ESSAYS ON THE ROLE OF GENDER AND GROUPS IN ECONOMIC DECISION MAKING

by

Jay Alan Schwarz

B.S., Texas A & M University, 2004

M.Eng., Texas A & M University, 2005

M.A., University of Pittsburgh, 2008

Submitted to the Graduate Faculty of

the Department of Economics in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

DEPARTMENT OF ECONOMICS

This dissertation was presented

by

Jay Alan Schwarz

It was defended on

April 13th, 2011

and approved by

Dr. Lise Vesterlund,University of Pittsburgh, Economics

Dr. Randall Walsh, University of Pittsburgh, Economics

Dr. Mark Hoekstra, University of Pittsburgh, Economics

Dr. Muriel Niederle, Stanford University, Economics

Dissertation Advisors: Dr. Lise Vesterlund,University of Pittsburgh, Economics,

Dr. Randall Walsh, University of Pittsburgh, Economics

ESSAYS ON THE ROLE OF GENDER AND GROUPS IN ECONOMIC DECISION MAKING

Jay Alan Schwarz, PhD

University of Pittsburgh, 2011

Groups collectively make many economic decisions. Corporate boards select CEOs or develop strategic plans. School boards hire superintendents and develop policies. Here I study ways gender may influence such group decisions. In Chapter 1, I study the formation of elected groups in an environment where females are well represented. Using California school board elections data, I estimate the electorate's response to either an additional male or female serving on the school board. I find that a female (male) winning the current election - and thus sitting on the board - increases the number of males (females) elected in the next cycle. I hypothesize the electorate has a preferred board gender composition it seeks to maintain. This result highlights that electoral environments with high-female representation may be substantially different from the previously studied low-female representation environments because in the former case voters have greater experience with female officeholders. In Chapter 2, I use the same school board data and study whether increased female representation on a school board affects a district's academic performance. I find that increasing the number of female board members has a positive impact on academic performance for districts failing to meet state requirements. I also show gender in fact influences how a school board executes its duty of monitoring the superintendent. For failing school districts, boards with more females are less likely to experience a superintendent separation. Lastly, in Chapter 3, I use an experimental design to study gender interactions within groups evaluating males and females competing against one another in a task. I elicit individual group members' evaluations and then groups' collective evaluations of the competitors. Consistent with studies of observational data, I find only groups with either low or high numbers of females tend to favor the male competitor. By mapping individuals to group decisions, I find this is partially explained by females being less likely to influence the group decision if they disagree *a priori* with the majority opinion.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I thank my lovely wife, Mary, for her support which has been crucial to the successful completion of my degree. I also am grateful for the advice, patience, and support of my committee.

1.0 ELECTION BEHAVIOR WITH HIGH-FEMALE REPRESENTATION

1.1 INTRODUCTION

Decision-making institutions such as corporate boards and elected legislatures often have fewer females than males. Efforts worldwide have focused on understanding and reducing this gap. Recent evidence from Indian (Beaman et al. 2009) and Italian (De Paola et al. 2010) elections finds that under-representation results in part from voter ignorance about female leadership. That is, lack of exposure to female leaders allows the electorate to hold mistaken beliefs about how a female would govern. These inaccuracies tend to bias voters against females, causing statistical discrimination against even qualified female candidates. I refer to this as "ignorance-based" discrimination. Furthermore, these studies find the affirmative action policies of political reservations and ballot quotas can be effective. Such policies directly increase female representation, and the additional exposure voters have with female leadership serves to improve the accuracy of expectations concerning *future* female candidates. This results in more females elected even after termination of the affirmative action policies. This outcome is in line with Coate and Loury (1993) who discuss how affirmative action may be effective because it gives individuals the opportunity to demonstrate capabilities.

These findings however raise an additional question. Do voters or corporate boards choosing leaders behave differently in situations where they *are* familiar with female performance? Specifically, I ask how the gender of a previously elected official influences the genders of future winners. This is essentially the same question addressed by Beaman et al. (2009) and De Paola et al. (2010) but with a potentially important difference. In the electoral environment I study females are well represented, meaning voters are unlikely to be ignorant about how a female will lead. Such situations are becoming increasingly common. For example, in the United States females of both political parties hold a total of 96 Congressional seats. While only about 19% of the legislature, this still affords voters the opportunity to form reasonably accurate expectations of how females will govern. In American industry, females occupy roughly 70% of managerial positions in the health care and education fields (GAO 2010). In these fields, those making hiring and promotion decision likely have accurate beliefs about how an average woman will perform. In effect, there is little room for ignorance-based discrimination.

To address this issue, I study whether gender continues to matter in elections when females already are well-represented as officeholders. One possibility is that gender ceases to be a factor and the genders of those elected are essentially random. I call this "gender indifference." A second possibility is that gender continues to provide information that voters value. Voters may value gender *per se* or may value the policy signal gender provides. For example, a wide array of evidence shows that males and females make public policy decisions differently. Thomas (1991), Berkman and O'Connor (1993), Vega and Firestone (1995), and Thomas and Welch (2001) all find divergence in the behavior and priorities of men and women officials. A set of papers in the economics literature shows that female decision makers increase health spending (Rehavi 2007), invest in more public education (Clots-Figueras 2006), and provide for infrastructure important to their gender (Chattopadhyay and Duflo 2004). Also, experimental work, such as Andreoni and Vesterlund (2001), suggests women are more likely to be egalitarian, a relevant fact in public goods decision-making. If voters recognize actual differences in how male and female officeholders govern, they may use gender as a convenient screening device to practice "policy-based" statistical discrimination. That is, voters may prefer a certain policy outcome and will elect candidates of a certain gender to achieve in expectation the desired outcome.[1]

However, it is difficult to disentangle whether voters value gender *per se* or because of the policy information gender provides. Voters may hold taste-based prejudices or affinities for one gender or the other (i.e., a *per se* preference) apart from policy concerns. Yet, a voter's actions resulting from each preference motivation will tend to be similar. For the current discussion I do not try to distinguish the two motivations. Rather, I make a distinction between voters' having *well-informed* preferences over gender (regardless of the motivation) as opposed to *ignorant* or *indifferent* gender preferences.

Thus there are three possibilities for how past female elected officials may influence the genders of future officials. First, if greater experience with a female reduces negative voter ignorance this will *increase* a future female candidate's probability of winning (Hypothesis 1). Second, if voters

[1]This may be especially true in low-information elections such as I consider where voters know little about candidates aside from what is on the ballot (e.g., name). This is discussed by McDermott (1997).

are indifferent to gender then no gender effect is expected (Hypothesis 2). Third, if voters hold well-informed gender preferences then electing an additional female on average will generate demand for males in reaction to the additional female, and vice versa (Hypothesis 3). The idea is that the well-informed electorate has a preferred gender composition and will vote to achieve it. In this paper I seek first to determine which hypothesis holds in a high-female representation situation. The previous work in India and Italy attributes much of the under-representation of females to ignorance-based discrimination. Yet, as females fill more management and governing positions, ignorance should diminish while other types of discrimination may rise in relative importance.

I specifically study California school district governing board elections. Overall, females comprise 45% of school board members. This is near-parity and in stark contrast to the India (11%) and Italy (7%) cases previously studied. I estimate the effect adding an additional female to the board has on election outcomes and candidacy two years later as it regards gender. Each governing board has officials elected in the two most recent elections. For this reason, who is currently on the board is relevant information to voters in the current election. If additional experience with female officials primarily reduces ignorance-based discrimination toward female candidates (as in India and Italy), I expect an additional female officeholder to *increase* the likelihood females are subsequently elected (Hypothesis 1). If voters are indifferent to gender I expect a zero effect (Hypothesis 2). If voters hold well-informed gender preferences, I expect an additional female officeholder to decrease the likelihood females are subsequently elected (Hypothesis 3). Since others have shown the effects of added exposure to female officials in low-female representation environments, my study allows us to learn whether high-female and low-female institutions are fundamentally different.

To clearly demonstrate how gender can still be relevant even when voters have well-informed beliefs, I present a simple election model. This model describes the essence of Hypothesis 3. In this model the electorate has an ideal gender composition for the board. As already discussed, this ideal gender composition may be based on policy concerns or gender *per se*. Ultimately I will not be able to distinguish the two empirically. The important point is that voters are *not* ignorant about the distributions describing how candidates of either gender will behave if elected. In the model, the previous election outcome determines the gender composition for the cohort of board seats that will continue to serve with new members chosen in the current election. This is done because school boards are elected in staggered elections every two years with four year terms. If the gender composition of continuing members does not equal the electorate's ideal composition then the voters will have demand in the next election for one gender or the other. In the next election, voters seek to elect more candidates of the higher-demand gender in an attempt to achieve the ideal composition. Potential candidates will therefore respond to this demand by entering or not entering depending on their gender. Furthermore, candidates will successfully arbitrage to zero the electoral opportunity in equilibrium. This means an additional female on the board will on average move the electorate to being "too female." This increases (decreases) demand for males (females) which results in more (fewer) males (females) running and getting elected. In equilibrium there is no net effect on who is elected other than through the candidates entering.

The empirical results support Hypothesis 3. However, I first show that a simple correlation between the genders of past and current elected officials appears to support the ignorance hypothesis. In fact, it is critical that the past election results be exogenously changed since unobserved variables and serial correlation are likely to bias results in favor of the ignorance hypothesis. To do this I use a regression discontinuity (RD) design to exogenously add an additional female or male board member, thus changing the percent female composition of the board. Overall, the additional female board member seems to reduce the number of females elected and increase the number of males. This tells us that ignorance-reduction is not the primary force related to voters' gender decisions. Instead, voters want to elect more females when the board becomes more male, and vice versa. Also, I find that candidates of each gender respond by their entry decisions to changes in voters' demands. When I control for who runs in the current election there is no estimated effect from the additional female on the board. This demonstrates how in equilibrium candidates enter in response to changes in demand and arbitrage to zero the electoral opportunity.

These findings are important for how we think about policy changes. For example, one may be interested in what will happen if affirmative action policies are terminated in high-female representation environments. If the policy has effectively decreased ignorance then we might expect policy-based discrimination to play a larger role. Removing the affirmative action policy may increase or decrease female representation relative to the affirmative action level but should also allow voters to move closer to their ideal policy outcome. To reduce policy-based statistical discrimination, it may be useful to find ways for candidates to send better, individual-specific signals.

Section 2 presents a theoretical model of voters and candidates when the potential to reduce ignorance-based discrimination is absent. Section 3 discusses the data and empirical strategy used. Section 4 presents the main results, and Section 5 closes with a discussion.

1.2 THEORETICAL MODEL

1.2.1 Environment

Suppose an electorate of voters with single-peaked preferences over the female composition of the board. Each voter's preferred female fraction falls in the space $[0, 1]$. The sequence of events is as follows:

1. Election $t-1$ determines the female fraction of the board, G_{t-1}, which will continue to serve with members elected in t. G_{t-1} is exogenous to the model.
2. The fraction G_{t-1} generates a relative demand for males, $D_{mt} \in [0, 1]$, in the next election t described by $D_{mt} = f(G_{t-1})$. D_{mt} is monotonically increasing in G_{t-1}, and the relative demand for females is simply $D_{ft} = 1 - D_{mt}$.
3. Potential candidates observe D_{mt} and decide whether to enter election t at cost c.[2]
4. The electorate chooses a single candidate in election t.

Model the vote share, V_j^g, obtained by candidate j of gender $g \in \{m, f\}$ as

$$V_j^g = \frac{D_{gt}}{N_g} + \beta\epsilon_j \tag{1.1}$$

where $\beta \geq 0$ and ϵ_j is Type I Extreme Value distributed. Let N_g denote the number of candidates of gender g. The intuition is that j's vote share depends on both the aggregate demand for j's gender and how many of j's gender run. The ϵ_j term models idiosyncratic factors affecting j's election outcome. β serves only to scale the draw from the distribution ϵ_j to the D_{gt}/N_g term. From these vote shares, I can write individual j's probability of winning as[3]:

$$P_j^g = \frac{e^{D_{gt}/(\beta N_g)}}{N_m e^{D_{mt}/(\beta N_m)} + N_f e^{D_{ft}/(\beta N_f)}}. \tag{1.2}$$

The individual probabilities for all candidates of gender g are equal in expectation. This means that the probabilities *any* member of gender g wins can be written as

$$P_g = N_g P_j^g \tag{1.3}$$

[2] I assume all candidates have the same costs. This substantially simplifies the model and does not appear to change any of the key conclusions.

[3] This is the standard multinomial logit framework.

1.2.2 Entry

Potential candidate j's payoff if he or she chooses to run is

$$\Pi_j^g = \gamma P_j^g(N_m, N_f | D_{gt}) - c \tag{1.4}$$

Candidates are motivated by the benefits of winning, γ, but must pay cost c to run. In equilibrium, j's marginal benefit of running equals the marginal cost of running, resulting in the equilibrium conditions

$$\gamma P_j^m - c = 0 \tag{1.5}$$

$$\gamma P_j^f - c = 0 \tag{1.6}$$

By substituting in P_j^m and P_j^f and equating the equilibrium conditions, I can show

$$\frac{N_f}{N_m} = \frac{D_{ft}}{D_{mt}} = \frac{1 - D_{mt}}{D_{mt}} \tag{1.7}$$

which means the ratio of the number of female to male candidates is decreasing in the demand for male candidates. Using $P_m = N_m P_j^m$ and $P_f = N_f P_j^f$, Equation (7), and the fact that in equilibrium $P_j^m = P_j^f$ must hold, we have

$$\frac{N_f}{N_m} = \frac{1 - D_{mt}}{D_{mt}} = \frac{P_f}{P_m} \tag{1.8}$$

which shows $P_m = D_{mt}$ and $P_f = 1 - D_{mt}$. This highlights what is assumed in the model. An increase in relative demand for males (females) will result in a direct increase in the probability a male (female) wins.

1.2.3 Comparative Statics

The model has the following comparative statics:

1. The electorate will choose males (females) more often when demand for males increases (decreases). As the model is constructed, an increase in female representation generates an increase in demand for males, and vice versa. If the electorate is trying to achieve an ideal gender composition across elections, this is the pattern we expect. An increase in demand for males predicts a decrease in females elected.

2. Entry decisions should respond to changes in demand. Specifically, N_f/N_m declines as demand for males increases. This relationship suggests more males and fewer females run when demand for males increases, and vice versa.
3. Controlling for entry decisions, there should be no impact from a change in demand on the election outcome because of equilibrium effects. That is, in equilibrium the entrants successfully arbitrage to zero all electoral opportunities.

1.3 SCHOOLS BOARDS, DATA, AND EMPIRICAL APPROACH

In this section I discuss the institutional environment and data used. I also present the empirical challenge and identification approach.

1.3.1 School Boards and Elections

An elected board governs each California school district. In many districts board members are parents, teachers, or community activists rather than career politicians. Running for office is also relatively simple as there are no filing fees or signatures required (California Ed Code Sect 5030, 35107, 72022, 72103 and E.C. Sect 201). No primaries take place so all candidates compete for open seats in the same election. These institutional features make California school boards a good place to observe gender preferences since we can avoid many concerns about primaries, parties, or pre-election bargaining. For instance, in a congressional race, the candidates at the general election have already won a primary. Furthermore, some candidates who desired to run in the primary may not have filed because of insufficient financing or party pressure. Regarding gender, Fox and Oxley (2003) have shown that parties select candidates based in part on gender stereotypes about a position. This would make inference from observed behavior difficult. In the school board elections,however, I avoid these problems since I presumably observe all those desiring to run given a nominal running cost.

A typical school board has five or seven members. These board members are responsible to voters for running the school district though most boards hire a superintendent to handle day-to-day operations. Every two years, "half" the board is up for election. In a five (seven) member district, three (four) seats will be up for election in year t and two (three) seats will be up in year

$t + 2$. Each term is four years so that at any given time the board has members from the first and last half of their terms. Most districts - and all in my sample - are open or at-large. That is, there are no board seats associated with specific geographic areas within the school district. All candidates run against each other and voters vote for as many as there are open seats. The top vote getters, up to the number of open seats, are elected.

1.3.2 Data

I use California school district governing board elections from 2000 to 2007 provided in the California Elections Data Archive (CEDA). CEDA includes election dates, districts, candidate names, ballot designations, incumbency information, and vote tallies for most contested school district elections in the state. As necessary and possible, I supplemented this data with results published by county election offices. The CEDA data set contains over 19,000 (non-unique) candidates.

Using this data, I first classified each school board candidate by gender. To do this, I assigned male or female values to each candidate based on a suffix (e.g., Jr., Sr.) or a first name that is gender-specific with high probability. Some names were also classified based on a gender-specific ballot designation, such as "Businesswoman" or "Housewife." This procedure classified about 90 percent of candidates. Next, I searched the internet for information that would identify the gender of candidates with ambiguous names. For instance, I could often determine a candidate's gender from news article quotes, pictures, or references in board meeting minutes. This procedure allowed me to classify the gender of about 98 percent of candidates from the original data set.[4]

Each election observation in my constructed data set consists of the number of candidates of each gender running, the number of incumbents of each gender seeking re-election, the number of incumbents and non-incumbents of each gender elected, previous election information, and control variables. Since I do not have district-level demographic data for each year, I use district-wide student body demographic information from the California Department of Education as a proxy for the voting public's demographics.

[4] I was especially conservative in assuming the gender associated with a name. For example, the names "Jean" and "Terry" were always verified. Furthermore, because most people in my sample have some public "paper trail" due to their involvement in politics, my method is probably superior to matching names based on a probability distribution of names. Except for names with very high gender associations, I individually verified the genders. An additional issue is how voters perceive gender-ambiguous names. In my data if voters mistake a candidates' gender, this will serve to attenuate my estimates due to measurement error.

1.3.3 Descriptive Statistics

Table 1.1 provides summary statistics for each variable. The summary statistics show several noteworthy features of the data. First, approximately 45% of board members and 42% of candidates are female. Also, the average number of seats up for election is 2.562. Since five-member boards (by far the modal number) alternate between having 2 and 3 seats up for election, this suggest the sample nearly evenly comes from each type. Figure 1.1 shows time trends for the percentage of board members and candidates who are female. Note the relative stability of the series over time.

Figure 1.2 shows the distribution of elections by the percent of female candidates in the race. Races are skewed toward having more male than female candidates. Note also how approximately 9% of election have less than 10% female candidates (leftmost bar) while only about 3% have an equally low number of male candidates (rightmost bar). Figure 1.3 shows the distribution of board gender compositions. For a five-person board, each bar corresponds to the percent of boards with zero, one, two, three, four, or five females. What this shows is that most boards have males and females, though more boards are all-male (leftmost bar) than are all-female (rightmost bar). This also shows that over 75% of boards have at least two female members and 95% have at least one female member. This further emphasizes that most electorates will have recent experience with female members.

1.3.4 Identification Strategy

As the question is formulated, I want to estimate the effect of an exogenous change in the board gender composition on future elections. However, exogenously varying the outcomes of the previous election so I can estimate how candidates and voters will respond in the current election is an empirical challenge. A simple OLS regression of the number of male or female election winners and how many females won the previous election likely yields biased estimates due to unobservables. To illustrate this problem, I show in Table 1.2 what happens when I estimate the effect of the number of females elected two years prior on the number of candidates who run and win in the current election. Columns 3 and 4 show that districts electing more women two years prior elect more women and fewer men in the current election. While interesting, this is unsurprising and does little to answer the causal question of how changes in the gender composition of the current board influences the current

Table 1.1: Summary Statistics

	Means		
	Females	Males	
Total Elected	1.150	1.409	
	(0.820)	(0.868)	
Non-Incumbents Elected	0.506	0.635	
	(0.654)	(0.764)	
Incumbents Elected	0.645	0.774	
	(0.712)	(0.778)	
Total Candidates per Seat	0.795	1.096	
	(0.515)	(0.612)	
Incumbent Candidates per Seat	0.322	0.383	
	(0.310)	(0.330)	
Non-Incumbent Candidates per Seat	0.479	0.723	
	(0.470)	(0.578)	
Seats up for Election			2.562
			(0.668)
District Enrollment (1000's)			8.261
			(10.525)
% English Language Learners (Students)			20.32
			(17.68)
% Free or Reduced Price Lunch (Students)			44.42
			(26.28)
% Minority (Students)			53.73
			(27.66)

