 1999; Hughes 2009; Rosen 2011; Inglehart and Norris 2003; Burns, Schlozman and Verba 2001). This is because labor force participation and higher levels of educational attainment increase women’s professional skills and development, giving them the types of qualifications and experience that voters have come to expect in someone running for political office (Norris and Lovenduski 1993). As women become more educated and financially inde-

pendent, they, themselves, also desire to have access to more powerful leadership positions and political ambition grows (Inglehart and Norris 2003).

Existing cross-national studies of women's descriptive representation treat supply-side and demand-side factors as if they can be examined separately (Paxton 1997; Kenworthy and Malami 1999; Studlar and McAllister 2002; Paxton and Kunovich 2003; Tremblay 2007; Swiss 2009; Stockemer 2011; Rosen 2011; Ruedin 2012). In practice, this means estimating an additive model that includes both supply-side and demand-side variables. In many cases, the goal seems to be to examine whether supply or demand has the greatest influence on women's representation (Kenny 2013, 23). To a large extent, these studies have produced mixed and inconsistent results, especially as they relate to the impact of supply-side factors. One problem with all of these empirical studies is that they ignore the inherent interaction between supply and demand that exists at the theoretical core of the supply and demand framework (Lovenduski 2016).

The theoretical relationship between supply and demand is a conditional one in which the impact of supply-side factors on women's representation depends on the level of demand in a country and the impact of demand-side variables depends on the level of supply. An increase in demand for women legislators should have little effect on women's descriptive representation when the supply of qualified female candidates is low. While people may want to elect women legislators in these circumstances, there are few qualified female candidates for whom to vote. As the pool of qualified female candidates grows, though, increases in demand can be more accurately or efficiently translated into actual women's representation. Similarly, an increase in the supply of qualified female candidates should have little effect when the demand for women legislators is low. Although the pool of qualified women is high in these circumstances, people are unwilling to vote for female candidates and the number of elected women legislators remains low. As voter demand for women legislators grows, though, increases in the supply of qualified female candidates flows more seamlessly into higher levels of women's representation. The conditional implications of the supply and demand framework are shown graphically in Figure 4.1. We should only expect high

Figure 4.1: The Interaction between Supply and Demand with Respect to Women’s Descriptive Representation



levels of women’s representation when both supply and demand are sufficiently high. Levels will be low if supply, demand, or both supply and demand are too low. The following two hypotheses can be derived from the conditional supply and demand framework:

Supply Hypothesis: An increase in the supply of qualified female candidates has little effect on women’s representation when the demand for female representatives is low. However, it has an increasingly large positive effect as the demand for female representatives grows.

Demand Hypothesis: An increase in the demand for female representatives has little effect on women’s representation when the supply of qualified female candidates is low. However, it has an increasingly large positive effect as the supply of qualified female candidates grows.

4.2 Empirical Analysis

My dependent variable, *Women’s Representation*, captures the percentage of women representatives in lower house legislatures around the world. Data come from Paxton, Kunovich and

Hughes (2007) for the 1990-2003 period and the Inter-Parliamentary Union (via the World Bank Indicators) for the 2004-2018 period. *Women's Representation* ranges from a low of 0% to a high of 63.8%; its mean is 14.8 and its standard deviation is 10.9.

In terms of supply, I use two commonly-used alternative measures from the World Development Indicators (World Bank 2017) to proxy for the size of the pool of qualified female candidates in a country. The first measure, *Supply (Labor Market)*, captures the ratio of female to male labor force participation rates. The idea is that the supply of qualified female candidates grows as more women become employed and increase their participation in the public sphere (Norris and Lovenduski 1993; Manza and Brooks 1998; Fox and Lawless 2004). This variable takes on the value 100 when women participate in the labor market at the same rate as men. Values less than 100 indicate that women participate in the labor market less than men and values more than 100 indicate that women participate in the labor market more than men. *Supply (Labor Market)* has a mean of 67.88 and a standard deviation of 20.51; it varies from a low of 8.61 in Yemen in 2017 to a high of 111.01 in Mozambique in 2016.

The second measure of supply, *Supply (Tertiary Education)*, captures women's tertiary education gross enrollment ratio. Part of the idea here is that greater educational attainment, particularly a university education, is often important for a political candidate to be considered electorally competitive.¹ Many of the 'pipeline occupations' for members of Congress in the United States, such as law, business, and education, require a university degree (Fox and Lawless 2004). While the specific pipeline occupations for a career in politics may vary across countries, they are likely to be positively correlated with greater educational attainment. Part of the idea here is also that women's political ambition is likely to grow with higher levels of education (Inglehart and Norris 2003). *Supply (Tertiary Education)* has a mean of 25.82 and a standard deviation of 27.16; it varies from a low of 0 in Djibouti in 1991 to a high of 148.39 in Cuba in 2008.²

¹My upcoming inferences are robust to using secondary, instead of tertiary, education as a proxy for *Supply*.

²While unusual, it is possible for the value of *Supply (Tertiary Education)* to be larger than 100. This is because women's tertiary education gross enrollment ratio is calculated as the ratio of total female tertiary enrollment, *regard-*

Rather than use variables related to things like economic development, culture, and religion that are thought to be *determinants* of the level of demand for women's representation, I choose to measure voter demand directly by capturing mass preferences from survey data (Inglehart and Norris 2003; Paxton and Kunovich 2003). In particular, I use two alternative measures from the pooled 1981-2016 World and European Values Surveys (WVS 2015; EVS 2015). *Demand (Political Leaders)* captures the percentage of respondents in a country who disagree or strongly disagree with the claim that men make better political leaders than women. This variable has a mean of 52.01 and a standard deviation of 20.71; it varies from a low of 7.98% in Egypt in 2008 to a high of 91.82% in Sweden in 2006. *Demand (Jobs Scarce)* captures the percentage of respondents in a country who disagree or strongly disagree with the claim that men have more right to a job than women when jobs are scarce (Valdini 2012). This variable has a mean of 51.11 and a standard deviation of 22.21; it varies from a low of 0.4% in Egypt in 2001 to a high of 97.50% in Sweden in 2009. Both of these alternative measures of demand are scaled such that higher values indicate less support for traditional gender roles and hence greater demand for women's political representation.

I control for key aspects of a country's institutional structure by including variables for regime type and electoral rules. With respect to regime type, scholars have generally argued that democracies provide a more open opportunity structure for female legislators than dictatorships (Htun and Weldon 2010; Inglehart and Norris 2003; Inglehart, Norris and Welzel 2002; Walby 2004). *Democracy* is a continuous variable that runs from -10 to +10, with higher values indicating higher levels of democracy (Marshall, Gurr and Jaggers 2019). In terms of electoral rules, I focus on the use of gender quotas because they have the capacity to more directly and automatically influence the level of women's representation than other aspects of the electoral system such as the electoral formula (Dahlerup 2006; Frechette, Maniquet and Morelli 2008; Krook 2009; Edgell 2017; Hughes et al. 2019).³ *Effective Quota* is a dichotomous variable that equals 1 if a country has

less of age, to the size of the female population of the age group that officially corresponds to the tertiary level of education.

³As I demonstrate in Appendix C.3, my upcoming inferences are robust to also controlling for whether a country

an effective gender quota and 0 otherwise. A gender quota is considered effective if there is a 10% *de facto* threshold for either candidate or reserved seat quotas; there must also be strong sanctions for noncompliance with the quota or strong placement mandates on party lists (Hughes et al. 2017, 2019). More descriptive statistics on my sample and variables can be found in Appendix C.1.

I treat the dependent variable as continuous and estimate an ordinary least squares (OLS) regression with the following interactive model specification,

$$\begin{aligned} \text{Women's Representation} = & \beta_0 + \beta_1 \text{Supply} + \beta_2 \text{Demand} + \beta_3 \text{Supply} \times \text{Demand} \\ & + \beta_4 \text{Democracy} + \beta_5 \text{Effective Quota} + \epsilon. \end{aligned} \quad (4.2.1)$$

The interaction term $\text{Supply} \times \text{Demand}$ is included to capture the conditionality of the *Supply Hypothesis* and the *Demand Hypothesis*. The unit of analysis is the survey-country-year.⁴

The marginal effect of *Supply* on women's representation is

$$\frac{\partial \text{Women's Representation}}{\partial \text{Supply}} = \beta_1 + \beta_3 \text{Demand}. \quad (4.2.2)$$

According to the *Supply Hypothesis*, an increase in the pool of qualified female candidates should have little effect in the absence of any demand for female legislators. Thus, β_1 should be close to 0. However, an increase in supply should have an increasingly large positive effect as demand grows. This means that β_3 should be positive and that the positive effect of $\beta_1 + \beta_3 \text{Demand}$ should

employs a majoritarian or proportional electoral system (Rule 1987; Rule and Zimmerman 1994; Matland and Studlar 1996; Tremblay 2008; Kittilson and Schwindt-Bayer 2012).

⁴A potential issue with using OLS in the context of *Women's Representation* is that percentages are bounded between 0 and 100. I note, though, that my inferences are robust to using a two-sided censored tobit model. A second potential issue has to do with the fact that some countries provide more than one observation to the sample, raising the possibility that these observations are not independent. On this issue, I note that my inferences are robust to using robust standard errors clustered by country. Finally, I note that the structure of the data make it inadvisable to examine the temporal dynamics between the *Women's Representation*, *Supply*, and *Demand* variables. One reason for this is that the surveys that provide the information for the *Demand* variables do not always come at regular intervals either within countries or between countries. Another reason is that the panel nature of the data (small T , large N) means that the inclusion of lagged variables would significantly reduce the sample size and drop all singleton countries that only ever had one survey.

grow with higher levels of demand. The marginal effect of *Demand* on women’s representation is

$$\frac{\partial \text{Women's Representation}}{\partial \text{Demand}} = \beta_2 + \beta_3 \text{Supply}. \quad (4.2.3)$$

According to the *Demand Hypothesis*, an increase in demand for female legislators should have little effect when there’s no supply of qualified female candidates. As a result, β_2 should be close to 0. However, an increase in demand should have an increasingly large positive effect as supply grows. This means that β_3 should be positive and that the positive effect of $\beta_2 + \beta_3 \text{Supply}$ should grow with higher levels of supply.

The results from my analysis are shown in Table 4.1. As a point of comparison, the first four columns present the results from an additive model that includes all four possible combinations of my supply and demand variables. The next four columns present the results from the interactive model in Eq.4.2.1 for the same four combinations of my variables. In terms of the additive models, the main point to notice up front is that while there’s consistent support for the claim that an increase in demand for female legislators is associated with higher levels of women’s descriptive representation, there is, in line with existing studies, only mixed support for the claim that an increase in the supply of qualified female candidates boosts women’s descriptive representation. Much seems to depend on exactly how we measure supply. The results from the four interactive models are remarkably consistent and, as I will demonstrate, provide strong support for my hypotheses and the predicted conditionality at the core of the supply and demand framework. Importantly, all four models provide support for the claim that supply and demand interact to determine the level of women’s descriptive representation in a country. This is indicated by the positive and statistically significant coefficient on the interaction term *Supply*×*Demand* in all four models.

To graphically show that the results from the interactive models are largely similar, Figure 4.2 provides four panels, each of which corresponds to one of the interactive models, showing how the predicted level of women’s descriptive representation varies with the different possible

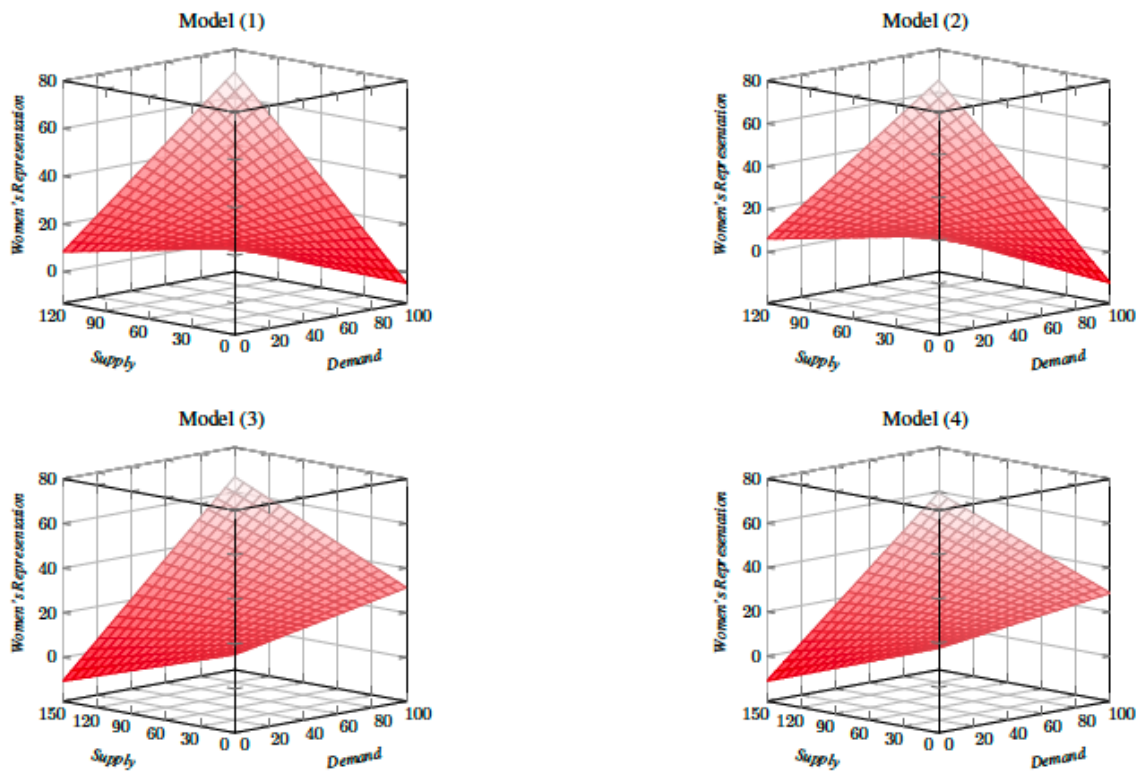
Table 4.1: Women's Legislative Representation, 1990-2018

Dependent Variable: <i>Women's Representation</i> , 0 – 100								
	Additive Models				Interactive Models			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
<i>Supply:</i>								
<i>Labor Market</i>	0.16** (0.04)	0.14** (0.04)			-0.12 (0.07)	-0.13* (0.06)		
<i>Tertiary Education</i>			0.05 (0.03)	0.04 (0.03)			-0.17 (0.09)	-0.19** (0.07)
<i>Demand:</i>								
<i>Political Leaders</i>	0.27** (0.04)		0.37** (0.05)		-0.28* (0.13)		0.16 (0.09)	
<i>Jobs Scarce</i>		0.23** (0.03)		0.30** (0.03)		-0.37** (0.10)		0.11 (0.06)
<i>Supply × Demand</i>					0.01** (0.002)	0.01** (0.001)	0.004* (0.002)	0.004** (0.001)
<i>Democracy</i>	-0.34* (0.13)	-0.16 (0.11)	-0.56** (0.17)	-0.33* (0.14)	-0.23 (0.13)	-0.06 (0.11)	-0.49* (0.17)	-0.25 (0.14)
<i>Effective Quota</i>	9.69** (1.64)	8.01** (1.48)	8.03** (2.00)	7.02** (1.70)	9.26** (1.57)	8.07** (1.39)	8.28** (1.96)	7.58** (1.66)
Constant	-7.84* (2.35)	-4.78* (2.17)	-2.67 (2.21)	1.23 (1.56)	11.57* (4.90)	13.58** (3.62)	7.15 (4.30)	10.26** (2.96)
Observations	186	283	134	219	186	283	134	219
R ²	0.47	0.40	0.45	0.41	0.52	0.48	0.47	0.44

Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$ (two-tailed).

combinations of values for *Supply* and *Demand*. Darker shades indicate lower levels of women's descriptive representation, while lighter shades indicate higher levels. The specific predicted values shown in Figure 4.2 are calculated for a fully dictatorial country that has an effective gender quota. Changing the level of democracy or the presence of an effective gender quota moves the surfaces shown in the 3-D plots up and down but does not change the nature of the surface. The key point to note is that women's descriptive representation is always highest when supply and demand are both high. In line with my predictions, the effect of *Demand* always becomes positive and increasingly

Figure 4.2: Visualizing the Conditional Relationship between *Supply*, *Demand*, and *Women's Representation*



Note: The surfaces show how the predicted value of *Women's Representation* varies across *Supply* and *Demand* and are based on the four interaction models shown in Table 4.1; I have assumed that *Democracy* = -10 and that there is an effective gender quota.

large as the value of *Supply* increases. This is seen in the way that the slope of the surface to the right becomes positive and steeper as the value of *Supply* becomes larger. Similarly, the effect of *Supply* always becomes positive and increasingly large as the value of *Demand* increases. This is seen in the way that the slope of the surface to the left becomes positive and steeper as the value of *Demand* becomes larger. These patterns, which are entirely consistent with the conditionality predicted by the supply and demand framework, are observed in all four plots.⁵

⁵Not too much should be read into the fact that the front two slopes of the surface are not flat and are sometimes negative. This is because there are no, or extremely few, real-world observations where *Supply* is 0 or *Demand* is 0.

While 3-D plots like those in Figure 4.2 are useful for visualizing the conditional relationship between *Supply*, *Demand*, and *Women's Representation* in a general sense, they do not provide us with the necessary information to fully evaluate my hypotheses. For example, they do not tell us if and when the effects of supply or demand are statistically significant. In what follows, I focus on the results from the first interactive model in Table 4.1 where supply is operationalized in terms of female participation in the labor market and demand is operationalized in terms of attitudes towards the gender of political leaders. As Appendix C.2 indicates, though, my upcoming inferences are qualitatively robust for all four interactive models.

Note that the coefficients on *Supply* and *Demand* tell us the effect of supply only when *Demand* is 0 and the effect of demand only when *Supply* is 0. As a result, these coefficients do not have much substantive relevance as the variables *Supply* and *Demand* are never 0 in the real world. To better evaluate my hypotheses, I use the results from the first interactive model in Table 4.1 to construct marginal effect plots for supply and demand (Brambor, Clark and Golder 2006). Figure 4.3(a) shows the marginal effect of *Supply* across the observed range of *Demand*. The dashed lines capture two-tailed 95% confidence intervals. The effect of supply is statistically significant whenever the upper and lower bounds of the confidence interval are both on the same side of the horizontal zero line. The gray histogram shows the percentage of respondents that exhibit different levels of demand. As predicted, the effect of supply on women's descriptive representation varies with the level of demand. The effect of supply is negative and statistically insignificant when $Demand < 16.36$ (3.4%), it is positive but statistically insignificant when $16.36 < Demand < 26.46$ (9.4%), and it is positive and statistically significant when $Demand > 26.46$ (87.2%). The numbers in parentheses indicate the percentage of observations that fall within each of these ranges. These results are consistent with the *Supply Hypothesis*, which predicts that the effect of supply on women's descriptive representation becomes positive once the level of demand in a country is sufficiently large and grows in magnitude as the level of demand increases further.

Figure 4.3(b) shows the marginal effect of *Demand* across the observed range of *Supply*. As

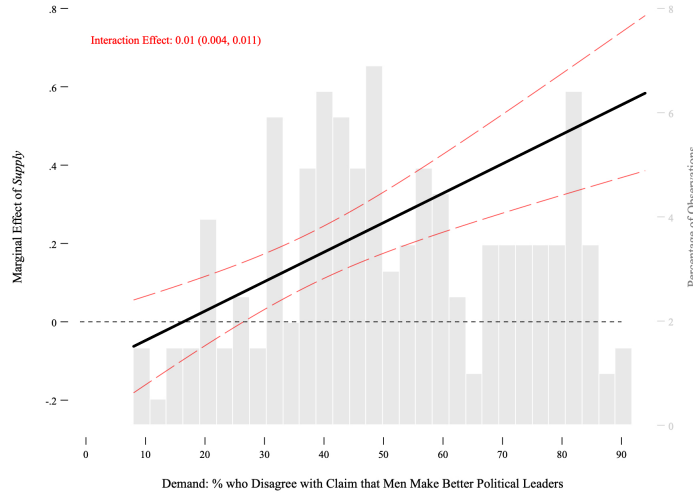
predicted, the effect of demand on women's descriptive representation varies with the supply of qualified female candidates in a country. *Demand* has a negative but statistically insignificant effect when *Supply* < 36.94 (10.9%), it has a positive but statistically insignificant effect when $36.94 < \textit{Supply} < 50.81$ (8.7%), and it has a positive and statistically significant effect when *Supply* > 50.81 (80.4%). These results are consistent with the *Demand Hypothesis*, which predicts that the effect of demand on women's descriptive representation becomes positive once the supply of qualified female candidates is sufficiently high and grows in magnitude as supply increases further.⁶

The conditional impact of supply and demand on women's descriptive representation is not just statistically significant, it is also substantively meaningful. Consider the effect of *Supply*. Increasing *Supply* by one standard deviation when *Demand* is at its mean value increases women's descriptive representation by 5.50 [3.81, 7.19] percentage points; two-tailed 95% confidence intervals are shown in brackets. This effect size is equivalent to a 25.3% increase in the 2018 mean level of women's legislative representation (21.7) around the world. Increasing *Supply* by the same amount when *Demand* is at its maximum observed value representation by 11.64 [7.69, 15.59] percentage points. This effect size is equivalent to a 53.6% increase in the 2018 mean level of women's legislative representation. These effect sizes are substantively large. The modifying effect of *Demand* is also substantively large. Indeed, the effect of a one standard deviation increase in *Supply* is 212% larger when *Demand* is at its maximum, as opposed to its mean, observed value. As predicted, the positive interaction effect indicates that the supply of qualified female candidates is more accurately translated into actual female representatives as the demand for female legislators

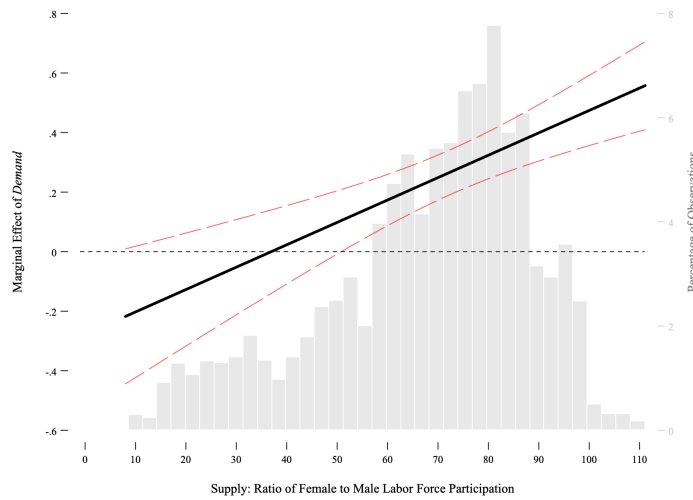
⁶That the effect of *Demand* is negative when the level of *Supply* is less than 36.94 could be read as inconsistent with the supply and demand framework. Note, though, that the negative effect of *Demand* is never statistically significant for observed values of *Supply* and relatively few observations in my sample have such low values of *Supply*. Note also that the negative effect of *Demand* at very low levels of *Supply* is likely an artefact of specifying an interactive model that assumes a *linear* interaction. To some extent, the supply and demand framework implies a threshold effect where the effect of demand (supply) should be positive only when supply (demand) is sufficiently high. This suggests that the effect of demand (supply) may be only *piecewise* linear across different values of supply (demand) (Berry, Golder and Milton 2012, 669-671). In Appendix C.4, I relax the assumption regarding the linearity of the interaction effect (Hainmueller, Mummolo and Xu 2019); my inferences remain robust.

Figure 4.3: The Conditional Effects of *Supply* and *Demand* on Women’s Descriptive Representation

(a) The Conditional Effect of *Supply*



(b) The Conditional Effect of *Demand*



Note: The plot in panel (a) shows the marginal effect of *Supply* on women’s descriptive representation across the observed range of *Demand*. It also shows the interaction effect between *Supply* and *Demand*. The plot in panel (b) shows the marginal effect of *Demand* on women’s descriptive representation across the observed range of *Supply*. The dashed lines indicate two-tailed 95% confidence intervals. The histograms show the percentage of observations at different values of each modifying variable. The two plots are based on the results from the first interactive model shown in Table 4.1.

increases. For example, a one point increase in supply leads to a 0.27 percentage point increase in women's descriptive representation when demand is at its mean and a 0.57 percentage point increase when it is at its maximum observed value.

Now consider the substantive effect of *Demand*. Increasing *Demand* by one standard deviation when *Supply* is at its mean value increases women's descriptive representation by 4.82 [3.21, 6.42] percentage points. This effect size is equivalent to a 22.2% increase in the 2018 mean level of women's legislative representation around the world. Increasing *Demand* by the same amount when *Supply* is at its maximum observed value increases women's descriptive representation by 11.54 [8.46, 14.62] percentage points. This effect size is equivalent to a 53.1% increase in the 2018 mean level of women's legislative representation. These effect sizes are substantively large. The modifying effect of *Supply* is also substantively large. Indeed, the effect of a one standard deviation increase in *Demand* is 239% larger when *Supply* is at its maximum, as opposed to its mean, observed value. As predicted, the positive interaction effect indicates that the demand for female representatives is more accurately translated into actual female representatives as the supply of qualified female candidates increases. For example, each percentage point increase in demand leads to a 0.23 percentage point increase in women's descriptive representation when supply is at its mean observed value and a 0.56 percentage point increase when supply is at its maximum observed value.

It is worth briefly comparing the results from the additive models, which are standard in the existing literature, with those from the more theoretically-appropriate interactive models. As previously noted, the additive models provide inconsistent evidence that the supply of qualified female candidates matters for women's descriptive representation. This is in line with much of the existing research (Paxton and Kunovich 2003; Paxton, Kunovich and Hughes 2007; Stockemer and Byrne 2012). As the interactive models demonstrate, though, this inconsistency is largely the result of mixing together countries where demand is low with countries where demand is high. In line with the conditionality at the theoretical core of the supply and demand framework, all of the interactive

models provide compelling evidence that the supply of qualified female candidates matters in a substantively meaningful way once the level of demand is sufficiently high.

The results from the additive models are also substantively misleading. For example, the first additive model indicates that a one-unit increase in supply is always associated with a 0.16 [0.09, 0.23] unit increase in women's descriptive representation. As the corresponding interactive model demonstrates, though, an increase in the supply of qualified female candidates has no significant effect when demand is less than 26.46 (12.8% of the sample).⁷ Indeed, the magnitude of the effect of supply is smaller than that reported by the additive model in 27.1% of the sample observations. Of course, this also means that the magnitude of the effect of supply is larger than that reported in the additive model in 72.9% of the sample observations. Indeed, when demand is at its maximum observed value, the effect of supply is more than three times larger than that indicated by the additive model. Similar differences occur with respect to the effect of demand. While the first additive model indicates that a one-unit increase in demand is associated with a 0.27 [0.19, 0.35] unit increase in women's descriptive representation, the corresponding interactive model indicates that the magnitude of the effect of demand ranges from a statistically insignificant low of -0.21 [$-.44, 0.01$] when supply is at its minimum observed value to a statistically significant high of 0.56 [0.41, 0.71] when supply is at its observed maximum. These types of difference really matter if we are trying to think about how evolving attitudes towards gender norms or policies designed to increase women's participation in the labor market or access to higher education are likely to affect women's descriptive representation in a particular country.

With respect to the institutional control variables, there is consistent evidence that effective gender quotas have a large and statistically significant positive effect on women's descriptive representation. There's no evidence that democracy promotes greater women's descriptive representation

⁷Countries with a *Demand* lower than 26.46 include Armenia (1997), Georgia (1996), Mali (2007), Ghana (2007, 2012), Nigeria (1995, 2000, 2012), Algeria (2014), Tunisia (2013), Libya (2014), Iran (2007), Iraq (2004, 2006, 2013), Egypt (2001, 2008, 2012), Jordan (2001, 2007, 2014), Saudi Arabia (2003), Kuwait (2014), Qatar (2010), Uzbekistan (2011), Pakistan (2012).

once we take account of supply-side and demand-side factors. Indeed, if anything, democracies are associated with less women's descriptive representation.