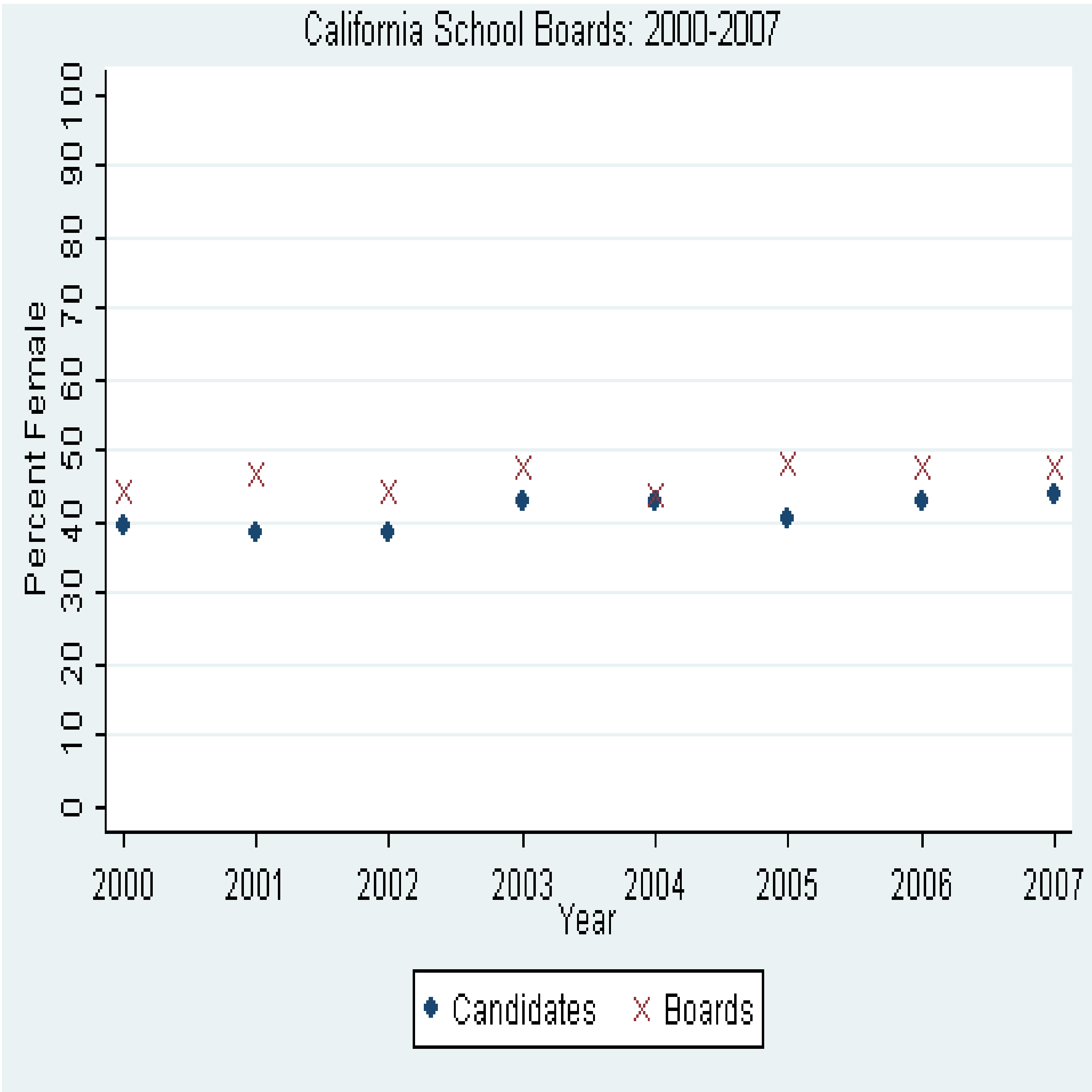

Figure 1.1: Percent Female Candidates and Board Members

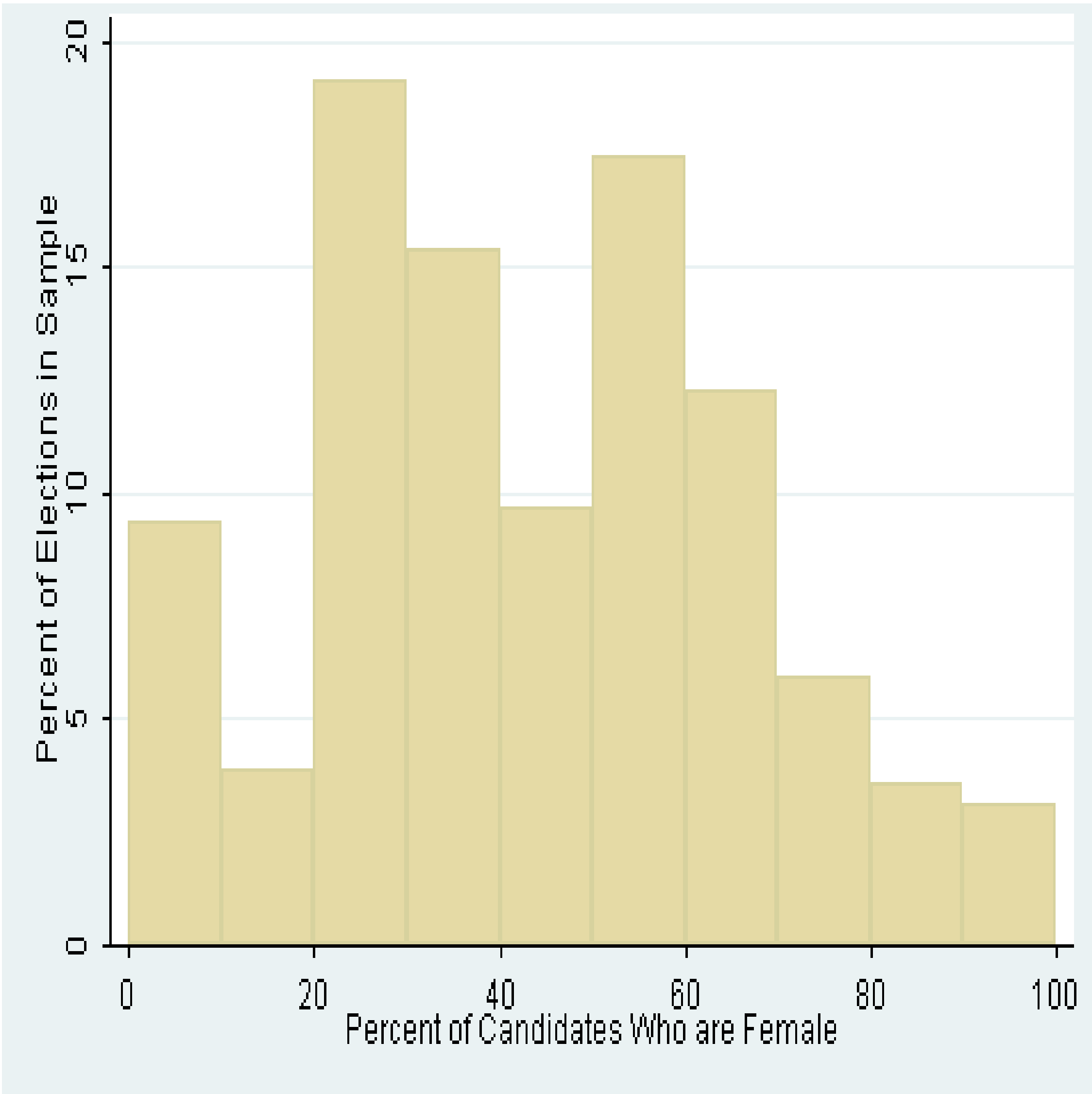

Figure 1.2: Histogram - Percent Female Candidates

Histogram shows the distribution of races in the sample by the percentage of female candidates. Sample includes elections in years 2000 to 2007.

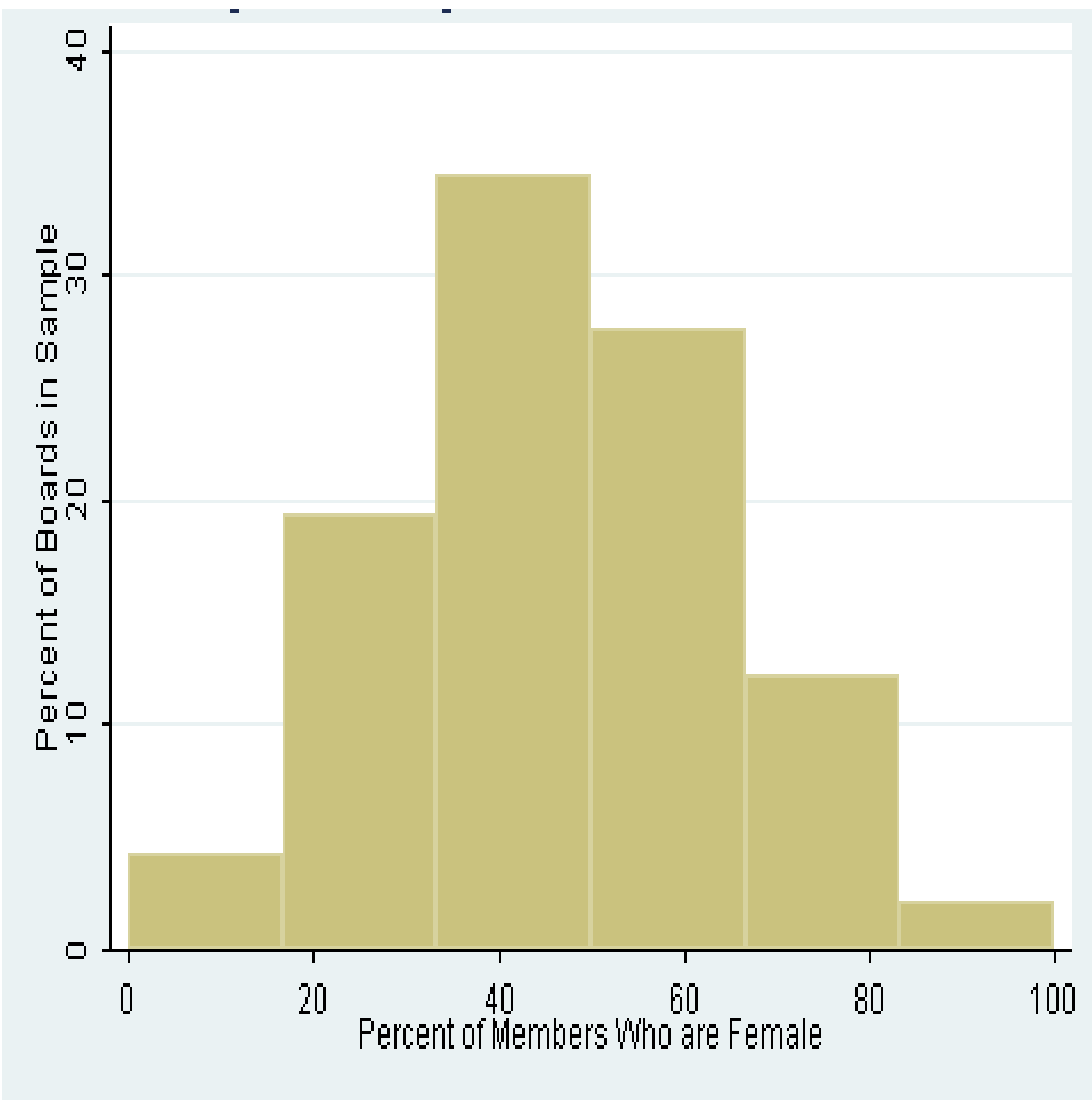

Figure 1.3: Histogram - Percent Female Members

Histogram shows the distribution of boards in the sample by the percentage of female members. Sample includes years 2000 to 2007.

election.[5] Taken literally, this estimate could falsely be seen as evidence for more female officeholders reducing ignorance-based discrimination even in this high-female representation scenario. District-level demographic and social variables such as a district's size, ethnic distribution, rural-ness, or political persuasion may influence both the number of candidates of a particular gender and which gender voters elect. Furthermore, is is likely unobserved district characteristics affecting a woman's likelihood of winning in the past election will also affect another woman's chances in the current election. Therefore, to credibly estimate a causal effect I need an exogenous source of variation in past results to isolate the variable of interest (i.e., female representation on the board).

To do this I use a regression discontinuity design. Essentially, I compare situations where one female closely defeats a male and where one female closely loses to a male in the previous election. In California, board members are elected to four-year, staggered terms on the first Tuesday of November[6]. For instance, in year $Y1$ three members are elected and in year $Y3$ two members are elected.[7] In year $Y5$ the cohort of seats filled in year $Y1$ are again up for election. This staggering allows me to exploit quasi-random variation in the board membership caused during the Y1 election to estimate effects on candidate behavior and election outcomes in Y3. Since the Y1 cohort will remain on the board two years past the Y3 election, it is reasonable to assume variation from the Y1 election is key information that Y3 candidates and voters use when considering election prospects.

To implement this approach, I construct a "running variable" for each election. The running variable is the vote difference between the female and male who were closest in vote percentages to each other but on opposite sides of the winning threshold. If three seats are open, the winning threshold is between the third and fourth highest candidates when ranked from most to least votes. I subtract from the female's vote percentage the male's vote percentage. If the male won, the running variable is negative. If the female won, it is positive. The distribution of the running variable used in my analysis is shown in Figure 1.4. Notice that most observations are near the zero point, and there are very few observations beyond a difference of twenty.

The discontinuity I use to identify my regressions is at the point where the running variable is zero. This is where a female just barely defeats a male, or vice versa. At this zero point I can

[5] Columns 1 and 2 in Table 1.2 show a similar (yet insignificant) pattern for candidate entry: women winning in the past increases (decreases) the number of women (men) currently running.

[6] Some districts have adopted more complicated arrangements and hold elections at other times of the year. These districts are not included in the sample.

[7] Most districts (94.8% of sample) have five members, but some have seven.

Table 1.2: Naive Effect of Females Winning on Gender of Candidates and Winners in Next Election

Dependent Variable	*Number Cand. Running*		*Number Elected*	
With Controls	Females	Males	Females	Males
Number of Fem. Win. Previous	0.048	-0.087	0.055*	-0.055*
	(0.049)	(0.054)	(0.032)	(0.032)
No. Enrolled Students (1000's)	0.001***	0.004***	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Pct. English Learners	0.006	-0.011**	0.007***	-0.007***
	(0.005)	(0.005)	(0.003)	(0.003)
Pct. Free or Reduced Lunch	0.003	0.004	-0.003**	0.003**
	(0.002)	(0.002)	(0.001)	(0.001)
Pct. Minority	-0.006	0.005	-0.004**	0.004**
	(0.003)	(0.003)	(0.002)	(0.002)
Seats Up for Election	0.664***	0.942***	0.432***	0.568***
	(0.066)	(0.070)	(0.039)	(0.039)
No. Observations	1468	1468	1468	1468

Standard errors clustered by district are shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Columns show results from regressing the number of female and male candidates (Columns 1 and 2) and the number of females and males elected (Columns 3 and 4) on the number of females elected in the previous election. This represents the naive approach to estimating the effect of past females winning on future election behavior.

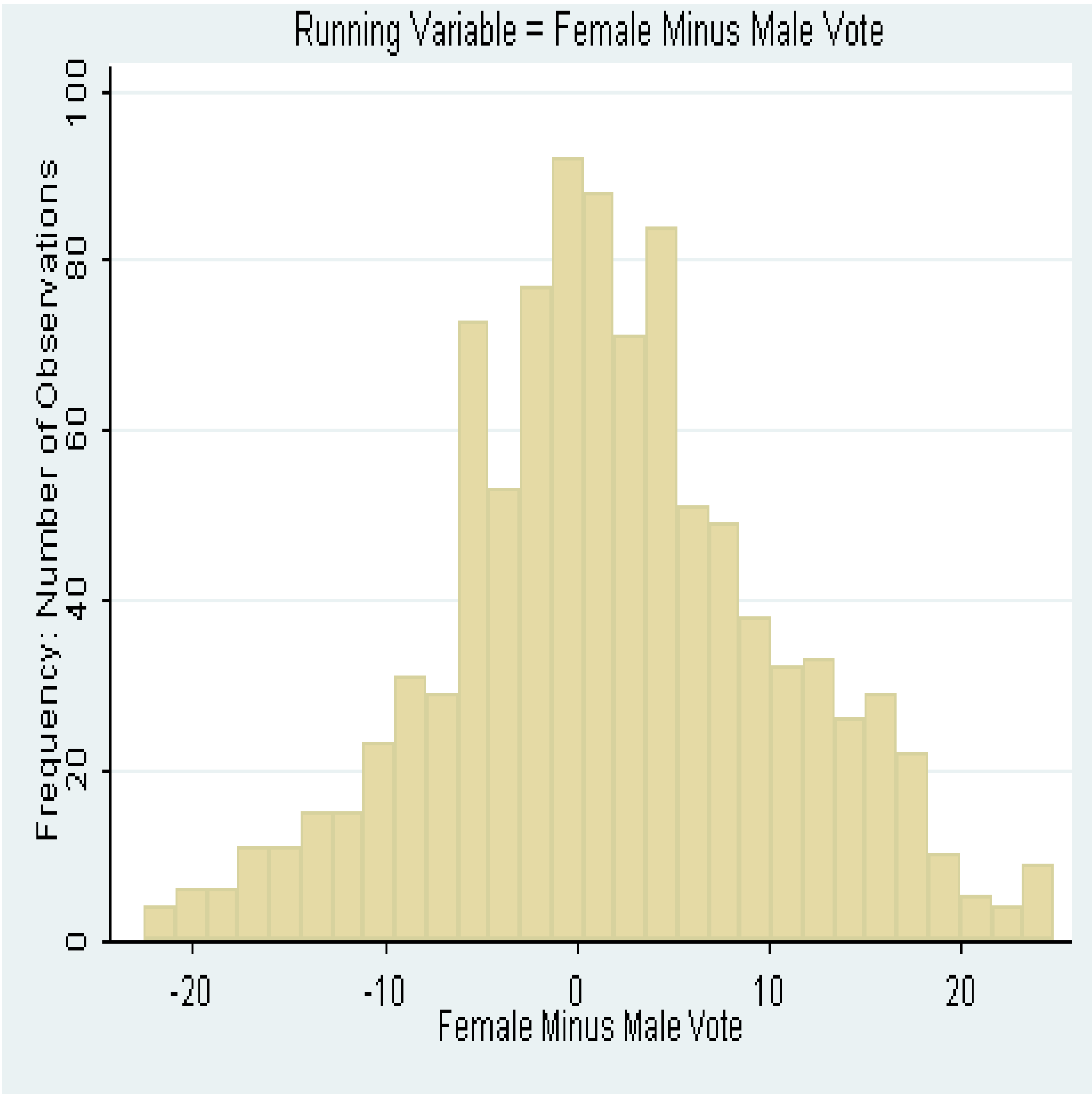

Figure 1.4: Distribution of Running Variable in Sample

The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. At 0, the running variable is interpreted as a male and female candidate tying. Positive (negative) values indicate a female (male) defeated a male (female).

estimate how the outcomes of interest two years later are affected. A small change in the vote tally near this point creates a binary change in whether a male or female wins and takes office. This creates the exogenous variation in the board composition needed to estimate causal effects. The intuition for this identification strategy is that very close elections are essentially random in whether a male or female wins, but the consequences are dramatic. The notion of using close elections to generate random variation is employed by DiNardo and Lee (2004), Lee (2008), Rehavi (2007), and Ferreira and Gyourko (2009).

To produce estimates and standard errors, I use the following regression:

$$Y_{it} = \alpha M_{jt-2} + T(M) + X_{jt}^{'}\gamma + \epsilon_{jt} \tag{1.9}$$

where Y_{it} is the outcome of interest, M_{jt-2} indicates if the female won the close election two years prior, $T(M)$ represents flexible polynomials fit separately to the running variable on each side of the discontinuity, and X_{jt} is a set of election and district controls and an intercept. The coefficient α is the main value of interest. The estimate of α is the average effect on the outcome of interest of a woman winning an election two years prior and thus sitting on the board.

In practice I consider several bandwidths. The bandwidth indicates how much of the data on each side of the discontinuity is included in the sample. The main results include all observations where the running variable is between -25 and 25, and I fit quadratic polynomials to the running variable on each side of the discontinuity. I also vary the degree of this polynomial and also the bandwidth. This adjusts how much observations far from the discontinuity are allowed to influence the estimates.

Before presenting results, it is important to check the key identifying assumption of the regression discontinuity approach. Since we assume observations near the discontinuity are randomly assigned to either side, there should be no difference in other variables near this discontinuity. Table 1.3 shows that for all the control variables this holds.

Table 1.3: Test of Identifying Assumptions

Dependent Variable	*District Enrollment*		*Pct. English Learners*		*Pct. Free/Reduced Lunch*		*Pct. Minority*	
Discontinuity at 0	-1034	-284	-0.778	-2.15	1.42	1.41	-2.84	-2.81
	(912)	(1420)	(1.91)	(2.74)	(2.77)	(3.89)	(2.84)	(3.99)
Polynomial Fit	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
No. Observations	997	997	997	997	997	997	997	997

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Each specification shows the effect on district control variables of a female candidate defeating a male candidate in the election two years prior. The discontinuity is at the point where prior female vote percentage minus male vote percentage is zero for the closest male and female on each side of the winning threshold. All specifications fit separate polynomials on each side of the discontinuity. Results, regardless of the polynomial shape fit to the data, indicate the identifying assumptions hold.

1.4 MAIN RESULTS

I now estimate whether an additional female on the board increases, decreases, or has zero effect on the number of females chosen in the current election. This will indicate which hypothesis is consistent with the data. An increase indicates the added exposure reduces voter ignorance (Hypothesis 1). A zero effect points to gender indifference (Hypothesis 2). A decrease suggests demand has shifted toward males because of the increase in females on the board (Hypothesis 3). This last hypothesis is modeled by the theory presented in Section 2.

As already mentioned, I find evidence from electoral outcomes supporting this third hypothesis. Therefore, I also proceed to see if candidates' entry decisions also are in line with my model. According to the model, N_f/N_m will decrease as demand for males increases. That is, an additional female exogenously added to the board will increase demand for males and will therefore decrease N_f/N_m. Estimates confirm this relationship. Finally, I test a key prediction of the model. The model predicts when the system is in equilibrium, candidates will arbitrage to zero any electoral opportunity. Empirically, if this is the case, we expect an additional female to have no effect on the genders of those elected once we control in the regressions for which types of candidates enter the race. Also, there should be no effect of a female winning in the past on an individual's probability of winning in equilibrium. This is indeed what I find.

These empirical results broadly support the hypothesis that voters have gender preferences but are not ignorant about candidates' expected behavior. Yet there is an important caveat: the impacts mostly fall on non-incumbents. The likely reason for this is that the election gender effects are marginal influences on potential candidates' decisions. Suppose incumbent candidates, who are highly likely to win re-election, make their decisions to run first. Then, once incumbents have decided whether to run, potential non-incumbent candidates decide. This means the marginal, non-incumbent candidates are the ones most affected by changes in gender composition.

1.4.1 Past Elections and Election Outcomes

In the main election outcome estimates, I examine the effect of past elections on six different outcome variables. First, the total number of females elected. Second, the number of non-incumbent females elected. Third, the number of female incumbents elected. The fourth, fifth, and sixth outcome variables are simply the number of total, non-incumbent, and incumbent *males* elected.

Figures 1.5 to 1.7 illustrate the discontinuities generated in the outcome variable (Y_{it}) at the point where a female barely defeated a male. I plot local averages and a simple quadratic fit on both sides of the discontinuity at zero. The horizontal axis, "Female Minus Male Vote," plots the difference in vote percentages of the lowest winning female and highest losing male candidate in the previous election, or vice versa. Thus, positive values indicate a female win and negative value indicate a male win. A movement from just left to just right of the zero point crosses the discontinuity, adding a female to the board in place of a male. Points near each side of zero represent extremely close elections while points at the left and right extremes are uncompetitive races between a male and female. Near the discontinuity we can see how the predicted fits "jump." This visually represents the effect of interest.

Figure 1.5 displays the effect of a female winning a close election two years prior on the number of males and females elected in the current election. In the top panel the fitted quadratics *suggest* an upward jump in the number of males elected when a female wins. The bottom panel indicates a drop in the number of females under the same conditions. Table 1.4, Panel A, presents regression estimates corresponding to the bottom plot in Figure 1.5 and estimates the size and significance of the discontinuity for females. The point estimates in Table 1.5, Panel A, show a positive impact on the number of males elected due to an increase in the female composition of the board. However, the estimates are not quite significant at conventional levels for the widest bandwidth. In part, this may be because observations far from the discontinuity are allowed to exert influence on the estimates. As Figure 1.4 showed, there are actually few observations far from the discontinuity which means there may be high variance in these observations. But in Table 1.4, Panel A, Columns 3 to 10, I trim these few observations to focus on only those close to the discontinuity. I vary the bandwidths between [-10,10], [-5,5], and [-3,3] and use quadratic and linear fits ($T(M)$ in the regression equation). These estimates are similar in direction to those in Columns 1 and 2 but statistically significant. In Table 1.5, Panel A, I do the same for males elected and also find much stronger estimates. Admittedly, the visual discontinuities in Figure 1.5 are not overwhelmingly clear. But, it is suggestive that the effect is in line with Hypothesis 3 and the Table 1.4 and 1.5 estimates support this. At minimum these results are evidence against Hypothesis 1 and weak evidence for Hypothesis 3.

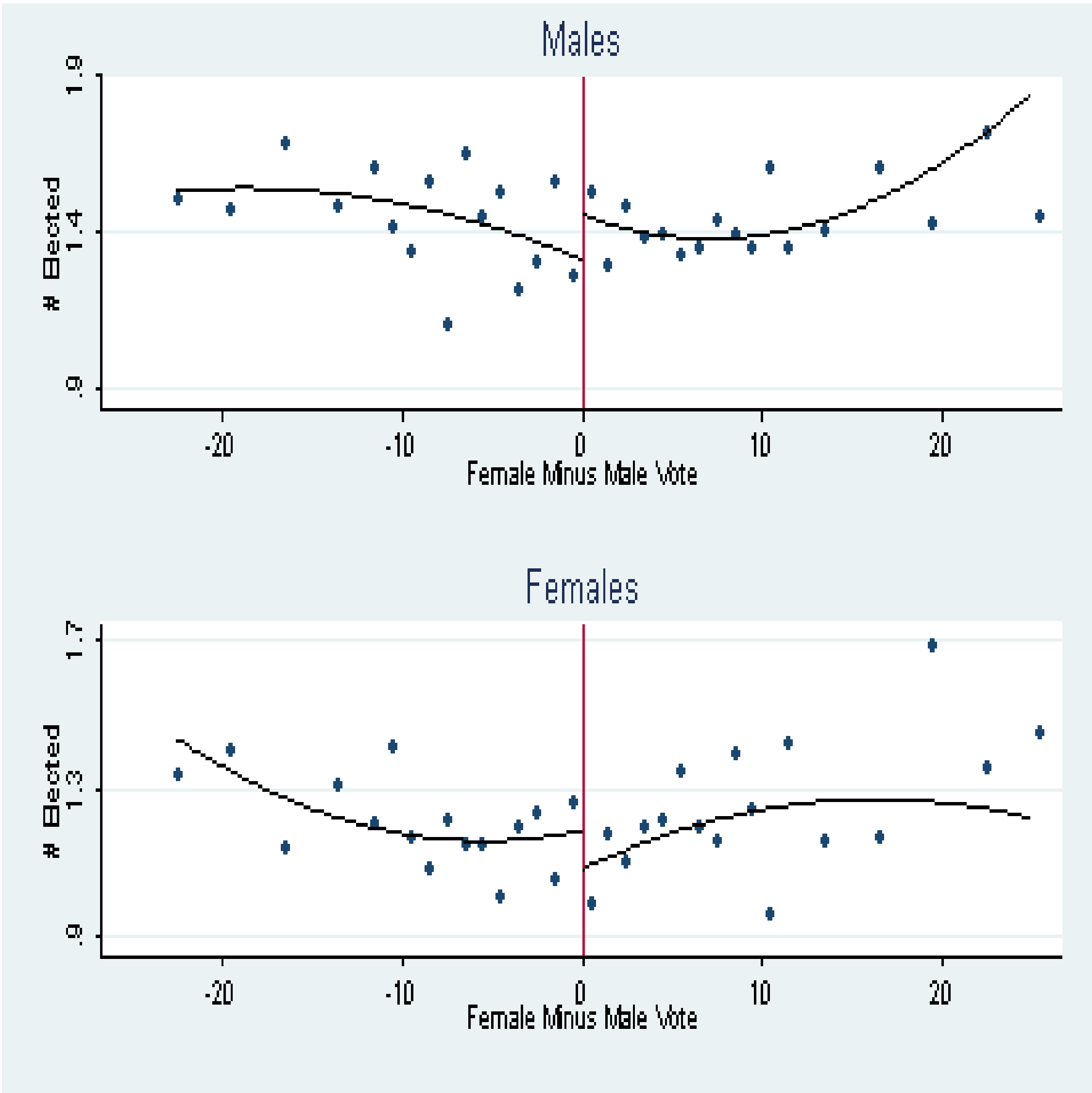

Figure 1.5: Total Number Elected by Gender

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

Table 1.4: Female Election Results

PANEL A										
Fem. Elected										
Bandwidth	[-25,25]		[-10,10]		[-10, 10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.148	-0.143	-0.349*	-0.340*	-0.144	-0.142	-0.346***	-0.346***	-0.299*	-0.288*
	(0.104)	(0.104)	(0.147)	(0.148)	(0.099)	(0.099)	(0.131)	(0.131)	(0.168)	(0.173)
PANEL B										
Non-Inc. Fem. Elected										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.219**	-0.228***	-0.241**	-0.254**	-0.186**	-0.194**	-0.258***	-0.265***	-0.161	-0.175
	(0.084)	(0.084)	(0.109)	(0.110)	(0.077)	(0.084)	(0.097)	(0.098)	(0.121)	(0.123)
PANEL C										
Inc. Fem. Elected										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.071	0.084	-0.108	-0.086	0.042	0.052	-0.088	-0.081	-0.137	-0.113
	(0.096)	(0.096)	(0.129)	(0.130)	(0.090)	(0.090)	(0.116)	(0.117)	(0.152)	(0.157)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial Fit	Quadratic	Quadratic	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Observations	997	997	737	737	737	737	470	470	307	307

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Panels organize the outcome variable and columns within each panel represent separate regressions. Each specification shows the effect on the number of females (Panel A), non-incumbent females (Panel B), or incumbent females (Panel C) who win in the current election of a female candidate defeating a male candidate in the election two years prior. The discontinuity is at the point where prior female vote percentage minus male vote percentage is zero for the closest male and female on each side of the winning threshold. Each cell shows a different combination of bandwidth, controls, and polynomial type. Moving left to right reduces bandwidth from [-25,25] to [-3,3]. Controls include number of seats up for election, district enrollment, percent minority, percent free/reduced lunch, and percent English language learners.

Table 1.5: Male Election Results

PANEL A										
Males Elected										
Bandwidth	[-25,25]		[-10,10]		[-10, 10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.148	0.143	0.349**	0.340**	0.144	0.142	0.346***	0.346***	0.299*	0.288*
	(0.104)	(0.104)	(0.147)	(0.147)	(0.099)	(0.099)	(0.131)	(0.131)	(0.168)	(0.172)
PANEL B										
Non-Inc. Males Elected										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.075	0.071	0.105	0.096	0.122	0.117	0.245*	0.246*	0.088	0.099
	(0.104)	(0.104)	(0.142)	(0.141)	(0.098)	(0.099)	(0.130)	(0.131)	(0.169)	(0.171)
PANEL C										
Inc. Males Elected										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.074	0.072	0.245*	0.244*	0.022	0.025	0.102	0.099	0.210	0.189
	(0.104)	(0.103)	(0.142)	(0.139)	(0.097)	(0.096)	(0.124)	(0.122)	(0.158)	(0.154)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial Fit	Quadratic	Quadratic	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Observations	997	997	737	737	737	737	470	470	307	307

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Panels organize the outcome variable and columns within each panel represent separate regressions. Each specification shows the effect on the number of males (Panel A), non-incumbent males (Panel B), or incumbent males (Panel C) who win in the current election of a female candidate defeating a male candidate in the election two years prior. The discontinuity is at the point where prior female vote percentage minus male vote percentage is zero for the closest male and female on each side of the winning threshold. Each cell shows a different combination of bandwidth, controls, and polynomial type. Moving left to right reduces bandwidth from [-25,25] to [-3,3]. Controls include number of seats up for election, district enrollment, percent minority, percent free/reduced lunch, and percent English language learners.

Figure 1.6 shows a similar graph for only non-incumbents. The bottom plot shows a substantial discontinuity in the number of non-incumbent females elected. Table 1.4, Panel B, estimates this as a highly significant effect. The election of a female in year $t-2$ reduces the number of non-incumbent females elected in t by about 0.22. This drop is 19% of the average total number of females elected and 43% of the average number of non-incumbent females elected. This means that most of the impact of an increase in female board composition falls on non-incumbent females. Conversely, the benefits of a male winning also go to the non-incumbent females.

While, as already discussed above, it appears the overall number of males elected increases when a female is added to the board, Table 1.5, Panels B and C, shows this effect is distributed between incumbent and non-incumbent male candidates. Furthermore, incumbent females (Table 1.4, Panel C) appear to also benefit somewhat. Mechanically, all the marginal reduction for non-incumbent females elected must be accounted for with positive values elsewhere. It appears much of the positive benefit goes to males though some goes to incumbent females. While these estimates are statistically insignificant, we know they together mirror the negative effect on non-incumbent females. It is reasonable to believe there is simply not enough power to precisely estimate each value separately. However, this also suggests gender may be functioning in an additional way not included in the model. It appears an additional female member causes voters to shift partially to all incumbents and also to males. This extra effect remains a puzzle. For this reason I assert the data is consistent with the model but also note there may be additional influences future researchers may want to consider. What is clear is the electorates appear to oscillate in large part around an ideal gender composition. When female representation increases, on average more males are chosen to balance this out. When male representation increases, the opposite happens.

1.4.2 Past Elections and Candidate Behavior

Having shown election results tend to be in line with the Hypothesis 3, I can now investigate whether candidates behave as described by the theoretical model. I estimate the impact of a female winning two years prior on candidate entry behavior. Figure 1.8 illustrates the discontinuities in the ratio of female to male candidates for all candidates and only non-incumbent candidates. Table 1.6 presents estimates of the effect. The estimates are in line with a key relationship posited by the model. As demand for males increases (due to an increase in female representation), the ratio N_f/N_m decreases. This effect falls completely on non-incumbent candidates (Table 1.6, Panel B)

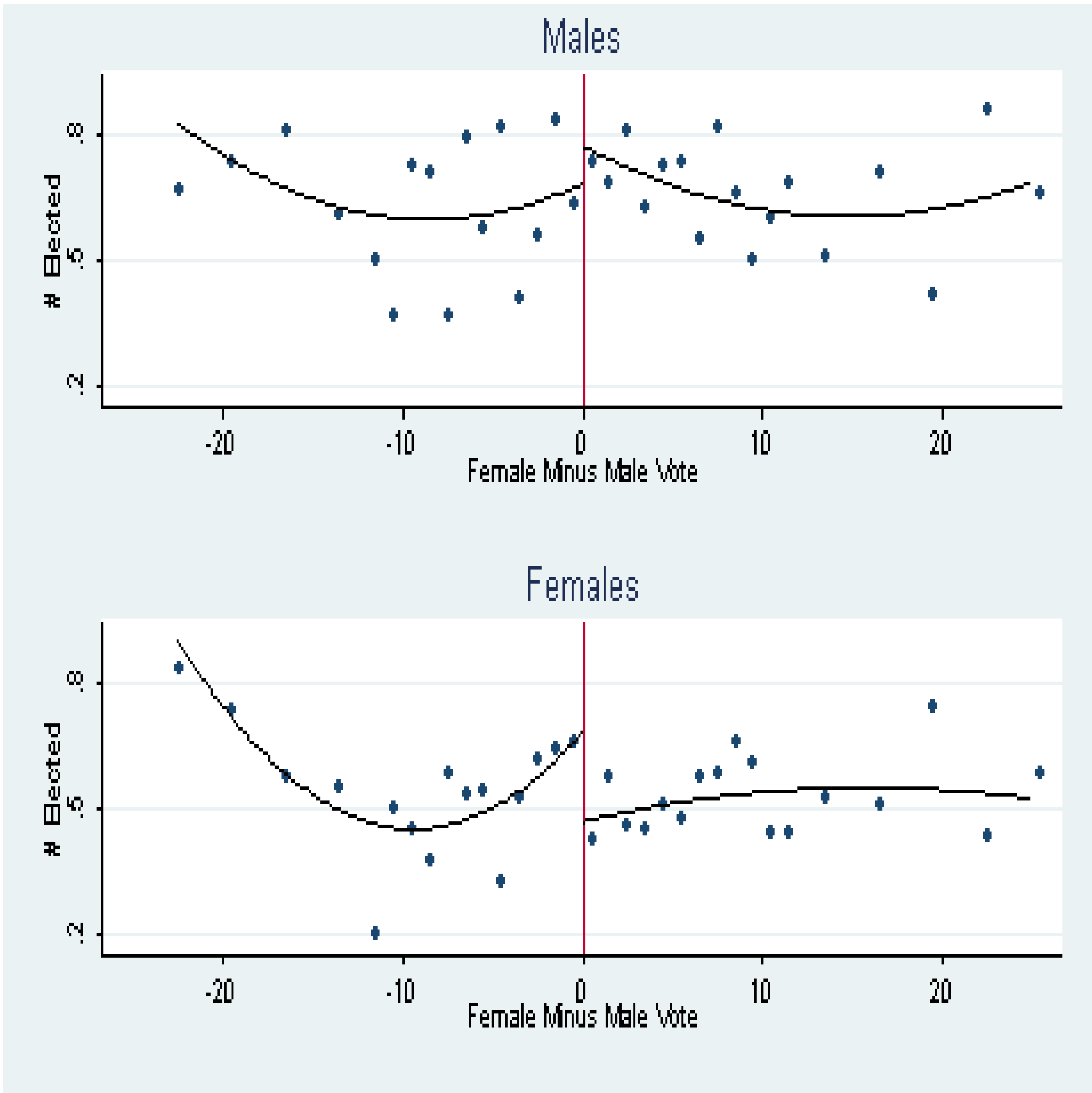

Figure 1.6: Non-Incumbents Elected by Gender

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

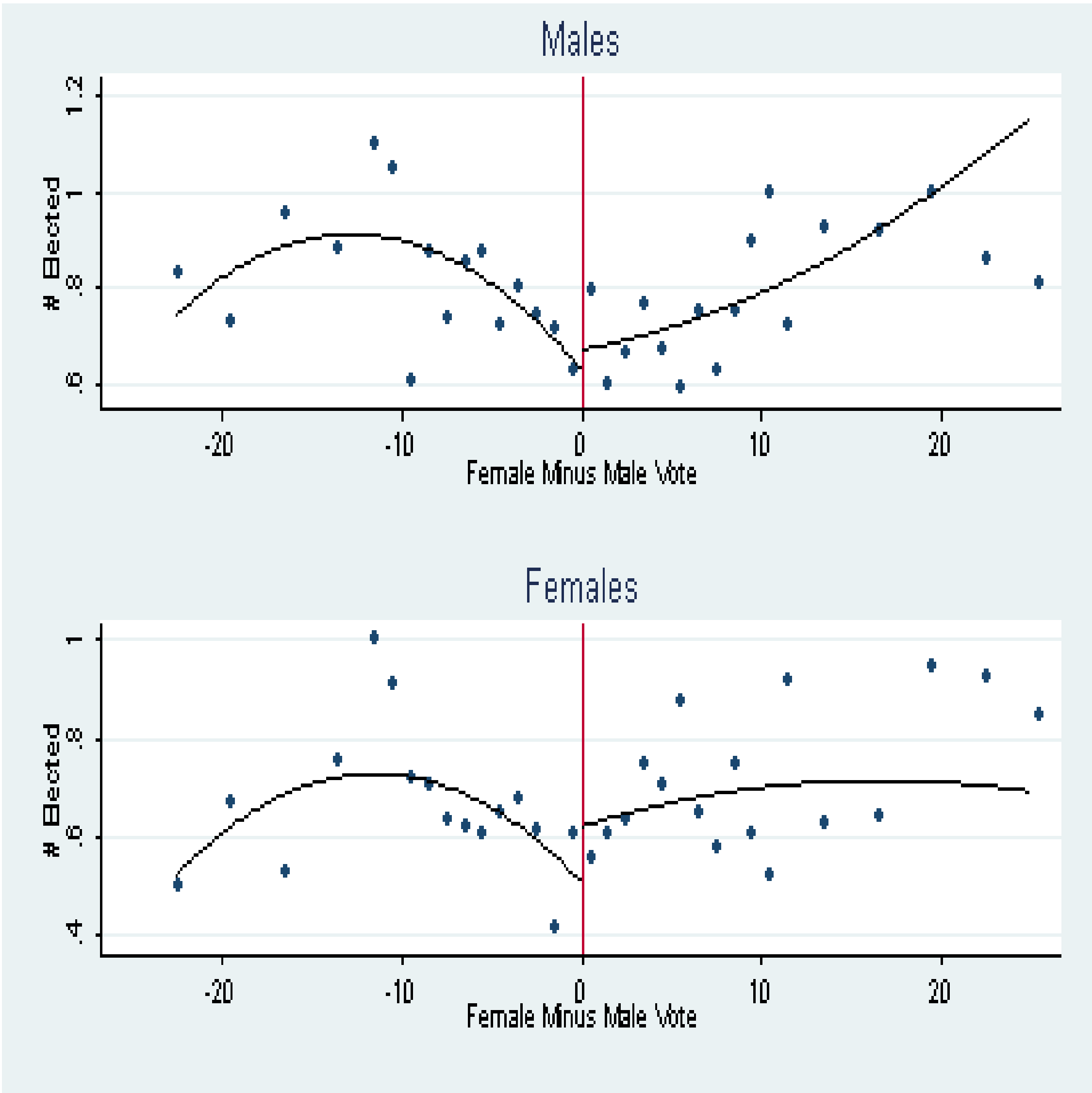

Figure 1.7: Incumbents Elected by Gender

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

rather than incumbent candidates (Table 1.6, Panel C).

Figure 1.9 alternately shows the discontinuities for the total number of male and female candidates who run. Clear jumps are visible: a female winning increases (decreases) the total number of males (females) who run. Tables 1.7 and 1.8 present estimates of this effect. From Table 1.7, Panel A, a female winning decreases the number of females who run by 0.233, or -29% of the average number running. From Table 1.8, Panel A, the effect on male candidates is an increase of 0.166, or 15% of the average number running.

In Tables 1.7 and 1.8, Panels B, estimates indicate this entry effect is mostly due to non-incumbents rather than incumbents changing their behavior. For non-incumbent females, the results are highly significant and robust. For non-incumbent males the results fall short of significance, though the point estimate is large relative to that for incumbent males and, as before, trimming extreme observations improves precision. Overall the entry results all comport with the model where candidates of both genders systematically respond to changes in demand.

1.4.3 Conditional Election Results

While Tables 1.4 and 1.5 show the overall impact of a previous female winner on who gets elected, the model says that in equilibrium the candidates arbitrage to zero any remaining electoral opportunity. Therefore, once I control for who runs for office there should be no effect on the election outcome. Table 1.9 shows estimates of the effect of adding a female to the board on election outcomes while controlling for the number of male incumbents, female incumbents, male non-incumbents, female non-incumbents, seats up for election, and district controls. In particular, I focus on non-incumbent females since this is the group for which I found a highly significant effect on election outcomes. As implied by the theoretical model, the estimated discontinuity disappears. This suggests that in fact candidates systematically respond to changes in demand to the point of arbitraging to zero any electoral opportunity. The lack of an estimated discontinuity in Table 1.9 is interpreted as a lack of any remaining electoral benefits that can be captured through candidate entry.

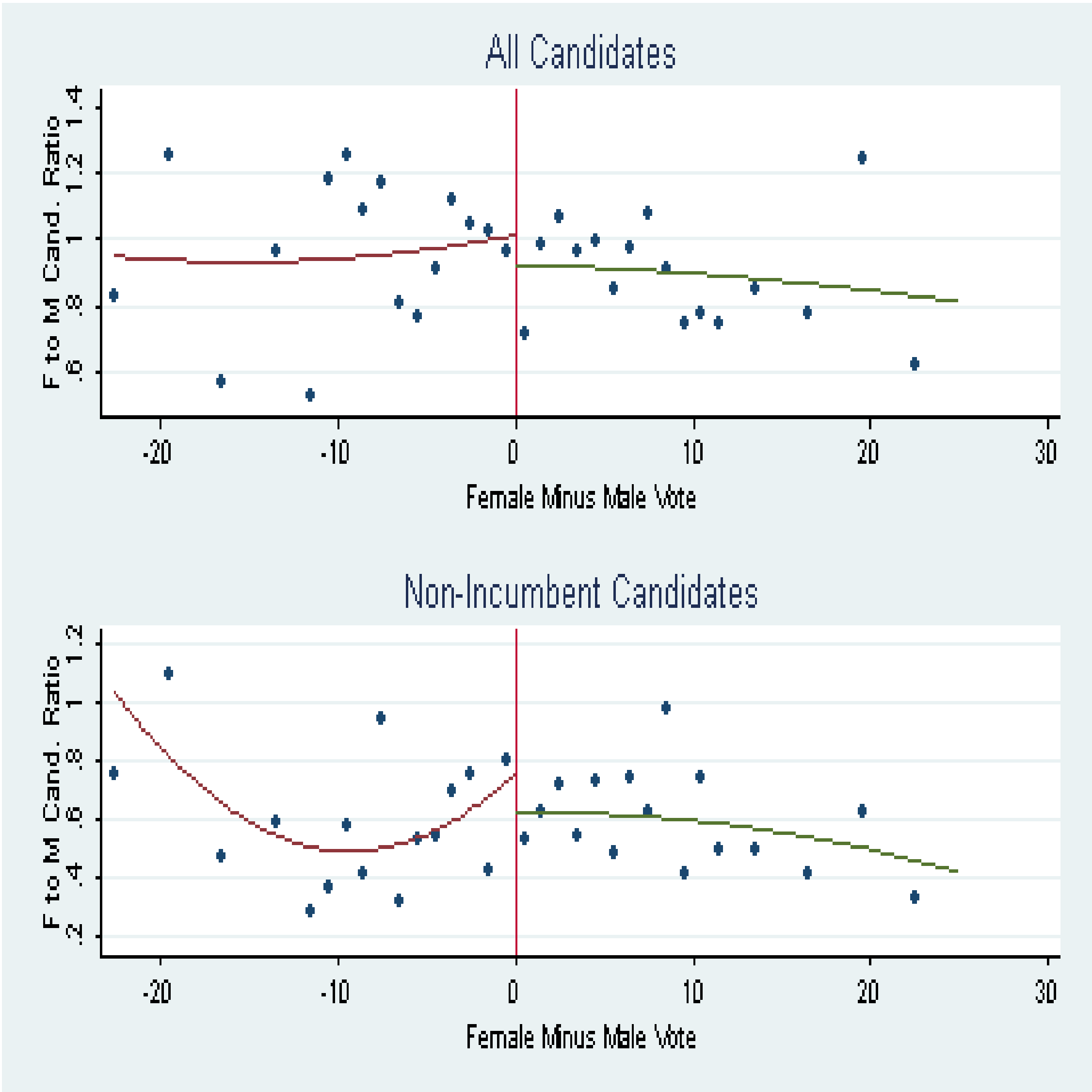

Figure 1.8: Ratio of Female to Male Candidates

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

Table 1.6: Effect on the Ratio of the Number of Female to Male Candidates

PANEL A										
All Cand. Ratio										
Bandwidth	[-25,25]		[-10,10]		[-10, 10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.175	-0.174	-0.393**	-0.430**	-0.129	-0.139	-0.347*	-0.355*	-0.394*	-0.460**
	(0.156)	(0.151)	(0.195)	(0.195)	(0.148)	(0.146)	(0.188)	(0.184)	(0.203)	(0.222)
PANEL B										
Non-Inc. Cand. Ratio										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.220**	-0.227**	-0.318**	-0.344***	-0.182*	-0.196**	-0.312**	-0.328***	-0.306*	-0.336**
	(0.111)	(0.111)	(0.128)	(0.125)	(0.097)	(0.097)	(0.126)	(0.125)	(0.157)	(0.153)
PANEL C										
Inc. Cand. Ratio										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.007	0.008	-0.087	-0.087	0.009	0.009	-0.067	-0.079	-0.102	-0.089
	(0.072)	(0.072)	(0.100)	(0.101)	(0.067)	(0.067)	(0.093)	(0.092)	(0.115)	(0.115)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial Fit	Quadratic	Quadratic	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Observations	997	997	737	737	737	737	470	470	307	307

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Panels organize the outcome variable and columns within each panel represent separate regressions. Each specification shows the effect on the ratio of the number of female to male candidates in the current election of a female candidate defeating a male candidate in the election two years prior. The discontinuity is at the point where prior female vote percentage minus male vote percentage is zero for the closest male and female on each side of the winning threshold. Results are shown with and without control. Each cell shows a different combination of bandwidth, controls, and polynomial type. Moving left to right reduces bandwidth from [-25,25] to [-3,3]. Controls include number of seats up for election, district enrollment, percent minority, percent free/reduced lunch, and percent English language learners.