4.3 Conclusion

Historically, many scholars have sought to explain women's descriptive representation in terms of supply-side and demand-side factors. Cross-national empirical studies, though, have been characterized by mixed results. This has resulted in growing scepticism in some quarters about the relevance or utility of the supply and demand framework for explaining cross-national variation in levels of women's descriptive representation. In this chapter, I argue that we should not be too quick to discard the insights from the supply and demand framework. One potential explanation for the inconsistent results in the existing literature is the unfortunate gap between theory and empirics that exists when it comes to testing the implications of the supply and demand framework. To be specific, all existing *empirical* studies have treated supply-side and demand-side factors separately and ignored the inherent interaction between supply and demand that exists at the center of the supply and demand framework. Theoretically, we should expect to see high levels of women's descriptive representation only when the supply of qualified female candidates and voter demand for women legislators are both high. Controlling for key aspects of a country's institutional structure, this is precisely what I find when I test the conditional implications of the supply and demand framework on a new global dataset of women's legislative representation. The supply of qualified female candidates positively modifies the extent to which an increase in mass demand for women legislators is reflected in the percentage of actual female representatives. Similarly, mass demand for female legislators positively modifies the accuracy with which an increase in the supply of qualified female candidates is translated into actual female representatives. Overall, my analysis indicates that the supply and demand framework provides important insights into cross-national variation in women's descriptive representation, once it is modeled appropriately.

I finish by noting that I have focused on the theoretical interaction between the supply of quali-

fied female candidates and *mass* demand for women legislators. I did so because it is this particular interaction that has been ignored by all previous empirical studies of the supply and demand framework. As many scholars recognize, though, *elite* demand is also important for women's representation. For example, political elites can be important in encouraging women to run for office (Kalla, Rosenbluth and Teele 2018; Golder, Crabtree and Dhima 2019; Valdini 2019; Dhima Forthcoming) and can act as significant gatekeepers when it comes to putting female candidates on the ballot (Niven 2006; Sanbonmatsu 2006; Schwindt-Bayer 2011). To the extent that institutional choices, such as the adoption of effective gender quotas, provide an indirect measure of elite preferences for female candidates (Paxton, Kunovich and Hughes 2007, 266; Paxton, Hughes and Barnes 2020, 113), my analysis controls for elite demand. Future research, though, should look to develop more direct measures of elite preferences and extend existing theoretical and empirical analyses of the supply and demand framework to take greater account of the potentially very complex interplay between mass and elite demand for women legislators.

5. SUMMARY AND CONCLUSIONS

In this dissertation, I focus on the causes of women's political under-representation. I presented three papers that all fit into an overarching supply, demand, and institutions framework. Each of the papers in my dissertation studies women's political under-representation by examining different components of this framework. In what follows, I provide a short summary of each chapter and discuss possible avenues for future research.

In Chapters Two and Three, I focus on demand for women representatives. In Chapter Two, I focus on *elite demand*. To examine whether elites exhibit gender discrimination when responding to political aspirants, I use one of the first gender-focused politics field experiments ever conducted outside of the United States. Overall, I find no evidence that Canadian legislators discriminate against female political aspirants who contact them about how to start a career in politics. These results suggest that contemporary political elites in Canada may be open to increasing female political representation and, thus, should serve as welcome encouragement for women to pursue their political ambitions. In work elsewhere, I conducted a similar experiment in New Zealand (Golder, Crabtree and Dhima 2019). I again find no evidence of gender discrimination against female political aspirants. Teele, Kalla and Rosenbluth (2018) find similar results in the United States (Teele, Kalla and Rosenbluth 2018). We now have three experimental studies in three different contexts, all showing that elected officials are equally encouraging of both male and female political aspirants. This line of research is important because it suggests that political elites are open to increasing female political representation, at least in this very early and informal stage of the political recruitment process, and the overall message should be seen as encouraging for female political aspirants.

In Chapter Three, I focus on *mass demand* in South Korea. Overall I find that voters, on average, prefer female candidates rather than male candidates. Yet, while voters indicate a preference

for female as opposed to male candidates, they think that women have less chance of actually winning the election. These results suggest that we need to be careful about the kind of inferences we make about voter *behavior* from studies that just focus on voter *preferences*. Worryingly, we have growing evidence that we don't really know how attitudes translate into behavior or the extent that they do across contexts (Fazio and Roskos-Ewoldsen 2005; Gross and Niman 1975; Schuman and Johnson 1976; Jerolmack and Khan 2014). Additionally, there has been relatively little research studying gender bias in Asia, with much of the existing literature on women and politics focused instead on the United States and Europe. This is surprising because gender role attitudes tend to be fairly traditional in these countries and in recent years, gender relations have become increasingly salient in the political sphere. In Chapter Three, I evaluated mass demand in South Korea. However, I have also conducted similar experiments in Japan, Taiwan, and China. In future work, I plan to examine whether the results in South Korea generalize to these other countries and how the results might vary across the different institutional and cultural contexts. One possible avenue for future research that could be a natural follow-up for the work presented in Chapters Two and Three of my dissertation is to examine how attitudes translate into behavior and the extent that they do across different contexts. Ultimately, we want to know whether people engage in discriminatory behavior and how to mitigate it.

In Chapter Four, I focused on the theoretical interaction between the supply of qualified female candidates and *mass* demand for women legislators. I did so because, it is this particular interaction, while present in the theoretical literature, that has been ignored by all previous empirical studies of the supply and demand framework in a cross-national setting. As I note, though, preferences for women's representation also depend on *elite* demand and *institutional context*. In future work, I plan to incorporate elite demand more explicitly. Importantly, the empirical analysis in Chapter Four only controls for institutional context. In reality, though, I expect that institutional context modifies the way that supply-side and demand-side factors combine to affect women's political representation. In other words, I plan to extend the existing theoretical and empirical analyses of

the supply and demand framework to take greater account of the potentially very complex interplay between supply, mass demand, elite demand, and institutions.

In my post-dissertation research, I plan to start doing this by examining how supply-side and demand-side factors affect women's descriptive representation across different institutional contexts defined by a country's regime type. Supply-side factors and mass demand should matter less in dictatorship than in democracies. This is because institutions in dictatorships are not designed to be responsive to the preferences and ambitions of the citizenry as they are, at least in principle, in democracies. Institutions in dictatorships are, instead, primarily designed for cooptation and repression. Thus, mass demand and supply should have little effect in dictatorships. In contrast, elite demand should be much more important for women's representation in dictatorships. This is because political elite often have tools that they can use to promote their preferences for women legislators. Several factors are thought to influence elite demand for women's descriptive representation in dictatorships (Tripp 2019; Donno, Fox and Kaasik 2021). Dictatorial elites may wish to increase women's representation in order to signal their commitment to democracy (Bush 2011), secure foreign aid (Edgell 2017), or increase their country's domestic and international reputation (Bush and Zetterberg 2021), especially when there is an upcoming election (Valdini 2019; Melinda 2007; Longman 2006; David and Nanes 2011).¹ While elite demand should matter in dictatorships, it should be less important in democracies when we control for mass demand. This is because democratic elites are electorally accountable and so there should not be a significant divergence between mass and elite demand when it comes to women legislators. Thus, once we control for mass demand, elite demand should have limited effect on women's political representation in democracies. The conditional implications of this more complex supply, demand, and institutions story are graphically shown in Figure 5.1.

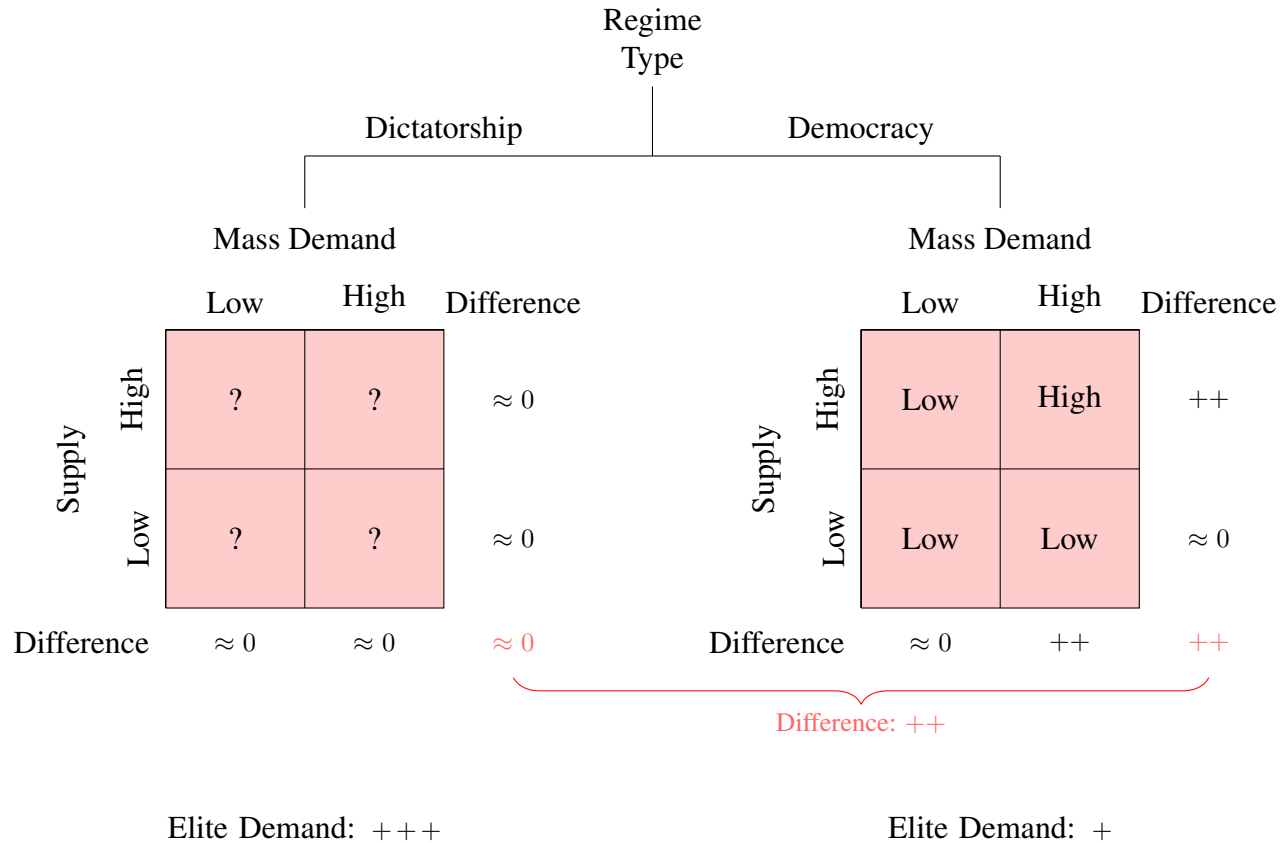
¹Note that the theoretical story that I have outlined simply assumes that we can identify some level of elite demand in democracies and dictatorships. It does not concern itself with whether this demand is driven by a sincere desire to increase women's descriptive representation or whether it is a strategic response to some external or internal pressure.

The conditional implications of the supply and mass demand story in democratic regimes are shown graphically in the right colored square in Figure 5.1.² The predicted effect of mass demand is the difference in the predicted level of women's descriptive representation as we move across the columns and is shown in the 'difference' column to the right of the colored square. We see that an increase in mass demand is expected to have a positive effect only when supply is high. The predicted effect of supply is the difference in the predicted level of women's descriptive representation as we move across the rows and is shown in the 'difference' row below the colored square. We see that an increase in supply is expected to have a positive effect only when mass demand is high. The interaction effect between mass demand and supply, which is predicted to be positive, is shown in light red to the bottom right of the colored square and is equivalent to the difference in the predicted effects shown in both the 'difference row' and the 'difference column'.

Unlike with mass demand, the effect of elite demand is arguably independent of the supply of qualified women. This is because elites are relatively free to place and appoint women to political positions as they see fit. For example, party leaders in closed list electoral systems can essentially guarantee the election of women candidates by placing them towards the top of their party lists. Similarly, party leaders in single-member district electoral systems can increase the number of women legislators by selecting more female candidates in 'safe' districts. Party leaders can also increase the number of women legislators by adopting gender quotas in party list systems or reserved seats. Theoretically, political elites in democracies are electorally accountable and, as a result, we should not see significant divergence in a country's level of elite and mass demand when it comes to women's descriptive representation. In effect, democratic political elites should be congruent with, and responsive to changes in, voter preferences for women legislators (Golder and Stramski 2010; Golder and Ferland 2018; Ferland and Golder 2020). To the extent that this is true, elite demand should have little impact on women's representation once we take account of

²I have dichotomized supply, mass demand, and regime type in Figure 5.1 purely for presentational purposes. In reality, supply, mass demand, and regime type are continuous.

Figure 5.1: Visualizing the Conditional Impact of Supply, Demand, and Institutions on Women’s Descriptive Representation



Note: The cells in the two colored squares indicate the predicted level of women’s descriptive representation for different combinations of supply, mass demand, and regime type. The predicted effect of *Mass Demand* is calculated as the difference in the predicted level of women’s descriptive representation as we move across the columns and is shown in the ‘difference’ column to the right of each colored square. The predicted effect of *Supply* is calculated as the difference in the predicted level of women’s descriptive representation as we move across the rows and is shown in the ‘difference’ row below each colored square. The predicted interaction effect between mass demand and supply is shown in light red to the bottom right of each colored square and is equivalent to the difference in the effects shown in both the ‘difference row’ and the ‘difference column’. The predicted modifying effect of regime type on the interaction effect between mass demand and supply is the difference in these two interaction effects and is shown in light red below the brace. The predicted effects of elite demand in dictatorships and democracies is shown below the respective colored squares.

mass demand. That said, existing studies suggest that elite and mass demand for women's descriptive representation are not always perfectly aligned (Broockman and Skovron 2018; Raines, Goodwin and Cutts 2017), at least in the short run. As a result, I expect an increase in elite demand to have a small positive effect on women's representation even after taking mass demand into account.

The conditional implications of the supply and mass demand story in authoritarian regimes are shown graphically in the left colored square in Figure 5.1. Since mass demand and supply are expected to be inconsequential in dictatorships, it is not possible to use them to predict when women's descriptive representation will be high or low. Since the level of women's descriptive representation in dictatorships has much more to do with whether political elites want to have women in the legislature than with whether the masses exhibit strong demand for women legislators, I expect that increases in elite demand should have a substantially larger positive effect on women's descriptive representation in dictatorships than democracies after taking account of mass demand.

Future research should thus look to extend existing theoretical and empirical analyses of the supply and demand framework to take greater account of the potentially very complex interplay between supply, mass demand, elite demand, and institutions. At minimum, such research could help us settle some of the inconsistent results in the existing literature when it comes to the impact of supply-side, demand-side, and institutional factors on women's descriptive representation.

Taken together, I hope that my dissertation research and these suggestions for future research avenues lead to important research that helps us better understand and overcome the causes of women's political under-representation.

Bibliography

- Abramson, Scott F., Korhan Kocak and Asya Magazinnik. N.d. “What do we learn about voter preferences from conjoint experiments?” *Working Paper*. Forthcoming.
- Aguilar, Rosario, Saul Cunow and Scott Desposato. 2015. “Choice Sets, Gender, and Candidate Choice in Brazil.” *Electoral Studies* 39(Supplement C):230–42.
- Ahler, Douglas J., Jack Citrin, Michael C. Dougal and Gabriel S. Lenz. 2017. “Face value? Experimental evidence that candidate appearance influences electoral choice.” *Political Behavior* 39(1):77–102.
- Ahmed, Ali M., Lina Anderson and Mats Hammarstedt. 2012. “Does age matter for employability? A field experiment on ageism in the Swedish labour market.” *Applied Economics Letters* 19(4):403–6.
- Ahmed, Ali M., Lina Anderson and Mats Hammarstedt. 2013. “Are gay men and lesbians discriminated against in the hiring process?” *Southern Economic Journal* 79(3):565–85.
- Alexander, Deborah and Kristi Andersen. 1993. “Gender as a factor in the attribution of leadership traits.” *Political Research Quarterly* 46(3):527–45.
- Anzia, Sarah F. and Christopher R. Berry. 2011. “The Jackie (and Jill) Robinson effect: Why do congresswomen outperform congressmen?” *American Journal of Political Science* 55(3):478–93.
- Baert, Stijn, Jennifer Norga, Yannick Thuy and Marieke Van Hecke. 2016. “Getting grey hairs in the labour market: An alternative experiment on age discrimination.” *Journal of Economic Psychology* 57:86–101.
- Ballard-Rosa, Cameron, Lucy Martin and Kenneth Scheve. 2017. “The structure of American income tax policy preferences.” *The Journal of Politics* 79(1):1–16.
- Bansak, Kirk, Jens Hainmueller, Daniel J. Hopkins and Teppei Yamamoto. 2019. *Cambridge*

- Handbook of Advances in Experimental Political Science*. New York: Cambridge University Press chapter Conjoint Survey Experiments.
- Bansak, Kirk, Jens Hainmueller and Dominik Hangartner. 2016. "How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers." *Science* 354(6309):217–22.
- Bansak, Kirk, Michael M. Bechtel and Yotam Margalit. 2021. "Why austerity? The mass politics of a contested policy." *American Political Science Review* 115(2):486–505.
- Barnes, Tiffany D. and Stephanie M. Burchard. 2012. "Engendering' Politics: The Impact of Descriptive Representation on Women's Political Engagement in Sub-Saharan African." *Comparative Political Studies* 46(7):767–90.
- Bashevkin, Sylvia. 1993. *Toeing the Lines: Women and Party Politics in English Canada*. Toronto: Oxford University Press.
- Bashevkin, Sylvia. 2009. Introduction. In *Opening Doors Wider: Women's Political Engagement in Canada*, ed. Sylvia Bashevkin. Vancouver: University of British Columbia Press pp. 2–18.
- Bashevkin, Sylvia. 2011. "Women's representation in the house of commons: A stalemate?" *Canadian Parliamentary Review* 34(1):17–22.
- Bauer, Nichole M. 2015. "Emotional, sensitive, and unfit for office? Gender stereotype activation and support for female candidates." *Political Psychology* 36:691–708.
- Baxter, Janeen and Erik Olin Wright. 2000. "The glass ceiling hypothesis: A comparative study of the United States, Sweden, and Australia." *Gender & Society* 14(2):275–94.
- Beckwith, Karen. 2000. "Beyond compare? Women's movements in comparative perspective." *European Journal of Political Research* 37(4):431–68.
- Bermeo, Nancy and Jasmine Bhatia. 2017. Strongmen or Technocrats? Experimental Evidence Testing Leadership Preferences in Afghanistan. Technical report Technical report International Growth Center.
- Berry, William D., Matt Golder and Daniel Milton. 2012. "Improving tests of theories positing

- interaction.” *Journal of Politics* 74(3):653–671.
- Bertrand, Marianne and Esther Duflo. 2017. Field Experiments on Discrimination. In *Handbook of Field Experiments*, ed. Esther Duflo and Abhijit Vinayak Banerjee. Amsterdam: Elsevier pp. 309–93.
- Bertrand, Marianne and Sendhil Mullainathan. 2004. “Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination.” *American Economic Review* 94(4):991–1013.
- Besley, Timothy and Stephen Coate. 1997. “An economic model of representative democracy.” *Quarterly Journal of Economics* 112(1):85–106.
- Bjarnegard, Elin. 2013. *Gender, Informal Institutions and Political Recruitment: Explaining Male Dominance in Parliamentary Representation*. London: Palgrave Macmillan.
- Bjarnegard, Elin and Meryl Kenny. 2015. “Revealing the “Secret Garden”: The informal dimensions of political recruitment.” *Politics & Gender* 11(4):748–53.
- Black, Jerome H. and Lynda Erickson. 2003. “Women candidates and voter bias: Do women politicians need to be better?” *Journal of Politics* 22(1):81–100.
- Blackman, Alexandra Domike and Marlette Jackson. 2021. “Gender Stereotypes, Political Leadership, and Voting Behavior in Tunisia.” *Political Behavior* 43(3):1037–66.
- Blais, André and Elisabeth Gidengil. 1991. “Making Representative Democracy Work: The Views of Canadians.” Volume 17. Research Studies of the Royal Commission on Electoral Reform and Party Financing. Toronto: Dundurn.
- Brambor, Thomas, William Roberts Clark and Matt Golder. 2006. “Understanding interaction models: Improving empirical analyses.” *Political Analysis* 14(1):63–82.
- Bratton, Kathleen A. and Leonard P. Ray. 2002. “Descriptive representation, policy outcomes, and municipal day-care coverage in Norway.” *Journal of Women, Politics, and Policy* 46:428–37.
- Brescoll, Victoria L., Erica Dawson and Eric Luis Uhlmann. 2010. “Hard won and easily lost: The fragile status of leaders in gender-stereotype-incongruent occupations.” *Psychological Science*

21(11):1640–2.

- Broockman, David E. 2013. “Black politicians are more intrinsically motivated to advance Black’s interests: A field experiment manipulating political incentives.” *American Journal of Political Science* 57(3):521–36.
- Broockman, David E. and Christopher Skovron. 2018. “Bias in perceptions of public opinion among political elites.” *American Political Science Review* 3(112):542–563.
- Brushman, Brad J. and Angelica M. Bonacci. 2004. “You’ve got mail: Using e-mail to examine the effect of prejudice on discrimination against Arabs.” *Journal of Experimental Political Science* 40(6):753–59.
- Burns, Nancy, Kay Lehman Schlozman and Sidney Verba. 2001. *The Private Roots of Public Action*. Cambridge, MA: Harvard University Press.
- Bush, Sarah. 2011. “International parties and the spread of quotas for women in legislatures.” *International Organization* 65(1):103–137.
- Bush, Sarah Sunn and Pär Zetterberg. 2021. “Gender quotas and international reputation.” *American Journal of Political Science* 65(2):326–341.
- Butler, Daniel M. 2014. *Representing the Advantaged: How Politicians Reinforce Inequality*. New York: Cambridge University Press.
- Butler, Daniel M. and David E. Broockman. 2011. “Do politicians racially discriminate against constituents?” *American Journal of Political Science* 55(3):463–77.
- Cameron, Adrian Colin and Pravin K. Trivedi. 2010. *Microeconometrics using Stata*. Vol. 2 College Station, TX: Stata Press.
- Carnes, Nicholas and Noam Lupu. 2016. “Do Voters Dislike Working-Class Candidates? Voter Biases and the Descriptive Underrepresentation of the Working Class.” *American Political Science Review* 110(4):832–844.
- Carroll, Susan J. and Kira Sanbonmatsu. 2013. *More Women Can Run: Gender and Pathways to the State Legislature*. Oxford: Oxford University of Press.

- Carty, Kenneth and Munroe Eagles. 2005. *Politics is local: National politics at the grassroots*. New York: Oxford University Press.
- Caul, Miki. 1999. "Women's representation in parliament: The role of political parties." *Party Politics* 5(1):79–98.
- Caul, Miki. 2001. "Political parties and the adoption of candidate gender quotas. A cross-national analysis." *Journal of Politics* 63(4):1214–29.
- Cheng, Tavits and Margit Tavits. 2011. "Informal influences in selecting female political candidates." *Political Research Quarterly* 64(2):460–71.
- Clark, Janet, Charles D. Hadley and R. Darcy. 1989. "Political ambition among men and women state party leaders: Testing the countersocialization perspective." *American Politics Research* 17(2):194–207.
- Clark, William Roberts and Matt Golder. 2015. "Big data, causal inference, and formal theory: Contradictory trends in political science." *PS: Political Science and Politics* 48(1):65–70.
- Clayton, Amanda, Amanda Lea Robinson, Martha C. Johnson and Ragnhild Muriaas. 2020. "(How) Do Voters Discriminate Against Women Candidates? Experimental and Qualitative Evidence From Malawi." *Comparative Political Studies* 53(3-4):601–30.
- Clayton, Amanda, Jennifer M. Piscopo and Diana Z. O'Brien. 2019. "All Male Panels? Representation and Democratic Legitimacy." *American Journal of Political Science* 63(1).
- Clayton, Katherine, Jeremy Ferwerda and Yusaku Horiuchi. 2021. "Exposure to immigration and admission preferences: Evidence from France." *Political Behavior* 43(1):175–200.
- Converse, Philip E., Angus Campbell, Warren E. Miller and Donald E. Stokes. 1961. "Stability and Change in 1960: A Reinstating Election." *American Political Science Review* 55(2):269–80.
- Coppock, Alexander. 2019. "Avoiding post-treatment bias in audit experiments." *Journal of Experimental Political Science* 6(1):1–4.
- Costa, Mia. 2017. "How responsive are political elites? A meta-analysis of experiments on public officials." *Journal of Experimental Political Science* 4:241–54.

- Costa, Mia. 2020. "Citizen evaluations of legislator-constituent communication." *British Journal of Political Science*. Published online January 30. <https://doi.org/10.1017/S0007123419000553>.
- Costa, Mia. 2021. "Ideology, not affect: What Americans want from political representation." *American Journal of Political Science* 65(2):342–58.
- Cotter, David A., Joan M. Hermsen, Seth Ovadia and Reeve Vanneman. 2001. "The glass ceiling effect." *Social Forces* 80(2):655–81.
- Crabtree, Charles. 2018. An Introduction to Email Audit Studies. In *Audit Studies: Behind the Scenes with Theory, Method, and Nuance.*, ed. S. Michael Gaddis. Cham, Switzerland: Springer pp. 103–19.
- Crabtree, Charles. 2019. "Measuring and Explaining Discrimination." Ph.D. Dissertation, University of Michigan.
- Cross, William Paul. 2002. Grassroots participation in candidate nomination. In *Citizen Politics: Research and Theory in Canadian Political Behavior*, ed. Joanna Everitt and Brenda O'Neill. New York: Oxford University Press.
- Cross, William Paul. 2004. *Political Parties*. Vancouver: University of British Columbia Press.
- Cross, William Paul. 2006. Candidate nomination in Canadian political parties. In *The Canadian General Election of 2006*, ed. Jon Pammett and Christopher Dornan. Toronto: Dundurn Press pp. 373–85.
- Cross, William Paul. 2016. "The importance of local party activity in understanding Canadian politics: Winning from the ground up in the 2015 federal election." *Canadian Journal of Political Science* 49(4):601–20.
- Cross, William Paul and Scott Pruysers. 2019. "The local determinants of representation: Party constituency associations, candidate nomination, and gender." *Canadian Journal of Political Science* 52(3):557–74.
- Dahl, Malte and Jacob Nyrup. 2020. "Confident and cautious candidates: Explaining underrepresentation of women in Danish municipal politics." *European Journal of Political Research*

60(1):199–224.

Dahlerup, Drude, ed. 2006. *Women, Quotas, and Politics*. New York: Routledge.

Daniel, William Wentworth. 1968. *Racial Discrimination in England: Based on the P.E.P. Report*. London: Penguin Books.

Dassonneville, Ruth and Ian McAllister. 2018. “Gender, Political Knowledge, and Descriptive Representation: The Impact of Long-Term Socialization.” *American Journal of Political Science* 62(2):249–65.

David, Assaf and Stefanie Nanes. 2011. “The women’s quota in Jordan’s municipal councils: international and domestic dimensions.” *Journal of Women, Politics & Policy* 32(4):275–304.

Desposato, Scott and Barbara Norrander. 2009. “The Gender Gap in Latin America: Contextual and Individual Influences on Gender and Political Participation.” *British Journal of Political Science* 39(1):141–62.