Table 1.7: Female Candidate Entry Results

PANEL A										
Total Fem. per Seat										
Bandwidth	[-25,25]		[-10,10]		[-10, 10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.234***	-0.233***	-0.411***	-0.409***	-0.170**	-0.169**	-0.308***	-0.303***	-0.420***	-0.425***
	(0.078)	(0.077)	(0.107)	(0.104)	(0.075)	(0.074)	(0.096)	(0.093)	(0.117)	(0.118)
PANEL B										
Non-Inc. Fem. per Seat										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.215***	-0.216***	-0.316***	-0.316***	-0.158**	-0.159**	-0.239***	-0.234***	-0.313***	-0.323***
	(0.072)	(0.071)	(0.097)	(0.095)	(0.068)	(0.067)	(0.088)	(0.086)	(0.110)	(0.110)
PANEL C										
Inc. Fem. per Seat										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	-0.019	-0.017	-0.094*	-0.093*	-0.012	-0.011	-0.069	-0.069	-0.107*	-0.103
	(0.040)	(0.040)	(0.055)	(0.056)	(0.039)	(0.039)	(0.051)	(0.051)	(0.063)	(0.064)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial Fit	Quadratic	Quadratic	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Observations	997	997	737	737	737	737	470	470	307	307

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Panels organize the outcome variable and columns within each panel represent separate regressions. Each specification shows the effect on female candidate entry in the current election of a female candidate defeating a male candidate in the election two years prior. The discontinuity is at the point where prior female vote percentage minus male vote percentage is zero for the closest male and female on each side of the winning threshold. Each cell shows a different combination of bandwidth, controls, and polynomial type. Moving left to right reduces bandwidth from [-25,25] to [-3,3]. Controls include district enrollment, percent minority, percent free/reduced lunch, and percent English language learners.

Table 1.8: Male Candidate Entry Results

PANEL A										
Total Males per Seat										
Bandwidth	[-25,25]		[-10,10]		[-10, 10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.167*	0.166**	0.205*	0.238*	0.173**	0.184**	0.291***	0.312***	0.206	0.234*
	(0.087)	(0.083)	(0.114)	(0.111)	(0.085)	(0.080)	(0.104)	(0.101)	(0.132)	(0.126)
PANEL B										
Non-Inc. Males per Seat										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.128	0.129	0.089	0.129	0.157*	0.168*	0.218**	0.242**	0.068	0.113
	(0.089)	(0.085)	(0.110)	(0.108)	(0.088)	(0.083)	(0.103)	(0.100)	(0.132)	(0.127)
PANEL C										
Inc. Males per Seat										
Bandwidth	[-25,25]		[-10,10]		[-10,10]		[-5,5]		[-3,-3]	
Discontinuity at 0	0.038	0.036	0.116*	0.109*	0.016	0.016	0.073	0.071	0.138*	0.121*
	(0.048)	(0.047)	(0.064)	(0.062)	(0.045)	(0.044)	(0.059)	(0.058)	(0.070)	(0.070)
Controls	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Polynomial Fit	Quadratic	Quadratic	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Observations	997	997	737	737	737	737	470	470	307	307

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Panels organize the outcome variable and columns within each panel represent separate regressions. Each specification shows the effect on male candidate entry in the current election of a female candidate defeating a male candidate in the election two years prior. The discontinuity is at the point where prior female vote percentage minus male vote percentage is zero for the closest male and female on each side of the winning threshold. Each cell shows a different combination of bandwidth, controls, and polynomial type. Moving left to right reduces bandwidth from [-25,25] to [-3,3]. Controls include district enrollment, percent minority, percent free/reduced lunch, and percent English language learners.

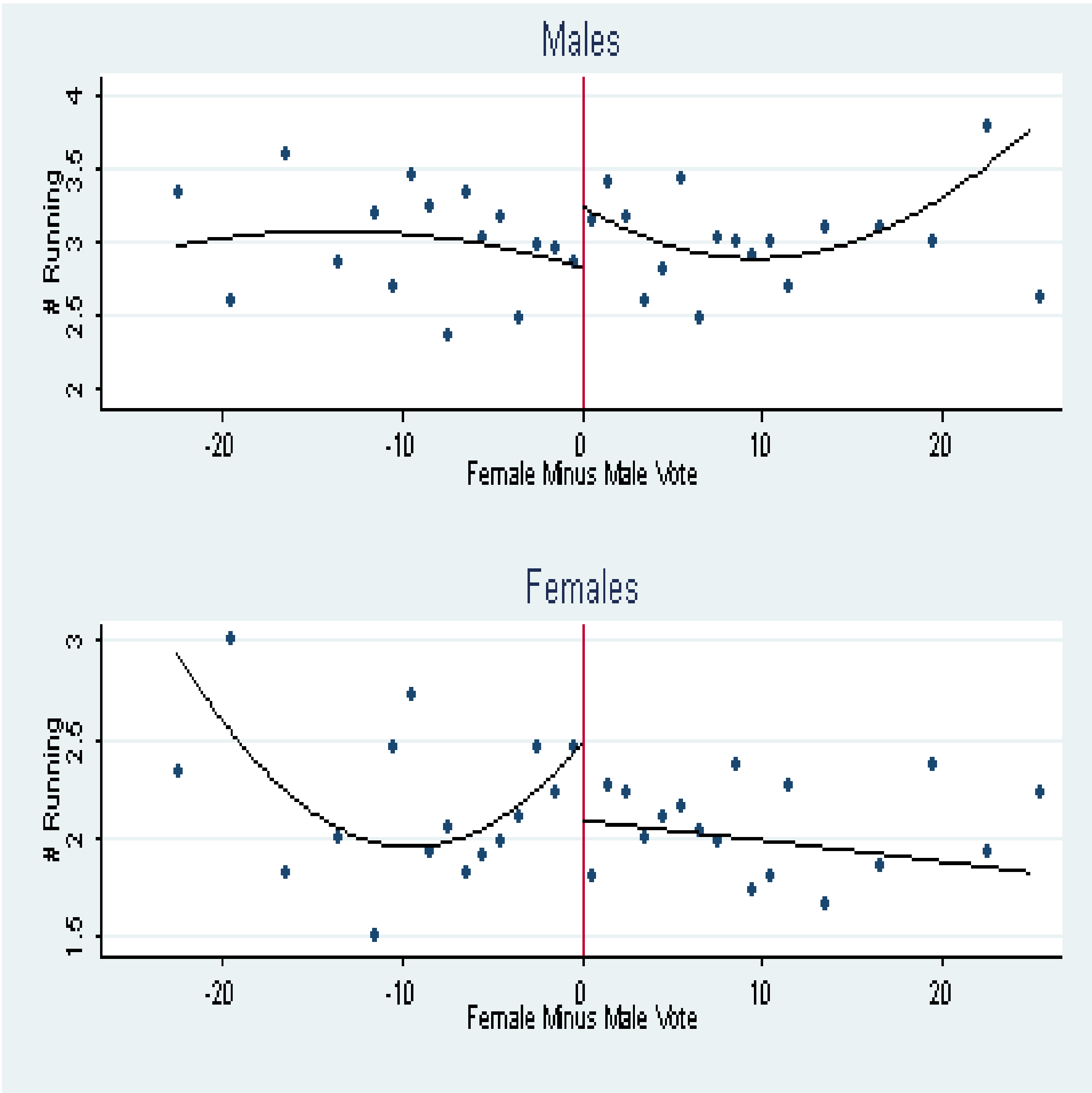

Figure 1.9: Total Number Running by Gender

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

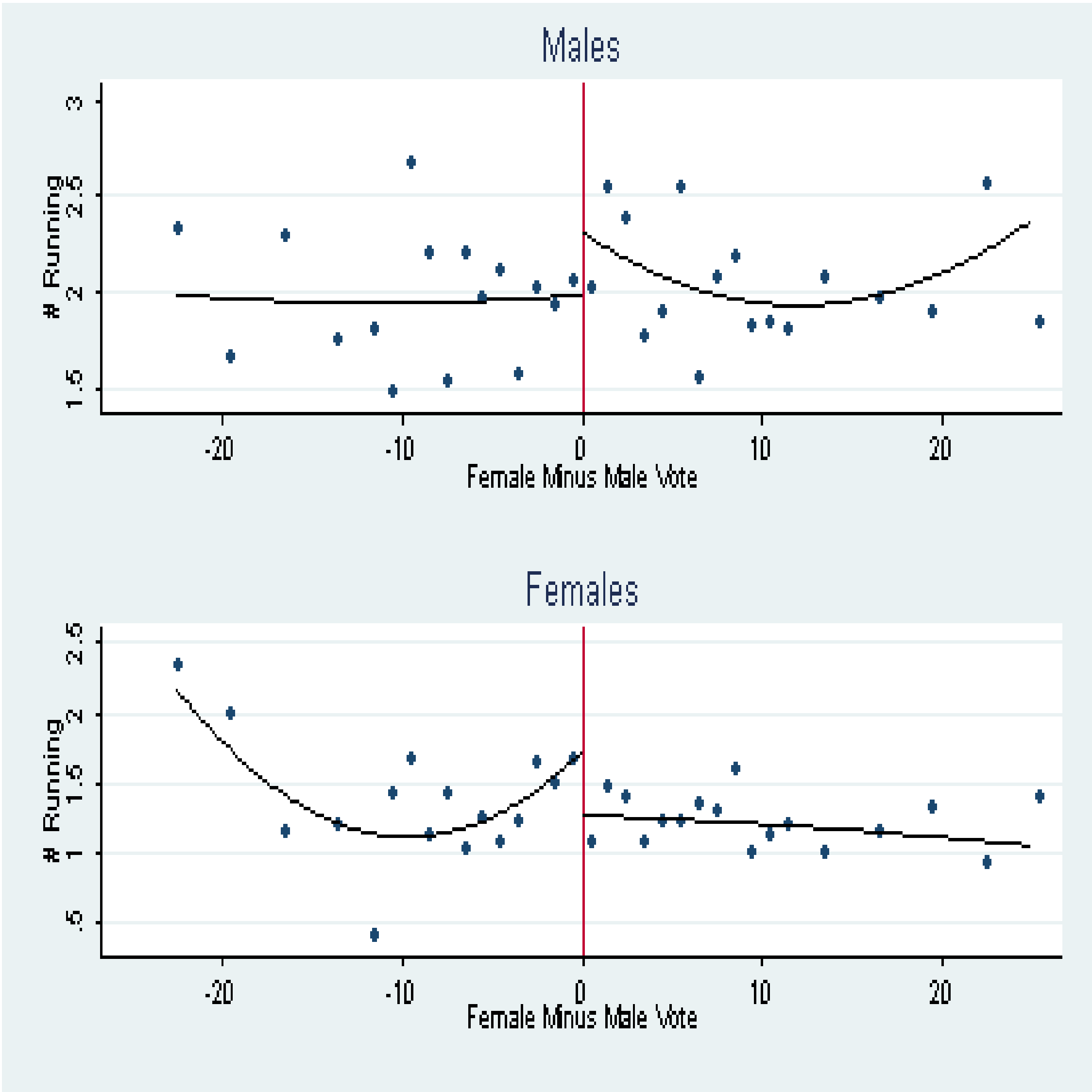

Figure 1.10: Non-Incumbents Running by Gender

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

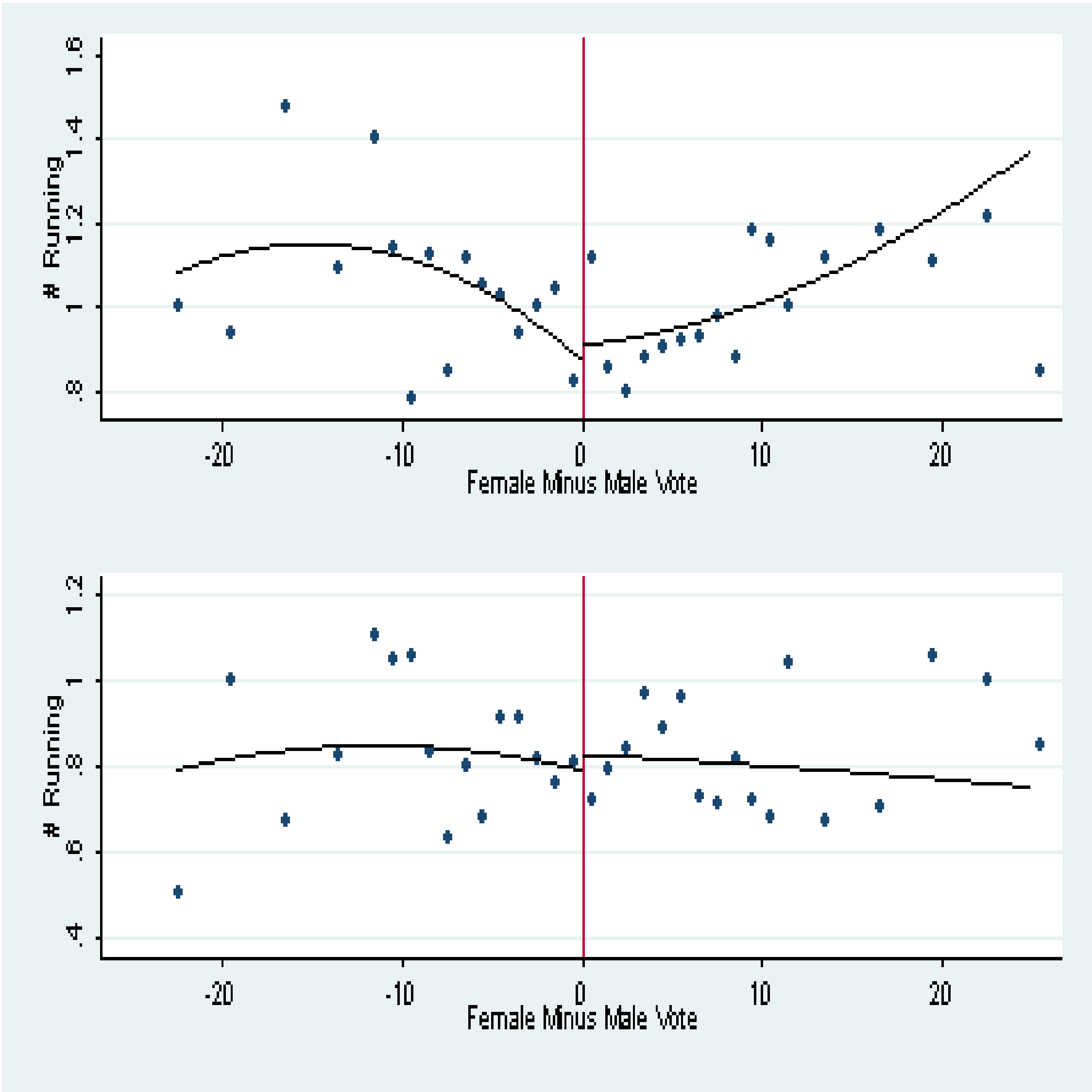

Figure 1.11: Incumbents Running by Gender

Each graph plots the local averages for the vertical axis variable against the running variable. The running variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A quadratic polynomial fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

Table 1.9: Election Results Conditional on Candidate Entry

Dependent Variable	*#. Elected*		*# Non-Inc Elected*	
	Females	Females	Females	Females
Discontinuity at 0	0.042	-0.027	-0.036	-0.053
	(0.068)	(0.096)	(0.060)	(0.066)
No. Female Non-Inc Running	0.224***	0.222***	0.265***	0.333***
	(0.017)	(0.020)	(0.017)	(0.016)
No. Male Non-Inc Running	-0.130***		-0.096***	
	(0.014)		(0.012)	
No. Female Inc Running	0.428***		-0.311***	
	(0.030)		(0.026)	
No. Male Inc Running	-0.305***		-0.279***	
	(0.028)		(0.026)	
Polynomial Fit	Quadratic	Quadratic	Quadratic	Quadratic
Controls	Yes	Yes	Yes	Yes
No. Observations	997	997	997	997

Standard errors clustered by district are shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance. Table shows the effects of past election outcomes, conditioning on who entered the race. All specifications include a full set of controls, including number of open seats, pct. minority, pct. English learners, pct. free/reduced lunch, and district enrollment size.

A different approach suggested by the model is to look at the effects on individual probabilities of winning. According to the model, in equilibrium the effect of a female winning in the past on any individual candidate's probability of winning should be zero. I check this by estimating linear probability and Probit models to see if a female winning in the past affects the outcome. Results are not shown, but as expected there is no effect for any type of candidate (i.e., females, non-incumbent females, etc.). In fact, estimates are not only statistically zero but often the point estimates are very small in magnitude. This provides further evidence for Hypothesis 3.

Taking all the major findings together, it seems clear gender still matters in this environment of high-female representation. Patterns in the data point to electorates trying to maintain a certain gender composition on the board rather than having ignorance reduced through experience. Thus in some cycles males are marginally favored and in others females are. Given this environment, candidates appear to rationally enter races.

1.4.4 Differences Between Incumbents and Non-Incumbents

As already mentioned, most of the effects appear to fall on non-incumbents. Indeed, non-incumbent candidate entry seems highly sensitive to changes in gender demand while incumbents are not affected. I suggest this is because the marginal candidates are typically non-incumbents since incumbents may choose to run first. Incumbents are highly likely to win re-election and so non-incumbents will factor into their chances of winning whether an incumbent is running. Here I present evidence for this assertion.

Table 1.10 shows that incumbents are much more sensitive to whether a male or female won in the previous election than incumbents are. The reason for this may be that the "incumbency effect" dominates any gender-related calculations by the candidates. That is, incumbents who run may be highly likely to win and this overwhelms most other concerns. Table 1.10, Columns 1 to 4, directly shows that both male and female incumbents win re-election with high frequency and this effect dominates any estimated effect at the discontinuity.

Table 1.10: Comparing Incumbent vs. Non-Incumbents

Dependent Variable	*Num. Inc. Elected*			
With Controls	Female	Male	Female	Male
Num. Female Inc. Running	0.778***		0.777***	
	(0.021)		(0.021)	
Num. Male Incumbents Running		0.752***		0.751***
		(0.021)		(0.021)
Discontinuity at 0	0.036	-0.008	0.089	0.018
	(0.041)	(0.044)	(0.057)	(0.062)
Polynomial Fit	Linear	Linear	Quadratic	Quadratic
Controls	Yes	Yes	Yes	Yes
No. Observations	997	997	997	997
Without Controls				
No. Female Inc. Running	0.777***		0.777***	
	(0.021)		(0.021)	
No. Male Incumbents Running		0.756***		0.756***
		(0.021)		(0.021)
Discontinuity at 0	0.029	-0.012	0.078	0.018
	(0.041)	(0.044)	(0.057)	(0.062)
Polynomial Fit	Linear	Linear	Quadratic	Quadratic
Controls	No	No	No	No
No. of Observations	997	997	997	997

Standard errors clustered by district shown in parenthesis. (***) for 1%, (**) for 5%, and (*) for 10% significance.

1.5 CONCLUSION

The empirical results presented show that environments where those choosing leaders have accurate beliefs about a potential leader's performance are very different from those where biases or ignorance about candidates predominate. On California school boards, females represent a large share of members, and voters are likely to be quite familiar with how both male and female candidates will govern on average. The evidence points to electorates trying to systematically maintain an ideal board gender composition. Candidates' entry decisions reveal rational responses to the electorate's desires. This is in contrast to studies in low-female representation environments where added exposure to female leadership reduces ignorance and therefore has the opposite net effect I find.

This also raises some unresolved questions. First, the nature of voters' preferences for a given gender composition remain unclear. Hypothesis 3 holds that gender matters to voters, yet it does not suggest the source of such preferences. Voters may have a desired policy outcome, will make accurate gender-based inferences about how each candidate will govern, and will choose candidates to achieve the desired policy. Or, voters may have preferences over gender *per se*. For example, voters could have taste-based prejudices or affinities for a certain gender. Regardless of whether voters care about policy or gender *per se*, both motivations will generate similar patterns in my data. It is reasonable to believe both play a role. Further study will be needed to parse-out the relative importance of each piece.

There is also the issue of candidate motivations. I assumed in the model candidates are motivated by the benefits of winning. These benefits for some candidates may entail reaping "the spoils of office" or the opportunity to implement preferred policy. But at least some candidates may also care about gender *per se*. If candidates enter to increase their gender's representation, entry would follow the pattern found in the data. This pattern is of course also consistent with candidates caring about winning. As with the question of the electorate's motivations, data limitations do not allow me to answer this question.

What the main findings do, however, show is that the specific environment matters. This means policy makers must consider carefully what types of discrimination are or are not at work when changing affirmative action policies. For example, if discrimination is primarily policy-based in a high-female representation environment, then removing affirmative action quotas while increasing candidate-specific information may be a reasonable approach which allows voters to achieve their optimal policy. While understandably most affirmative action research pertains to low-representation situations, it will become increasingly important to understand how people behave once high-representation is achieved.

2.0 CAN BOARD DIVERSITY IMPROVE THE DELIVERY OF EDUCATIONAL SERVICES?

2.1 INTRODUCTION

How corporate board diversity affects firm performance is extensively discussed in the business management and organizational behavior literature. Several theories seek to explain how diversity might positively affect organizational performance while other theories posit a negative diversity effect. Empirical findings are also mixed concerning the role of board diversity on such outcomes as a firm's share price and profitability (Smith et al. 2006). Concerning board gender diversity, studies from various nations exist yet do not reach a clear consensus (see Smith et al. 2006; Carter et al. 2010; Shrader 1997; Kochan 2003; Adler 2001; DuRietz and Henrekson 2000; Rose 2004). For good reasons, this literature has primarily focused on private-sector firm performance. I however study whether governing board gender diversity matters in the performance of an important public organization, namely, school districts. I estimate how gender diversity on California school districts' governing boards affects academic performance. This provides some of the first evidence that board gender diversity can influence objective measures of performance in the public sector.

Two dominant theories propose how governing board diversity may improve organizational performance. While typically applied to corporate boards, each is useful for considering gender diversity on school boards. These theories, presented throughout the literature, are here referred to as the "resource dependent" and "agency" theories. The resource dependent theory, in the present context, posits that having a mixture of male and female board members increases the quantity of unique information and thus improves strategic decision making. That is, diversity offers new perspectives and fosters communication on topics not frequently addressed (Page 2007; Carter et al. 2010; Smith et al. 2006; Erhardt et al. 2003). For example, in a school district, female members who are mothers or teachers may understand specific issues better than the average male member.

This input may affect board policies and ultimately improve the delivery of educational services.

The agency theory says that diversity helps a board in its monitoring role by increasing its independence. A school board serves as an agent of the public and is responsible for monitoring district operations. Diversity of members may reduce the possibility the board becomes beholden to the superintendent, teachers' unions, corporations, or any segment of voters. Furthermore, the board's monitoring capabilities may be improved if certain members are able to form better relationships with certain school employees. This may increase the likelihood the board succeeds in implementing its policies.

However, some authors state that homogeneity will produce better outcomes (Hambrick et al. 1996; Smith et al. 2006). For example, it is possible a board could be torn by infighting among diverse members and consequently fail to implement first-best policies. Or, factions on the board might create poor relationships within the schools and hamper the delivery of quality education. School boards with reputations for infighting may also be unable to hire competent superintendents and teachers. Given that *a priori* this theory is as plausible as those implying positive effects on educational outcomes, I empirically estimate whether increasing gender diversity on a school board helps to improve academic performance. Specifically, the main results show the impact on key academic outcomes of adding an additional female to the school board.

When estimating whether board gender diversity *causally* influences an organization's performance, potential unobserved variables and endogeneity are a chief concern. This issue is well documented in the corporate governance literature, and some studies seek a cleaver solution (Smith et al. 2006; Carter et al. 2010). With school boards, this also is a major concern. District-specific unobserved variables which are simultaneously correlated with a district's propensity to choose female board members and a district's academic performance will bias estimates. Thus, finding a correlation between gender diversity and academic performance will offer little insight into whether gender diversity actual affects academic outcomes.

To address this causality problem, I employ a regression discontinuity (RD) design. By comparing situations where a female barely defeated a male, and vice versa, in school board elections, I can quasi-exogenously add either a female or male to the board. This exogenous change in the number of female board members allows for causal estimates of the effect on academic performance. The intuition for the RD approach is that the outcome of very close elections between a male and female candidate are "as good as random" because in such cases idiosyncratic factors influence a small number of decisive votes. Such an approach has been used by DiNardo and Lee (2004), Lee

(2008), Rehavi (2007), and Ferreira and Gyourko (2009).

Using the RD design, I estimate the effect of increasing the number of female board members on a district's academic performance. I consider the effect on whether a district achieves Adequate Yearly Progress (AYP) within the next three years and on the total change in the Academic Performance Index (API) over three years.[1] For both measures of performance, I find evidence that increasing the female board composition increases a district's academic performance on average for districts that are currently failing to make AYP. While the data also suggests performance increases over a shorter time frame, these estimates are of smaller magnitude and less significant. But, it appears performance steadily improves each year which is reasonable given that it may take several years for board policies to be implemented. I find no diversity effect for districts currently achieving AYP. Together, this suggests that times of crises (i.e. a district failing in its education mission) may be when diversity is most valuable.

Having shown an average net benefit of gender diversity on school boards, I would like to know whether the data supports either the resource dependent or agency theory. Unfortunately, I cannot conclusively distinguish the two. However, I can observe one outcome related to the agency theory. Using the same RD approach I find an increase in the number of female board members decreases the probability the district superintendent separates from the position in the districts that were failing AYP. Unfortunately, the data does not reveal the reason for the separation (i.e. "firing" or voluntary). However, this does point to gender differences in a board's monitoring capacity. As an agent of the public, the board must manage school personnel, most notably the superintendent. Apparently having more females on the board reduces the likelihood either party chooses to end the superintendent's tenure. This may partially explain the performance improvements if retaining a superintendent ultimately is better than hiring a new executive and undergoing the costs of such a transition. However, this effect on superintendent separations could also be a proxy for the general quality of board-staff relationships. If there is an acrimonious air about the board, not only is the superintendent likely to leave but the board will have difficulty working well with other staff to implement beneficial reforms. While I cannot identify exactly how the variation in superintendent separations relates to academic outcomes, it does appear female and male board members differ in how they choose to exercise their monitoring role. This comports with other evidence finding male and female officeholders tend to make different decisions (Thomas 1991;

[1]By looking at performance measures over several years I mitigate the risk of simply finding spurious correlation in a single year.

Berkman and O'Connor 1993; Vega and Firestone 1995; Thomas and Welch 2001; Rehavi 2007; Clots-Figueras 2006; Chattopadhyay and Duflo 2004).

Section 2 discusses California school districts, boards, and superintendents. I also explain the data construction and provide some basic summary statistics. In Section 3 I describe the identification strategy. Section 4 presents the main results, and Section 5 concludes.

2.2 CALIFORNIA SCHOOL DISTRICTS, BOARDS, AND SUPERINTENDENTS

2.2.1 Organizational Features

An elected board governs each California school district and within State and Federal guidelines has substantial latitude in how it provides educational services. Typically a board has five members, though some have more. Seats are usually "at-large", meaning members do not represent a specific area of the district but instead are chosen by all voters in the jurisdiction.[2] Board members are elected to four-year terms in elections staggered every two years. This way, some members of the board are always in the first half of their term and some are in the second half. There are no term limits, primaries, or political parties. Many if not most board members are simply parents, teachers, or community members. In "at large" elections all candidates appear on the same ballot. Voters can choose up to N candidates, where N is the number of open board seats. When candidates are ranked by their vote tally, the top N vote-getters are elected to the board. In a five-member district, three seats will be filled in election year t and two seats will be filled in election year $t+2$.

As elected officials, board members are accountable to voters for the district's academic performance as well as the district's finances, personnel, facilities, curriculum, disciplinary policies, etc. Some boards choose to be heavily involved in management issues while others will delegate much of this work. In practice, school boards hire an executive, or superintendent, to run the day-to-day operations of the district. This is typically one of the most crucial decisions a board will make. The relationship between the board and superintendent can be a key factor in how smoothly district operations run.

As an agent elected by the public, the board expects the superintendent to achieve a certain level of performance. While many performance measures are intangible, a key performance indicator is

[2] Typically the large urban districts have larger boards and geographic member districts. These represent a small number of all districts and are not included in my sample.

academic achievement. In California during the sample period, the top-level measure of a district's performance was whether it achieved Adequate Yearly Progress (AYP).[3] As the chief executive, the superintendent is likely to be judged in part by this performance measure.

Usually each summer, the board decides to renew or not renew the current superintendent's contract. However, superintendents also may choose to retire or leave for other positions (usually at a different district). Consequently the decision to separate is jointly determined. Some superintendents may leave for idiosyncratic reasons mostly unrelated to the current position (e.g. spouse takes a job in a different area). Many will seek better positions. Some will be unceremoniously "fired." Still others will be "asked to resign" or choose to resign due to the employment relationship no longer being desired by one or both parties.

2.2.2 Data

The analysis relies on several data sources. Academic performance data in each district for the sample years is provided by the California Department of Education (CDE). This data reports on the district-level whether AYP is achieved in a given year as well as a district's Academic Performance Index (API) score. This allows me to not only see if adequate performance was achieved in a given year but how much the key performance measure changed over time.

I also use California school district governing board elections from 1999 to 2007 provided in the California Elections Data Archive (CEDA) to determine who won election and therefore served on the board. CEDA includes election dates, districts, candidate names, ballot designations, incumbency information, and vote tallies for most contested school district elections in the state. This election data provides information essential to my identification strategy. As necessary and possible, I supplemented the CEDA with results published by county election offices. Using this data, I first classified each school board candidate by gender. To do this, I assigned male or female values to each person based on a suffix (e.g. Jr., Sr.) or a first name that is gender-specific with high probability. Some candidate names were also classified based on a gender-specific ballot designation, such as "Businesswoman" or "Housewife." This procedure classified about 90 percent of individuals. Next, I searched the internet for information that would identify the gender of those with ambiguous names. For instance, I could often determine a person's gender from news article quotes, pictures, or references in board meeting minutes. This procedure allowed me to classify the

[3] AYP is a measure that of academic achievement given to each school district. A formula incorporating past and current academic performance is used to determine if a district, school, or the state as a whole achieves AYP.

gender of about 98 percent of persons from the original data set.[4]

District superintendent information was collected from annual editions of the *California School Directory* published by the California Department of Education (CDE). By observing consecutive years, I constructed a panel of information about whether a superintendent was hired or separated in a given year. As with the board members, I then classified each superintendent by his or her gender.

Each annual, district observation contains AYP and API information, the board's current gender composition, the most recent election information, information on the superintendent, and district controls. District control variables include district size, revenue per average daily attendance, and the percent of students who are English learners, free and reduced price lunch recipients, and minorities.

Table 2.1 provides information about the sample. About 69% of superintendents are male and about 47% of board members are female. Over the sample period the board gender composition has remained steady, as has the percentage of candidates who are female. It should be noted that unlike most other elected bodies in the United States (and worldwide), California school boards have an extremely high proportion of females - nearly reaching parity. Table 2.1 also shows that in the sample 54.9% of districts failed to make AYP. While this may appear especially high, it is consistent with passing rates published on the CDE web site. The table also shows there is little difference in superintendent gender percentages between districts making and failing AYP.

2.3 IDENTIFICATION STRATEGY

The key challenge in estimating the effect of board member characteristics (i.e. gender) on district performance is that many district-specific, unobserved variables may be correlated with both the independent and dependent variables.[5] A district prone to elect a high number of females may

[4] I was especially conservative in assuming the gender associated with a name. For example, the names "Jean" and "Terry" were always verified. Furthermore, because most people in my sample have some public "paper trail" due to their involvement in politics or school administration, my method is probably superior to matching names based on a probability distribution of names. Except for names with very high gender associations, I individually verified the genders.

[5] This is just as much a problem with data in a corporate situation. Factors that make a firm likely to have more female board members are also likely to influence other corporate decisions that could affect profitability.

Table 2.1: Summary Statistics

Key Variables	**Mean**
% of Districts with Male Superintendent	68.76
	(46.4)
% of Board that is Female	47.04
	(21.6)
% if Districts Failing AYP	54.91
	(49.8)
% of Failing with Male Superintendent	69.67
	(46.0)
% of Passing with Male Superintendents	67.63
	(46.8)
API for Passing Districts	776.1
	(90.4)
API for Failing Districts	708.4
	(71.1)
Controls	
District Enrollment (1000's)	9.619
	(10.51)
% English Language Learner	22.31
	(17.83)
% Free and Reduced Lunch	45.07
	(26.56)
% Minority	58.78
	(27.04)
% Revenue per ADA ($)	8056
	(2747)

also be in a higher income area which likely also affects academic performance. Such a situation would generate correlation between board gender composition and academic performance even if the board members' genders did not causally influence academic performance.

To determine whether board gender composition causally affects academic performance, I need a plausibly exogenous source of variation in the number of males and females sitting on the school board. To do this I use a regression discontinuity (RD) design. Essentially, I compare school board elections where a female closely defeats a male and where a female closely loses to a male in the most recent election. The winner will of course serve on the school board.

In California, board members are elected to four-year, staggered terms on the first Tuesday of November[6]. For instance, in year t three members are elected and in year $t + 2$ two members are elected.[7] In year $t + 4$ the cohort of seats filled in year t are again up for election. After the November election, board members take office in December. These members will participate in district decisions for four years.

I use quasi-random outcomes between male and female candidates in the board elections to generate variation in the gender composition of the board members. The intuition is that very close elections between the lowest winning female (male) and highest losing male (female) is "as good as random." On election day numerous idiosyncratic factors can cause a very close election to go one way or the other. This randomness allows me to estimate the causal effect of having an additional male or female on the board.

To implement the RD approach, I construct a "forcing variable" for each election. The forcing variable is the vote difference between the female and male who were closest in vote percentages to each other but on opposite sides of the winning threshold. If three seats are open, the winning threshold is between the third and fourth highest candidates when ranked from most to least votes. For the closest male/female pair across the winning threshold, I subtract from the female's vote percentage the vote percentage of the male. Consequently, the forcing variable is positive if the female won and negative if the male won. A forcing variable of zero would indicate the male and female tied. The distribution of the forcing variable used in my analysis is shown in Figure 2.1. Notice that most observations are around the zero point and there are relatively few observations beyond a difference of ten. Since the estimate will consider what happens when we switch from just below to just above where the forcing variable equals zero, it is important to have a large number

[6] Some districts have adopted more complicated arrangements and hold elections at other times of the year. These districts are not included in the sample.

[7] Most districts have five members, but some have seven.

of observations here.

As one moves across the zero point, a small change in the vote tally creates a sharp, binary change in whether a male or female wins and serves on the board. The intuition for this identification strategy is that very close elections are essentially random in whether a male or female wins but the consequences are dramatic. This approach of using close elections to generate random variation is employed by DiNardo and Lee (2004), Lee (2008), Rehavi (2007), and Ferreira and Gyourko (2009).

To produce estimates and standard errors, I use the following regression:

$$Y_{it} = \beta M_{jt-1} + T(M) + X_{jt}^{'}\gamma + \epsilon_{jt} \tag{2.1}$$

where Y_{it} is the outcome of interest, M_{jt-1} indicates if the female won the most recent close election, $T(M)$ represents flexible polynomials fit separately to the forcing variable on each side of the discontinuity, and X_{jt} is a set of district controls and an intercept. The coefficient β is the main value of interest. The estimate of β is the average effect on the outcome of interest of a woman relative to a man sitting on the board . In estimating β, I consider several bandwidths. The bandwidth indicates how much of the data on each side of the discontinuity is included in the sample. Since there are very few observations in the extremes, these points can vary wildly and may have undue influence on the estimates. Consequently I show estimates using a range of bandwidths which serve to demonstrate the overall robustness of the results.

Before presenting my findings, it is important to check the key identifying assumption of the RD approach. Since RD assumes observations near the discontinuity are randomly assigned to either side, there should be no correlation between the discontinuity and other variables. Table 2.2 shows that for all the control variables this holds (Col. 3- 6). Furthermore, the table shows that the assumed exogenous change in board gender composition has no effect on whether the district currently makes AYP, which might also indicate selection problems (Col. 1-2).

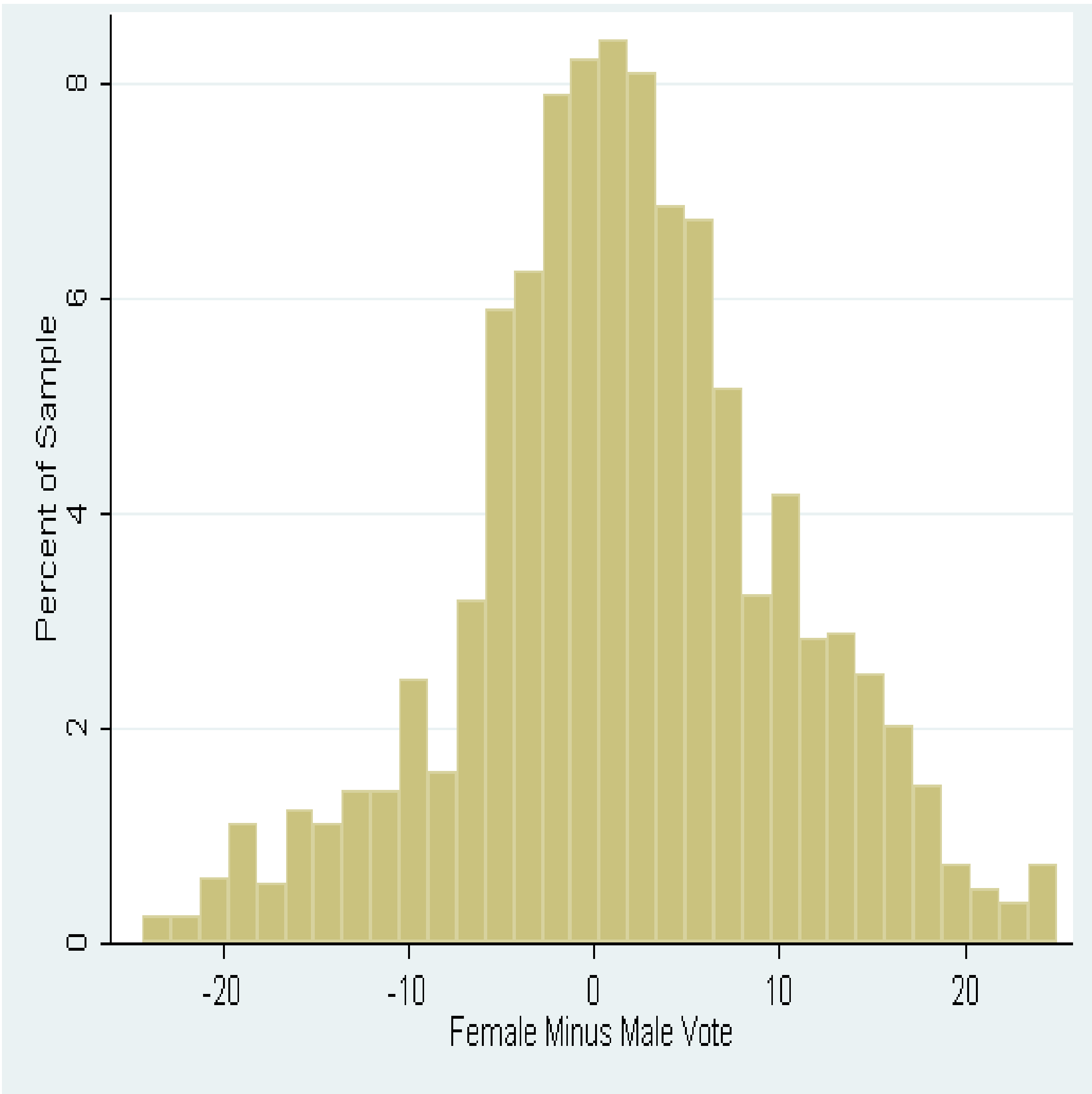

Figure 2.1: Distribution of Forcing Variable

The forcing variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. At 0, the forcing variable is interpreted as a male and female candidate tying. Positive (negative) values indicate a female (male) defeated a male (female).

Table 2.2: Checking Regression Discontinuity Assumptions

Outcome	Made AYP	Made AYP	Enrollment	Pct. Eng. Learner	Pct. FRP Lunch	Pct. Minority
	(1)	(2)	(3)	(4)	(5)	(6)
Discontinuity at 0	0.034	0.031	0.296	-1.282	0.135	-2.551
	(0.050)	(0.042)	(1.077)	(1.957)	(2.968)	(3.062)
Trend Fit	Linear	Linear	Linear	Linear	Linear	Linear
Controls	No	Yes	N/A	N/A	N/A	N/A
Observations	1104	1104	1104	1104	1104	1104

Significance: (*): 10%, (**): 5%, (***): 1%. Each column presents results checking the RD assumptions. Columns 1 and 2 estimate the effect of a female board member on the district making AYP. Columns (3) to (6) show the treatment variable is uncorrelated with the controls. All specifications use a bandwidth encompassing the entire sample.

2.4 RESULTS

2.4.1 Academic Performance

Estimates show that the gender composition of the school board can have a substantial impact on some districts' medium-run academic performance. I find that on average for districts currently failing to make AYP, an additional female on the school board helps to improve performance over the next three years. For districts already making AYP, I find no significant impacts on performance. Specifically, I measure the following outcomes for failing districts: whether the district achieves AYP within the next three years and the total API change over the next three years.

Table 2.3 presents estimates of the effect increasing the number of females on the school board has on whether the district makes AYP within the next three years. Note the sample only includes districts that were failing AYP. Point estimates show that an additional female on the school board increases by about 12 to 25% the likelihood a district makes AYP in the next three years. This table shows a range of bandwidth and estimates with and without controls. Furthermore, I also vary the trend type fit through the forcing variable. While the point estimates are all positive, the magnitudes and significance levels vary. Since the outcome is discrete, it is a blunt measure and will not pick up improvements that do not change the overall AYP standing. This will make it more difficult to find an effect. However, the estimates do suggest an additional female board member has a positive average marginal effect on a district's attempt to return to good standing. In Figure 2.2 the estimated effect is illustrated. The horizontal axis, plots the forcing variable. The vertical axis plots whether the district achieved AYP within three years. Each point plots a local average of the observations, and I fit a linear trend through the data. Moving from left to right of zero on the x-axis means switching from a male to female serving on the board. A slight upward jump in the likelihood the district achieves AYP within the next three years may be discernible, but it is not immediately apparent.

Though AYP is of great importance, some districts may improve their API score but not actually make AYP. For this reason, I consider a continuous rather than discrete measure of performance. In Tables 2.4, I present estimates showing the effect of an additional female serving on the school board on the total API change over three years. Not only will this allow me to observe API changes not resulting in a discrete change in the AYP state, but considering three years will reduce the

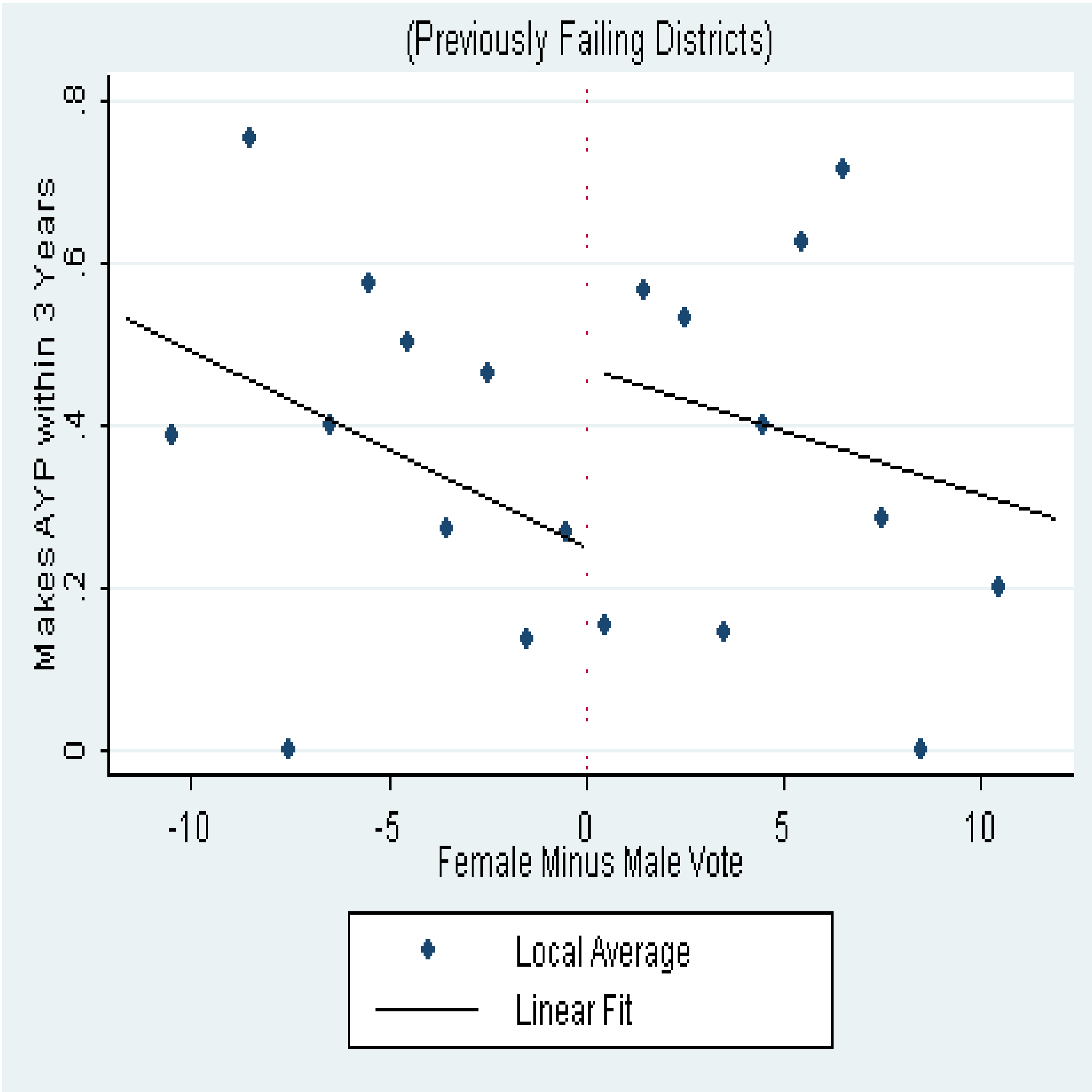

Figure 2.2: Makes AYP within 3 Years

Graph plots the local averages for the vertical axis variable against the forcing variable. The forcing variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A linear trend fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

Table 2.3: Effect of Additional Female on Making AYP within 3 Years in Failing Districts

	Makes AYP within 3 Years							
	[-25,25]		[-15,15]		[-10,10]		[-5,5]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Discontinuity at 0	0.212	0.240	0.122	0.134	0.221*	0.239*	0.166	0.199
	(0.139)	(0.125)	(0.117)	(0.105)	(0.132)	(0.125)	(0.191)	(0.168)
Trend Fit	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	228	228	201	201	170	170	121	121

Significance: (*): 10%, (**): 5%, (***): 1%. Each column presents results from a linear probability model regressing whether or not the district returned to making AYP within three years (1 = made AYP) on an indicator for being right of the discontinuity representing a female win, a trend line, and in some specifications district controls. Specifications adjust the bandwidth and trend fit. Coefficients are interpreted as the increase/decrease in the probability of making AYP within three years due to an additional female sitting on the board in districts that are currently failing to make AYP. Standard errors are clustered by district.

possibility I pick up a statistical anomaly. Table 2.4 shows that an extra female on the board increases the total API gain over three years by about 7 to 17 API points. This represents an increase of about one-third to three-fourths of a standard deviation in the average API change in those three years. Results are significant and robust to changes in bandwidth and controls. This effect is shown graphically in Figure 2.3. This continuous measure of performance is strong evidence of a positive, average impact of female members on performance.