Dhima, Kostanca. Forthcoming. “Do elites discriminate against female political aspirants? Evidence from a field experiment.” *Politics & Gender* .

Dhima, Kostanca and Matt Golder. 2021. “Secularization theory and religion.” *Politics & Religion* 14(1):37–53.

Dhima, Kostanca, Sona Golder, Laura Stephenson and Karine Van der Straeten. 2022. “Permissive electoral systems and descriptive representation.” *Electoral Studies* .

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

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

Dolan, Kathleen. 2014. *When Does Gender Matter?: Women Candidates and Gender Stereotypes in American Elections*. New York: Oxford University of Press.

Donno, Daniela, Sara Fox and Joshua Kaasik. 2021. “International incentives for women’s rights

- in dictatorships.” *Comparative Political Studies* p. 00104140211024306.
- Dyck, Rand. 1998. *Canadian Politics: Concise Edition*. Scarborough: Nelson.
- Eagly, Alice H. 2007. “Female leadership advantage and disadvantage: Resolving the contradictions.” *Psychology of Women Quarterly* 21(1):1–12.
- Eagly, Alice H. and Steven J. Karau. 2002. “Role congruity theory of prejudice toward female leaders.” *Psychological Review* 109(3):573–98.
- Edgell, Amanda B. 2017. “Foreign aid, democracy, and gender quota laws.” *Democratization* pp. 1–21.
- Eggers, Andrew C., Nick Vivyan and Markus Wagner. 2018. “Corruption, Accountability, and Gender: Do Female Politicians Face Higher Standards in Public Life?” *The Journal of Politics* 80(1):321–326.
- Erickson, Lynda. 1991. *Women and Candidates for the House of Commons*. Toronto: Dundurn Press.
- Escobar-Lemmon, Maria C. and Michelle M. Taylor-Robinson. 2009. “Getting to the top: Career paths of women in Latin American cabinets.” *Political Research Quarterly* 62(4):685–99.
- EVS. 2015. “European Values Survey Longitudinal Data File 1981-2018 (EVS 1981-2008).” GESIS Data Archive, Cologne. ZA4804 Data File Version 3.0.0.
- Fazio, Russell H. and David R. Roskos-Ewoldsen. 2005. *Persuasion: Psychological Insights and Perspectives*. Thousand Oaks, CA: SAGE Publications chapter Acting as We Feel: When and How Attitudes Guide Behavior, pp. 41–62.
- Fenno, Richard F. 1978. *Home Style: House Members in Their Districts*. Boston: Little, Brown.
- Ferland, Benjamin and Matt Golder. 2020. “Citizen representation and electoral systems.” *The Oxford Research Encyclopedia*. Oxford University Press.
- Ferree, Myra Marx and Bandana Purkayastha. 2000. “Equality and cumulative disadvantage: Response to Baxter and Wright.” *Gender & Society* 14(6):809–13.
- Fiorina, Morris P. 1989. *Congress, Keystone of the Washington Establishment*. New Haven, CTY:

Yale University Press.

- Folke, Olle and Johanna Rickne. 2016. "The glass ceiling in politics: Formalization and empirical tests." *Comparative Political Studies* 49(5):567–99.
- Fox, Richard L. and Jennifer L. Lawless. 2004. "Entering the arena? Gender and the decision to run for office." *American Journal of Political Science* 48(2):264–80.
- Fox, Richard L. and Jennifer L. Lawless. 2010. "If only they'd ask: Gender, recruitment, and political ambition." *Journal of Politics* 72(2):310–26.
- Fox, Richard L. and Jennifer L. Lawless. 2011. "Gendered perceptions and political candidacies: A central barrier to women's equality in electoral politics." *American Journal of Political Science* 55(1):59–73.
- Fox, Richard L. and Jennifer L. Lawless. 2014. "Uncovering the origins of the gender gap in political ambition." *American Political Science Review* 108(3):499–519.
- Fox, Richard L. and Zoe M. Oxley. 2003. "Gender stereotyping in state executive elections: Candidate selection and success." *Journal of Politics* 65(3):833–50.
- Franchino, Fabio and Francesco Zucchini. 2014. "Voting in a Multi-dimensional Space: A Conjoint Analysis Employing Valence and Ideology Attributes of Candidates." *Political Science Research and Methods* 3(2):221–41.
- Franzmann, Majella. 2000. *Women and Religion*. New York: Oxford University of Press.
- Frechette, Guillaume R., Francois Maniquet and Massimo Morelli. 2008. "Incumbents' interests and gender quotas." *American Journal of Political Science* 52(4):891–909.
- Fulton, Sarah A. 2012. "Running backwards and in high heels: The gendered quality gap and incumbent electoral success." *Political Research Quarterly* 65(2):303–14.
- Fulton, Sarah A., Cherie D. Maestas, L. Sandy Maisel and Walter J. Stone. 2006. "The sense of a woman: Gender, ambition and the decision to run for Congress." *Political Research Quarterly* 59(2):235–48.
- Fulton, Sarah A. and Kostanca Dhima. 2020. "The gendered politics of congressional elections."

- Political Behavior*. Published online March 26. <https://doi.org/10.1007/s11109-020-09604-7>.
- Gaddis, S. Michael. 2018a. An introduction to Audit Studies in the Social Sciences. In *Audit Studies: Behind the Scenes with Theory, Method, and Nuance.*, ed. S. Michael Gaddis. Cham, Switzerland: Springer pp. 3–44.
- Gaddis, S. Michael, ed. 2018b. *Audit Studies: Behind the Scenes with Theory, Method, and Nuance*. Cham, Switzerland: Springer.
- Gaskins, Ben, Matt Golder and David A. Siegel. 2013a. “Religious participation and economic conservatism.” *American Journal of Political Science* 57:823–840.
- Gaskins, Ben, Matt Golder and David A. Siegel. 2013b. “Religious participation, social conservatism, and human development.” *Journal of Politics* 75:1125–1141.
- Gell-Redman, Neil Visalvanich, Charles Crabtree and Christopher J. Fariss. 2018. “It’s all about race: How state legislators respond to immigrant constituents.” *Political Research Quarterly* 71(3):517–31.
- Gerber, Alan S. and Donald P. Green. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W. W. Norton.
- Goggin, Stephen N, John A. Henderson and Alexander Theodoridis G. 2020. “What goes with red and blue? Mapping partisan and ideological associations in the minds of voters.” *Political Behavior* 42(4):985–1013.
- Golder, Matt and Benjamin Ferland. 2018. *The Oxford Handbook of Electoral Systems*. New York: Oxford University Press chapter Electoral systems and citizen-elite ideological congruence.
- Golder, Matt and Jacek Stramski. 2010. “Ideological congruence and electoral institutions.” *American Journal of Political Science* 54(1):90–106.
- Golder, Sona N., Charles Crabtree and Kostanca Dhima. 2019. “Legislative representation and gender (bias).” *Political Science* 71(1):1–16.
- Golder, Sona N., Laura B. Stephenson, Karine Van der Straeten, Andre Blais, Damien Bol, Philipp Harfst and Jean-Francois Laslier. 2017. “Votes for women: Electoral systems and support for

- female candidates.” *Politics & Gender* 13(1):107–31.
- Goodyear-Grant, Elizabeth. 2010. Who votes for women candidates and why? Evidence from recent Canadian elections. In *Voting Behaviour in Canada*, ed. Cameron D. Anderson and Laura B. Stephenson. Vancouver: University of British Columbia Press pp. 43–64.
- Grose, Christian R. 2011. *Congress in Black and White: Race and Representation in Washington and at Home*. New York: Cambridge University Press.
- Gross, Steven Jay and Michael C. Niman. 1975. “Attitude-behavior consistency: A review.” *Public Opinion Quarterly* 39(3):358–68.
- Gutek, Barbara A. and Bruce Morasch. 1982. “Sex-Ratios, Sex-Role Spillover, and Sexual Harassment of Women at Work.” *Journal of Social Issues* 38(4):55–74.
- Hainmueller, Jens, Daniel J. Hopkins and Teppei Yamamoto. 2014. “Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments.” *Political Analysis* 22(1):1–30.
- Hainmueller, Jens, Dominik Hangarten and Teppei Yamamoto. 2015. “Validating vignette and conjoint survey experiments against real-world behavior.” *Proceedings of the National Academy of Sciences* 112(8):2395–400.
- Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. “How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice.” *Political Analysis* 27(2):163–192.
- Hawkesworth, Mary. 2003. “Congressional enactments of race-gender: Toward a theory of raced-gendered institutions.” *American Journal of Political Science* 97(4):529–50.
- Heath, Michelle, Leslie A. Schwindt-Bayer and Michelle M. Taylor-Robinson. 2005. “Women on the sidelines: Women’s representation on committees in Latin American legislatures.” *American Journal of Political Science* 49(2):420–36.
- Hogan, Bernie and Brent Berry. 2011. “Racial and ethnic biases in rental housing: An audit study of online apartment listings.” *City & Community* 10(4):351–72.

- Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2011. "Sex, stereotypes, and security: A study of the effects of terrorist threat on assessments of female leadership." *Journal of Women, Politics, and Policy* 32(3):173–192.
- Holman, Mirya R., Jennifer L. Merolla and Elizabeth J. Zechmeister. 2016. "Terrorist threat, male stereotypes, and candidate evaluations." *Political Research Quarterly* 69(1):134–147.
- Horiuchi, Yusaku, Daniel M. Smith and Teppei Yamamoto. 2020. "Identifying voter preferences for politicians' personal attributes: A conjoint experiment in Japan." *Political Science Research and Methods* 8(1):75–91.
- Htun, Mala. 2016. *Inclusion without Representation: Gender Quotas and Ethnic Reservations in Latin America*. New York: Cambridge University Press.
- Htun, Mala and Laurel S. Weldon. 2010. "When do governments promote women's rights? A framework for the comparative analysis of sex equality policy." *Perspectives on Politics* 8(1):207–216.
- Huddy, Leonie and Nayda Terkildsen. 1993. "Gender stereotypes and the perception of male and female candidates." *American Journal of Political Science* 37(1):119–47.
- Hughes, Melanie. 2009. "Armed conflict, international linkages, and women's parliamentary representation in developing nations." *Social Problems* 56(1):174–204.
- Hughes, Melanie M., Pamela Paxton, Amanda Clayton and Pär Zetterberg. 2017. *Quota Adoption and Reform Over Time (QAROT), 1947-2015*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2017-08-16. <https://doi.org/10.3886/E100918V1>. .
- Hughes, Melanie M., Pamela Paxton, Amanda Clayton and Pär Zetterberg. 2019. "Global gender quota adoption, implementation, and reform." *Comparative Politics* 51(2):2019–238.
- Inglehart, Ronald and Christian Welzel. 2005. *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge University Press.
- Inglehart, Ronald and Pippa Norris. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World*. New York: Cambridge University Press.

- Inglehart, Ronald, Pippa Norris and Christian Welzel. 2002. "Gender quality and democracy." *Comparative Sociology* 1(3):321–345.
- Inter-Parliamentary Union. 2022. "Women in National Parliaments." February, .
- Jerolmack, Colin and Shamus Khan. 2014. "Talk is cheap: Ethnography and the attitudinal fallacy." *Sociological Methods & Reserach* 43(2):178–209.
- Johnson, Daniel A., Richard J. Porter and Patricia L. Mateljan. 1971. "Racial discrimination in apartment rentals 1." *Journal of Applied Social Psychology* 1(4):364–77.
- Jones, Mark P. and Patricio Navia. 1999. "Assessing the effectiveness of gender quotas in open-list proportional representation electoral systems." *Social Science Quarterly* 80(2):341–55.
- Jones, Mark P.s. 1998. "Gender quotas, electoral laws, and the election of women: Evidence from the Latin American vanguard." *Comparative Political Studies* 31(1):3–21.
- Judkins, David R. 2016. "Robustness of ordinary least squares in randomized clinical trials." *Statistics in Medicine* 35(11):1763–73.
- Kage, Rieko, Frances Rosenbluth and Seiki Tanaka. 2018. "What Explains Low Female Political Representation? Evidence from Survey Experiments in Japan." *Politics & Gender* Forthcoming.
- Kahn, Kim Fridkin. 1994. "Does gender make a difference? An experimental examination of sex stereotypes and press patterns in statewide campaigns." *American Journal of Political Science* 38:162–95.
- Kahn, Kim Fridkin. 1996. *The Political Consequences of Being a Woman: How Stereotypes Influence the Conduct and Consequences of Political Campaigns*. Columbia University Press.
- Kalla, Joshua L., Frances Rosenbluth and Dawn Teele. 2018. "Are you my mentor? A field experiment on gender, ethnicity, and political self-starters." *Journal of Politics* 80(1):337–41.
- Kanthak, Kristin and Jonathan Woon. 2015. "Women don't run? Election aversion and candidate entry." *American Journal of Political Science* 59(3):595–612.
- Karpowitz, Christopher F., J. Quin Monson and Jessica Robinson Preece. 2017. "How to elect more women: Gender and candidate success in a field experiment." *American Journal of Political*

- Science* 61(4):927–943.
- Kathlene, Lyn. 1994. “Power and influence in state legislative policymaking: The interaction of gender and position in committee hearing debates.” *American Political Science Review* 88(3):560–76.
- Kenny, Meryl. 2013. *Gender and Political Recruitment: Theorizing Institutional Change*. New York: Palgrave Macmillan.
- Kenworthy, Lane and Melissa Malami. 1999. “Gender inequality in political representation: A worldwide comparative analysis.” *Social Forces* 78(1):235–68.
- Kim, Jeong Hyun and Yesola Kweon. 2022. “Why Do Young Men Oppose Gender Quotas? Group Threat and Backlash to Legislative Gender Quotas.” *Legislative Studies Quarterly* .
- King, Gary and Langche Zeng. 2006. “The dangers of extreme counterfactuals.” *Political Analysis* 14(2):131–59.
- Kirkland, Patricia A. and Alexander Coppock. Forthcoming. “Candidate Choice Without Party Labels: New Insights from Conjoint Survey Experiments.” *Political Behavior* 40(3):571–91.
- Kittilson, Miki Caul. 2006. *Challenging Parties, Changing Parliaments: Women and Elected Office in Contemporary Western Europe*. Ohio State University Press.
- Kittilson, Miki Caul. 2008. “Representing women: The adoption of family leave policies in comparative perspective.” *Journal of Politics* 70(2):323–334.
- Kittilson, Miki Caul and Leslie A. Schwindt-Bayer. 2010. “Engaging citizens: The role of power-sharing institutions.” *Journal of Politics* 72(4):990–1002.
- Kittilson, Miki Caul and Leslie A. Schwindt-Bayer. 2012. *The Gendered Effect of Electoral Institutions: Political Engagement and Participation*. New York: Oxford University Press.
- Kloppenborg, Ria and Wouter J. Hanegraaff, eds. 1995. *Female Stereotypes in Religious Traditions*. New York: E.J. Brill.
- Koch, Jeffrey W. 1999. “Candidate gender and assessments of senate candidates.” *Social Science Quarterly* 80(1):84–96.

- Koch, Jeffrey W. 2000. "Do citizens apply gender stereotypes to infer candidates' ideological orientations?" *The Journal of Politics* 62(2):414–429.
- Koch, Jeffrey W. 2002. "Gender stereotypes and citizens' impressions of house candidates' ideological orientations." *American Journal of Political Science* 46(2):453–62.
- Koenig, Anne M., Alice H. Eagly, Abigail A. Mitchell and Riina Ristikari. 2011. "Are leader stereotypes masculine? A meta-analysis of three research paradigms." *Psychological Bulletin* 137(4):616–42.
- Krook, Mona Lena. 2009. *Quotas for Women in Politics: Gender and Candidate Selection Reform Worldwide*. New York: Oxford University Press.
- Krook, Mona Lena and Pippa Norris. 2014. "Beyond quotas: Strategies to promote gender equality in elected office." *Political Studies* 62(1):2–20.
- Krupnikov, Yanna and Adam Seth Levine. 2014. "Cross-sample comparisons and external validity." *Journal of Experimental Political Science* 1(1):59–80.
- Kunovich, Sheri and Pamela Paxton. 2005. "Pathways to power: The role of political parties in women's national political representation." *American Journal of Sociology* 111:505–552.
- Lawless, Jennifer L. 2004. "Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era." *Political Research Quarterly* 57(3):479–90.
- Lawless, Jennifer L. and Kathryn Pearson. 2008. "The primary reason for women's underrepresentation? Reevaluating the conventional wisdom." *Journal of Politics* 70(1):67–82.
- Lawless, Jennifer L. and Richard L. Fox. 2005. *It Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.
- Lawless, Jennifer L. and Richard L. Fox. 2010. *It Still Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.
- Leeper, Thomas J., Sarah B. Hobolt and James Tilley. 2020. "Measuring Subgroup Preferences in Conjoint Experiments." *Political Analysis* 28(2).
- Lenz, Gariel S. and Cappell Lawson. 2011. "Looking the part: Television leads less informed citi-

- zens to vote based on candidates' appearance." *American Journal of Political Science* 55(3):574–89.
- Lin, Winston. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *Annals of Applied Statistics* 7(1):295–318.
- Loewen, Peter John and Michael Kenneth MacKenzie. 2019. "Service representation in a federal system. A field experiment." *Journal of Experimental Political Science* 6(2):93–107.
- Longman, Timothy. 2006. *Women in African Parliaments*. Boulder: Lynne Rienner chapter Rwanda: Achieving Equality or Serving an Authoritarian State.
- Lovenduski, Joni. 2016. "The supply and demand model of candidate selection: Some reflections." *Government and Opposition* 51(3):513–28.
- Luhiste, Maarja. 2015. "Party gatekeepers' support for viable female candidacy in PR-list systems." *Politics & Gender* 11(1):89–116.
- Mansbridge, Jane. 1999. "Should blacks represent blacks and women represent women? A contingent "yes"." *Journal of Politics* 61(3):628–657.
- Manza, Jeff and Clem Brooks. 1998. "The gender gap in US presidential elections: When? Why? Implications?" *American Journal of Sociology* 103(5):1235–1266.
- Marshall, Monty G., Ted Robert Gurr and Keith Jagers. 2019. *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2018, Dataset Users' Manual*. <http://www.systemicpeace.org/inscr/p4manualv2018.pdf>: .
- Matland, Richard E. 1998. "Women's legislative representation in national legislatures: Developed and developing countries." *Legislative Studies Quarterly* 23(1):109–125.
- Matland, Richard E. 2005. *Women in Parliament: Beyond Numbers*. Stockholm: International IDEA chapter Enhancing women's political representation: Legislative recruitment and electoral systems.
- Matland, Richard E. and Donley T. Studlar. 1996. "The contagion of women candidates in single-member district and proportional representation electoral systems: Canada and Norway." *Jour-*

- nal of Politics* 58(3):707–33.
- Mayer, Lawrence and Ronald E. Smith. 1985. “Feminism and religiosity: Female electoral behavior in Western Europe.” *West European Politics* 8(4):38–49.
- McDermott, Monika L. 1997. “Voting cues in low-information elections: Candidate gender as a social information variable in contemporary United States elections.” *American Journal of Political Science* 41(1):270–283.
- McDermott, Monika L. 2005. “Candidate occupations and voter information shortcuts.” *Journal of Politics* 67(1):201–19.
- McDermott, Rose. 2011. Internal and External Validity. In *Handbook of Experimental Political Science*, ed. James N. Druckman, Donald P. Green, James H. Kuklinski and Arthur Lupia. New York: Cambridge University Press pp. 27–40.
- McPherson, Miller, Lynn Smith-Lovin and James M. Cook. 2001. “Birds of a feather: Homophily in social networks.” *Annual Review of Sociology* 27:415–44.
- Medeiros, Mike, Benjamin Forest and Chris Erl. 2019. “Where women stand: Parliamentary candidate selection in Canada.” *Politics, Groups, and Identities* 7(2):389–400.
- Melinda, Adams. 2007. “National Machineries’ and Authoritarian Politics: The Case of Cameroon.” *International Feminist Journal of Politics* pp. 176–197.
- Mendelberg, Tali, Christopher F. Karpowitza and Nicholas Goedert. 2014. “Does Descriptive Representation Facilitate Women’s Distinctive Voice?” *American Journal of Political Science* 58(2):291–306.
- Milkman, Katherine L, Modupe Akinola and Dolly Chugh. 2015. “What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations.” *Journal of Applied Psychology* 100(6):1678–712.
- Mo, Cecilia Hyunjung. 2015. “The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters.” *Political Behavior* 37(2):357–95.
- Montgomery, Jacob M., Brendan Nyhan and Michelle Torres. 2018. “How conditioning on post-

- treatment variables can ruin your experiment and what you can do about it.” *American Journal of Political Science* 62(3):760–775.
- Moore, Gwen and Gene Shackman. 1996. “Gender and authority: A cross-national study.” *Social Science Quarterly* 77(2):273–88.
- Moore, Ryan T. and Keith Schnakenberg. 2012. “blockTools: Blocking, assignment, and diagnosing interference in randomized experiments.”
- Morton, Rebecca B. and Kenneth C. Williams. 2008. Experimentation in Political Science. In *The Oxford Handbook of Political Methodology*, ed. Janet M. Box-Steffensmeier, Henry E. Brady and David Collier. New York: Oxford University Press pp. 339–56.
- Neumark, David, Roy J. Bank and Kyle D. Van Nort. 1996. “Sex discrimination in restaurant hiring: An audit study.” *Quarterly Journal of Economics* 111(3):915–41.
- Niven, David. 1998. *The Missing Majority: The Recruitment of Women as State Legislative Candidates*. Westport, CT: Praeger.
- Niven, David. 2006. “Throwing your hat out of the ring: Negative recruitment and the gender imbalance in state legislative candidacy.” *Politics & Gender* 2(4):473–89.
- Norris, Pippa and Joni Lovenduski. 1993. “‘If only more candidates came forward’: Supply-side explanations of candidate selection in Britain.” *British Journal of Political Science* 23(3):373–408.
- Norris, Pippa and Joni Lovenduski. 1995. *Political Recruitment: Gender, Race and Class in the British Parliament*. New York: Cambridge University Press.
- O’Brien, Diana Z. 2018. “Righting conventional wisdom: Women and right parties in established democracies.” *Politics & Gender* 14(1):27–55.
- O’Brien, Diana Z. 2019. “Female Leaders and Citizens’ Perceptions of Political Parties.” *Journal of Elections, Public Opinion & Parties* 29(4):465–489.
- Ono, Yoshikuni and Masahiro Yamada. 2020. “Do voters prefer gender stereotypic candidates? Evidence from a conjoint survey experiment in Japan.” *Political Science Research and Methods*

8(3):1–16.

- Palmer, Barbara and Dennis Simon. 2001. “The political glass ceiling: Gender, strategy, and incumbency in US house elections, 1978-1998.” *Women & Politics* 23(1-2):59–78.
- Palmer, Barbara and Dennis Simon. 2010. *Breaking the Political Glass Ceiling: Women and Congressional Elections*. London: Routledge.
- Paxton, Pamela. 1997. “Women in national legislatures: A cross-national analysis.” *Social Science Research* 26(4):442–464.
- Paxton, Pamela, Melanie M. Hughes and Tiffany Barnes. 2020. *Women, Politics, and Power: A Global Perspective*. New York: Rowman & Littlefield Publishers.
- Paxton, Pamela and Sheri Kunovich. 2003. “Women’s political representation: The importance of ideology.” *Social Forces* 82(1):87–114.
- Paxton, Pamela, Sheri Kunovich and Melanie M. Hughes. 2007. “Gender in politics.” *Annual Review of Political Science* 33:263–84.
- Piscopo, Jennifer M. and Susan Franceschet. 2014. “Sustaining gendered practices? Power and elite networks in Argentina.” *Comparative Political Studies* 47(1):86–111.
- Pruysers, Scott and Julie Blais. 2019. “Narcissistic women and cash-strapped men: Who can be encouraged to consider running for office?” *Political Research Quarterly* 72(1):229–42.
- Pruysers, Scott and William Paul Cross. 2016. “Candidate selection in Canada: Local autonomy, centralization, and competing democratic norms.” *American Behavioral Scientist* 60(7):781–98.
- Putnam, Robert. 1976. *The Comparative Study of Political Elites*. Englewood Cliffs, NJ: Prentice Hall.
- Quillian, Lincoln, Devah Pager, Ole Hexel and Arnfinn H. Midtøen. 2017. “Meta-analysis of field experiments shows no change in racial discrimination in hiring over time.” *Proceedings of the National Academy of Sciences* 114(41):364–77.
- Raines, Thomas, Matthew Goodwin and David Cutts. 2017. “The Future of Europe.” *Comparing Public and Elite Attitudes*. Chatham House 47.

- Rosen, Jennifer. 2011. "The effects of political institutions on women's political representation: A comparative analysis of 168 countries from 1992 to 2010." *Political Research Quarterly* 66(2):306–321.
- Ruedin, Didier. 2012. "The representation of women in national parliaments: A cross-national comparison." *European Sociological Review* 28(1):96–109.
- Rule, Wilma. 1987. "Electoral systems, contextual factors, and women's opportunities for election to parliament in twenty-three democracies." *Western Political Quarterly* 40:477–98.
- Rule, Wilma and Joseph Francis Zimmerman. 1994. *Electoral systems in comparative perspective: Their impact on women and minorities*. Westport, CT: Greenwood Press.
- Saha, Sparsha and Ana Catalano Weeks. 2020. "Ambitious Women: Gender and Voter Perceptions of Candidate Ambition." *Political Behavior* .
- Salmond, Rob. 2006. "Proportional representation and female parliamentarians." *Legislative Studies Quarterly* 31(2):175–204.
- Sanbonmatsu, Kira. 2002. "Political parties and the recruitment of women to state legislatures." *Journal of Politics* 64(3):791–809.
- Sanbonmatsu, Kira. 2006. "Do parties know that 'women win'? Party leader beliefs about women's electoral chances." *Politics & Gender* 2(4):431–50.
- Sanbonmatsu, Kira and Kathleen Dolan. 2009. "Do gender stereotypes transcend party?" *Political Research Quarterly* 62(3):485–95.
- Sayers, Anthony Michael. 1999. *Parties, Candidates, and Constituency Campaigns*. Vancouver: University of British Columbia Press.
- Schmidt, Gregory D. 2009. "The election of women in list PR systems: Testing the conventional wisdom." *Electoral Studies* 28(2):190–203.
- Schneider, Monica C. and Angela L. Bos. 2014. "Measuring stereotypes of female politicians." *Political Psychology* 35(2):245–264.
- Schuman, Howard and Michael P. Johnson. 1976. "Attitudes and behavior." *Annual Review of*

- Sociology* 2(1):161–207.
- Schwarz, Susanne and Alexander Coppock. 2020. “What have we learned about gender from candidate choice experiments? A meta-analysis of 30 factorial survey experiments.” *The Journal of Politics* .
- Schwindt-Bayer, Leslie A. 2011. “Women who win: Social backgrounds, paths to power, and political ambition in Latin American legislatures.” *Politics & Gender* 7(1):1–33.
- Schwindt-Bayer, Leslie A., Michael Malecki and Brian F. Crisp. 2010. “Candidate gender and electoral success in single transferable vote systems.” *British Journal of Political Science* 40(3):693–709.
- Schwindt-Bayer, Leslie A. and William Mishler. 2005. “An integrated model of women’s representation.” *Journal of Politics* 67(2):407–428.
- Senninger, Roman and Daniel Bischof. Forthcoming. “Do voters want domestic politicians to scrutinize the European Union?” *Political Science Research and Methods* .
- Sevi, Semra, Vincent Arel-Bundock and André Blais. 2019. “Do women get fewer votes? No.” *Canadian Journal of Political Science* 52(1):201–10.
- Shadish, William R., Thomas D. Cook and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth.
- Shugart, Matthew S. 1994. Minorities Represented and Unrepresented. In *Electoral Systems in Comparative Perspective*, ed. Wilma Rule and Joseph F. Zimmerman. Westport, CT: Greenwood Press pp. 31–41.
- Smith, Janet L. 2015. Public Housing Transformation: Evolving National Policy. In *Where are Poor People to Live? Transforming Public Housing Communities*, ed. Larry Bennett, Janet L. Smith and Patricia A. Wright. London: Routledge pp. 19–40.
- Stockemer, Daniel. 2011. “Women’s parliamentary representation in Africa: The impact of democracy and corruption on the number of female deputies in national parliaments.” *Political Studies* 59:693–712.