Overall, the two estimates agree qualitatively. Having an additional female on the school board increases academic performance over the next three years. While some effects appear to exist in the shorter run (one or two years), the magnitude tends to be smaller and estimates are not as significant. For example, the 1-year and 2-year point estimates for changes in total API scores are about 3.2 (p-value: 0.20) and 6.2 (p-value:0.21), respectively. This suggests that while an impact may exist in earlier years, it takes a while for change to occur. The larger effect occurs after several years of the female serving on the board. This is plausible since board policies and personnel changes may not take effect immediately.

2.4.2 Superintendent Separations

As already mentioned, several theories try to explain why board diversity may yield benefits. Given the available data, it is not possible to distinguish the two here. But, estimating whether board gender composition affects the likelihood of a superintendent separation may provide insight concerning the boards' monitoring roles. Boards have responsibility for how district employees provide educational services. In practice this often means holding the district superintendent responsible for academic performance. With the data available, I can observe superintendent separations. This may indicate whether the gender composition of the board influences how the board executes its monitoring role as an agent of the public.

I use a similar RD approach but with the outcome being whether a superintendent separation occurred. While I do not know if the separation was voluntary or not, any significant estimates would suggest board gender composition influences how a district monitors the superintendent's performance. Furthermore, an impact on the superintendent could be a proxy for how well a board is able to work with all district staff to implement policies.

As shown in Table 2.5, an additional female on the school board decreases the likelihood a

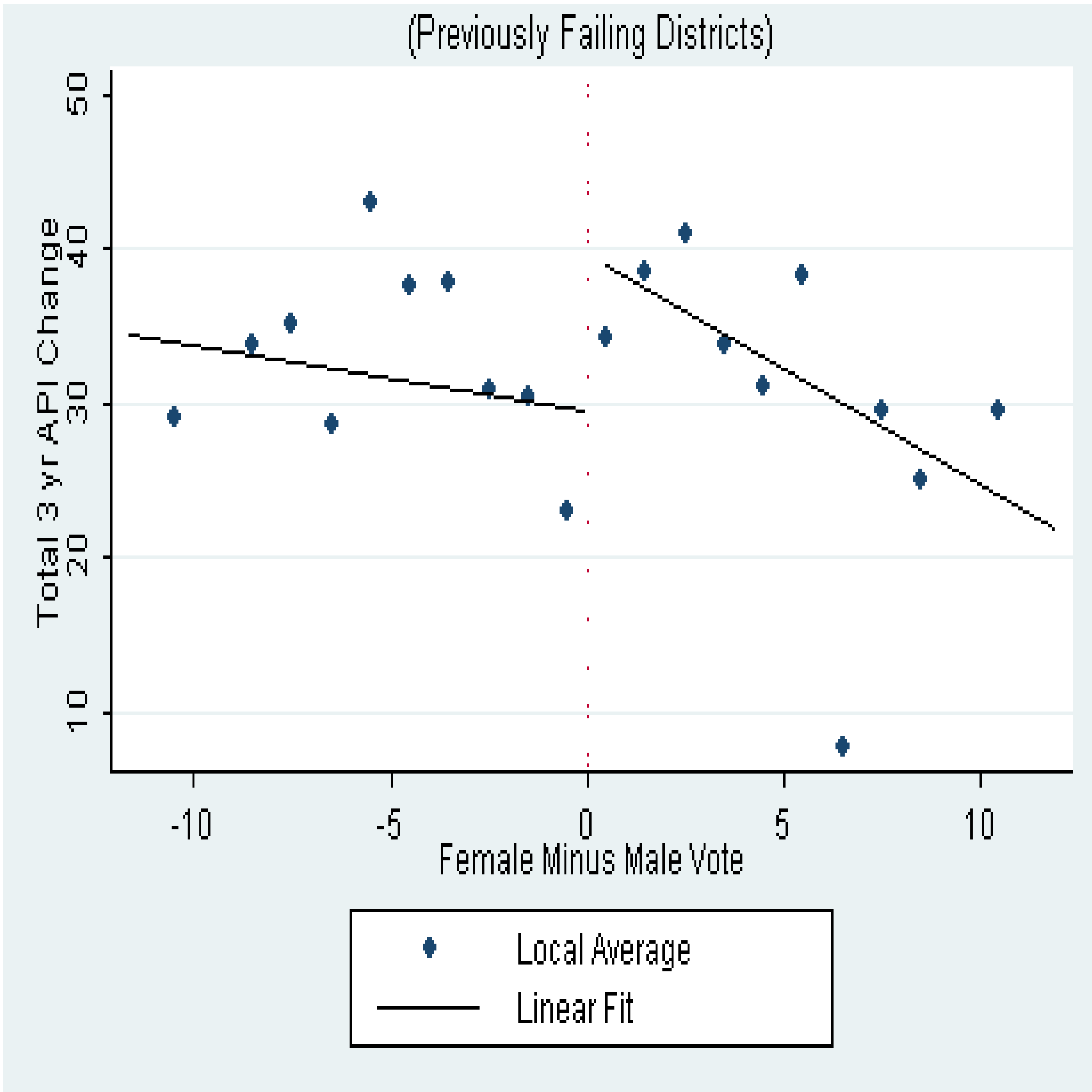

Figure 2.3: Total API Change in 3 Years

Graph plots the local averages for the vertical axis variable against the forcing variable. The forcing variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A linear trend fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

Table 2.4: Effect of Additional Female on Total 3-Year AYP Change in Failing Districts

	Total 3-Year AYP Change							
	[-25,25]		[-15,15]		[-10,10]		[-5,5]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Discontinuity at 0	14.78**	11.67**	8.28*	7.00	14.14***	9.96*	17.32**	15.37**
	(5.81)	(5.85)	(4.73)	(4.69)	(5.37)	(5.39)	(6.84)	(6.93)
Trend Fit	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	228	228	201	201	170	170	121	121

Significance: (*): 10%, (**): 5%, (***): 1%. Each column presents results from regressing the total 3-year API change on an indicator for being right of the discontinuity representing a female win, a trend line, and in some specifications district controls. Specifications adjust the bandwidth and trend fit. Coefficients are interpreted as the increase/decrease in the total API change over three years due to an additional female sitting on the board. Standard errors are clustered by district.

superintendent separation occurs. An additional female reduces the probability the superintendent separates by about 10 percent. Figure 2.4 illustrates this effect. While I cannot clearly connect positive academic performance with the lower rate of separation, this does show gender can influence how a board monitors. This is in line with Adams and Ferreira (2009) who also find gender differences in corporate monitoring. More generally it connects with the recent literature showing group gender composition can influence personnel decisions (Bagues and Esteve-volart 2009).

2.5 CONCLUSION

How board gender diversity affects organizational performance remains an open question in the corporate governance literature. However, much less attention has been given to how diversity can influence public sector performance. I use objective measures of school performance and find that increasing the number of females on a school board on average helps failing schools perform better. Several mechanisms have been proposed to explain how diversity might improve organizational performance, referred to as the resource dependent and agency theories. The agency theory says that diversity helps a board better monitor organizational personnel. I find that in fact the gender composition of the school board does influence whether the district superintendent separates from his or her position. Though the connection with academic performance is unclear, it does appear that board gender composition influences *how* a board functions as an agent. Exploring the relative importance of the resource dependent and agency theories in explaining organizational performance improvements is therefore an important direction for future research.

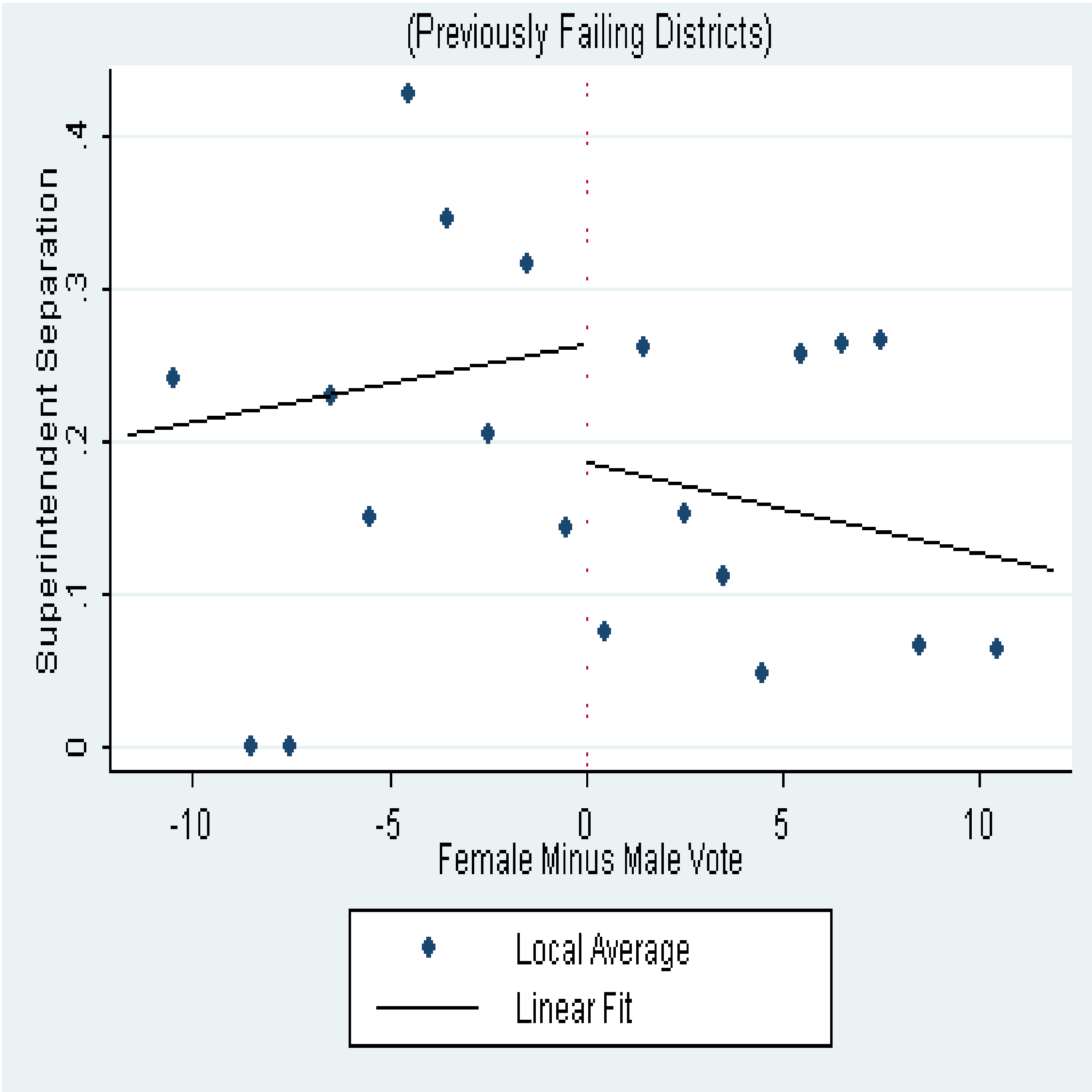

Figure 2.4: Superintendent Separations

Graph plots the local averages for the vertical axis variable against the forcing variable. The forcing variable is the female minus male vote percentage for the closest candidates of each gender on each side of the winning threshold in the previous election. A linear trend fit to the data illustrates any discontinuity around extremely close races. A move from left to right across the zero represents a switch from a male to female winning a close election. Thus, the discontinuity represents the effect on the vertical axis variable of this change.

Table 2.5: Effect of Additional Female on Separations in Failing Districts

	Separations in Districts Failing AYP							
	[-25,25]		[-15,15]		[-10,10]		[-5,5]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Discontinuity at 0	-0.089*	-0.091*	-0.100*	-0.100*	-0.080**	-0.082**	-0.124***	-0.124***
	(0.050)	(0.050)	(0.056)	(0.057)	(0.038)	(0.038)	(0.045)	(0.046)
Trend Fit	Quadratic	Quadratic	Linear	Linear	Linear	Linear	Linear	Linear
Controls	No	Yes	No	Yes	No	Yes	No	Yes
Observations	587	587	528	528	451	451	308	308

Significance: (*): 10%, (**): 5%, (***): 1%. Each column presents results from a linear probability model regressing whether or not a separation occurred (1 = separated) on an indicator for being right of the discontinuity representing a female win, a trend line, and in some specifications district controls. Specifications adjust the bandwidth and trend fit. Coefficients are interpreted as the increase/decrease in the probability a separation occurs due to an additional female sitting on the board. Standard errors are clustered by district.

3.0 GENDER AND GROUP DECISION MAKING: EXPERIMENTALLY CONNECTING INDIVIDUAL AND GROUP BELIEFS

3.1 INTRODUCTION

Economic decisions are frequently made by groups rather than individuals. Decisions to hire and promote in the labor market are good examples. Corporate boards hire new CEOs, school boards select district superintendents, and recruiting committees choose new faculty members. Even for lower-ranking positions, the decision to hire an employee is often made with input from multiple individuals. When considering such collective decisions, a natural question is how characteristics of individual decision makers might influence a group's collective choice. Addressing this question, recent work has considered how the gender composition of a hiring committee influences its evaluation of job candidates (Bagues & Esteve-Volart 2010). This research, carefully exploiting observational data and a natural experiment, provides interesting insights but lacks the ability to show *how* gender interactions within the committees shape the final decision. That is, the data reports a groups' collective decision but lacks information on individual group members' beliefs. This means it is impossible to map individual members' beliefs into the collective decision or see how members' genders (or any other characteristics) interact in the deliberation process.

Using an experimental design, I seek to avoid this problem while studying how a group's gender composition affects its evaluation of "job candidates." In the experiment, the "job candidates" presented to subjects are male/female pairs that have previously competed in a skill task. I first elicit each subject's personal beliefs by having them evaluate male/female candidate pairs after viewing the candidates' pictures side-by-side. To do this subjects must report in percentage terms how likely they believe each job candidate is to have performed best in the previously held skill competition. Then, I randomly assign subjects to groups of varying gender compositions. These groups must collectively evaluate the same side-by-side candidate pairs. In this way, I can study how individuals' beliefs and gender influence the groups' collective decisions. In particular, I can see whether a "group shift" occurs. A group shift occurs when the group deviates significantly from the outcome that would occur if all members' individual beliefs were aggregated through simple majority voting without discussion (Ambrus, Greiner, & Pathak 2009). A group shift indicates the outcome cannot be simply explained by individual pre-deliberative beliefs. If group shifts in fact occur, I can then try to understand whether within-group gender interactions might account for such shifts.[1] By studying the decision in this way I can provide new insights into how mixed-gender groups interact to evaluate job candidates.

My findings fit well with the most relevant empirical field study of this issue. Bagues and Esteve-Volart (2010) consider whether the gender composition of randomly-assigned, Spanish judicial candidate evaluation committees affects the committees' decisions regarding male and female job candidates. As far as I am aware, this is the only study to use a large sample and credible identification strategy to isolate the effect in question.[2] Overall, Bagues & Esteve-Volart (2010) find a positive relationship between more female committee members and better evaluation of male candidates. However, a careful look at their results shows this pro-male effect occurs only on the committees when females are a small (but non-zero) minority and when females are the majority.[3] Furthermore, it is important to realize this effect is identified off of decisions for marginal candidates, presumably those where members were not in substantial agreement. This is important because it will inform how my experimental results should be appropriately compared to results from the field. In fact, I find a remarkably similar pattern as Bagues & Esteve-Volart (2010) in situations where group members' individual beliefs indicate strong pre-deliberation disagreement within a group. In these cases, five-member groups with one female or with four females both favor male candidates on average, while two- and three-female groups do not. This is qualitatively similar to the pattern found in the field.

Several mechanisms might explain why more female decision-makers in a group could lead to better evaluations for male candidates.[4] The obvious possibility is that females individually evaluate

[1] Ambrus, Greiner & Pathak (2009) present a survey of past work on group shifts in general and discuss the leading theories from social psychology to explain such shifts.

[2] Several papers do address the relationship between hiring committee gender composition and choice of job candidate (Ehrenberg, Jakubson, Martin, Main, & Eisenberg 2009; Kurtulus & Tomaskovic-Devey 2009). These studies though fail to answer the question at hand primarily because an organization's gender composition can also influence the supply of job applicants of each gender. In Bagues & Esteve-Volart (2010), the candidate pool is independent of the committee gender compositions.

[3] For instance, Table 3.5, Column 4, of Bagues & Esteve-Volart (2010) shows that having females compose between 10% and 30% of the committee significantly increases the male success rate. However, the effect is statistically zero for committees with 30% to 50% females. Then, for majority female committees the positive effect returns.

[4] Caution must be made in accepting this statement without nuance. As discussed above, Bagues and Esteve-

male candidates relatively more favorably (or female candidates less favorably) and thus more females in a group will produce a more pro-male outcome. Such individual-level gender differences may exist in the labor market because females (males) over-evaluate male (female) candidates, because females (males) have a stronger affinity for working with males (females), because male (female) candidates interview better with females (males), because male (female) members are more aggressive toward male (female) interviewees, or because evaluators consciously seek to counteract same-gender bias. Using information on individual beliefs, I can test whether individual beliefs explain the group decisions.[5] I do this by testing for a "group shift" - calculated as the difference between the group decision and the median member's decision. I find that 4-female/1-male groups and 1-female/4-male groups both significantly shift in favor of the male for pairs where there is substantial pre-group disagreement (i.e. the type of cases Bagues & Esteve-Volart (2010) identify off of). As indicated by a "group shift" in these cases, simple aggregations of individual members' beliefs does not fully account for the groups' choices. That is, something about gender *interactions* must contribute to the group outcome in some cases.

Bagues and Esteve-Volart (2010) suggest one such explanation which relies on group members' interactions. They posit that the presence of women in the committees affects the choices of male committee members such that male members increasingly favor male candidates. However, another explanation not explored by Bagues & Esteve-Volart (2010) is that a given member's influence on the decision depends on his or her gender and perhaps the gender composition of other members. For example, a male may be more likely than a female to challenge the committee if he disagrees with the majority opinion. His willingness to do so might even depend on how many other males and females are in the group. In the present study, I offer evidence that gender may explain a member's differential influence in the collective decision and result in the observed group shifts.

I find that in groups which shift significantly, where there is substantial pre-group disagreement among members, that males who disagree with the majority are much more successful than similarly situated females in moving the group toward their belief. In a four-female, one-male group the pre-group beliefs (according to the individual elicitation) will favor the male on average. When the male is the strongest supporter of the female candidate, this male is significantly more likely (relative

Volart (2010) find a pro-male effect for majority female groups *and* for small minority female groups (less than 30%, but not zero). Bagues & Esteve-Volart (2010) highlight this non-linearity and suggest their results are consistent with two hypotheses: (i) women in committees favor male candidates; (ii) men in committees favor male candidates when sitting in mixed-gender committees.

[5]Of course, I cannot say as much about why a preference exists. Note however that some of the possibilities (e.g. preference for working together, performance in interviews) are explicitly ruled out by my experiment since evaluators and candidates never interact.

to a female member) to move the group toward the female candidate. In one-female, four-male groups the pre-group beliefs favor the female. However, a male who strongly dissents and is most favorable toward the male candidate, is much more capable than a similarly situated female in moving the group's decision toward his belief. So, in both situations, a male member who *a priori* strongly dissents with most group members influences the outcome much more than a female in such a position.

This evidence paints an interesting picture. When part of the minority opinion, males manage to "make their opinion heard" better than females. If females perhaps do not challenge the majority opinion by staying silent, the group will only incorporate a fraction of the members' beliefs into the decision which effectively "pushes" the collective decision further toward the majority opinion. If males do express their dissent from the majority opinion, they are able to "pull" the collective decision toward their beliefs. This process - one of differential influence in the collective decision - provides a viable explanation for why groups with more extreme gender compositions are observed to shift toward the male candidate.

3.2 EXPERIMENTAL DESIGN AND PROCEDURES

The experiment consisted of two types of sessions: a "contender" session and a "decider" session. The contender session, conducted first, generated the data on "job candidates" which were then used in the decider sessions. Subjects in the decider session were given information about subjects from the contender session for purposes of evaluating how likely it was each contender performed best in the skill task. For this reason, it was important the deciders were unfamiliar with the contenders, outside of the experiment so I conducted the contender session at the University of Pittsburgh - Johnstown and the decider session at the University of Pittsburgh - Main Campus. While in the same system, these campuses operate separately and there is no evidence deciders recognized any subjects in the contender session. Contender session subjects were recruited from introductory economics classes near the experiment site. Decider session subjects were recruited through the Pittsburgh Experimental Economics Lab (PEEL) recruiting system and sessions were conducted on-campus in a space that allowed for the isolated meeting of multiple groups. Sessions were conducted in the Spring and Summer of 2010. Two contender sessions and three decider sessions were conducted. Instructions used in the sessions are provided in the Appendix.