- Stockemer, Daniel and Maeve Byrne. 2012. "Women's representation around the world: The importance of women's participation in the workforce." *Parliamentary Affairs* 65(4):802–821.
- Studlar, Donley T. and Ian McAllister. 2002. "Does a critical mass exist? A comparative analysis of women's legislative representation since 1950." *European Journal of Political Research* 41(2):233–253.
- Swiss, Liam. 2009. "Decoupling values from action: An event-history analysis of the election of women to parliament in the developing world, 1945-90." *International Journal of Comparative Sociology* 50(1):69–95.
- Taagepera, Rein. 1994. Beating the Law of Minority Attrition. In *Electoral Systems in Comparative Perspective*, ed. Wilma Rule and Joseph F. Zimmerman. Westport, CT: Greenwood Press pp. 235–45.
- Tajfel, Henri and John C. Turner. 1979. An Integrative Theory of Intergroup Conflict. In *The Social Psychology of Intergroup Relations*, ed. William G. Austin and Stephen Worchel. Monterey, CA: Brooks/Cole pp. 33–47.
- Teele, Dawn, Joshua L. Kalla and Frances Rosenbluth. 2018. "The ties that double bind: Social roles and women's underrepresentation in politics." *American Political Science Review* 112(3):525–541.
- Thames, Frank C. and Margaret S. Williams. 2010. "Incentives for personal votes and women's representation in legislatures." *Comparative Political Studies* 43(12):1575–600.
- Thomas, Malanee and Marc Andre Bodet. 2013. "Sacrificial lambs, women candidates, and district competitiveness in Canada." *Electoral Studies* 32(1):153–66.
- Thorlakson, Lori. 2009. "Patterns of party integration, influence, and autonomy in seven federations." *Party Politics* 15(2):157–77.
- Tolley, Erin. 2011. "Do women 'do better' in municipal politics? Electoral representation across three levels of government." *Canadian Journal of Political Science* 44(3):573–94.
- Tremblay, Manon. 2007. "Democracy, representation, and women: A comparative analysis." *De-*

- mocratization* 14(4):533–553.
- Tremblay, Manon. 2008. *Women and Legislative Representation: Electoral systems, Political Parties, and Sex Quotas*. New York: Palgrave Macmillan.
- Tremblay, Manon and Réjean Pelletier. 2000. “More feminists or more women? Descriptive and substantive representations of women in the 1997 Canadian federal elections.” *International Political Science Review* 21(4):381–405.
- Tremblay, Manon and Réjean Pelletier. 2001. “More women constituency party presidents: A strategy for increasing the number of women candidates in Canada?” *Party Politics* 7(2):157–90.
- Tripp, Aili Mari. 2019. *Seeking Legitimacy: Why Arab Autocracies Adopt Women’s Rights*. New York: Cambridge University Press.
- Valdini, Melody E. 2012. “A deterrent to diversity: The conditional effect of electoral rules on the nomination of women candidates.” *Electoral Studies* 31(4):740–749.
- Valdini, Melody E. 2019. *The Inclusion Calculation: Why Men Appropriate Women’s Representation*. New York: Oxford University Press.
- Walby, Sylvia. 2004. “The European Union and gender equality: Emergent varieties of gender regime.” *Social Politics: International Studies of Gender, State, and Society* 11(1):4–29.
- Wolbrecht, Christina and David E. Campbell. 2007. ““Leading by example: Female members of parliament as political role models.” *American Journal of Political Science* 51(4):921–39.
- Wooldridge, Jeffrey M. 2003. “Cluster-sample methods in applied econometrics.” *American Economic Review* 93(2):133–38.
- World Bank. 2017. “World Development Indicators.” Washington, D.C. :The World Bank, <https://datacatalog.worldbank.org/dataset/world-development-indicators>.
- Wright, Matthew, Morris Levy and Jack Citrin. 2016. “Public Attitudes Toward Immigration Policy Across the Legal/Illegal Divide: The Role of Categorical and Attribute-Based Decision-Making.” *Political Behavior* 38(1):229–53.

- WVS. 2015. "World Values Survey 1981-2015 Official Aggregate v.20150418." World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: JDSystems, Madrid.
- WVS. 2020. "World Values Survey 2017-2020." World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: JDSystems, Madrid.
- Young, Lisa. 2000. *Feminists and Party Politics*. Vancouver: University of British Columbia Press.
- Young, Lisa. 2006. *Women's Representation in the Canadian House of Commons*. London: Routledge.
- Young, Lisa and William Paul Cross. 2003. *Women and Electoral Politics in Canada*. Toronto: Oxford University Press.

APPENDIX A

This appendix provides supplementary information for Chapter Two “Do Elites Discriminate Against Female Political Aspirants? Evidence from a Field Experiment.”

A.1 What about Quebec?

The email message from the hypothetical student in my audit experiment was written in English. I chose to send the emails in English to replicate as closely as possible the empirical strategy employed by Kalla, Rosenbluth and Teele (2018). This has implications for Quebec with its large French-speaking population. I now address some of these implications here.

One implication is that the overall response rate for the legislators in Quebec will be lower than the response rate for the legislators elsewhere in Canada. This is borne out in the data. There are 302 Quebec legislators in my sample, 89 (29%) of whom responded to my email. This response rate is significantly lower than the overall response rate in Canada as a whole, which was 49%. Notably, though, the response rate in Quebec was still higher than the response rate of 26% in the United States.

While the overall response rate in Quebec may well have been higher had I sent my emails in French, it’s unclear whether this would affect any of the inferences that I draw with respect to the *directionality* of the responses. One way to evaluate this is to look to see if the pattern of responses from Quebec is different from that found elsewhere in Canada. In Table A.1, I provide information about the response rates by treatment name and legislator gender *for Quebec only*. The results indicate that legislators in Quebec exhibit a pro-female bias of 7.1 percentage points. This is almost identical to the 7 percentage points pro-female bias observed in the full sample. As with the full sample, the female legislators in Quebec exhibit a higher pro-female bias (12 percentage points) than their male counterparts (5 percentage points). The magnitudes of the pro-female bias exhibited by the female and male legislators in Quebec are again almost identical to those observed

Table A.1: Response Rates by Treatment Name and Legislator Gender: Only Quebec

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
All Representatives	26% [19%, 33%] <i>N</i> = 154	33% [25%, 41%] <i>N</i> = 148	7.1 [-5, 19]	0.22
Female Legislators	24% [12%, 36%] <i>N</i> = 54	36% [21%, 50%] <i>N</i> = 45	12 [-6, 29]	0.19
Male Legislators	27% [18%, 36%] <i>N</i> = 100	32% [23%, 41%] <i>N</i> = 103	5 [-9, 19]	0.47

Note: The first two columns of Table A.1 show the response rates to the email messages sent from male and female students for different sets of elected representatives. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

in the full sample, where female legislators exhibited a pro-female bias of 11 percentage points and the male legislators exhibited a pro-female bias of 6 percentage points. Unlike with the full sample, the pro-female biases in Quebec do not reach conventional levels of statistical significance. However, this is likely due to the much smaller sample size used in the Quebec-only comparisons. The bottom line here is that the legislators in Quebec who did respond to my email message did so in an almost identical way to the legislators in the other provinces.

A second way to evaluate whether my inferences are affected by my decision to send the email message in English is to see what happens if I exclude the responses from Quebec. In Table A.2, I provide information about the response rates by treatment name and legislator gender in the full

Table A.2: Response Rates by Treatment Name and Legislator Gender: Excluding Quebec

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
All Legislators	49% [45%, 53%] <i>N</i> = 734	56% [53%, 60%] <i>N</i> = 738	7.2 [1, 13]	0.02
Female Legislators	47% [40%, 54%] <i>N</i> = 205	58% [51%, 64%] <i>N</i> = 213	11 [1, 19]	0.03
Male Legislators	50% [45%, 54%] <i>N</i> = 529	56% [51%, 60%] <i>N</i> = 525	6 [-1, 12]	0.07

Note: The first two columns of Table A.2 show the response rates to the email messages sent from male and female students for different sets of elected representatives. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

sample *with Quebec excluded*. The results indicate that the legislators in the full sample with Quebec excluded exhibit a pro-female bias of 7.2 percentage points. This is almost identical to the 7 percentage points pro-female bias observed in the full sample. Again, the female legislators exhibit a higher pro-female bias (11 percentage points) than their male counterparts (6 percentage points). The magnitudes of these pro-female biases exhibited by the female and male legislators are identical to those observed in the full sample, where female legislators exhibited a pro-female bias of 11 percentage points and the male legislators exhibited a pro-female bias of 6 percentage points. The pattern of statistical significance across the three groups – All Legislators, Female Legislators, and Male Legislators — is also identical to that found in the full sample.

We can also look at how the response rates differ across the different levels of government when we exclude the responses from Quebec. This information is shown in Table A.3. The main

Table A.3: Response Rates by the Treatment Name, Level of Office, and Legislator Gender: Excluding Quebec

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
<i>Municipal:</i>				
All Legislators	58% [53%, 64%] <i>N</i> = 339	64% [59%, 69%] <i>N</i> = 342	6 [-0.2, 12]	0.06
Female Legislators	53% [43%, 63%] <i>N</i> = 94	64% [54%, 74%] <i>N</i> = 89	11 [-1, 22]	0.07
Male Legislators	60% [54%, 67%] <i>N</i> = 245	64% [58%, 70%] <i>N</i> = 253	4 [-2, 10]	0.20
<i>Provincial:</i>				
All Legislators	50% [44%, 57%] <i>N</i> = 227	58% [52%, 65%] <i>N</i> = 240	8 [-2, 18]	0.12
Female Legislators	49% [37%, 61%] <i>N</i> = 69	63% [52%, 74%] <i>N</i> = 79	14 [-4, 32]	0.11
Male Legislators	51% [43%, 59%] <i>N</i> = 158	56% [48%, 64%] <i>N</i> = 161	5 [-6, 16]	0.34
<i>Federal:</i>				
All Legislators	29% [22%, 35%] <i>N</i> = 168	35% [28%, 43%] <i>N</i> = 156	6 [-4, 17]	0.20
Female Legislators	31% [16%, 46%] <i>N</i> = 42	36% [21%, 50%] <i>N</i> = 45	5 [-14, 23]	0.61
Male Legislators	28% [20%, 36%] <i>N</i> = 126	35% [26%, 44%] <i>N</i> = 111	7 [-4, 18]	0.18

Note: The first two columns of Table A.3 show the response rates to the email messages sent from male and female students for different sets of elected representatives separated by level of office. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

thing to note here is that there is always a pro-female bias at each level of government. As in the full sample that includes Quebec, this pro-female bias is larger for female legislators than male legislators at the municipal and provincial levels. At the municipal level, the pro-women bias exhibited by female legislators (11 percentage points) is 2.8 times larger than that exhibited by male legislators (4 percentage points). At the provincial level, the pro-women bias exhibited by female legislators (14 percentage points) is 2.8 times larger than that exhibited by male legislators

(5 percentage points). There is no substantive difference in the magnitude of the pro-women bias across female and male legislators at the federal level.

In summary, these additional analyses show that my inferences are robust to looking only at Quebec or excluding Quebec from the full sample. As a result, my decision to send my email message only in English is not problematic.

A.2 Sample of Legislators

I obtained my sample of Canadian legislators from the Represent Civic Information API. The original sample consisted of 1,936 legislators. After dropping legislators for whom an email address was not provided and ‘duplicate’ legislators who held multiple official roles, I obtained a sample of 1,779 unique legislators. In this sample, there were 854 municipal legislators (28.6% women), 591 provincial legislators (31.6% women), and 334 federal legislators (26.7% women). Ultimately, five of these legislators had an invalid email address and so my results are based on a final sample of 1,774 legislators who received an email.

Although my sample does not represent the full number of legislators across the three levels of government in Canada, the API states that it’s “the most comprehensive source in Canada for elected officials and electoral districts.” My sample of Canadian legislators is over nine times larger than that used in a recent audit study looking at the responsiveness of Canadian legislators to constituent requests for assistance (Loewen and MacKenzie 2019). The API does not indicate why it has information about some elected representatives and not others. It appears that there is no information for legislators at the municipal and ‘provincial’ level in the three Canadian territories; the federal-level legislators are, however, included in the sample. I see no obvious reason why the legislators with contact information provided by the API would be qualitatively different when it comes to exhibiting a gender bias from the legislators for whom the API does not provide contact information.

My sample includes elected representatives from the following provincial legislatures: Legislative Assembly of Alberta, Legislative Assembly of British Columbia, Legislative Assembly of Manitoba, Nova Scotia House of Assembly, Legislative Assembly of Ontario, Legislative Assembly of Prince Edward Island, Assemblée nationale du Québec, and Legislative Assembly of Saskatchewan. It includes elected representatives from the following municipal councils: Ajax Town Council, Belleville City Council, Brampton City Council, Brantford City Council, Burling-

ton City Council, Burnaby City Council, Caledon Town Council, Calgary City Council, Cambridge City Council, Cape Breton Regional Council, Charlottetown City Council, Conseil municipal de Beaconsfield, Conseil municipal de Dorval, Conseil municipal de Gatineau, Conseil municipal de Kirkland, Conseil municipal de Laval, Conseil municipal de Lévis, Conseil municipal de Montréal, Conseil municipal de Montréal-Est, Conseil municipal de Pointe-Claire, Conseil municipal de Saint-Jean-sur-Richelieu, Conseil municipal de Saint-Jérôme, Coquitlam City Council, County of Grande Prairie No. 1 Council, Edmonton City Council, Fredericton City Council, Georgina Town Council, Grande Prairie City Council, Greater Sudbury City Council, Grimsby Town Council, Guelph City Council, Haldimand County Council, Halifax Regional Council, Hamilton City Council, Kelowna City Council, King Township Council, Kingston City Council, Kitchener City Council, Lambton County Council, Langley Township Council, Lethbridge City Council, Lincoln Town Council, London City Council, Mississauga City Council, Moncton City Council, Newmarket Town Council, Niagara Regional Council, North Dumfries Township Council, Oakville Town Council, Oshawa City Council, Ottawa City Council, Peel Regional Council, Peterborough City Council, Pickering City Council, Richmond City Council, Richmond Hill Town Council, Saanich District Council, Saint John City Council, Saskatoon City Council, St. Catharines City Council, Strathcona County Council, Surrey City Council, Thunder Bay City Council, Toronto City Council, Uxbridge Township Council, Vancouver City Council, Victoria City Council, Waterloo City Council, Waterloo Regional Council, Welland City Council, Whitby Town Council, Windsor City Council, and Winnipeg City Council.

A.3 Staff Responses

As I noted in the main text, there's no guarantee that the legislator is the person to receive and respond to the student's email message. This is true for all audit experiments of this type and not just the one discussed here. Thus, the unit of analysis is technically the email address of the legislator and not the legislator themselves.

It's often possible to identify when a staff member has sent the email response rather than the legislator. This is because the response either comes from a different email address or the person responding introduces themselves as the person responsible for communicating on behalf of the official. On this basis, it appears that about 38% of the email responses come from staff members. In general, we would not expect staff members to express their own views or opinions when engaging in official business. Instead, we'd expect them to express views that are consistent with those of the legislator for whom they work. Of course, if this is true, then we should see similar results to those reported in the main text when we exclude the email responses that have been identified as coming from staff members. This is exactly what I find. In Table A.4, I provide information about the response rates by treatment name and legislator gender *when the staff responses are excluded*. The key thing to note is that there's still always a statistically significant pro-female bias. The overall pro-female bias is 6 percentage points in this smaller sample and 7 percentage points in the 'full' sample. Email accounts associated with both female and male legislators continue to exhibit a pro-female bias in this smaller sample. In contrast to the full sample, the pro-female bias exhibited by email accounts associated with female legislators is not larger than that exhibited by email accounts associated with male legislators.

As one might expect, the extent to which staff members send the email responses increases as we move from the local to the national level. While 17% of the responses at the municipal level came from staff, 79% of the responses at the federal level did so. This may help to explain why my results in the full sample, particularly with respect to the gender of the legislator, are weakest at the

Table A.4: Response Rates by Treatment Name and Legislator Gender: Excluding Staff Responses

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
All Representatives	34% [30%, 37%] <i>N</i> = 735	40% [37%, 44%] <i>N</i> = 708	6 [2, 12]	0.01
Female Legislators	33% [26%, 39%] <i>N</i> = 221	40% [33%, 46%] <i>N</i> = 197	7 [-1, 15]	0.08
Male Legislators	34% [30%, 38%] <i>N</i> = 514	41% [36%, 45%] <i>N</i> = 511	7 [0.2, 13]	0.04

Note: The first two columns of Table A.4 show the response rates to the email messages sent from male and female students for different sets of elected representatives. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

federal level — there’s evidence of a pro-female bias for all legislators at the federal level, female and male legislators but it’s never statistically significant. However, I can’t rule out the possibility that the lack of statistical significance at the federal level is simply a result of the significantly smaller sample size at this level of government.

A.4 Including Responses After the 2 Week Cutoff

In the main text, I focus on only those email responses that came during the first two weeks after my initial email was sent. In this appendix, I show that my inferences are robust to the inclusion of the email responses that came after the two week cutoff. In total, 50 responses came after the two week cutoff: 11 were from legislators at the municipal level, 24 from legislators at the provincial level, and 15 from legislators at the federal level.

In Table A.5, I provide information about the response rates by treatment name and legislator gender in the sample *with later responses included*. The results continue to show a pro-women bias of 7 percentage points, which is identical to the 7 percentage points pro-women bias observed in the original sample. Female legislators continue to exhibit a higher pro-women bias (10 percentage points) than their male counterparts (5 percentage points). The magnitudes of these pro-women

Table A.5: Response Rates by Treatment Name and Legislator Gender: Including Later Responses

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
All Legislators	48% [45%, 52%] <i>N</i> = 888	55% [51%, 58%] <i>N</i> = 886	7 [1, 12]	0.02
Female Legislators	47% [41%, 53%] <i>N</i> = 259	57% [51%, 63%] <i>N</i> = 258	10 [2, 18]	0.02
Male Legislators	49% [45%, 53%] <i>N</i> = 629	54% [50%, 58%] <i>N</i> = 628	5 [-1, 11]	0.10

Note: The first two columns of Table A.5 show the response rates to the email messages sent from male and female students for different sets of legislators. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

biases are almost identical to those observed in the original sample, where female legislators exhibited a pro-women bias of 11 percentage points and the male legislators exhibited a pro-women bias of 6 percentage points.

In Table A.6, I show how the response rates differ across the different levels of government when I include responses that were received after two week cutoff. Again, we see a pro-women bias at each level of government. As before, this pro-women bias is larger for female legislators than male legislators at the municipal and provincial levels.

In Table A.7, I look at how the response rates differ across the different types of political parties. As in the original sample, legislators from the left-leaning parties respond at higher rates (61%) to female political aspirants than legislators from the right-leaning parties (52%). The pro-women bias exhibited by the legislators in the left-leaning parties is again much larger — over three times as large — as that exhibited by the legislators in the right-leaning parties.

In sum, my results are robust to the inclusion of the email responses that came after the two week cutoff used in the main text.

Table A.6: Response Rates by the Treatment Name, Level of Office, and Legislator Gender: Including Later Responses

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
<i>Municipal:</i>				
All Legislators	55% [51%, 60%] <i>N</i> = 421	62% [57%, 66%] <i>N</i> = 429	7 [2, 11]	0.01
Female Legislators	50% [41%, 59%] <i>N</i> = 123	61% [52%, 70%] <i>N</i> = 119	12 [0.4, 23]	0.04
Male Legislators	58% [52%, 63%] <i>N</i> = 298	62% [57%, 67%] <i>N</i> = 310	4 [-1, 10]	0.12
<i>Provincial:</i>				
All Legislators	47% [41%, 53%] <i>N</i> = 295	54% [48%, 60%] <i>N</i> = 295	7 [-3, 17]	0.19
Female Legislators	48% [38%, 59%] <i>N</i> = 93	60% [50%, 70%] <i>N</i> = 93	12 [-4, 28]	0.15
Male Legislators	47% [40%, 53%] <i>N</i> = 202	51% [44%, 58%] <i>N</i> = 202	4 [-7, 16]	0.44
<i>Federal:</i>				
All Legislators	33% [26%, 40%] <i>N</i> = 172	38% [30%, 45%] <i>N</i> = 162	4 [-7, 15]	0.42
Female Legislators	37% [22%, 52%] <i>N</i> = 43	39% [24%, 54%] <i>N</i> = 46	2 [-20, 24]	0.86
Male Legislators	32% [24%, 40%] <i>N</i> = 129	37% [28%, 46%] <i>N</i> = 116	5 [-5, 16]	0.32

Note: The first two columns of Table A.6 show the response rates to the email messages sent from male and female students for different sets of legislators separated by level of office. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

Table A.7: Response Rates by the Treatment Name, Party Ideology, and Legislator Gender: Including Later Responses

Received Email	Male Sender	Female Sender	Difference	<i>p</i> -value of Difference
<i>Left Party:</i>				
All Legislators	51% [41%, 62%] <i>N</i> = 90	61% [51%, 70%] <i>N</i> = 99	10 [-4, 23]	0.17
Female Legislators	54% [37%, 70%] <i>N</i> = 39	59% [45%, 73%] <i>N</i> = 49	5 [-16, 27]	0.62
Male Legislators	49% [35%, 63%] <i>N</i> = 51	62% [48%, 76%] <i>N</i> = 50	13 [-8, 34]	0.21
<i>Center Party:</i>				
All Legislators	40% [32%, 48%] <i>N</i> = 141	47% [39%, 55%] <i>N</i> = 144	7 [-5, 20]	0.24
Female Legislators	38% [22%, 53%] <i>N</i> = 40	56% [40%, 71%] <i>N</i> = 45	18 [2, 34]	0.03
Male Legislators	41% [31%, 50%] <i>N</i> = 101	43% [33%, 53%] <i>N</i> = 99	3 [-12, 18]	0.69
<i>Right Party:</i>				
All Legislators	49% [41%, 58%] <i>N</i> = 138	52% [44%, 61%] <i>N</i> = 124	3 [-11, 17]	0.69
Female Legislators	65% [44%, 86%] <i>N</i> = 23	63% [42%, 83%] <i>N</i> = 24	-2 [-26, 20]	0.81
Male Legislators	46% [37%, 55%] <i>N</i> = 115	50% [40%, 60%] <i>N</i> = 100	4 [-11, 19]	0.60

Note: The first two columns of Table A.7 show the response rates to the email messages sent from male and female students for different sets of legislators separated by party ideology. The third column indicates the percentage-point-difference in response rates, with positive differences indicating a pro-female bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The *p*-values come from a linear probability model where *Email Response* is the dependent variable and the model includes strata fixed effects (gender of legislator) and robust standard errors clustered at the email account level.

A.5 Analyzing the Content of the Email Responses

In this section of the appendix, I further analyze the content of the email responses from the legislators. As I note in the main text, there are different ways to analyze the content of the email responses. We can think of these approaches as the *'micro'-mentorship approach*, the *quality of response approach*, and the *length of response approach*. Below, I describe each of the approaches in more detail and then present the results. It's important to note, though, that to avoid possible post-treatment bias, the analyses presented in this appendix are *not* conditional on having received an email response. As Coppock (2019, 1) explains, an email “response is a post-treatment outcome” and “conditioning on post-treatment outcomes ‘de-randomizes’ an experiment in the sense that the resulting treatment and control groups no longer have potential outcomes that are in expectation equivalent.” To avoid conditioning on having received a response, I redefine the outcome variables I'm about to create to include non-responses in the 0 category. What this means in practice is that an outcome variable that, say, indicates whether helpful advice was provided is coded as 0 if either (i) no response was received or (ii) a response was received but did not provide helpful advice; the outcome variable is coded 1 only if helpful advice was provided.

A.5.1 'Micro'-mentorship Approach

The *'micro'-mentorship approach* to evaluating the content of the email responses comes from Kalla, Rosenbluth and Teele (2018). They created several indicators of micro-mentorship by qualitatively coding whether email responses (1) were meaningful, (2) provided praise, (3) offered help, (4) provided a warning, or (5) provided advice.

1. *Meaningful Response*: The email response was coded 1 if it was determined to “contain real content”.
2. *Praise*: The email response was coded 1 if it either “praises student for an interest in a political career” or provides vague praise such as “good luck with everything” or “hope this helps.”
3. *Offer to help*: The email response was coded 1 if it indicated a willingness on the part of the

legislator to meet, talk on the phone, or email further, or a general offer to follow up such as “If you have any other specific questions, please let me know.”

4. *Warning*: The email response was coded 1 if it contained an explicit statement not to run, an encouragement to consider other career paths, or a warning of time commitment, work-life balance challenges, the difficulty of finding time for family, the challenges of fundraising, or the loss of privacy.
5. *Substantive Advice*: The email response was coded 1 if it contained either practical advice (e.g., motivational advice, get a business job, go to law school, get a different type of job, become involved in local community groups, attend local party or political meetings, volunteer, get a mentor, fundraising advice, run for student government, learn about the issues, get a good education, always put your values first, stay loyal to your political party) or personality/image advice (e.g., always have a professional appearance, have thick skin, learn to be extroverted, learn to deal with conflict).

A.5.2 Quality of Response Approach

The *quality of response approach*, which has recently been developed by Costa (2020), classifies “quality” and “satisfying” responses as those that are “not automated, answers the question, arrives promptly, is at least 400 characters in length, and includes a named greeting, invitation to follow up, link to a website, and a sign-off” (Costa 2020, 15). Responses that have all of these criteria are coded as 1. If a response is missing one of these criteria, then a specified amount is subtracted or ‘discounted’ from 1. The specific discount formula used by Costa (2020) is

$$\begin{aligned} \text{Discount} = & \text{Automated} \times 0.209 + \text{No named greeting} \times 0.048 + \text{No invite follow up} \times 0.0613 \\ & + \text{Did not answer question} \times 0.116 + \text{No website link} \times 0.0533 + \text{No sign-off} \times 0.0333 \\ & + \text{Characters} < 400 \times 0.0005 + \text{Days until response} \times 0.0033. \end{aligned}$$

The components of the discount formula are defined as:

- *Automated*: Since I classify automated responses as non-responses, any automated response in my sample actually receives a discount factor of 1. In other words, the overall quality of response for automated responses is always 0.
- *No named greeting*: This is coded 1 if the response is not personalized and 0 otherwise.

- *No invite follow up*: This is coded 1 if the response did not invite the student to follow up with further questions and 0 otherwise.
- *Did not answer question*= This is coded 1 if the response did not answer the student's question and 0 otherwise.
- *No website link*: This is coded 1 if the response did not include a website link and 0 otherwise.
- *No sign-off*: This is coded 1 if the response did not include a sign-off (eg. Sincerely, best, regards, etc) and 0 otherwise.
- *Characters < 400*: This is coded 1 if the response was fewer than 400 characters and 0 otherwise.
- *Days until response*: This is a count of the number of days until a response came; this is capped at 30 days.

To calculate the overall *quality* of an email response, I simply subtract the discount scores from 1 for all responses that were received. To avoid post-treatment bias, I code “no replies” as having a discount factor of 1 and, thus, an overall quality score of 0.

A.5.3 Length of Response Approach

The *length of response approach* assumes that longer responses are more substantively meaningful than shorter ones. For this metric, I evaluate (i) the word count, (ii) the log word count, and (iii) the number of characters in a response. To avoid post-treatment bias, I code “no replies” as having 0 words/characters and as $\log(1)$ for the log word count.

A.5.4 Results

Having described the three different approaches, we can now turn to the results. In Table A.8, I provide information about the content of the email responses by treatment name. Overall, I find a pro-women bias across all nine of the different metrics – the numbers in the *Difference* column are always positive. This pro-women bias is statistically significant in eight of these nine cases. The only metric on which the pro-women bias is not statistically significant is when we look at whether the email responses provides praise.

Table A.8: Quality of Response by Treatment Name

	Male Name	Female Name	Difference	<i>p</i> -value of Difference
	<i>N</i> = 888	<i>N</i> = 886		
<i>“Micro”-mentorship Approach</i>				
Meaningful Response	12% [9%, 14%]	16% [14%, 18%]	4 [1, 7]	0.01
Praise	8% [7%, 10%]	10% [8%, 12%]	2 [-1, 4]	0.35
Offer to help	23% [20%, 26%]	29% [26%, 32%]	6 [0.5, 11]	0.03
Warning	1% [0.4%, 2%]	3% [2%, 4%]	2 [0.3, 4]	0.03
Substantive advice	11% [9%, 13%]	14% [12%, 17%]	3 [1, 6]	0.01
<i>Quality of Response Approach</i>				
Quality Response	0.31 [0.29, 0.34]	0.37 [0.34, 0.39]	0.06 [0.02, 0.09]	0.00
<i>Length of Response Approach</i>				
Word count	36 [30, 41]	46 [39, 52]	10 [3, 17]	0.01
Log word count	0.76 [0.71, 0.82]	0.90 [0.84, 0.96]	0.14 [0.05, 0.22]	0.00
Character count	161 [137, 184]	206 [176, 235]	45 [14, 76]	0.01

Note: The first two columns of Table A.8 evaluate the content of the email responses depending on whether the original email message was sent by a male or female student. The third column indicates the difference in the content of the email response, with positive differences indicating a pro-women bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant.

In Table A.9, I provide information about the content of the email responses by treatment name and legislator gender. In regards to the *‘micro’-mentorship approach*, I find a pro-women bias exists among both female and male legislators across all metrics except for *Praise*. This pro-

women bias is larger and only statistically significant among male legislators. When it comes to the *quality of response approach*, I find a statistically significant pro-women bias among both female and male legislators. On the *length of response approach*, I also find a pro-women bias, with email responses to female students being significantly longer than those to male students. This pro-women bias is only statistically significant for male legislators.

Table A.9: Quality of Response by Treatment Name and Legislator Gender

	Male Legislators				Female Legislators			
	Male Sender	Female Sender	Difference	<i>p</i> -value of Difference	Male Sender	Female Sender	Difference	<i>p</i> -value of Difference
	<i>N</i> = 629	<i>N</i> = 628			<i>N</i> = 259	<i>N</i> = 258		
<i>“Micro”-mentorship Approach</i>								
Meaningful Response	11% [9%, 14%]	16% [13%, 19%]	5 [1, 9]	0.01	13% [9%, 17%]	15% [10%, 19%]	2 [-4, 8]	0.49
Praise	7% [5%, 10%]	9% [7%, 11%]	2 [-1, 5]	0.28	11% [7%, 15%]	11% [7%, 15%]	0 [-6, 6]	0.99
Offer to help	25% [21%, 28%]	30% [27%, 34%]	5 [0.3, 11]	0.04	19% [14%, 24%]	26% [21%, 31%]	7 [-2, 16]	0.12
Warning	1% [0.2%, 2%]	4% [2%, 5%]	3 [0.5, 5]	0.02	1% [-0.2%, 2%]	2% [0.03%, 3%]	1 [-2, 3]	0.71
Substantive advice	10% [8%, 13%]	15% [12%, 18%]	5 [0.4, 8]	0.03	12% [8%, 16%]	13% [9%, 17%]	1 [-5, 6]	0.76
<i>Quality of Response Approach</i>								
Quality Response	0.32 [0.29, 0.34]	0.36 [0.34, 0.39]	0.04 [0.004, 0.09]	0.03	0.30 [0.26, 0.35]	0.38 [0.34, 0.42]	0.08 [0.02, 0.14]	0.01
<i>Length of Response Approach</i>								
Word count	33 [28, 38]	43 [36, 50]	10 [2, 18]	0.02	42 [31, 53]	53 [39, 67]	11 [-5, 27]	0.19
Log word count	0.77 [0.70, 0.84]	0.89 [0.82, 0.96]	0.12 [0.01, 0.22]	0.03	0.75 [0.63, 0.86]	0.94 [0.83, 1.05]	0.19 [0.04, 0.35]	0.02
Character count	148 [123, 173]	194 [162, 226]	46 [8, 84]	0.02	193 [140, 246]	235 [170, 300]	42 [-32, 117]	0.25

Note: The first two columns of Table A.9 evaluate the content of the email responses from male legislators. The third column indicates the difference in the content of the responses, with positive differences indicating a pro-women bias. 95% confidence intervals are shown in square brackets. The fourth column indicates whether the differences are statistically significant. The fifth and sixth columns evaluate the content of the email responses from female legislators. The seventh column indicates the difference in the content of the responses, with positive differences indicating a pro-women bias. The 95% confidence intervals are shown in square brackets. The eighth column indicates whether the differences are statistically significant.

APPENDIX B

This appendix provides supplementary information for Chapter Three “Do Voters Discriminate Against Female Candidates? Evidence from an Experiment in South Korea.”

B.1 Experimental Design in Korean

Figure B.1: Experimental Design in Korean

귀하가 선호하는 정당이 국회의원 선거에 출마할 다음 두 후보자들을 고려하고 있다고 가정합니다.

다음 두 후보자의 이력서를 검토하십시오:

	후보자 1	후보자 2
성별	남성	여성
이념 성향	진보적	보수적
전문 분야	국가 안보	보건 의료
정치 스타일	인정 많은 정치에의 접근법	강경함 정치에의 접근법
정치 경험	3 년	8 년
나이	35 세	45 세
결혼 여부	싱글	기혼
자녀의 수	무자녀	1 명

위의 정보를 바탕으로 두 사람 중 한 명을 선택해야 한다면 국회의원으로 어떤 후보를 원하십니까?

후보자 1 <input type="radio"/>	후보자 2 <input type="radio"/>
--------------------------------	--------------------------------

위의 정보에 근거하여, 어느 후보가 선거에서 이길 확률이 더 높다고 생각하십니까?

후보자 1 <input type="radio"/>	후보자 2 <input type="radio"/>
--------------------------------	--------------------------------

Note: Figure B.1 shows an example of one set of candidate profiles that was presented to a respondent in the conjoint experiment in Korean.

B.2 Descriptive Information About the Survey & Sample

Here, I provide additional information about my survey and sample.

B.2.1 Survey

Introduction: Respondents were asked to imagine a race for National Assembly with the following introduction:

In the next part of the survey, we are going to present you with six pairs of candidates that your preferred party is considering to run in a district that has no incumbent candidate running for National Assembly.

For each pair of candidates, we are going to ask you (1) who you would prefer as a member of National Assembly and (2) who you think has a better chance of winning the election.

Please act as if you were a voter in their district.

We will provide you with several pieces of information about each candidate who might run for the seat. For each pair of candidates, please indicate which of the two candidates you would personally prefer.

This exercise is purely hypothetical. But even if you are not entirely sure, please indicate which of the two you prefer.

The Outcome Variables:

- *Preferred Candidate:* Based on the information above, if you had to choose between them, which candidate would you prefer as a member of National Assembly?
- *Winning Candidate:* Based on the information above, which candidate do you think has a better chance of winning the elections?

B.2.2 Sample

My sample consists of 2,258 South Korean citizens. In Table B.1, I provide descriptive statistics of my sample.

Table B.1: Demographics of Survey Sample

	Number	Percentage
<i>Sex</i>		
<i>Male</i>	1,091	48.3
<i>Female</i>	1,167	51.7
<i>Gender Role Attitudes</i>		
<i>'Less Traditional'</i>	1,059	46.9
<i>'More Traditional'</i>	1,199	53.1
<i>Partisanship</i>		
<i>'Left'</i>	626	27.7
<i>'Right'</i>	217	9.6
<i>No Preferred Party</i>	1,415	62.7
<i>Annual Income</i>		
<i>Lower Level</i>	365	16.2
<i>Middle Level</i>	1,490	66.0
<i>Upper Level</i>	403	17.8
<i>Age Group</i>		
18-29	781	34.6
30-39	541	24.0
40-49	407	18.0
50-59	296	13.1
60 above	233	10.3
<i>Education</i>		
<i>No BA Degree</i>	918	40.7
<i>BA or Above</i>	1,340	59.3

B.3 Full Regression Results – Preferred Candidate

Here, I report the full regression results for the regression used to compute the average marginal component effects (AMCEs) visualized in Figure 3.3 in the main text. The dependent variable is *Preferred Candidate*, coded 1 if the respondent chose that candidate profile as their preferred profile and 0 if not. This outcome is regressed on the various candidate attributes (omitting one reference category as the baseline level). I also report the marginal means for all candidate attributes.

B.3.1 Regression Results

Table B.2: Effects of Candidate Attributes on Being Preferred

Average Marginal Component Effects (AMCE)	
Sex (baseline: <i>Male</i>)	
<i>Female</i>	0.026*** (0.006)
Personality Trait (baseline: <i>Tough Approach to Politics</i>)	
<i>Compassionate Approach to Politics</i>	0.006 (0.006)
Issue Specialization (baseline: <i>National Security</i>)	
<i>Foreign Policy</i>	0.011 (0.008)
<i>Healthcare</i>	0.004 (0.009)
<i>Social Welfare</i>	0.017 (0.009)
Ideological Placement (baseline: <i>Right-Leaning</i>)	
<i>Left-Leaning</i>	0.065*** (0.007)
Number of Years in Politics (baseline: <i>None</i>)	
<i>Three Years</i>	0.062*** (0.008)
<i>Eight Years</i>	0.109*** (0.008)
Age (baseline: <i>35 years old</i>)	
<i>45 years old</i>	0.015* (0.008)
<i>65 years old</i>	-0.072*** (0.008)
Marital Status (baseline: <i>Single</i>)	
<i>Divorced</i>	-0.033*** (0.007)
<i>Married</i>	0.045*** (0.007)
Number of Children (baseline: <i>No Children</i>)	
<i>One Child</i>	0.030*** (0.007)
<i>Three Children</i>	0.030*** (0.008)
Number of Observations	27,096
Number of Respondents	2,258

Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$ *** $p < 0.001$ (two-tailed).

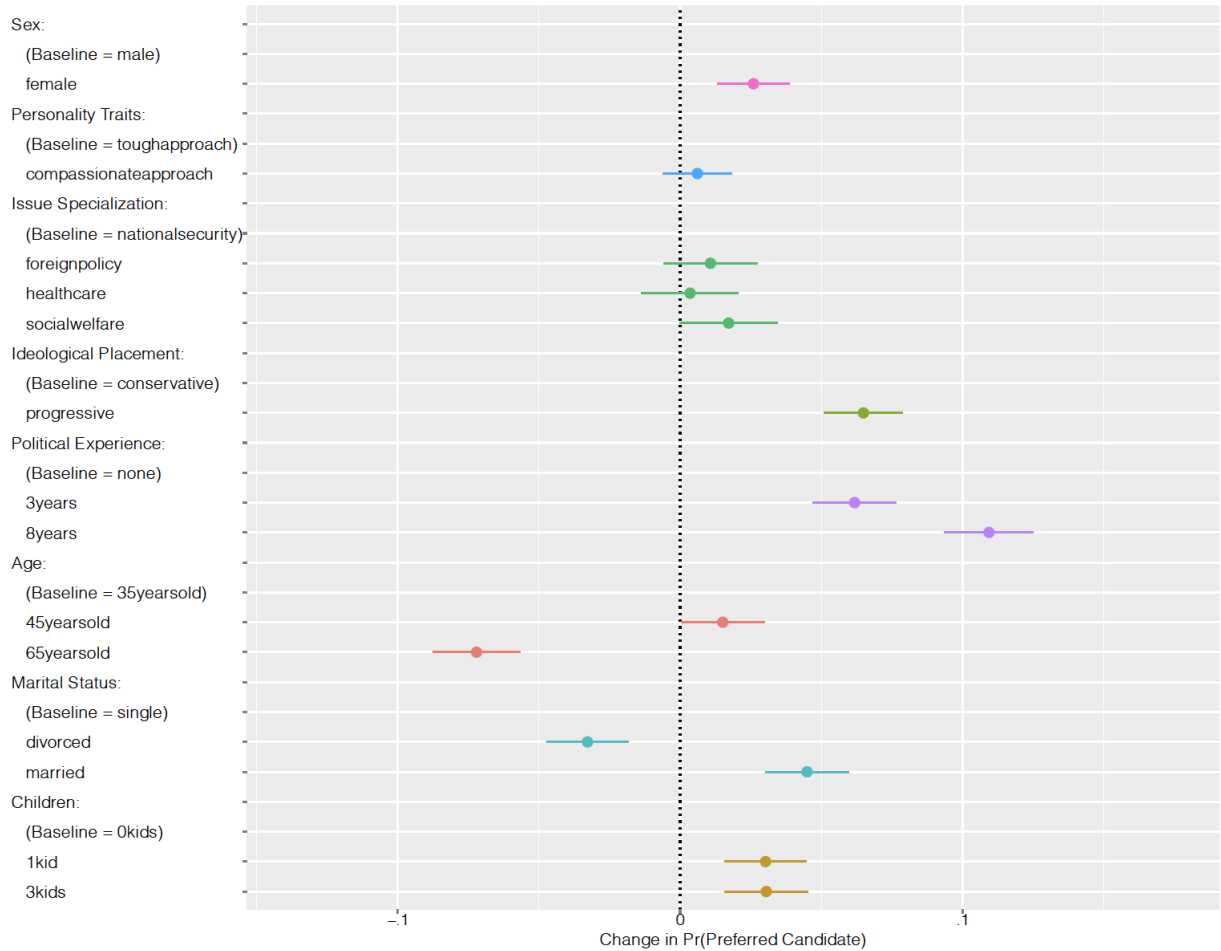
B.3.2 AMCE of Candidate Attributes

Figure B.2 shows estimates of the effects of the randomly assigned candidate attribute values on the probability of being preferred as a representative in the National Assembly. As explained by Hainmueller, Hopkins and Yamamoto (2014), the *average marginal component effect (AMCE)* represents the marginal effect of a specific attribute over the joint distribution of all other attributes. I am able to estimate this simple linear regression because each attribute was randomly assigned independently of all other attributes. I use cluster-robust standard errors at the respondent level.

As explained in the main text, I find that respondents do not exhibit any bias against female political candidates. In fact, the average effect of being a woman (versus a man) is associated with an approximately 2.6 percentage point increase (S.E. = 0.65) in being chosen as the *Preferred Candidate*.

In addition, specializing in foreign policy, healthcare, or social welfare (versus national security) has no effect on being chosen as the *Preferred Candidate*. Presenting a hypothetical candidate as having a progressive as opposed to a conservative ideology increased the probability of respondents choosing the profile as their preferred candidate, all else constant. Likewise, the average effect of having political experience (versus none) is associated with an increase in the probability of being chosen as the *Preferred Candidate*, when respondents are also given information about the other seven attributes. Being older (65 versus 35 years old) has a negative effect, all else equal. Family status also affects preference: Having children (versus not having children) has a positive effect, all else constant. Being married (versus being single) has a positive effect and being divorced has a negative effect (versus being single), all else equal.

Figure B.2: Effects of Candidate Attributes on Being Preferred – AMCEs

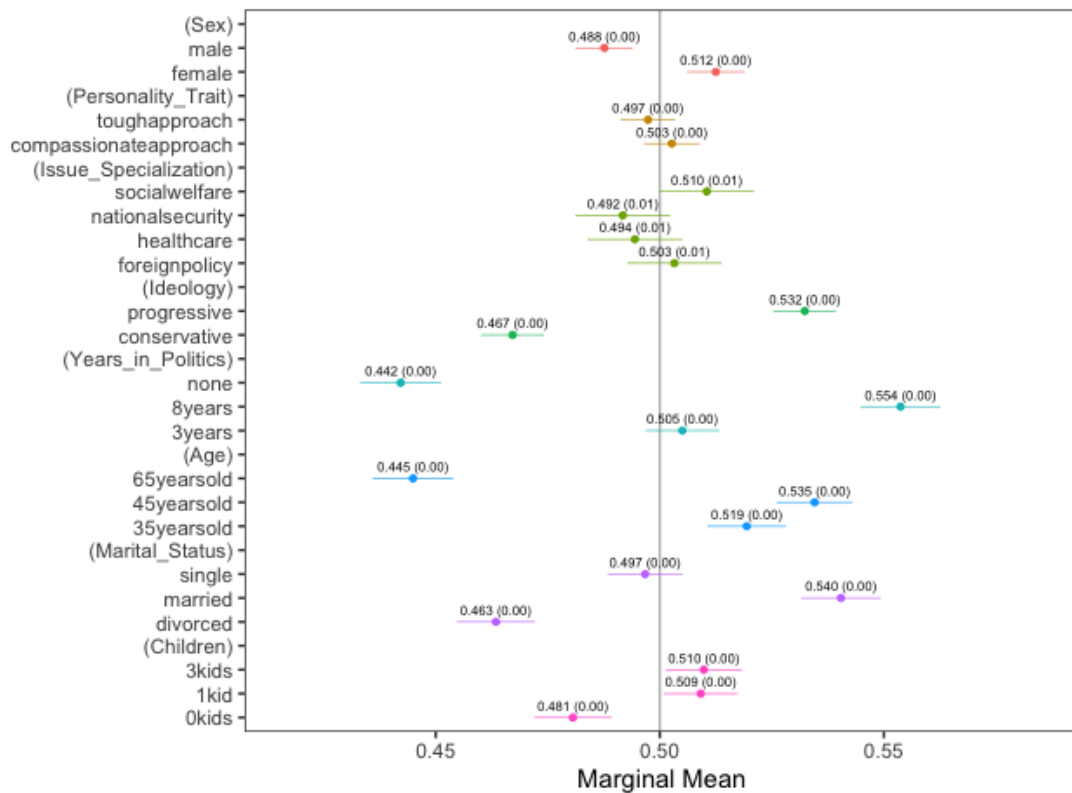


Note: Figure B.2 shows estimates of the effects of the randomly assigned candidate attribute values on the probability of being the preferred candidate. The attribute values without estimates denote the attribute value that is the reference category for each attribute.

B.3.3 Marginal Means of Candidate Attributes

Following the advice of Leeper, Hobolt and Tilley (2020), I also report marginal means (MM). While AMCEs restrict the AMCE for the reference category to zero (or undefined), here there is information conveyed about the preferences of respondents for *all* attribute levels. The results are consistent. The AMCE for a candidate sex reported in the main text is 0.026, indicating a 2.6 percentage points increase in favorability. Here, that effect is 0.024 (or 2.4 percentage points), reflected in the marginal means for female (0.51) and male candidates (0.49). The results for the other attributes are similar to those reported above but now we have effects for the baseline categories as well.

Figure B.3: Effects of Candidate Attributes on Being Preferred – MMs



Note: Figure B.3 shows the marginal means (MMs) of the randomly assigned candidate attribute values on the probability of being the preferred candidate.

B.4 Full Regression Results – *Winning Candidate*

Here, I report the full regression results for the regression used to compute the average marginal component effects (AMCEs) visualized in Figure 3.6 in the main text. The dependent variable is *Winning Candidate*, coded 1 if the respondent chose that candidate profile as the winning profile and 0 if not. This outcome is regressed on the various candidate attributes (omitting one reference category as the baseline level). I also report the marginal means for all candidate attributes.

B.4.1 Regression Results

Table B.3: Effects of Candidate Attributes on Being Winner

Average Marginal Component Effects (AMCE)	
Sex (baseline: <i>Male</i>)	
<i>Female</i>	-0.045*** (0.006)
Personality Trait (baseline: <i>Tough Approach to Politics</i>)	
<i>Compassionate Approach to Politics</i>	-0.001 (0.006)
Issue Specialization (baseline: <i>National Security</i>)	
<i>Foreign Policy</i>	0.015 (0.008)
<i>Healthcare</i>	0.005 (0.009)
<i>Social Welfare</i>	-0.002 (0.009)
Ideological Placement (baseline: <i>Right-Leaning</i>)	
<i>Left-Leaning</i>	0.046*** (0.007)
Number of Years in Politics (baseline: <i>None</i>)	
<i>Three Years</i>	0.087*** (0.008)
<i>Eight Years</i>	0.158*** (0.008)
Age (baseline: <i>35 years old</i>)	
<i>45 years old</i>	0.043*** (0.008)
<i>65 years old</i>	-0.021** (0.008)
Marital Status (baseline: <i>Single</i>)	
<i>Divorced</i>	-0.040*** (0.007)
<i>Married</i>	0.050*** (0.007)
Number of Children (baseline: <i>No Children</i>)	
<i>One Child</i>	0.033*** (0.007)
<i>Three Children</i>	0.036*** (0.008)
Number of Observations	27,096
Number of Respondents	2,258

Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$ *** $p < 0.001$ (two-tailed).

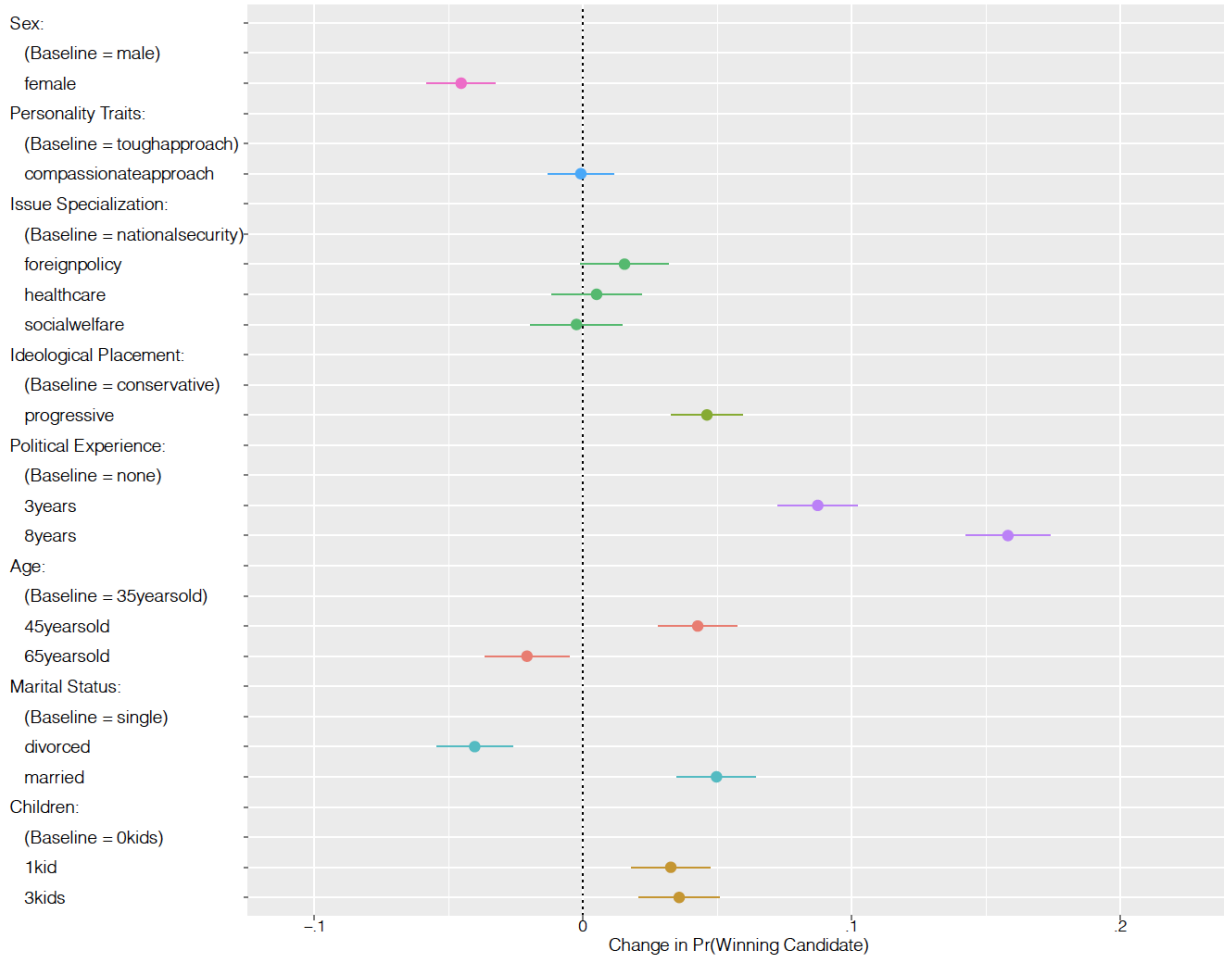
B.4.2 AMCE of Candidate Attributes

Figure B.4 shows estimates of the effects of the randomly assigned candidate attribute values on the probability of winning the election. As explained by Hainmueller, Hopkins and Yamamoto (2014), the *average marginal component effect (AMCE)* represents the marginal effect of a specific attribute over the joint distribution of all other attributes. I am able to estimate this simple linear regression because each attribute was randomly assigned independently of all other attributes. I use cluster-robust standard errors at the respondent level.

As explained in the main text, I find that respondents exhibit bias against female political candidates. In fact, the average effect of being a woman (versus a man) is associated with an approximately 4.5 percentage point decrease (S.E. = 0.65) in being chosen as the *Winning Candidate*.

In addition, approach to politics has no effect on being chosen as the *Winning Candidate*. Likewise, issue specialization – specializing in foreign policy, healthcare, or social welfare (versus national security) – has no effect on being chosen as the *Winning Candidate*. Presenting a hypothetical candidate as having a progressive as opposed to a conservative ideology increased the probability of respondents choosing the profile as the winning candidate, all else constant. Likewise, the average effect of having political experience (versus none) is associated with an increase in the probability of being chosen as the *Winning Candidate*, when respondents are also given information about the other seven attributes. Being older (65 versus 35 years old) has a negative effect, all else equal. Family status also affects preference: Having children (versus not having children) has a positive effect, all else constant. Being married (versus being single) has a positive effect and being divorced has a negative effect (versus being single), all else equal.