3.2.1 Contender Session

After subjects (henceforth, "contenders") signed the consent form, an experimental assistant took a photo of each from the shoulders up. This occurred prior to explaining the experiment procedure. Contenders were asked to remove any sunglasses for the pictures. Contenders were told the pictures and information about their performance may be shown in other experiment sessions at a different school.

The experimenter then explained that subjects would be paid $10 for participating in the experiment but had the opportunity to earn more through their performance on the task. The session consisted of two stages. At the end of the session, one stage was randomly selected for payment. In Stage 1, contenders had five minutes to sum sets of five two-digit numbers (e.g. 13+15+67+90+45) without a calculator. All contenders had the same sets of problems and were provided with pencils and a sheet with sets to sum and space to answer. If Stage 1 was selected for payment, each correct sum earned $0.50 and incorrect answers did not decrease payoffs.

In the Stage 2, contenders were randomly matched with someone else in the room. Contenders were told they would compete in a one-on-one competition with whomever they were matched with, though no one was told who they were competing against (even after the session). Using a similar adding task as in the previous stage, if Stage 2 was selected for payment, contenders earned $1 for each correct sum *if they won the head-to-head competition.* The loser earned nothing. In the event of a tie, each contender was paid for half of their correct sums. After completion of Stage 2, selection of the stage to be paid, grading of answers, and determination of winners and payoffs, contenders were paid and excused. The contender sessions lasted approximately thirty minutes each.

3.2.2 Decider Session

Prior to the decider sessions, the contender session photos and Stage 1 (i.e. non-competitive piece rate) performance data were used to make "contender pairs." Including Caucasian contenders only, subjects were separated based on their quartile of performance in Stage 1, then randomly paired with a contender of the opposite gender. The photos of each contender were placed next to each other and labeled "A" or "B." The position of the contenders (A or B) was randomly assigned. In the pictures, contenders' bodies were covered so only the contenders' heads were visible. Below each pair of contender photos, a label indicated the pair's non-competitive performance quartile (First being poorest, Fourth being best). Same-gender pairs were also created so the study's investigation of gender was not obvious. Through this process, a set of 27 contender pairs (18 opposite gender) was created for display in the decider sessions.

Each decider session was designed for forty subjects, twenty of each gender. Upon entry to the decider session, deciders were given a ticket with a subject number (1 through 40). Unbeknownst to the subjects, males received odd numbers and females received even numbers. Subjects were told they would receive $7 for participating in the session but could earn more depending on their decisions in the session. Subjects were told there would be four stages. To determine payments, one decision from Stage 1 and and one decision from Stage 2, 3, or 4 would be randomly selected.

Deciders were then told about the contender sessions and how they would be presented with pictures of contender pairs that also indicated the quartile those contenders fell in during the non-competitive task. It was clear the pictures were taken before the tasks. In Stage 1 of the decider sessions, subjects were asked to individually decide which member of a contender pair scored the most points in the head-to-head competition. Reporting their beliefs about how likely each contender was to win was a two-step process. First, using a sheet provided, deciders chose either contender "A" or "B" as the highest scorer. Second, deciders reported in increments of five from 50% to 100% how likely their favored contender was to have solved the most sums. A higher percentage meant the favored contender was more likely to be the higher scorer. Reporting 50% meant the decider believed the contenders were equally likely to score highest. In Stage 1, all deciders simultaneously evaluated all 27 contender pairs as they were displayed at the front of the room.

I implemented an elicitation mechanism designed to mitigate risk-averse subjects' tendency to bias downward their reported beliefs. If a decision was selected for payment, the payment amount would be determined as follows: A number between 50 and 100 would be publicly drawn. If the random number was greater than or equal to the percentage chosen for one's favored contender, the decider earned $10 with a probability corresponding to that random number (i.e. if 60 was drawn, the decider would face a lottery with 40% chance of winning nothing and 60% chance of winning $10). In this case the payoff does not depend on which contender actually scored more points. However, if the random number drawn was strictly less than the decider's reported percentage, then the decider earned $10 if their favored contender won and $0 if their favor contender lost. If the contenders tied and the payoff depended on the outcome, the decider earned $5. Deciders were reminded that this elicitation mechanism means they should always choose a percentage for their favored contender equal to what they actually believe.

Once Stage 1 finished, deciders were sent to nearby, but visually and verbally separate, rooms listed on the ticket they were given upon entry. Stages 2, 3, and 4 took place in these rooms. For each stage subjects were reassigned to different groups as directed by the room numbers on their tickets. By design, deciders' room assignments generated for each stage eight groups of five with varying gender compositions. Group gender compositions were (1 female, 4 males), (2 females, 3 males), (3 females, 2 males), and (4 females, 1 male). In each stage, there were two groups of each gender composition. Furthermore, groups were assigned so that no two individuals were ever in the same group in different stages.[6]

Once in the group, an assistant told the group they would see some of the same contender pairs from Stage 1. This time however, they must make a collective decision concerning which contender scored highest using the same method of reporting. The assistant then one at a time displayed some of the contender pairs. In each stage, 9 of the 27 pairs were displayed. No instructions were given about how the group must decide and assistants were instructed to not participate in the discussion or point to either contender. If the group took two minutes to deliberate on a pair, the assistant prompted the group to make a decision. The assistant then recorded on paper and out of sight the decision verbally reported by the group members. Each group deliberation was video recorded and subjects were told to ignore the camera. The same procedure took place in each group in Stages 2, 3, and 4.

After Stage 4, subjects returned to the main room. Decisions were selected for payment, random numbers were drawn to determine payoffs, and subjects completed a questionnaire about themselves. Subjects were paid for their decisions and excused. Average earners were $21.50. Each session lasted approximately ninety minutes.

3.3 RESULTS

Results are presented in three parts. First, individuals' decisions over contender pairs from Stage 1 are presented. Second, group decisions are presented. In particular, I show which groups shift significantly. I discuss how these experimental results relate to the Bagues and Esteve-Volart (2010)

[6] One session did not fill completely. In this case, several groups were not full. These groups are not used in the analysis.

study which relied on a natural experiment and field data. For the appropriate comparison groups, the pattern is remarkably similar. Third, I explore how individual group members' genders and Stage 1 beliefs influence the collective decisions in Stage 2, 3, and 4. Results point to a key gender dynamic as partially explaining the observed group shifts. Also, Table 3.1 summarizes deciders' characteristics.

Table 3.1: Subject Characteristics

Decider Characteristics	
Variable	Mean
Percent Male	48.3
Percent White	71.5
Age (Years)	21.35
	(3.89)
Years of College	3.47
	(1.13)
Number of Subj.	116

3.3.1 Individuals' Decisions

Table 3.2 shows deciders' Stage 1 decisions. Decisions are displayed by gender and contender pairs' pre-competitive performance quartiles (Q1, Q2, Q3, Q4). Reported values represent the deciders' "wagers" on the male contender (ranging from 0 to 100), representing deciders' beliefs about how likely the male is to win.[7] In all, deciders evaluated six, four, five, and three contender pairs from Q1, Q2, Q3, and Q4, respectively. This data presents a striking picture about how males and females perceive the contenders. For the lowest performing quartile, Q1, on average males prefer the female contender and females prefer the male contender. That is, male and female deciders disagree on average about who will score highest. For the second quartile, Q2, male and female deciders do not disagree at conventional levels. For the third quartile, Q3, male and female deciders

[7] All decisions were reported by subjects as described in the procedure. For analysis, all decisions are reported as wagers on the male contender in the pair.

agree on average in favor of the male contender. Finally, for the highest performing quartile, Q4, both male and female deciders favor the male contenders but males more strongly prefer the males than do the females. Considering all quartiles together, the point estimates are not statistically different and in fact obscure substantial differences of opinion between male and female deciders depending on the pair quartile. Figure 3.1 plots kernel densities of individual wagers on the male for each quartile. The patterns described above are also apparent in these distributions. For Q1, the female decider distribution is shifted toward higher wagers on the male while the male deciders shift toward higher wagers on the female. Substantial agreement appears in Q2 and Q3, but male deciders' greater enthusiasm for Q4 males is also apparent.

Table 3.2: Individual Wagers on Male

Individual Wagers on Male Contender				
Quartile	Both	Males	Females	M - F Difference
1	49.3	47.2	51.7	-4.55**
	(0.8)	(1.0)	(1.2)	(1.58)
2	50.2	51.3	48.9	2.35
	(0.9)	(1.3)	(1.3)	(1.83)
3	56.7	55.4	57.7	-2.28
	(0.9)	(1.3)	(1.2)	(1.77)
4	56.2	59.2	52.7	6.51**
	(1.4)	(1.8)	(2.1)	(2.80)
All	52.8	52.4	53.2	0.87
	(0.5)	(0.7)	(0.7)	(0.96)
Number of Subj.	116	56	60	

Table shows, by pre-competitive performance quartile and decider gender, the average individual (Stage 1) wager on the male contender. The wager on the male contender is the deciders' reported belief about how likely it is the male outscored the female in the head-to-head skill competition. The far right column estimates the difference in means between male and female deciders within a pair quartile. Standard errors clustered by individual are in parenthesis. Statistically significant differences are indicated: (*): 10%, (**): 5%, (***): 1%.

In general, all deciders prefer males more often in higher quartiles. This presumably is due to perceptions about men and women's math skill and competitiveness (Niederle & Vesterlund 2007;

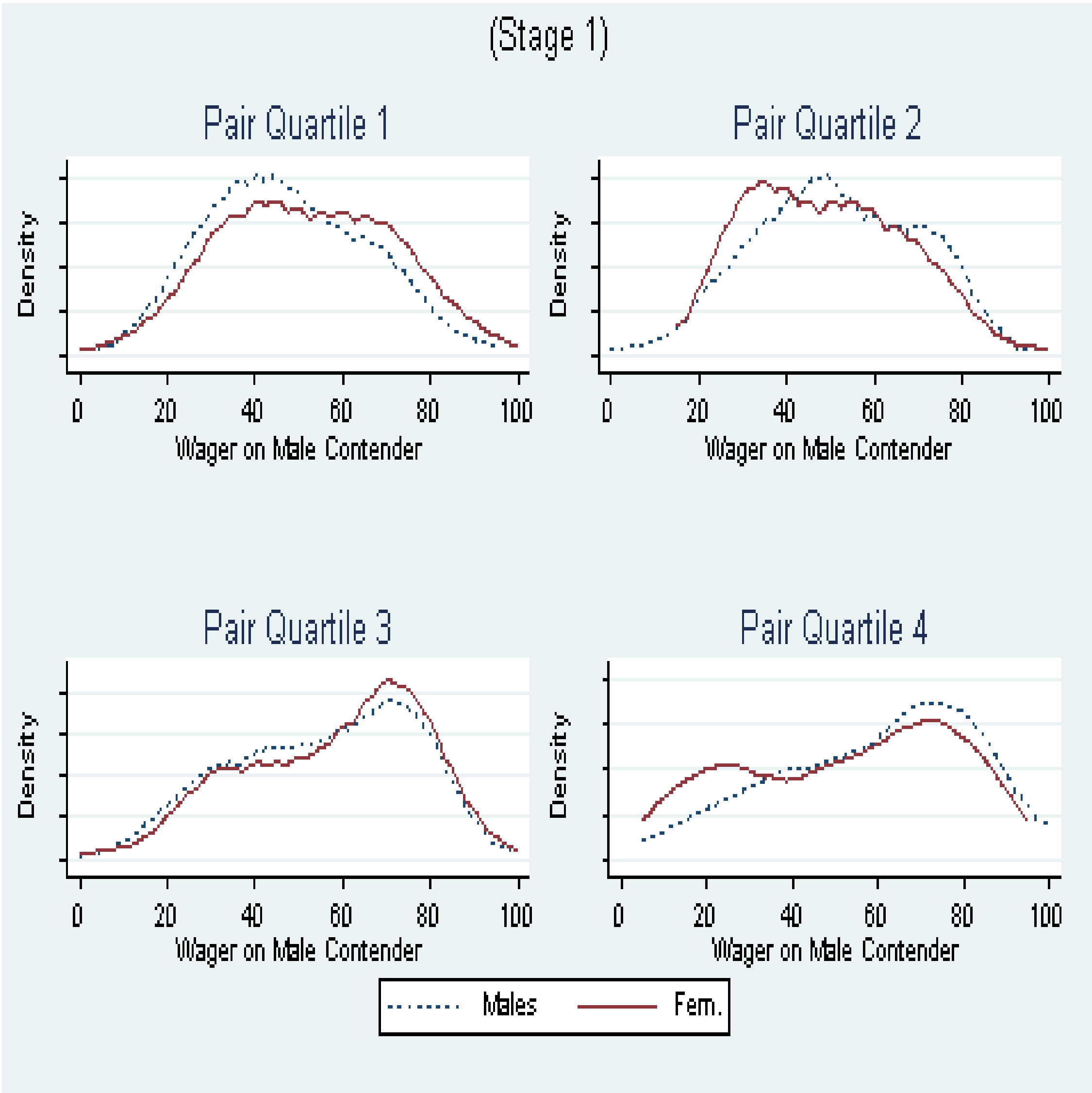

Figure 3.1: Distribution of Individual Wagers on Male Contender

Each frame shows the male and female decider distributions of individuals' Stage 1 wagers on the male contender for each pair quartile. (Quartile 1 = lowest performance). The wager on the male contender variable is the deciders' reported beliefs about how likely it is the male outscored the female in the head-to-head skill competition.

Gneezy *et al* 2003). What is relevant for this study is that for low-performing quartiles, male and female deciders disagree on average about which contender is better. In the group meetings, this disagreement should provide the greatest opportunity for gender interactions to appear. That is, the pairs for which members are most in disagreement should reveal whether one gender is more likely to hold sway in the collective decision. For pairs where members agree more on average, gender effects may still occur but could be less likely and smaller in magnitude. If *a priori* most or all members agree on the male contender, then perhaps there will be less need for deliberation or give-and-take.

This point is especially important when qualitatively comparing the results with field studies such as Bagues and Esteve-Volart (2010). That study estimates how different evaluation committee gender compositions affects male and female job candidates success. Since candidates are randomly assigned to committees this presents a nice natural experiment for isolating the effect of gender composition. Of course, the estimates are identified off marginal candidates (as opposed to high-quality, sure passes or low-quality, sure fails). Presumably a marginal candidate from the committee's point of view is one where board members disagree substantially. In the present experiment, the individual results in Table 3.2 indicate that lowest quartile pairs (Q1) are the ones where the most disagreement exists across genders. Thus, results from these quartiles are what should appropriately be compared with the applied results in Bagues & Esteve-Volart (2010).

3.3.2 Groups' Decisions

Tables 3.3 and 3.4 report, by gender composition and pre-competitive performance quartiles, group members' and groups' decisions. Table 3.3 reports average wagers (0 to 100) while Table 3.4 reports the fraction choosing the male. In both tables, the top panel reports the average of all the group members' Stage 1 decisions. This shows what group members believed prior to deliberations. The bottom panels show the average collective decisions of the groups. In these panels one can compare raw group decisions by pair quartile and group gender composition. For the high-performing quartiles (Q3 and Q4), groups on average favor the male though there is variation across gender compositions. With low-performing quartiles (Q1 and Q2), all groups except those with two females favor the male though typically to a lesser extent than for high-performing quartiles. By comparing the top and bottom panels in each table, it appears some differences exist between group members' *a priori* beliefs and the reported group decisions.

Table 3.3: Individual and Collective Wagers on Male

Group Member's Avg. Wagers on Male Contender				
Quartile	1 Fem.	2 Fem.	3 Fem.	4 Fem.
1	47.6	48.5	51.5	50.6
	(2.0)	(1.7)	(2.5)	(1.7)
2	49.5	49.9	51.4	50.1
	(1.8)	(2.0)	(1.4)	(1.8)
3	54.1	57.5	56.7	57.2
	(0.9)	(1.1)	(1.0)	(1.2)
4	54.3	58.8	55.7	56.5
	(2.1)	(3.0)	(3.5)	(2.7)

Group Wagers on Male Contender				
Quartile	1 Fem.	2 Fem.	3 Fem.	4 Fem.
1	52.1	48.5	50.2	53.5
	(2.0)	(3.0)	(3.0)	(1.9)
2	52.9	46.2	51.2	54.2
	(2.4)	(2.9)	(3.1)	(3.1)
3	52.7	58.2	57.2	53.8
	(2.3)	(3.3)	(3.6)	(3.5)
4	59.4	64.7	58.8	56.4
	(4.7)	(5.2)	(3.8)	(4.6)

Top panel reports, by pre-competitive performance quartile and group gender composition, the mean wager on the male contender of all members in a group. Bottom panel reports the same groups' mean collective wagers on the male contender. Standard errors clustered by group and individual are in parenthesis.

To formally test whether groups produce different results from that suggested by members' beliefs, I compute for each pair the difference between each collective decision and the median member's decision. This difference represents any "group shift" that occurred and will indicate to what extent individual preferences account for the group decision. Table 3.5 shows these results. All standard errors cluster on both the group and individual level. The top panel reports shifts in wagers while the bottom panel reports discrete shifts from the median member's favored contender. Figure 3.2 shows box plots of the same group shifts for each quartile and gender composition. In only a small number of cases is there a statistically significant shift. Specifically, shifts occur in one or four female groups evaluating lowest quartile pairs (Q1). Also, two- and four-female groups facing second lowest quartile pairs (Q2) also indicate a shift. Since low quartiles are where the most disagreement exists (based on Stage 1 individual beliefs), it is reasonable gender effects would show up here, if any where.

However, the gender compositions for which shifts occur are equally interesting. For Q1 pairs, where the strongest effect would be expected, the extreme gender compositions (1 female/4 males; 4 females/1 males) show a significant shift toward the male contender. For Q2 pairs, the most female group also shifts toward the male contender while the second most male group shifts toward the female contender. Other group compositions in the lower quartiles do not indicate a group shift. Groups evaluating higher-quartile pairs also do not shift significantly. This leads to the first key finding: many group decisions can be fairly well explained by individual preferences (i.e. those with no significant shifts), yet some groups shift substantially.

The pattern of shifting, particularly for the Q1 pairs, is qualitatively similar to Bagues and Esteve-Volart's (2010) findings. In their study, groups with small, non-zero female representation *and* groups with majority female representation both favored males. Groups with middling female representation showed no differential preference for one candidate gender. So, the experimental results - when focusing on the marginal cases analogous to those in Bagues and Esteve-Volart (2010) - strongly agree with findings in the field. In both cases a pro-male shift occurs at similar group gender compositions.

Table 3.4: Fractions of Individuals and Groups Choosing Male

Pct. Of Group Members Choosing Male Contender				
Quartile	1 Fem.	2 Fem.	3 Fem.	4 Fem.
1	41.4	45.0	50.4	52.1
	(4.7)	(4.4)	(6.0)	(3.8)
2	50.5	49.0	51.2	46.3
	(5.0)	(5.3)	(4.4)	(4.7)
3	60.7	66.3	61.7	64.7
	(1.8)	(1.1)	(1.9)	(2.6)
4	52.8	72.7	60.0	66.4
	(4.7)	(6.6)	(7.8)	(6.5)

Pct. Of Groups Choosing Male Contender				
Quartile	1 Fem.	2 Fem.	3 Fem.	4 Fem.
1	52.8	41.7	50.0	66.7
	(8.9)	(7.5)	(10.8)	(8.9)
2	57.1	33.3	52.9	65.8
	(9.5)	(10.2)	(11.8)	(9.9)
3	63.3	70.0	60.0	55.0
	(6.7)	(7.3)	(11.5)	(8.8)
4	66.7	83.3	69.2	59.1
	(9.3)	(9.7)	(12.8)	(11.4)

Top panel reports, by pre-competitive performance quartile and group gender composition, the average proportion of group members favoring the male contender. Bottom panel reports the average proportion of groups' collective decisions which favor the male contender. Standard errors clustered by group and individual are in parenthesis.

Table 3.5: Group Shifts

Group Minus Median Member Wager on Male				
Quartile	1 Fem.	2 Fem.	3 Fem.	4 Fem.
1	5.69*	0.50	0.00	3.96**
	(3.29)	(2.34)	(2.69)	(1.58)
2	1.19	-5.48**	0.00	6.84*
	(3.68)	(2.70)	(4.13)	(3.83)
3	-4.00	-1.67	1.83	-4.50
	(4.04)	(4.20)	(3.71)	(4.00)
4	6.11	4.33	0.77	0.91
	(5.00)	(6.94)	(6.34)	(3.31)

Group Choice Minus Median's Choice of Male				
Quartile	1 Fem.	2 Fem.	3 Fem.	4 Fem.
1	0.18	0	0.02	0.17*
	(0.12)	(0.05)	(0.07)	(0.08)
2	0	-0.21**	0	0.21
	(0.16)	(0.10)	(0.19)	(0.14)
3	-0.03	-0.08	0.02	-0.12
	(0.13)	(0.07)	(0.12)	(0.10)
4	0.14	0.07	0	-0.09
	(0.11)	(0.14)	(0.16)	(0.11)

Top panel reports, by pair pre-competitive performance quartile and group gender composition, the average group shift in the wager on the male contender. For each decision, the group shift is the collective wager minus the Stage 1 wager of the median group member on the pair in question. Positive (negative) values mean the group shifted toward the male (female) contender, relative to the median member. The bottom panel displays directional shifts on the extensive margin. The shift is calculated as the collective choice (1 = male favored) minus the median member's choice. Standard errors clustered by group and individual are in parenthesis. (*): 10%, (**): 5%, (***): 1%.