Figure B.4: Effects of Candidate Attributes on Being Winner – AMCEs

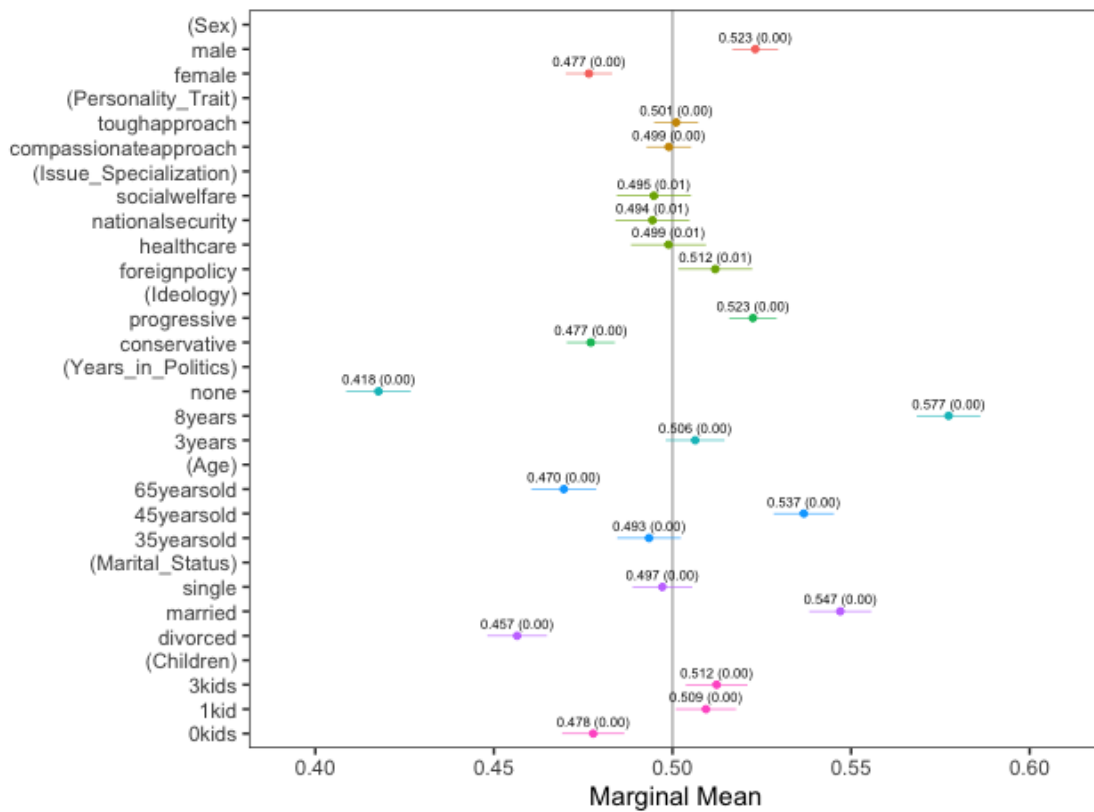


Note: Figure B.4 shows estimates of the effects of the randomly assigned candidate attribute values on the probability of winning the election. The attribute values without estimates denote the attribute value that is the reference category for each attribute.

B.4.3 Marginal Means of Candidate Attributes

Following the advice of Leeper, Hobolt and Tilley (2020), I also report marginal means (MM). While AMCEs restrict the AMCE for the reference category to zero (or undefined), here there is information conveyed about the preferences of respondents for *all* attribute levels. The results are consistent. The AMCE for a candidate sex in the main text is 0.045, indicating a 4.5 percentage points decrease in winning the election. Here, that effect is 0.046 (or 4.6 percentage points), reflected in the marginal means for female (0.52) and male candidates (0.48). The results for the other attributes are similar to those reported above but now we have effects for the baseline categories as well.

Figure B.5: Effects of Candidate Attributes on Being Winner – MMs



Note: Figure B.5 shows the marginal means (MMs) of the randomly assigned candidate attribute values on the probability of winning the election.

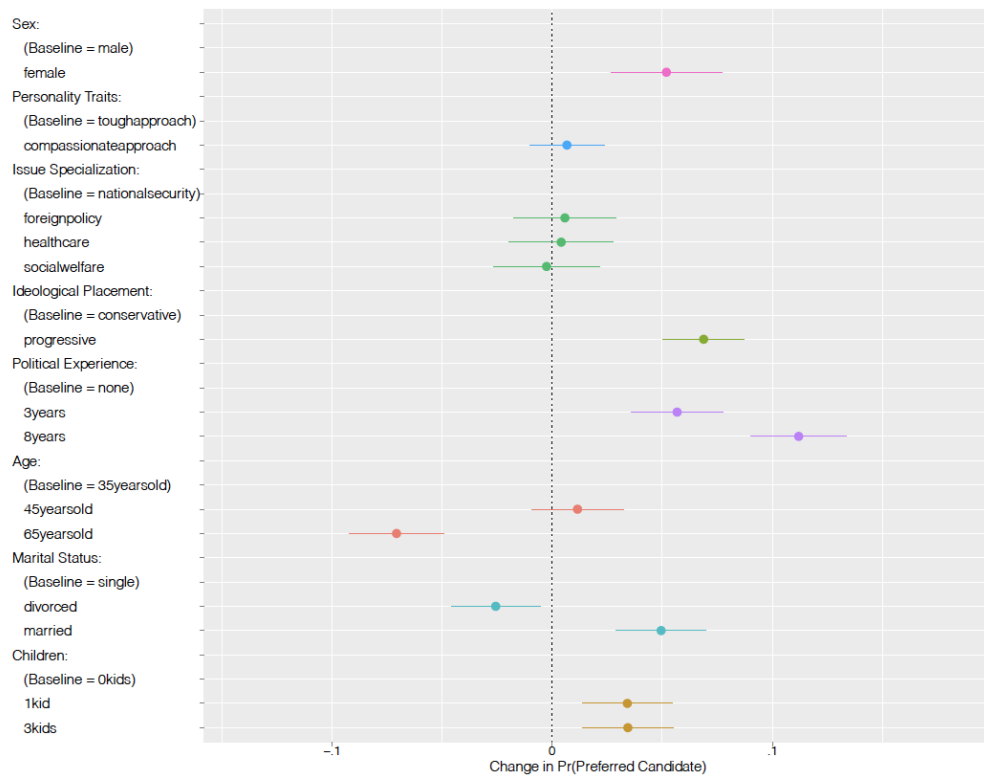
B.5 Mixed-Sex Pairings

In this appendix, I evaluate the effect of candidate sex on candidate evaluations for mixed-sex pairings (one female and one male candidate).

B.5.1 AMCE of Candidate Attributes – Mixed-Sex Pairings

In the main text, the average effect of being a woman (versus a man) is associated with an approximately 2.6 percentage point increase (S.E. = 0.65) in being the preferred candidate. This average effect is across both mixed-sex and same-sex (two male candidates or two female candidates). A concern is that the effect of candidate sex may be underestimated since it includes both of these types of candidate pairs. Among the 13,548 evaluated candidate pairings, 6,805 (50.23 %) are same-sex pairings and 6,743 (49.77 %) are mixed-sex pairings. When I exclude same-sex pairings and evaluate only mixed-sex candidate pairings, I find that the boost for female candidates is 5.2 percentage points (S.E. = 1.3).

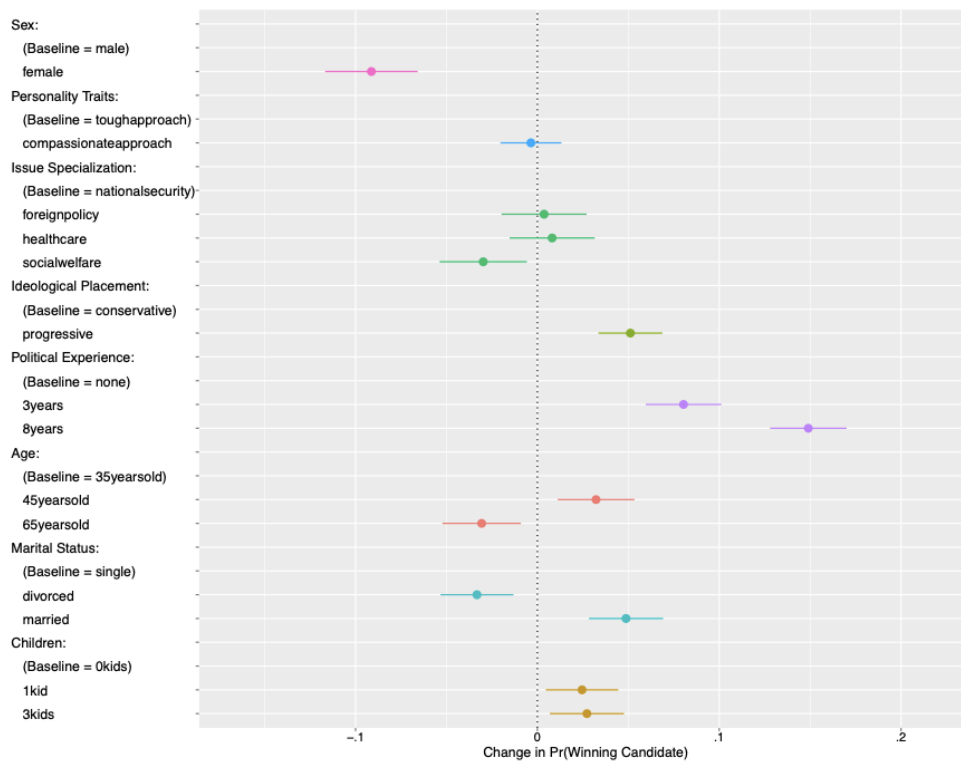
Figure B.6: Effects of Candidate Attributes on Being Preferred – AMCEs for Mixed-Sex Pairings



Note: Figure B.6 shows estimates of the effects of the randomly assigned candidate attribute values on the probability of being preferred for only mixed-sex pairings. The attribute values without estimates denote the attribute value that is the reference category for each attribute.

In terms of *Winning Candidate*, as reported in the main text, when both mixed-sex and same-sex candidate pairings are included, the average effect of being a woman (versus a man) is associated with an approximately 4.5 percentage point decrease (S.E. = 0.65) in being selected as the winning candidate. When I exclude same-sex pairings and evaluate only mixed-sex candidate pairings, I find that the bias against female candidates increases to 9.1 percentage points (S.E. = 1.3).

Figure B.7: Effects of Candidate Attributes on Being Winner – AMCEs for Mixed-Sex Pairings

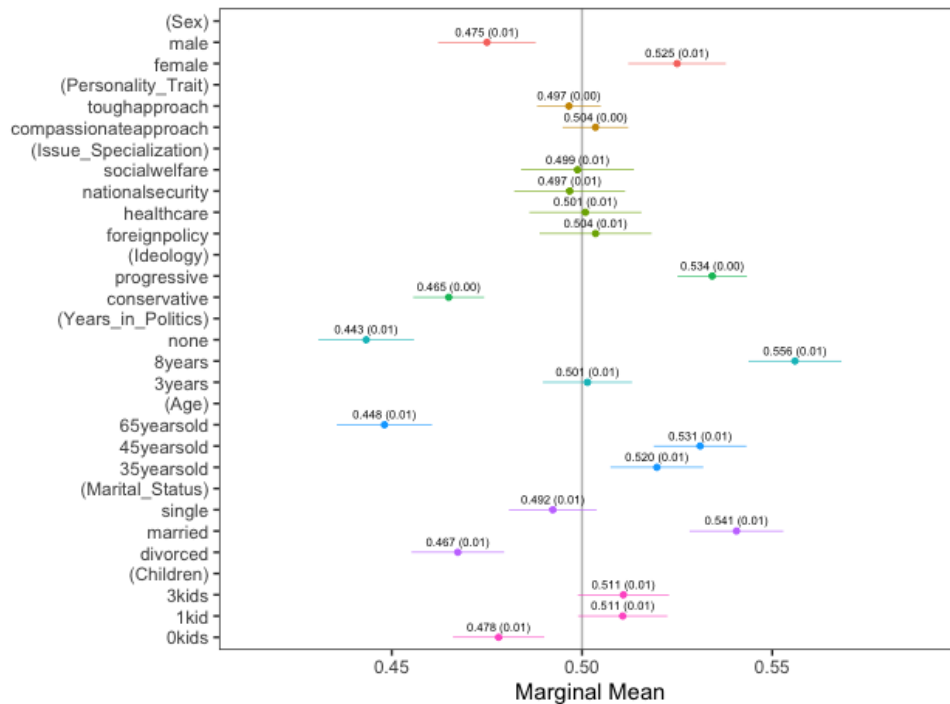


Note: Figure B.7 shows estimates of the effects of the randomly assigned candidate attribute values on the probability of being winning candidate for only mixed-sex pairings. The attribute values without estimates denote the attribute value that is the reference category for each attribute.

B.5.2 Marginal Means of Candidate Attributes – Mixed-Sex Pairings

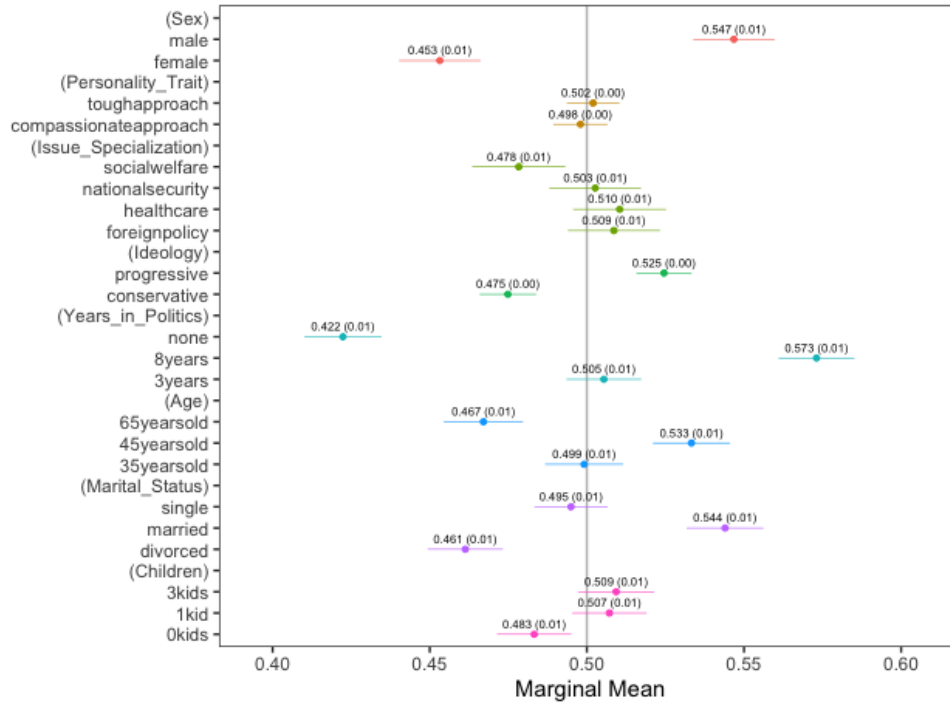
Following the advice of Leeper, Hobolt and Tilley (2020), I also report marginal means (MM) for *Preferred Candidate* in Figure B.8 and for *Winning Candidate* in Figure B.9. The results remain consistent.

Figure B.8: Effects of Candidate Attributes on Being Preferred – MMs for Mixed-Sex Pairings



Note: Figure B.8 shows the marginal means (MMs) of the randomly assigned candidate attribute values on the probability of being the preferred candidate.

Figure B.9: Effects of Candidate Attributes on Being Winner – MMs for Mixed-Sex Pairings

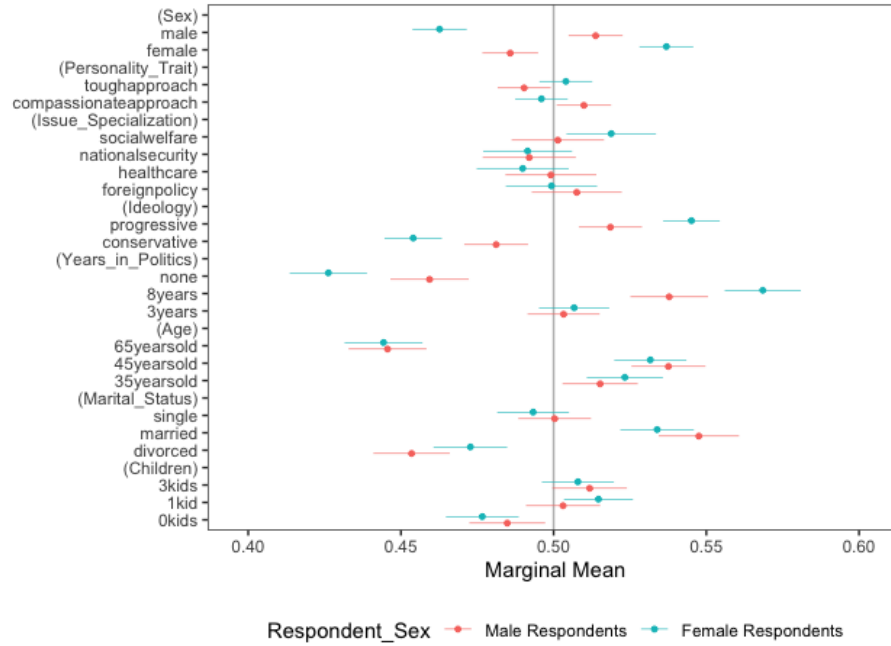


Note: Figure B.9 shows the marginal means (MMs) of the randomly assigned candidate attribute values on the probability of being the winning candidate.

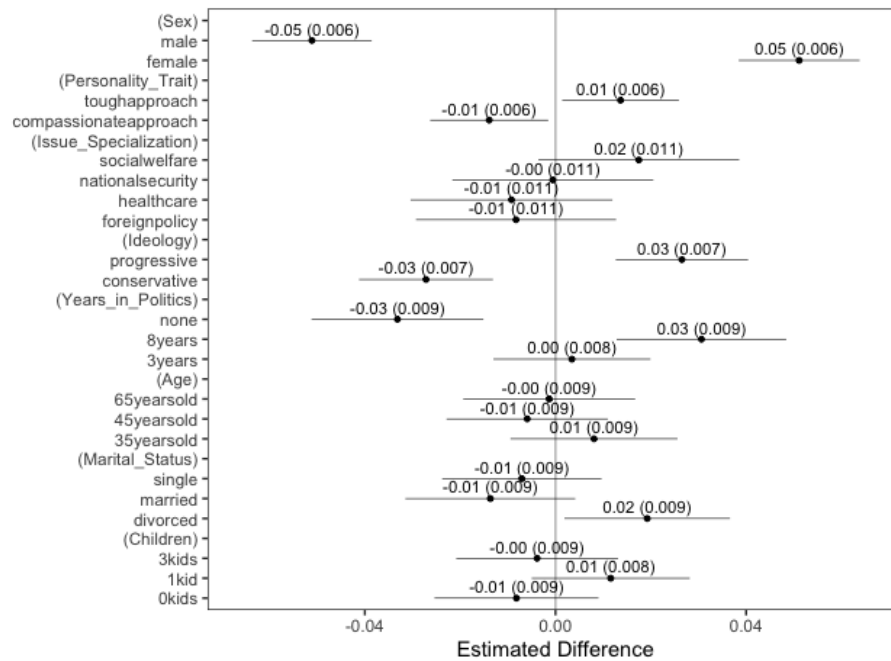
B.6 Effects of Candidate Attributes by Respondent Sex

In this appendix, I report the effect of all candidate attributes on choice for male and female respondents. Figure B.10 shows the effect of candidate sex on *Preferred Candidate* across female and male respondents. The top figure shows the conditional marginal mean by respondent sex, and the bottom figure shows the differences in conditional marginal means. Figure B.11 shows the results for *Winning Candidate*.

Figure B.10: Effects of Candidate Sex on Being Preferred by Respondent Sex



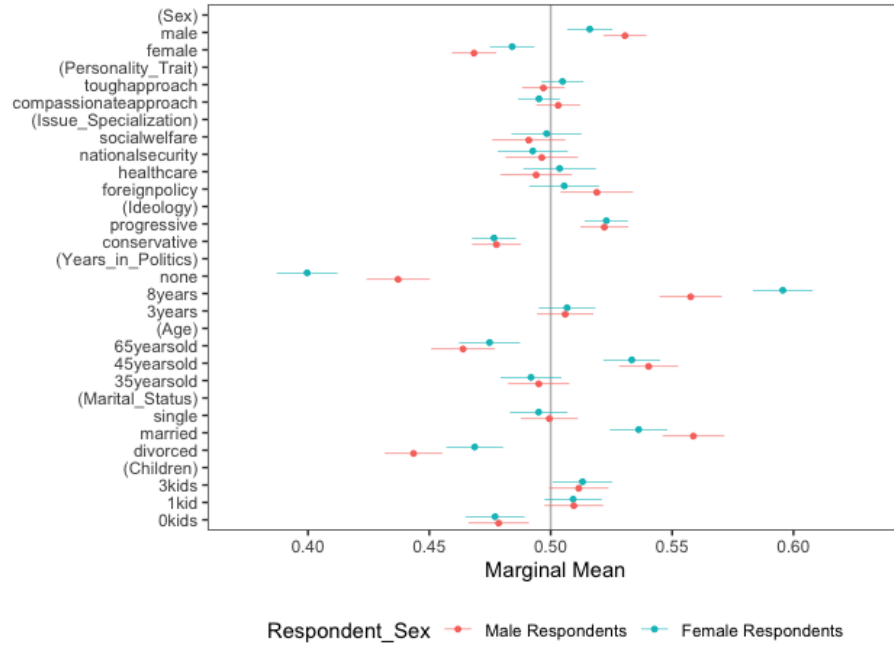
a) Marginal Means by Respondent Sex



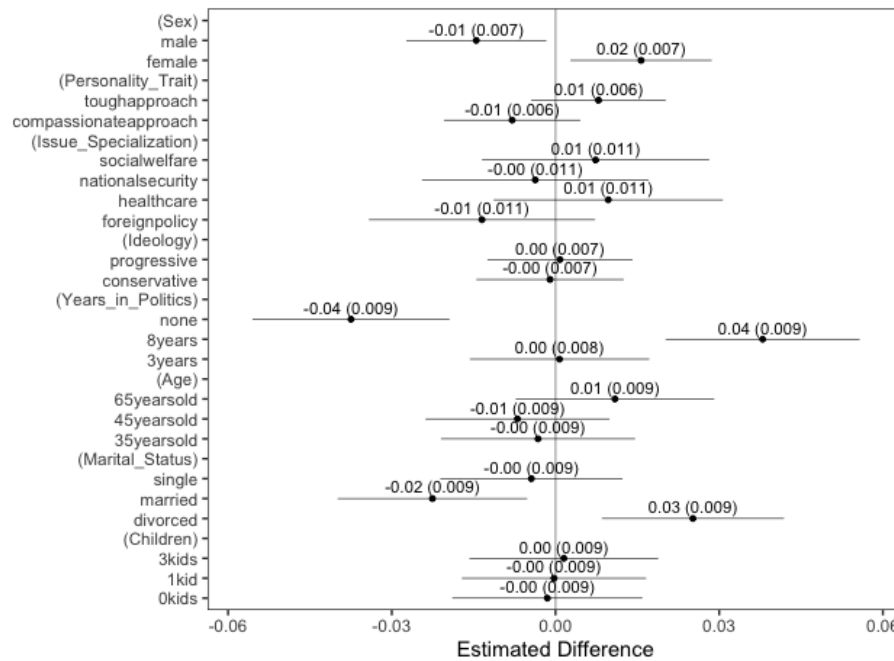
b) Difference in Marginal Means across Respondent Sex

Note: Figure B.10 shows the effect of candidate's sex on *Preferred Candidate* across respondent sex. Top plot shows marginal means and bottom plot shows the differences in marginal means. These plots show estimates of the effects of the randomly assigned candidates' sex on the probability of being chosen as the preferred candidate by respondent sex. Estimates are based on the regression estimators with clustered standard errors; bars represent 95% confidence intervals.

Figure B.11: Effects of Candidate Sex on Being Winner by Respondent Sex



a) Marginal Means by Respondent Sex



b) Difference in Marginal Means across Respondent Sex

Note: Figure B.11 shows the effect of candidate's sex on *Winning Candidate* across respondent sex. Top plot shows marginal means and bottom plot shows the differences in marginal means. These plots show estimates of the effects of the randomly assigned candidates' sex on the probability of being chosen as the winning candidate by respondent sex. Estimates are based on the regression estimators with clustered standard errors; bars represent 95% confidence intervals.

APPENDIX C

This appendix provides supplementary information for Chapter Four “Reexamining the Supply and Demand Framework for Explaining Women’s Descriptive Representation.”

C.1 Sample and Descriptive Statistics

In this section of the appendix, I provide descriptive statistics about my variables and additional information about my sample.

C.1.1 Descriptive Statistics

In Table C.1, I provide descriptive statistics for my variables. Recall that *Women’s Representation* captures the percentage of women representatives in lower house legislatures. It ranges from a low of 0 in countries like Micronesia, Papua New Guinea, Vanuatu, and Yemen to a high of 63.6% in Rwanda. *Supply (Labor Market)* captures the ratio of female to male labor force participation and varies from a low of 8.61 in Yemen in 2017 to a high of 111.01 in Mozambique in 2016. *Supply (Tertiary Education)* captures women’s tertiary education gross enrollment ratio and ranges from a low of 0 in Djibouti in 1991 to a high of 148.39 in Cuba in 2008. *Demand (Political Leaders)* captures the percentage of respondents in a country who disagree or strongly disagree with the claim that men make better political leaders than women and ranges from a low of 7.98% in Egypt in 2008 to a high of 91.82% in Sweden in 2006. *Demand (Jobs Scarce)* captures the percentage of respondents in a country who disagree or strongly disagree with the claim that men have more right to a job than women when jobs are scarce and ranges from a low of 0.4% in Egypt in 2001 to a high of 97.50% in Sweden in 2009.

C.1.2 Sample

In Table C.1, I present the results from the interactive model shown in Eq.4.2.1 for all four possible combinations of my supply and demand variables. The unit of analysis is the survey-country-year. For those who are interested, I now indicate the sample observations used in each of the four interactive models.

Table C.1: Descriptive Statistics, 1990-2018

Variable	Mean	Standard Deviation	Minimum	Maximum
<i>Dependent Variable</i>				
<i>Women's Representation</i>	14.79	10.87	0	63.8
<i>Independent Variables</i>				
<i>Supply (Labor Market)</i>	67.88	20.51	8.61	111.01
<i>Supply (Tertiary Education)</i>	25.82	27.16	0	148.39
<i>Demand (Political Leaders)</i>	52.01	20.71	7.98	91.82
<i>Demand (Jobs Scarce)</i>	51.11	22.21	0.4	97.50
<i>Control Variables</i>				
<i>Democracy</i>	3.20	6.59	-10	10
<i>Effective Quota</i>	0.11	0.32	0	1

C.1.2.1 Model (1): Supply (Labor Market) and Demand (Political Leaders)

Albania [1998, 2002]; Algeria [2002, 2014]; Argentina [1995, 1999, 2006]; Armenia [1997, 2011]; Australia [1995, 2005, 2012]; Azerbaijan [1997, 2011]; Bangladesh [1996, 2002]; Belarus [2011]; Brazil [1997, 2006, 2014]; Bulgaria [1997, 2006]; Burkina Faso [2007]; Canada [2000, 2006]; Chile [1996, 2000, 2006, 2012]; China [1995, 2001, 2007, 2013]; Colombia [1998, 2005, 2012]; Croatia [1996]; Cyprus [2006, 2011]; Czech Republic [1998]; Dominican Republic [1996]; Ecuador [2013]; Egypt [2001, 2008, 2012]; El Salvador [1999]; Estonia [1996, 2011]; Ethiopia [2007]; Finland [1996, 2005]; France [2006]; Georgia [1996, 2009, 2014]; Germany [1997, 2006, 2013]; Ghana [2007]; Guatemala [2004]; Hungary [1998, 2009]; India [1995, 2001, 2006, 2012]; Indonesia [2001, 2006]; Iran [2000, 2007]; Iraq [2013]; Italy [2005]; Japan [1995, 2000, 2005, 2010]; Jordan [2001, 2007, 2014]; Kazakhstan [2011]; Korea [1996, 2001, 2005, 2010]; Kuwait [2014]; Kyrgyzstan [2003, 2011]; Latvia [1996]; Lebanon [2013]; Libya [2014]; Lithuania [1997]; Macedonia [1998, 2001]; Malaysia [2006, 2012]; Mali [2007]; Mexico [1996, 2000, 2005, 2012]; Moldova [1996, 2002, 2006]; Morocco [2001, 2007, 2011]; Netherlands [2006, 2012]; New Zealand [1998, 2004, 2011]; Nigeria [2000, 2012]; Norway [1996, 2007]; Pakistan [1997, 2012]; Peru [1996, 2001, 2006, 2012]; Philippines [1996, 2001, 2012]; Poland [1997, 2005, 2012]; Qatar [2010]; Romania [1998, 2005, 2012]; Russia [1995, 2006, 2011]; Rwanda [2007, 2012]; Serbia [2006]; Singapore [2002, 2012]; Slovakia [1998]; Slovenia [1995, 2005, 2011]; South Africa [1996, 2001, 2006, 2013]; Spain [1995, 2000, 2007, 2011]; Sweden [1996, 1999, 2006, 2011]; Switzerland [2007]; Tanzania [2001]; Thailand [2007, 2013]; Trinidad and Tobago [2006, 2010]; Tunisia [2013]; Turkey [1996, 2001, 2007, 2012]; Uganda [2001]; Ukraine [1996, 2006, 2011]; United Kingdom [2005]; United States [1995, 1999, 2006, 2011]; Uruguay [1996, 2006, 2011]; Uzbekistan [2011]; Venezuela [1996, 2000]; Vietnam [2001, 2006]; Zambia [2007]; Zimbabwe [2001, 2012].

C.1.2.2 Model (2): Supply (Labor Market) and Demand (Jobs Scarce)

Albania [1998, 2002, 2008]; Algeria [2002, 2014]; Argentina [1991, 1995, 1999, 2006]; Armenia [1997, 2008, 2011]; Australia [1995, 2005, 2012]; Austria [1990, 1999, 2008]; Azerbaijan [1997, 2011]; Bangladesh [1996, 2002]; Belarus [2000, 2008, 2011]; Belgium [1900, 1999, 2009]; Brazil [1991, 1997, 2006, 2014]; Bulgaria [1991, 1997, 1999, 2006, 2008]; Burkina Faso [2007]; Canada [1990, 2000, 2006]; Chile [1990, 1996, 2000, 2006, 2012]; China [1995, 2001, 2007, 2013]; Colombia [1998, 2012]; Croatia [1996, 1999, 2008]; Cyprus [2006, 2008, 2011]; Czech Republic [1998, 1999, 2008];

Denmark [1990, 1999, 2008]; Dominican Republic [1996]; Ecuador [2013]; Egypt [2001, 2008, 2012]; El Salvador [1999]; Estonia [1996, 1999, 2008, 2011]; Ethiopia [2007]; Finland [1990, 1996, 2000, 2005, 2009]; France [1990, 1999, 2006, 2008]; Georgia [1996, 2008, 2009, 2014]; Germany [1990, 1997, 1999, 2006, 2008, 2013]; Ghana [2007]; Greece [1999, 2008]; Guatemala [2004]; Hungary [1991, 1998, 1999, 2008, 2009]; India [1990, 1995, 2001, 2006, 2012]; Indonesia [2001, 2006]; Iran [2013]; Iraq [2013]; Ireland [1990, 1999, 2008]; Italy [1990, 1999, 2005, 2009]; Japan [1990, 1995, 2000, 2005, 2010]; Jordan [2001, 2007, 2014]; Kazakhstan [2011]; Korea [1990, 1996, 2001, 2005, 2010]; Kuwait [2014]; Kyrgyzstan [2003, 2011]; Latvia [1996, 1999, 2008]; Lebanon [2013]; Libya [2014]; Lithuania [1997, 1999, 2008]; Luxembourg [1999, 2008]; Macedonia [1998, 2001]; Malaysia [2006, 2012]; Mali [2007]; Mexico [1990, 1996, 2000, 2005, 2012]; Moldova [1996, 2002, 2006, 2008]; Morocco [2001, 2007, 2011]; Netherlands [1990, 1999, 2006, 2008, 2012]; New Zealand [1998, 2004, 2011]; Nigeria [2000, 2012]; Norway [1990, 1996, 2007, 2008]; Pakistan [1997, 2012]; Peru [1996, 2001, 2006, 2012]; Philippines [1996, 2001, 2012]; Poland [1990, 1997, 1999, 2005, 2008, 2012]; Portugal [1990, 1999, 2008]; Qatar [2010]; Romania [1993, 1998, 1999, 2005, 2008, 2012]; Russia [1995, 1999, 2006, 2008, 2011]; Rwanda [2007, 2012]; Serbia [2006, 2008]; Singapore [2002, 2012]; Slovakia [1998, 1999, 2008]; Slovenia [1992, 1995, 1999, 2005, 2008, 2011]; South Africa [1990, 1996, 2001, 2006, 2013]; Spain [1990, 1995, 1999, 2000, 2007, 2008, 2011]; Sweden [1990, 1996, 1999, 2006, 2009, 2011]; Switzerland [1999, 2007, 2008]; Tanzania [2001]; Thailand [2007, 2013]; Trinidad and Tobago [2006, 2010]; Tunisia [2013]; Turkey [1990, 1996, 2001, 2007, 2009, 2012]; Uganda [2001]; Ukraine [1996, 1999, 2006, 2008, 2011]; United Kingdom [1990, 1998, 1999, 2005, 2009]; United States [1990, 1995, 1999, 2006, 2011]; Uruguay [1996, 2006, 2011]; Uzbekistan [2011]; Venezuela [1996, 2000]; Vietnam [2001, 2006]; Zambia [2007]; Zimbabwe [2001, 2012].

C.1.2.3 Model (3): Supply (Tertiary Education) and Demand (Political Leaders)

Albania [1998, 2002]; Algeria [2014]; Argentina [1999, 2006]; Armenia [2011]; Australia [1995]; Azerbaijan [1997, 2011]; Bangladesh [2002]; Belarus [2011]; Brazil [2014]; Bulgaria [1997, 2006]; Burkina Faso [2007]; Chile [1996, 2000, 2006, 2012]; China [2007, 2013]; Colombia [1998, 2005, 2012]; Croatia [1996]; Cyprus [2006, 2011]; Czech Republic [1998]; Egypt [2001, 2008, 2012]; El Salvador [1999]; Estonia [1996, 2011]; Ethiopia [2007]; Finland [1996, 2005]; France [2006]; Georgia [1996, 2009, 2014]; Germany [1997, 2013]; Ghana [2007]; Hungary [1998, 2009]; India [1995, 2001, 2006]; Indonesia [2001, 2006]; Iran [2000, 2007]; Iraq [2013]; Italy [2005]; Japan [1995, 2000,

2005, 2010]; Jordan [2007]; Kazakhstan [2011]; Korea [1996, 2001, 2005, 2010]; Kyrgyzstan [2003, 2011]; Lebanon [2013]; Lithuania [1997]; Macedonia [1998, 2001]; Mexico [1996, 2000, 2005, 2012]; Moldova [1996, 2002, 2006]; Morocco [2001, 2007, 2011]; Netherlands [2006, 2012]; New Zealand [1998, 2011]; Norway [1996, 2007]; Pakistan [2012]; Peru [2001, 2006]; Philippines [1996, 2001, 2012]; Poland [1997, 2005, 2012]; Qatar [2010]; Romania [1998, 2005, 2012]; Russia [2006, 2011]; Rwanda [2012]; Serbia [2006]; Slovakia [1998]; Slovenia [1995, 2005, 2011]; South Africa [2013]; Spain [1995, 2000, 2007, 2011]; Sweden [1996, 1999, 2006, 2011]; Switzerland [2007]; Tanzania [2001]; Thailand [2007, 2013]; Tunisia [2013]; Turkey [2001, 2007, 2012]; Uganda [2001]; Ukraine [2006, 2011]; United Kingdom [2005]; United States [1995]; Uruguay [2006]; Uzbekistan [2011]; Venezuela [2000]; Vietnam [2001, 2006]; Zimbabwe [2012].

C.1.2.4 Model (4): Supply (Tertiary Education) and Demand (Jobs Scarce)

Albania [1998, 2002, 2008]; Algeria [2014]; Argentina [1999, 2006]; Armenia [2008, 2011]; Australia [1995]; Austria [1990]; Azerbaijan [1997, 2011]; Bangladesh [2002]; Belarus [2000, 2008, 2011]; Belgium [1900, 1999, 2009]; Brazil [2014]; Bulgaria [1991, 1997, 1999, 2006, 2008]; Burkina Faso [2007]; Canada [1990]; Chile [1996, 2000, 2006, 2012]; China [2007, 2013]; Colombia [1998, 2012]; Croatia [1996, 1999, 2008]; Cyprus [2006, 2008, 2011]; Czech Republic [1998, 1999, 2008]; Denmark [1990, 1999, 2008]; Egypt [2001, 2008, 2012]; El Salvador [1999]; Estonia [1996, 1999, 2008, 2011]; Ethiopia [2007]; Finland [1990, 1996, 2000, 2005, 2009]; France [1990, 1999, 2006, 2008]; Georgia [1996, 2008, 2009, 2014]; Germany [1997, 2013]; Ghana [2007]; Greece [1999]; Hungary [1991, 1998, 1999, 2008, 2009]; India [1990, 1995, 2001, 2006]; Indonesia [2001, 2006]; Iran [2000, 2007]; Ireland [1990, 1999, 2008]; Italy [1990, 1999, 2005, 2009]; Japan [1990, 1995, 2000, 2005, 2010]; Jordan [2007]; Kazakhstan [2011]; Korea [1990, 1996, 2001, 2005, 2010]; Kyrgyzstan [2003, 2011]; Latvia [1999, 2008]; Lebanon [2013]; Lithuania [1997, 1999, 2008]; Luxembourg [2008]; Macedonia [1998, 2001]; Mexico [1996, 2000, 2005, 2012]; Moldova [1996, 2002, 2006, 2008]; Morocco [2001, 2007, 2011]; Netherlands [1990, 1999, 2006, 2008, 2012]; New Zealand [1998, 2011]; Norway [1990, 1996, 2007, 2008]; Pakistan [2012]; Peru [2001, 2006]; Philippines [1996, 2001, 2012]; Poland [1989, 1990, 1997, 1999, 2005, 2008, 2012]; Portugal [1999, 2008]; Qatar [2010]; Romania [1993, 1998, 1999, 2005, 2008, 2012]; Russia [2006, 2008, 2011]; Rwanda [2012]; Serbia [2006, 2008]; Slovakia [1998, 1999, 2008]; Slovenia [1992, 1995, 1999, 2005, 2008, 2011]; South Africa [1990, 2013]; Spain [1990, 1995, 1999, 2000, 2007, 2008, 2011]; Sweden [1990, 1996, 1999, 2006, 2009, 2011]; Switzerland

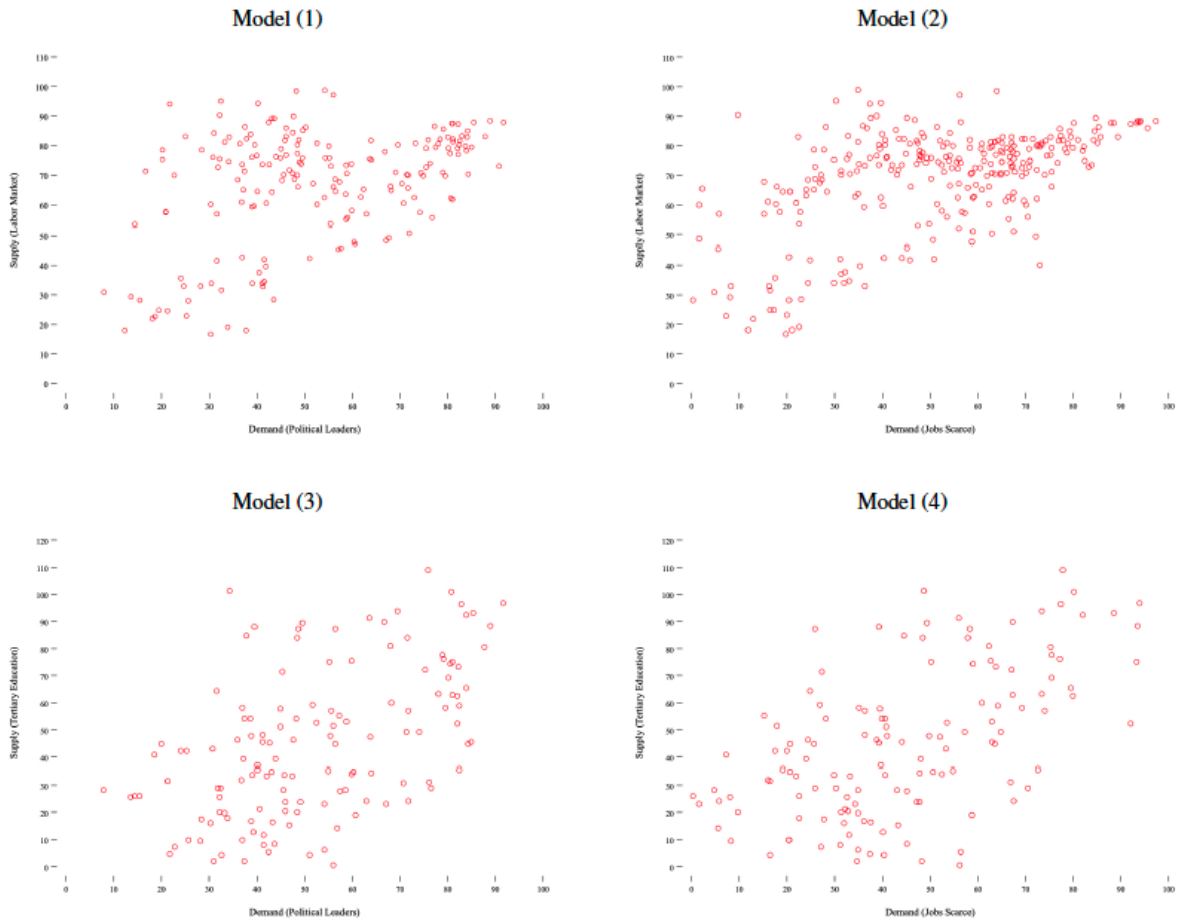
[1999, 2007, 2008]; Tanzania [2001]; Thailand [2007, 2013]; Tunisia [2013]; Turkey [1990, 2001, 2007, 2009, 2012]; Uganda [2001]; Ukraine [1999, 2006, 2008, 2011]; United Kingdom [1990, 1998, 1999, 2005, 2009]; United States [1990, 1995]; Uruguay [2006]; Uzbekistan [2011]; Venezuela [2000]; Vietnam [2001, 2006]; Zimbabwe [2012].

C.1.3 Visualizing Supply and Demand

For those who are interested, Figure C.1 graphically shows the relationship between my supply and demand variables. While there appears to be a loose positive relationship between the supply and demand variables (correlations range from 0.39 to 0.59), the key thing to note for the purpose of the analyses conducted here is that there is substantively meaningful variation in *Supply* across the observed range of *Demand* as well as substantively meaningful variation in *Demand* across the observed range of *Supply*.¹ As noted in footnote 4 in the main text, the structure of the data make it impractical to examine any temporal dynamics between *Women's Representation*, *Supply*, and *Demand*. One reason for this is that the surveys that provide the information for the *Demand* variables do not always come at regular intervals either within countries or between countries, making the use of lags inadvisable. The second reason is that the panel nature of the data (small T , large N) means that the inclusion of lagged variables would significantly reduce the sample size and drop all singleton countries that only ever had one survey.

¹Further evidence for this comes from the 'raw data plots' that appear later in Figure C.5 in section D.

Figure C.1: Scatter Plots of *Supply* and *Demand*



Note: The plots show the scatter plot of *Supply* and *Demand* based on the samples of the four interaction models shown in Table 4.1.

C.2 Different Operationalizations for Supply and Demand

In this chapter, I use two alternative measures to proxy for the size of the pool of qualified female candidates in a country. *Supply (Labor Market)* captures the ratio of female to male labor force participation. *Supply (Tertiary Education)* captures women's tertiary education gross enrollment ratio. I use two alternative measures to capture mass demand for women legislators. *Demand (Political Leaders)* captures the percentage of respondents in a country who disagree or strongly disagree with the claim that men make better political leaders than women. *Demand (Jobs Scarce)* captures the percentage of respondents in a country who disagree or strongly disagree with the claim that men have more right to a job than women when jobs are scarce.

In Table 4.1, I show the results from four interactive models that each use one of the four possible combinations of my supply and demand variables. In the main text, I noted that the results from all four models are remarkably consistent. All four models, for example, provide empirical support for the claim that there is an interaction between supply and demand when it comes to women's descriptive representation. This is demonstrated by the positive and statistically significant coefficient on the interaction term $Supply \times Demand$ in each model. To further support my assertion, I used Figure 4.2 to show, for each of the interactive models, how the predicted level of women's descriptive representation in a country varies with different possible combinations of values for *Supply* and *Demand*. The key point here was that the 3-D plots associated with each interactive model were qualitatively similar.

When it came to a more detailed analysis in the main text, I focused on the first interactive model where supply was operationalized as *Supply (Labor Market)* and demand was operationalized as *Demand (Political Leaders)*. In what follows, I demonstrate that the inferences that I drew from this particular model are qualitatively similar to the inferences we can draw from the other three interactive models. In particular, I now show that, while the cut-points may differ slightly from one model to another, we always find that an increase in demand is associated with higher

levels of women's descriptive representation once the supply of qualified female candidates is sufficiently high and that an increase in the supply of qualified female candidates is also associated with greater female representation once the demand for women legislators is sufficiently high.

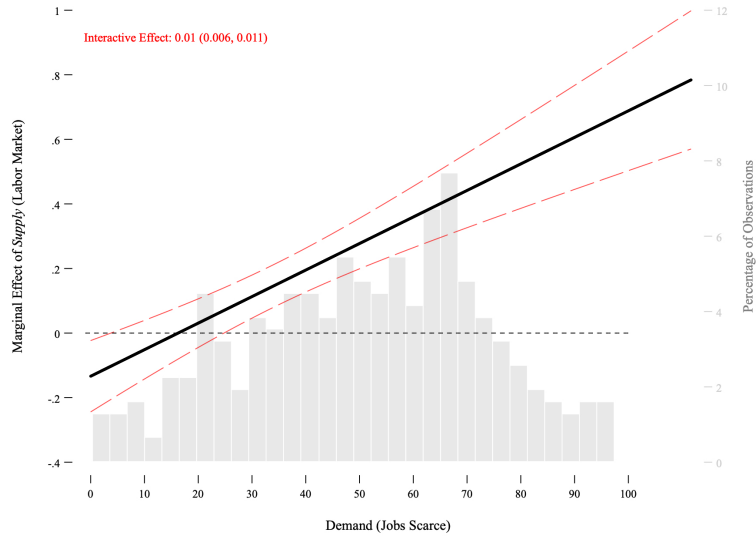
C.2.1 Model (2): *Supply (Labor Market) and Demand (Jobs Scarce)*

Using the results from interactive model (2) in Table 4.1, Figure C.2 shows the marginal effect of supply on women's descriptive representation across the observed range of demand (top panel) and the marginal effect of demand across the observed range of supply (bottom panel). As predicted, the marginal effect plot in panel (a) shows that the effect of supply varies with the level of demand in a country. To be specific, the effect of supply is negative and statistically significant when $Demand < 3.76$ (0%), it has a statistically insignificant effect when $3.76 < Demand < 24.79$ (4.7%), and it has a positive and statistically significant effect when $Demand > 24.79$ (95.3%). The numbers in parentheses indicate the percentage of observations that fall within each of these ranges. These results are entirely consistent with the *Supply Hypothesis*, which predicts that the effect of supply on women's descriptive representation becomes positive once the level of demand in a country is sufficiently large and grows in magnitude as the level of demand increases further.

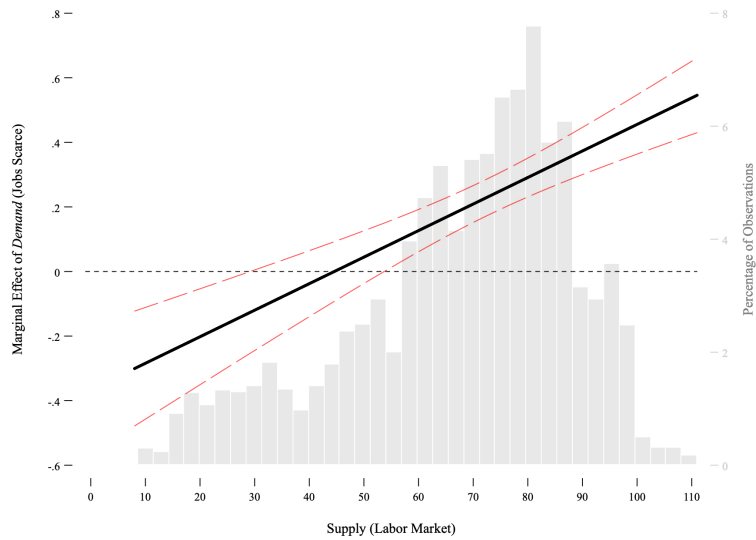
As predicted, the marginal effect plot in panel (b) shows that the effect of demand varies with the supply of qualified female candidates. To be specific, *Demand* has a negative and statistically significant effect when $Supply < 29.18$ (6.7%), it has a statistically insignificant effect when $29.18 < Supply < 53.71$ (15.8%), and it has a positive and statistically significant effect when $Supply > 53.71$ (77.5%). These results are consistent with the *Demand Hypothesis*, which predicts that the effect of demand on women's descriptive representation becomes positive once the supply of qualified female candidates is sufficiently high and grows in magnitude as supply increases further. As noted in the main text, not too much should be made of the fact that the effect of demand is negative when supply is very low. This is because there are so few real-world observations where supply is low enough for this negative effect to be observed. Indeed, fully 93.3% of the sample

Figure C.2: The Conditional Effects of *Supply* and *Demand* on Women’s Descriptive Representation – Model (2)

(a) The Conditional Effect of *Supply*



(b) The Conditional Effect of *Demand*



Note: Panel (a) shows the effect of a one-unit increase in *Supply* on women’s descriptive representation across the observed range of *Demand*. It also shows the interaction effect between *Supply* and *Demand*. Panel (b) shows the effect of a one-unit increase in *Demand* on women’s descriptive representation across the observed range of *Supply*. The dashed lines indicate two-tailed 95% confidence intervals. The histograms show the percentage of observations at different values of each modifying variable. The two plots are based on the results from the interaction model shown in Model 2 in Table 4.1.

observations are consistent with the *Demand Hypothesis*.

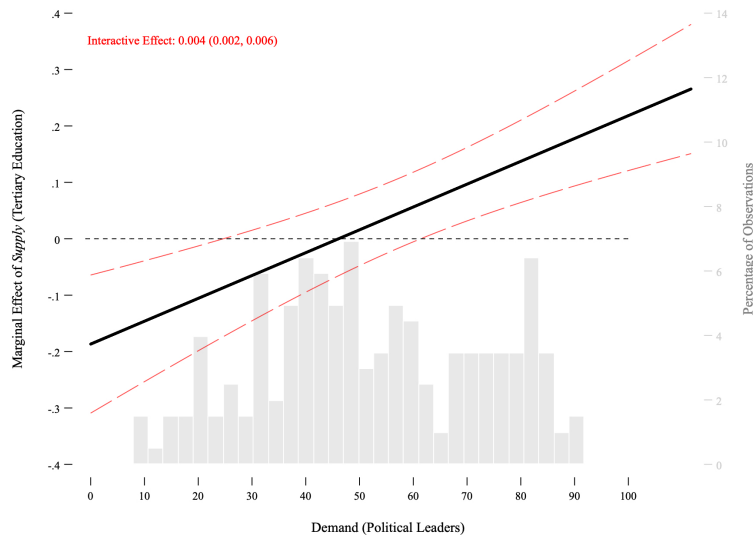
C.2.2 Model (3): *Supply (Tertiary Education) and Demand (Political Leaders)*

Using the results from interactive model (3) in Table 4.1, Figure C.3 shows the marginal effect of supply on women's descriptive representation across the observed range of demand (top panel) and the marginal effect of demand across the observed range of supply (bottom panel). As predicted, the marginal effect plot in panel (a) shows that the effect of supply on women's descriptive representation varies with the level of demand in a country. More specifically, *Supply* has a negative and statistically significant effect when $Demand < 24.67$ (3.4%), it has a statistically insignificant effect when $24.67 < Demand < 61.38$ (19.2%), and it has a positive and statistically significant effect when $Supply > 61.38$ (87.2%). These results are consistent with the *Supply Hypothesis*, which predicts that the effect of supply on women's descriptive representation becomes positive once the level of demand in a country is sufficiently large and grows in magnitude as the level of demand increases further. Fully 96.6% of the sample observations are consistent with the *Supply Hypothesis*.

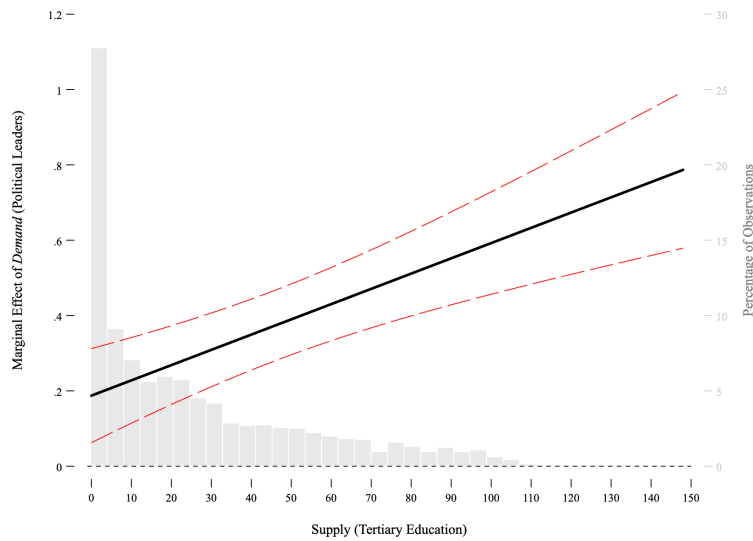
As predicted, the marginal effect plot in panel (b) shows that the effect of demand on women's descriptive representation varies with the supply of female candidates. While the positive effect of demand is substantively small when supply is low, it grows in magnitude as the supply of qualified female candidates grows.

Figure C.3: The Conditional Effects of *Supply* and *Demand* on Women’s Descriptive Representation – Model (3)

(a) The Conditional Effect of *Supply*



(b) The Conditional Effect of *Demand*



Note: Panel (a) shows the effect of a one-unit increase in *Supply* on women’s descriptive representation across the observed range of *Demand*. It also shows the interaction effect between *Supply* and *Demand*. Panel (b) shows the effect of a one-unit increase in *Demand* on women’s descriptive representation across the observed range of *Supply*. The dashed lines indicate two-tailed 95% confidence intervals. The histograms show the percentage of observations at different values of each modifying variable. The two plots are based on the results from the standard interaction model shown in Model 3 in Table 4.1.

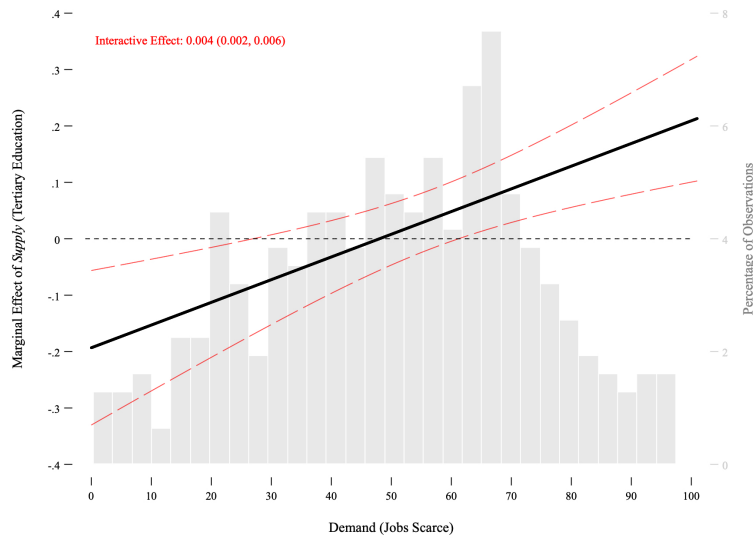
C.2.3 Model (4): *Supply (Tertiary Education) and Demand (Jobs Scarce)*

Using the results from interactive model (4) in Table 4.1, Figure C.4 shows the marginal effect of supply on women's descriptive representation across the observed range of demand (top panel) and the marginal effect of demand across the observed range of supply (bottom panel). As predicted, the marginal effect plot in panel (a) shows that the effect of supply varies with the level of demand in a country. To be specific, *Supply* has a negative and statistically significant effect when $Demand < 26.94$ (16.9%), it has a statistically insignificant effect when $26.94 < Demand < 61.20$ (45.7%), and it has a positive and statistically significant effect when $Demand > 61.20$ (37.4%). These results are consistent with the *Supply Hypothesis*, which predicts that the effect of supply on women's descriptive representation becomes positive once the level of demand in a country is sufficiently large and grows in magnitude as the level of demand increases further. Fully 83.1% of the sample observations are consistent with the *Supply Hypothesis*.

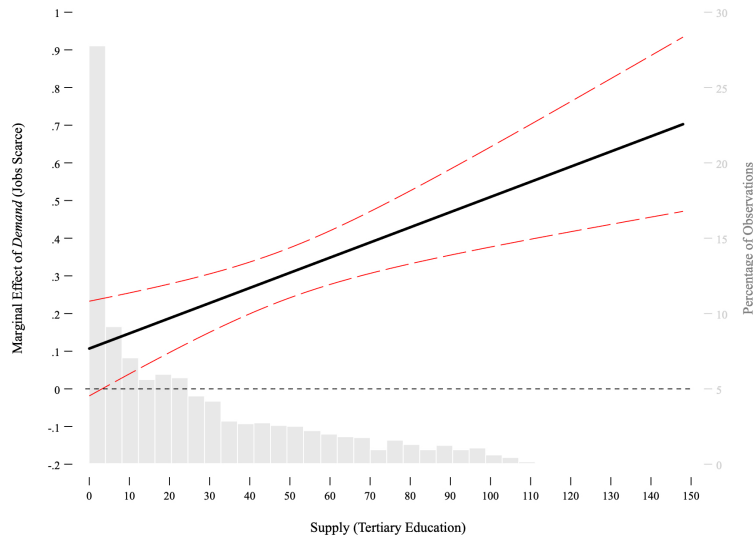
As predicted, the marginal effect plot in panel (b) shows that the effect of demand varies with the supply of female candidates. More specifically, *Demand* has a positive but statistically insignificant effect when $Supply < 3.16$ (24.5%) and a positive and statistically significant effect when $Supply > 3.16$ (75.5%). These results are entirely consistent with the *Demand Hypothesis*, which predicts that the effect of demand on women's descriptive representation becomes positive once the supply of qualified female candidates is sufficiently high and grows in magnitude as supply increases further.

Figure C.4: The Conditional Effects of *Supply* and *Demand* on Women’s Descriptive Representation – Model (4)

(a) The Conditional Effect of *Supply*



(b) The Conditional Effect of *Demand*



Note: Panel (a) shows the effect of a one-unit increase in *Supply* on women’s descriptive representation across the observed range of *Demand*. It also shows the interaction effect between *Supply* and *Demand*. Panel (b) shows the effect of a one-unit increase in *Demand* on women’s descriptive representation across the observed range of *Supply*. The dashed lines indicate two-tailed 95% confidence intervals. The histograms show the percentage of observations at different values of each modifying variable. The two plots are based on the results from the standard interaction model shown in Model 4 in Table 4.1.

C.3 Controlling for electoral system

The empirical analysis conducted in the main text controls for key aspects of a country's institutional structure by including variables for regime type and electoral rules. When it came to electoral rules, I focused on gender quotas because these quotas have the capacity to more directly and automatically influence the level of women's descriptive representation in both democracies and dictatorships than other aspects of the electoral system such as the electoral formula. However, I did note at the time that my inferences were qualitatively robust to also controlling for whether a country employs a majoritarian or proportional electoral system. I provide the empirical basis for this claim here. In Table C.2, I provide the results for the same four interactive models as shown in the main text except that I also now control for whether a country employs a proportional electoral system as opposed to a majoritarian one. In line with much of the existing research, I find that women's descriptive representation is significantly higher in countries that employ proportional electoral systems. This is indicated by the positive coefficient on *Proportional Electoral System* in all four models. Importantly for my discussion of the supply and demand framework, I find that the coefficient on the interaction term *Supply* \times *Demand* is always positive and statistically significant. This provides the key evidence supporting my claim that supply and demand interact to determine women's descriptive representation.

Table C.2: Determinants of Women's Legislative Representation

	Dependent Variable: <i>Women's Representation</i> , 0 – 100			
	(1)	(2)	(3)	(4)
<i>Supply:</i>				
Labor Market	-0.50** (0.14)	-0.49** (0.08)		
Tertiary Education			-0.50** (0.13)	-0.40** (0.09)
<i>Demand:</i>				
Political Leaders	-0.79** (0.21)		-0.20 (0.12)	
Jobs Scarce		-0.79** (0.13)		0.01 (0.07)
<i>Supply × Demand</i>	0.01** (0.003)	0.01** (0.001)	0.01** (0.002)	0.01** (0.001)
<i>Democracy</i>	-0.11 (0.36)	0.42 (0.29)	0.03 (0.42)	0.30 (0.36)
<i>Effective Quota</i>	6.67** (1.99)	5.91** (1.54)	5.92* (2.32)	4.97* (1.83)
<i>Proportional Electoral Systems</i>	3.26 (1.69)	5.16** (1.30)	7.29** (2.31)	7.59** (1.75)
Constant	35.95** (10.25)	28.47** (5.98)	16.02** (2.58)	6.29 (4.34)
Observations	101	175	73	140
R ²	0.61	0.63	0.64	0.59

Standard errors in parentheses. * $p < 0.05$; ** $p < 0.01$ (two-tailed).

C.4 Relaxing the Linear Interaction Assumption

In the main text, I showed marginal effect plots for supply and demand based on the results from the first interactive model shown in Table 4.1. As the marginal effect plot for supply indicates, the effect of supply is negative, although statistically insignificant, when demand is very low. Similarly, the marginal effect plot for demand indicates that the effect of demand is negative, although again statistically insignificant, when supply is very low. Technically, these estimated *negative* effects are inconsistent with the supply and demand framework, which predicts that the effects of supply and demand should be non-negative. As I mentioned in footnote 6, not too much should be read into these estimated negative effects. This is because, in addition to being statistically insignificant, there are very few observations in the sample with a value for supply or demand where we would predict the effects of supply and demand to be negative. As I discussed in Appendix C.1.3, this general pattern is consistent with the results from the other interactive models in Table 4.1.

As I also mentioned in footnote 6, the negative effect for supply (demand) at very low levels of demand (supply) may well be an artefact of the fact that I have assumed that there is a linear interaction between supply and demand. Put simply, if I expect that the effect of supply (demand) substantially increases in a linear way with demand (supply), then the marginal effect line for supply (demand) may end up being ‘forced’ below the zero line when demand (supply) is very low. To some extent, it is arguable that the supply and demand framework implies a threshold effect, where the effect of supply (demand) becomes positive only when demand (supply) is sufficiently high. Once past the threshold, we would expect the positive effect of supply (demand) to increase linearly with demand (supply). This suggests that the effect of supply (demand) on women’s descriptive representation is only *piecewise* linear across different values of demand (supply). I now examine this possibility by relaxing the assumption of a linear interaction.

To examine my assumption of a linear interaction effect, I follow the advice of Hainmueller,

Mummolo and Xu (2019) and use (1) a ‘binning estimator’ and (2) a ‘kernel estimator’ that allow for heterogeneity both in how the conditional marginal effect of *Demand* changes with *Supply* and in how the conditional marginal effect of *Supply* changes with *Demand*. As we will see, both estimation strategies indicate that my assumption of a linear interaction effect is reasonable and hence that my inferences in the main text are robust to relaxing the linearity assumption for the interaction effect. While supportive of a linear interaction effect, the additional flexibility of the binning and kernel estimators does provide some support for a threshold effect and suggests that the estimated marginal effects of supply and demand are non-negative.

C.4.1 Linear Interaction Diagnostic Plot

Before using the binning and kernel estimators, I attempt to visually diagnose if there is any non-linearity in the interaction effect between supply and demand. I first do this with the raw data plots shown in Figure C.5. Since *Demand* and *Supply* are both continuous variables, I adopt a binning approach for graphing the relationship between women’s descriptive representation and supply (demand) across the observed range for demand (supply). Following Hainmueller, Mummolo and Xu (2019), I use three ‘bins’ or groups. With respect to the raw data plots in panel (a), this means splitting the sample into three roughly equal sized groups based on the value of *Demand* – low demand (first tercile), medium demand (second tercile), and high demand (third tercile). I then plot the sample observations in each of these three groups in terms of their values for *Women’s Representation* (vertical axis) and *Supply* (horizontal axis) as solid black circles. For each group, I then overlay the observations with the ‘best fit’ lines from a linear (blue) and a non-parametric LOESS (red) regression. The raw data plots in panel (a) indicate that it is reasonable to assume that there is a linear relationship between *Women’s Representation* and *Supply* for all three levels of *Demand*. This is because the blue linear regression lines and the red LOESS regression lines do not diverge too much in any of the three plots. The plots also indicate that there is a positive relationship between *Women’s Representation* and *Supply* for all three levels of demand and that

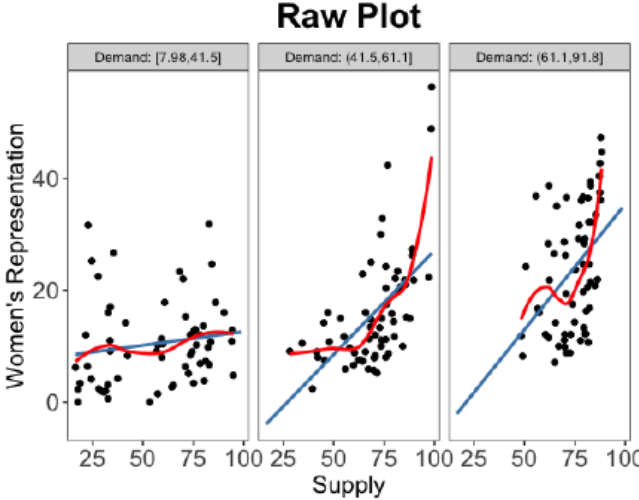
this positive relationship is greater when we have medium or high demand than low demand. This is indicated by the steeper blue lines in the two plots on the right.

The raw data plots in panel (b) are similar except that I now split the sample into three roughly equal sized groups based on the value of *Supply* – low supply (first tercile), medium supply (second tercile), high supply (third tercile). I then plot the sample observations in each of these three groups in terms of their values for *Women's Representation* and *Demand*. The raw data plots in panel (b) indicate once again that it is reasonable to assume that there is a linear relationship between *Women's Representation* and *Demand* for all three levels of *Supply*. The plots also indicate that there is a positive relationship between *Women's Representation* and *Demand* for all three levels of supply and there is some evidence that the magnitude of this positive effect increases with higher levels of supply.

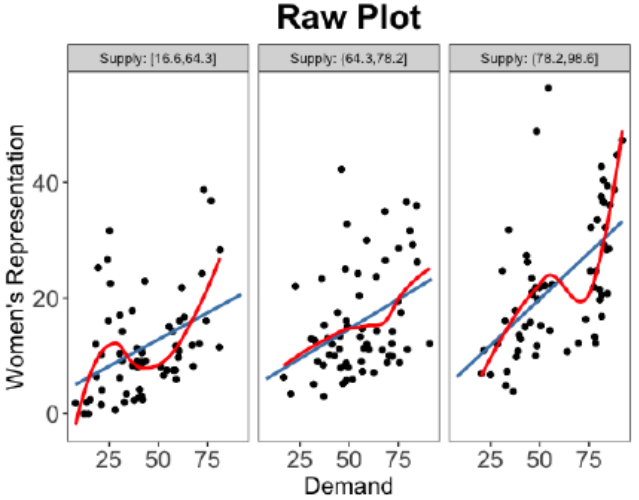
Since both *Demand* and *Supply* are continuous, we can also visually evaluate the assumption of a linear interaction effect by using a generalized additive model (GAM) to plot how women's descriptive representation varies with different combinations of values for *Supply* and *Demand*. Such a plot is shown in Figure C.6. Darker shades indicate lower levels of women's descriptive representation, while lighter shades indicate higher levels. The 3-D GAM plots can be compared to the corresponding 3-D plots shown in Figure B.1. While the 3-D GAM plots allow for a non-linear interaction effect, the corresponding 3-D plots in Figure B.1 assume a linear interaction effect. As we can see, the 3-D GAM plots are very similar to the ones shown in Figure B.1. For example, we see that women's descriptive representation is highest when both supply and demand are high. We also see that the marginal effect of demand (the slope of the surface to the right) increases as we move from low supply to high supply. Similarly, we see that the marginal effect of supply (the slope of the surface to the left) increases as we move from low demand to high demand. Unlike the 3-D plots in Figure B.1, we see that the marginal effect of demand when supply is low (the slope to the right at the front) is always positive and that the marginal effect of supply when demand is low (the slope to the left at the front) is also always positive. Importantly, the surfaces of the

Figure C.5: Linear Interaction Diagnostic Plots

(a) Relationship between Women’s Representation and Supply when Demand is Low, Medium, and High

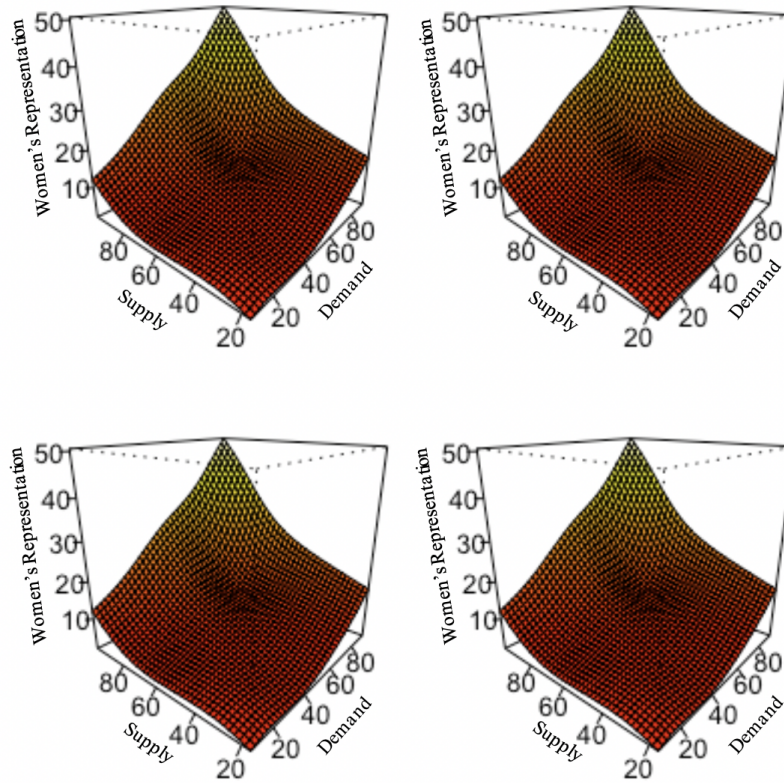


(b) Relationship between Women’s Representation and Demand when Supply is Low, Medium, and High



Note: The plot in panel (a) shows the relationship between women’s representation and demand when demand is low, medium, and high. The plot in panel (b) shows relationship between women’s representation and supply when supply is low, medium, and high. A linear fit line is displayed in blue and Loess fit line is shown in red. The histograms show the percentage of observations at different values of each modifying variable.

Figure C.6: GAM Plots of the Relationship between Women’s Representation, Supply, and Demand



3-D GAM plots are fairly smooth, with a gentle curvature in the middle but devoid of any drastic humps, wrinkles, or holes that might indicate a non-linear interaction effect.

The bottom line is that we do not see any visual evidence that would cause us to be concerned with the assumption of a linear interaction effect.

C.4.2 Binning Estimator

I now turn to the binning estimator proposed by Hainmueller, Mummolo and Xu (2019) to estimate the marginal effect of demand (supply) on women’s descriptive representation across the observed range of supply (demand). The binning estimator requires that we (1) split the continuous modifying variables (*Demand* or *Supply*) into several bins (usually three roughly equal sized groups

– low, medium, and high), (2) create dichotomous indicator variables for each bin, and then (3) interact each dichotomous variable with the relevant ‘treatment indicator’ (*Supply* or *Demand*).² This binning estimator gives us the marginal effect of supply when demand is low, medium, and high. It also gives us the marginal effect of demand when supply is low, medium, and high. Unlike the interaction model in Eq.4.2.1, the binning estimator does not require that the marginal effect of supply (demand) changes in a linear way with supply (demand).

In Figure C.7, I overlay the marginal effects from the binning estimator on top of the marginal effect plots for supply (top panel) and demand (bottom panel) that we constructed from the interaction model shown in Eq.4.2.1 and that were originally shown in Figure B.1. The estimates of the marginal effects from the binning estimator are shown as red dots with vertical 95% confidence intervals. The marginal effects from the binning estimator are labeled L, M, and H, to indicate that they show the marginal effect of supply (demand) when demand (supply) is Low, Medium, or High.

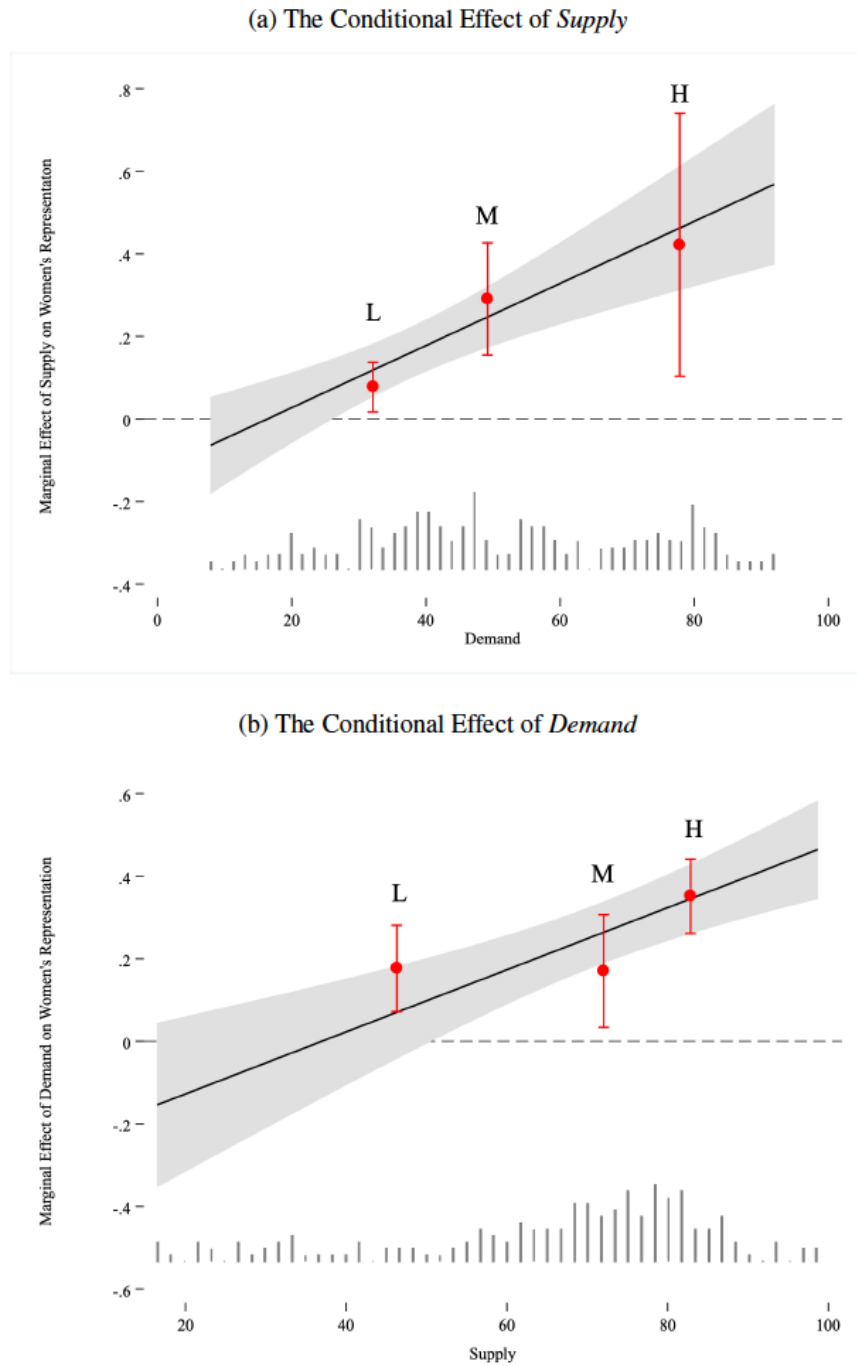
The key thing to note from the top panel is that the estimates for the marginal effect of supply from the binning estimator line up very closely with the marginal effect line for supply from the original interaction model in Eq.4.2.1 that assumes a linear interaction effect. The three marginal effects from the binning estimator indicate that the marginal effect of supply, which is never negative, increases with higher levels of demand. This is consistent with my inferences in the main text. In the bottom panel, the three estimates for the marginal effect of demand from the binning estimator again line up closely with the marginal effect line for demand from the original interaction model in Eq.4.2.1 that assumes a linear interaction effect. This time, though, we see that there is no significant difference in the marginal effect of demand when we have medium as opposed to low supply. However, the marginal effect of demand is significantly higher when supply is high than when supply is low or medium. This provides suggestive evidence that there is some kind of threshold effect when it comes to the marginal effect of demand on women’s descriptive

²For more information, see (Hainmueller, Mummolo and Xu 2019, 170).

representation.

Overall, the results from the binning estimator show that the linear interaction model shown in Eq.4.2.1 provides a reasonable approximation of the conditional relationship between women's descriptive representation, supply, and demand.

Figure C.7: Conditional Marginal Effects from Binning Estimator



Note: Figure C.7 displays the conditional marginal effect estimates of the binning estimator (red point estimates), labeled L, M, and H, superimposed on the estimates from the multiplicative interaction model shown in Eq.4.2.1 (black line). The plot in panel (a) shows the marginal effect of supply when demand is low, medium, and high. The plot in panel (b) shows the marginal effect of demand when supply is low, medium, and high. The histograms show the percentage of observations at different values of each modifying variable.

C.4.3 Kernel Estimator

I now turn to the kernel estimator proposed by Hainmueller, Mummolo and Xu (2019) to estimate the marginal effect of demand (supply) on women's descriptive representation across the observed range of supply (demand). The kernel estimator allows the functional form of the marginal effect of supply (demand) to vary across the observed range of demand (supply) by estimating a series of local effects using kernel weighted locally linear regressions.³ I present the estimated marginal effects for supply and demand from the kernel estimator in Figure C.8. The confidence intervals are generated using 1,000 iterations of a non-parametric bootstrap where we resample the data with replacement. The histograms show the percentage of observations at different values of each modifying variable.

Panel (a) displays the marginal effect of supply across the observed range of demand based on the kernel smoothing estimator. The key thing to note is that the marginal effect of supply is essentially zero when demand is at its lowest observed value. It then becomes positive and increases in magnitude in a fairly linear way with higher levels of demand. At very high levels of demand, the magnitude of the positive effect of supply increases sharply, suggesting a possible threshold effect. On the whole, though, the information in the marginal effect plot in panel (a) is consistent with the *Supply Hypothesis* and the inferences that I made in the main text.

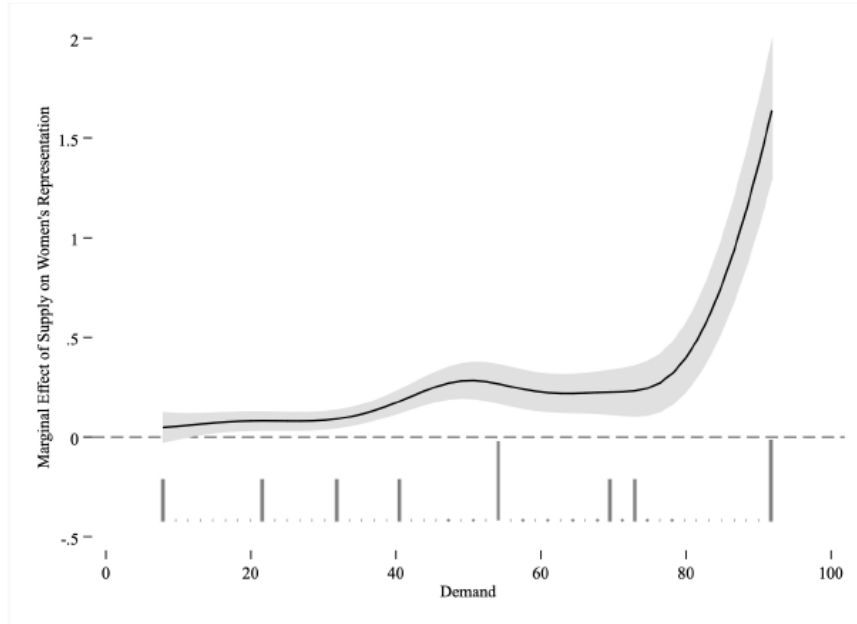
Panel (b) displays the marginal effect of demand across the observed range of supply based on the kernel smoothing estimator. The key thing to note here is that the marginal effect of demand is essentially zero when supply is at its lowest observed value. It then becomes positive and increases in magnitude in a fairly linear way with higher levels of supply. This is entirely consistent with the *Demand Hypothesis* and the inferences that I made in the main text.

The bottom line is that the inferences that I make in the main text are robust to relaxing the assumption of a linear interaction effect between supply and demand.

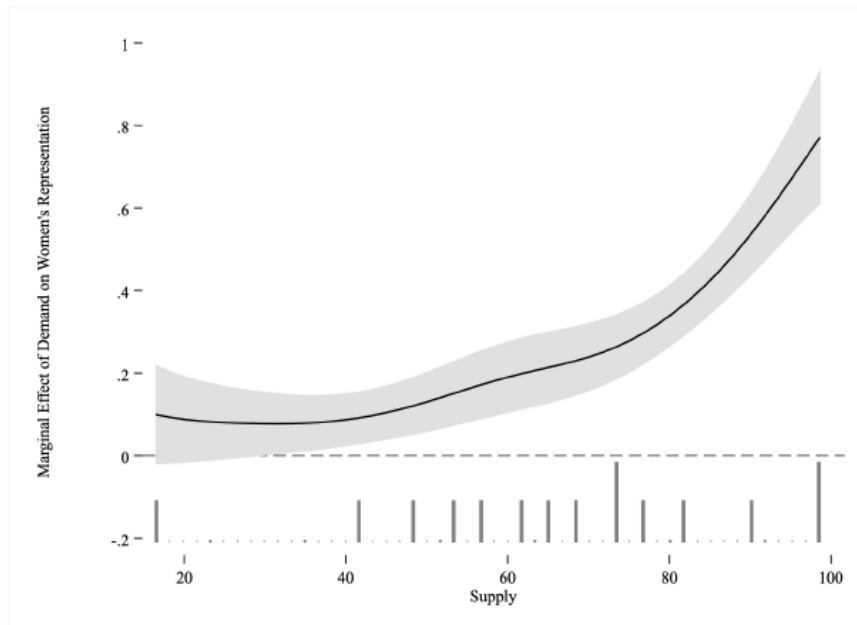
³For more information, see (Hainmueller, Mummolo and Xu 2019, 173).

Figure C.8: Conditional Marginal Effects from Kernel Estimator

(a) The Conditional Effect of *Supply*



(b) The Conditional Effect of *Demand*



Note: The plot in panel (a) shows the marginal effect of supply. The plot in panel (b) shows relationship the marginal effect of demand. The histograms show the percentage of observations at different values of each modifying variable.