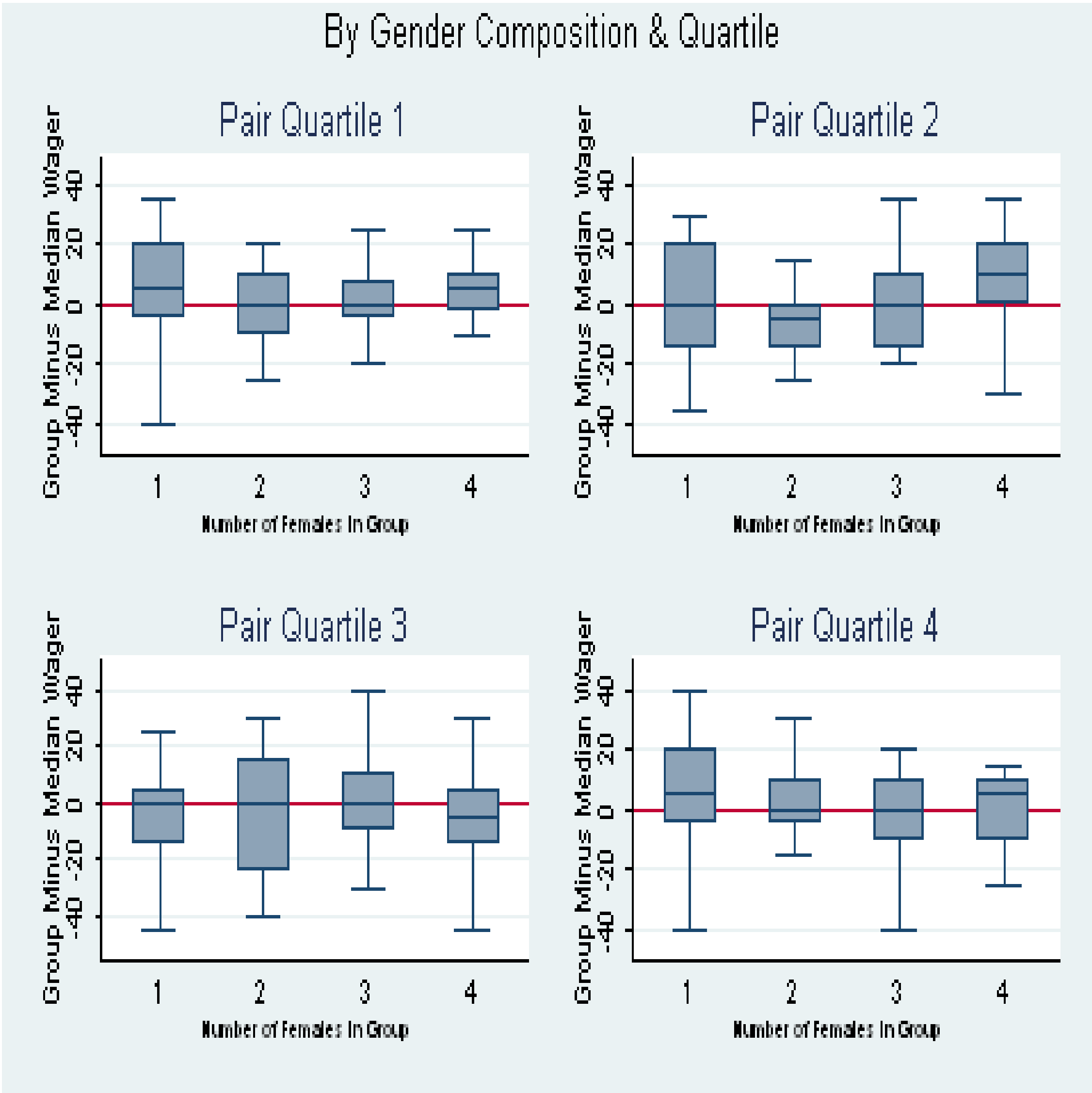

Figure 3.2: Group Minus Median Member's Wager

Each frame shows group shift data for a pair quartile. Within each quartile, the range of shifts (calculated as the difference between a group's wager on the male minus the median individual's wager on male) is shown for each gender composition. The shaded box represents the range of the middle two data quartiles.

Of course, results for Q2 pairs are more perplexing. Based on the individual beliefs, we would expect to find less of an effect since the intra-group disagreement is less stark for Q2 pairs. Furthermore, the two-male groups' decisions on Q2 pairs may be a statistical outlier, only choosing the male contender 33% percent of the time. This is by far the lowest average rate of any group and thus far defies simple explanation. Henceforth, analysis will mostly focus on Q1 pairs since these results are most comparable with field studies this experiment is meant to complement.

In summary, the group decisions provide important pieces of information. First, by carefully considering the most relevant cases (Q1 pairs), I am able to replicate in the lab the results of the most relevant field study. Second, I show there are group shifts related to gender which cannot be explained solely by individual preferences over contenders. This points to the possibility that gender interactions within the group are partially driving the collective decisions. In the next sub-section I present a possible explanation for these shifts by mapping individual group members' decisions into the groups' decisions.

3.3.3 Individual Influence and Interactions on Group Outcomes

Since some groups' decisions cannot be fully explained by members' beliefs, I investigate whether gender interactions may explain the shifts. Bagues and Esteve-Volart (2010) propose two hypotheses: (i) women in committees favor male candidates and (ii) men in committees favor male candidates when sitting in mixed-gender committees. As revealed by a sizable group shift, the first hypothesis does not fully explain the Q1/4-female decisions. The second hypothesis also is less than satisfying because two-female groups do not shift at all toward the male. It is hard to explain with the second hypothesis why one-female groups would shift while two-female groups would not. This suggests another explanation is needed. In this section, I offer such a theory and provide evidence to help explain both of the observed shifts for the Q1 pairs.

In mapping individual group members' beliefs (elicited in Stage 1) to groups' decisions, I am particularly interested in whether male and female group members have differential impacts on moving the group toward their personal belief about the contenders. That is, in a given situation is a male more or less likely relative to a female to be able to affect the group's decision? To do this I rank each group member based on how he or she wagered on the pair in question. Rank 1 is most pro-male in the group while Rank 5 is most pro-female in the group. I then estimate the effect of a male relative to a female in a given rank being able to move the group toward his position. I construct as an outcome the absolute difference between the group's decision and the member in the given ranked position. I then regress this constructed variable on a binary variable (1 = male) indicating the gender of the member occupying the rank, conditional on all the wagers of all group members. The coefficient on this indicator variable estimates for a given position in the group how likely a male relative to a female is to move the group near his or her Stage 1 belief. If the estimated coefficient is negative, this means a male is more able than a similarly situated female to reduce the distance between his individual belief and the group's decision. A positive coefficient means a female is more able than a similarly situated male to bring the group decision toward her belief.

Table 3.6 shows coefficient estimates and standard errors for each gender composition/quartile/rank combination. All standard errors are clustered on the group and individual levels. Significant estimates appear for several cases, but given the small sample sizes some of these are likely due to statistical variation. However, directing attention to the lowest quartile (Q1) cases in the topmost panel where significant group shifts occurred (1 female and 4 female groups), an interesting and consistent pattern emerges.

In the 1 female/Q1 case (top left corner of Table 3.6) a male decider in the most pro-male position (Rank 1) manages to pull the group toward his belief a significant amount as evidenced by the -8.77 coefficient. In these one-female groups, the *a priori* beliefs of the mostly male group favor the female because for Q1, male deciders on average favor the female (recall the Stage 1 beliefs presented in Table 3.2 and Table 3.3, Top Panels). So, the most pro-male decider in these groups will usually be in the minority opinion. In this case, the coefficients estimate a male in the minority opinion moves the group closer to his Stage 1 belief than a similarly situated female.

In the 4 female/Q1 case (top right corner of Table 3.6) a male decider in either the most pro-female (Rank 5) or second-most pro-female (Rank 4) positions manages to pull the group decision toward his Stage 1 belief more than a female at either rank. In these groups, the *a priori* beliefs of the mostly female group favor the male because for Q1, female deciders on average favor the male (again, refer to Tables 3.2 and 3.3, Top Panels). In these groups the most pro-female decider will tend to be in the minority opinion. As before, the males in the minority opinion positions are more successful in pulling the group toward their beliefs than similarly situated females.

Taken together, these results point to a potentially important element of group gender inter-

Table 3.6: Gender and Rank Effects on Group Shifts

		Effect of Male Being in Rank, Cond'l on All Wagers			
Quartile	Rank	1 Female	2 Female	3 Female	4 Female
	1	-8.77 (4.22)**	-4.94 (3.14)	4.59 (6.03)	-5.24 (5.41)
	2	-4.48 (4.74)	5.01 (5.22)	-6.16 (3.62)*	-0.49 (4.32)
1	3	-3.57 (5.68)	-1.37 (3.21)	5.29 (5.53)	-3.94 (6.21)
	4	1.06 (3.95)	-7.4 (2.48)***	1.50 (3.99)	-8.32 (3.47)**
	5	-1.95 (7.35)	5.10 (5.56)	-0.58 (5.34)	-12.96 (5.12)**
	1	3.82 (7.86)	0.60 (4.13)	-5.82 (4.99)	-1.66 (3.39)
	2	2.34 (5.72)	-2.06 (4.10)	-8.17 (3.52)**	0.75 (6.58)
2	3	-1.30 (6.40)	4.64 (4.53)	-9.30 (4.46)**	19.37 (9.96)*
	4	21.18 (6.16)***	1.19 (2.38)	2.21 (9.41)	-4.89 (6.53)
	5	-3.78 (22.2)	-2.21 (3.448)	7.91 (11.43)	6.31 (9.98)
	1	18.20 (6.36)***	-6.66 (8.60)	-4.69 (5.86)	6.09 (6.39)
	2	22.20 (2.26)***	-4.05 (4.46)	4.79 (4.80)	-2.12 (6.55)
3	3	-5.44 (11.03)	-6.08 (6.88)	9.69 (4.54)**	5.88 (6.43)
	4	-4.87 (7.29)	4.40 (5.68)	-11.92 (3.62)***	-6.51 (7.67)
	5	24.20 (44.34)	-5.34 (7.49)	2.36 (4.95)	8.64 (4.86)*
	1	4.67 (20.61)	0.42 (4.53)	4.32 (3.94)	-2.4 (28.06)
	2	-19.41 (3.49)***	1.96 (5.34)	-8.39 (2.18)***	-2.90 (13.48)
4	3	-2.33 (4.15)	-8.00 (1.27)***	4.57 (3.60)	18.89 (1.02)***
	4	48.39 (12.14)***	4.71 (4.94)	-12.77 (7.73)*	0.74 (1.36)
	5	-15.16 (11.08)	-6.44 (5.39)	3.69 (9.05)	7.82 (8.63)

Each coefficient estimate is the result of a regression, for the specified quartile and gender composition, where the outcome variable is the absolute difference between the group's wager and the Stage 1 wager of the member in the given rank. Reported coefficients are on an indicator for whether the member occupying the rank is male (=1) or female (=0). Regressions also control for Stage 1 wagers of all members by rank, pair fixed effects, and session fixed effect. Rank 1 is the most pro-male member. A positive (negative) coefficient means a female (male) was more effective in moving the group toward her (his) Stage 1 belief. For Q1 estimates, note that the groups' *a priori* preferences favor on average females in the one-female groups and males in the four-female groups. This means the males found to have influence in both cases are in the minority opinion. Standard errors clustered by groups and individuals. (*): 10%, (**): 5%, (***): 1%.

actions. Males and females who find themselves (randomly) part of the minority opinion differ in how likely they are to pull the group toward their belief. This actually ties-in nicely with one of the leading theories for why group shifts occur: the "persuasive argument theory" (Burnstein et al., 1973; Brown, 1974). The theory states that the pool of arguments in one direction are more persuasive. One possibility, is that females are less likely to challenge the majority opinion which effectively makes the majority opinion appear even more persuasive. So, in both cases a male in the minority opinion is more effective in influencing the decision.

While the possibility that females might have less influence in some decisions is an important finding, this can also help explain the pattern of group decisions observed in this study and Bagues & Esteve-Volart (2010). If females are on average less likely to influence the group when in the minority position, this should produce systemic patterns. In the 4-female/1-male groups the most pro-female members (Rank 4 or Rank 5) will usually be a female (even though males on average are more pro-female for Q1 pairs) and will therefore most of the time *not* challenge the majority's pro-male priors. Thus, the Rank 4 or Rank 5 female's opinion is not incorporated into the collective decision and the outcome is skewed toward the male more often than not. In the 1-female/4-male groups, the most pro-male member (Rank 1) will usually be a male (even though females on average are more pro-male for Q1 pairs) and will therefore most of the time challenge the majority's pro-female priors. This causes the group more often than not to incorporate the most pro-male opinion and move toward the male contender. Furthermore, if the female in the group tends to not challenge *any* strong opinions then this effect may exacerbated. Finally, in 2- or 3-female groups no effect occurs. In these cases it is less likely a clear majority prior exists, and it is more likely there will be males of opposing opinions. These males may challenge each other with the result being roughly the median member's wager.

3.4 DISCUSSION AND CONCLUSION

This experiment provides important insight into gender interactions within evaluating committees. A key question motivated by the empirical literature is whether individual committee members' preferences can fully account for the collective decision. Carefully-executed work (Bagues and Esteve-Volart 2010) finds that the gender composition of evaluation committees influences how the committees evaluate male and female job candidates. But, these studies are unable to say how much of these differences are due to individual committee members' preferences and how much are due to interactions within the group. Using an experimental design where I sequentially elicit individuals' then groups' evaluations of male-female candidate pairs, I observe how individual beliefs map into the group decisions. First, I find a pattern similar to that found by Bagues and Esteve-Volart (2010) where relevant groups with a small minority of females or a majority of females favor male candidates. Second, I show this effect is not fully explainable by group members' individual preferences. Third, I show evidence that males who disagree with other group members are much more successful than a similarly situated female in moving the collective decision toward their belief. This effect can partially explain why some groups shift in systematic ways from the individual members' beliefs.

This study's findings also point toward further research questions. For instance, it is not clear from the data why exactly dissenting males' opinions matter. Perhaps these males assert themselves more? Or, perhaps dissenting males and females speak-up at similar rates and groups place less weight on the female's opinion? Is there evidence that only some members are participating? Analysis of group discussions may provide useful information about these issues. Furthermore, it may be interesting to study differences in group deliberations when most members agree *a priori* versus when there is disagreement. It would be interesting to learn whether one gender is more likely to lead discussions or make firm statements about their beliefs. Do members ever clearly change their public beliefs? While my study points with data to one mechanism, a richer picture may be developed through detailed analysis of group discussions.

Finally, these findings can guide policy related to group decision-making procedures. One of the arguments for diversity on committees and boards is that better outcomes result (Carter et al. 2007; Smith et al. 2006; Erhardt et al. 2003). However, the present analysis indicates some members' opinions are systematically incorporated less. This means some of the benefits from diversity may not be realized. If some group members' opinions are systematically excluded from the collective decision, a slight procedural changes could help regain the benefits of having a diverse group of decision makers. For example, group members could independently and anonymously record their evaluations before a group discussion. Then the anonymized evaluations would be made public to the group. This might increase the likelihood each members' position is fully considered.

APPENDIX

EXPERIMENT INSTRUCTIONS

NOTE: In the instructions below, all bracketed (i.e. [ ]) sentences are actions to be executed by the researcher and/or his assistants. All other text is to be read in the session in the order presented.

A.1 CONTENDER SESSION INSTRUCTIONS

Good morning/afternoon. Thank you for coming to this experiment. This session should last less than 30 minutes, and you will be paid in cash at the end of the experiment. Please read the consent form along with me and sign it if you still want to participate. [Researcher reads the consent form aloud.] These pictures and information about how you perform in the tasks in this experiment will be displayed in later sessions of this experiment at the University of Pittsburgh. This information will not be used at your school here or displayed outside the experiment. Are there any questions?

[After answering subjects' questions, proceed.] [Take 2 head-shot pictures of each subject.]

Now I will explain the first task and how you will earn money.

Stage 1

This experimental session has two stages. For completing both stages you will be paid $10. However, you can earn more money through your performance in the experiment. At the end of both stages, one stage will be randomly selected. Your performance in that stage, and only that stage, of the experiment will be used to determine how much money you earn in addition to the $10 you are sure to earn for completing both stages of the session.

In stage 1 you can earn money by adding a series of five randomly-chosen two-digit numbers (e.g. 13+ 15+ 67 + 90+45). Calculators are not allowed. You will have five minutes to add as many sets of numbers as possible. If stage 1 is randomly selected at the end of the session, you will get $0.50 for every correct sum. Your payment will not decrease if you provide an incorrect answer. Here is an example of the task you will have. [Display the example task.] You will place your answer from adding up the five numbers in the empty space. Your score will be the number of correct sums you get in total. I will inform you each minute how much time remains. If you have any questions, please raise your hand, otherwise on my signal the task will begin. You will have five minutes to add as many sets as possible.
[Have subjects complete first timed task.]

Stage 2

For stage 2 of the session, you have been randomly paired with someone else in this room. You will now compete against this person in a one-on-one competition. As before, you will have five minutes to add five randomly-selected two digit numbers. This time, however, you will only earn money if you have more correct answers than the person you are paired with. You will not learn who you are paired with if this stage is selected for payment.

The person in each pair who correctly adds the most sums in five minutes will earn $1 for each correct answer if stage 2 is randomly selected. The person in each pair with the lowest number of points will not earn any extra money if stage 2 is selected for payment. In the event of a tie, each person in the pair will receive half of the number of correct answers. Enter your answers exactly as before. If you have any questions, please raise your hand, otherwise on my signal the competitions will begin. You will have five minutes to add as many sets as possible.

[Have subjects complete second timed task.]

Stage Selection and Payoffs

Both stages are now complete. Will someone volunteer to select a stage out of this hat? [Volunteer selects card from hat.] Please announce which stage you selected and verify the other slip of paper contains the other stage number.

Please proceed to the exit and I will pay your earnings one at a time.

[Make payments for each subject based on their performance in the randomly selected stage.]

A.2 DECIDER SESSION INSTRUCTIONS

[Upon entry, each subject is given a ticket with their subject number and a consent form.]

Good afternoon. Thank you for coming to this experiment. This session should last about 90 minutes, and you will be paid in cash at the end of the experiment. I will explain in a moment how you earn money. Before we begin, please read and sign this Consent Form. The consent form explains that your personal information will be protected and you can leave at any time and this will not be held against you.

[Collect consent forms]

[Pass out INSTRUCTIONS and INDIVIDUAL ANSWER SHEETS]

Stage 1

This experimental session has four stages. For completing all the stages you will be paid $7. However, you can earn more money through your performance in the experiment. At the end of all four stages, we will randomly select one of your decisions from Stage 1 and one of your decision from either Stage 2, 3, or 4. This means you will be paid for two of your decisions today plus the fixed

$7. We will randomly select these decisions right here in everyone's view after the last stage. Since only two stages will be chosen, you want to do your best for every decision.

In this experiment, you will be asked in different situations to make decisions about how people performed in an actual competition. The decisions you are involved in and the actual outcomes of the competition will be the basis for calculating how much you earn.

Here's how you earn money: You will be shown pictures of a series of contender pairs. The contenders are students at Pitt-Johnstown and participated in an experiment on March 17, 2010. During the experiment they were asked to solve a series of small problems. First, contenders had five minutes to add as many sets of five two-digit numbers as they could. They earned $0.50 for each correct sum. The second time they were in a head-to-head competition with someone else in the room. The contenders did not know who they were competing against - just that they were competing. The contender who solved the most problems won $10 while the loser won nothing. Both contenders had the same problems. Your task will be to guess which of two contenders solved the most problems in the competition. **Note that the pictures were taken at the beginning of the experiment before the competition.**

You will see pictures of contender pairs one by one. Here is an example. On the left is contender "A" while on the right is contender "B". At the bottom, you are given information about how both contenders performed in the initial, non-competitive task compared to a larger group of contenders. In this example, both contenders' scores in the first, individual task were in the "second highest quartile." Quartiles break up a population into the lowest 25%, second lowest 25%, second highest 25%, and highest 25%. Being in the highest quartile means the contenders scored better on the initial task than at least 75% of other contenders. Being in the second highest means a contender scored better than at least 50% but worse than at least the best 25% of contenders.

As I show each pair, you will decide which contender scored the most points in the head-to-head competition. This is a two step process. First, report who you believe won (A or B). Then choose a percentage between 50% and 100% to describe your belief. Only choose increments of five: 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100. Choosing a higher number means you think your favored contender is more likely to win.

Let's say you think Contender A is certain to win in the competitive task. Then you will mark on your sheet "A" in the first space for this pair. In the second space you will then put 100% because you are sure A will win. But, let's say instead you think B is slightly more likely to win. You would write B in the first space and then might choose 65% for the second spot. This means "Overall, I think B is 65% likely to win the competition." If you think the contenders tied, then write TIE in the first space and 50% in the second space. Notice that all the numbers you write should be 50% or higher because you are describing how likely your favored contender is to win.

Here's how we will determine your payment for a given choice. Once the experiment is complete and we have selected pairs for payment, we will publicity draw a random number between 50 and 100. **If this random number is greater than or equal to the percentage you chose for your favored contender, then you will earn $10 with a probability corresponding to that random number.** For example, if you chose 70% on Contender A to describe Contender A's likelihood of winning, and a random number of 80 is drawn then you would have an 80% chance of winning $10. If 85, 90, 95 or 100 are drawn as the random number then you have an 85, 90, 95 or 100% chance, respectively, of winning $10. Notice that this does not depend on which contender won the competition. Also notice that if you choose the contenders to tie your payoff will for sure be determined based on the random number because the random number will always be greater than or equal to 50%.

But, **if the random number chosen is less than the number you chose, then you will win $10 if your favored contender wins and $0 if your favored contender loses the competition.** So, returning to our example, if you chose 70% on Contender A and the random number is 50, 55, 60 or 65 then you would win $10 only if A wins. But, if B wins then you win nothing. If the contenders tied and your payoff depends on the outcome you will win $5.

This situation means you should always choose a percentage for your favored contender according to what you actually believe. If you report a percentage above or below what you actually believe this will only lower your expected payoff. If you choose a percentage lower than what you believe you can expect to earn less than if you reported what you believe. Keep this in mind and try to choose percentages closest to how likely you actually believe each contender is to win.

Also, as you make your decisions, keep in mind that while the contenders in a pair scored in the same quartile initially their competitive performance was often very different. For instance, frequently a contender jumped up or down one or two quartiles in the competitive task relative to the individual task. Overall there is substantial variation in how a contender performed competitively compared to individually. Try to use your best judgment to determine which contender scored when competing. Remember it is always best to report what you actually believe and reporting a tie will guarantee your payoff is determined randomly regardless of which contender scored more points.Are there any questions?

At this time, please verify the sheet in front of you has the subject number from your ticket.

[Show first pair of contenders.]

Using the sheet in front of you, please write down your wager for Pair A. Hide your wager behind your folder. Remember to choose your favored contender and then a number between 50% and 100% which indicates how likely you believe that contender is to win. After you finish, please quietly wait for the next pair.

[Show all 27 pairs.]

This is the end of Stage 1. Now you will go to your assigned rooms. Please turn in your answer sheet as you leave. Follow the assistant for your room as I call you.

Stages 2, 3, and 4

[Experiment assistant reads script aloud to the group.]

Now I will show you some of the same contender pairs you already saw and wagered on in Stage 1. In this stage however you will all discuss together how you would like to wager on each contender pair. As I show you a pair, you will take as much time as necessary to agree on a wager together. Report your wager exactly as before: state the group's favored contender (A or B or TIE) and what percentage between 50 and 100% you place on that contender. You may only report increments of 5, such as 50, 55, 60, etc.

Payoffs will be determined exactly as described earlier. The only difference is that if a decision from this round is randomly selected for payment everyone in this group's payment will be determined by the group's decision.

As I show each pair, carefully consider all the information available to you and discuss how you would like to wager as a group. When you have agreed, please announce your decision to me and I will record it. Remember, just like before, it always benefits you to report a wager equal to what you actually believe.

Here is the first pair. Please discuss among yourselves how you want to wager and report your decision to me.

BIBLIOGRAPHY

Adams, R.B. and D. Ferreira, "Women in the boardroom and their impact on governance and performance," *Journal of Financial Economics*, 2009, *94*, 291–309.

Adler, R.D., "Women in the executive suite correlate to high profits," *Harvard Business Review*, 2001.

Ambrus, Attila, Ben Greiner, and Parag Pathak, "Group versus individual decision-making: Is there a shift?," *working paper*, 2009.

Andreoni, J. and L. Vesterlund, "Which is the fair sex? Gender differences in altruism," *Quarterly Journal of Economics*, 2001, *116* (1), 293–312.

Bagues, Manuel and Berta Esteve-Volart, "Can Gender Parity Break the Glass Ceiling? Evidence from a Repeated Randomized Experiment," *Review of Economic Studies*, 2009, *forthcoming.*

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova, "Powerful Women: Does Exposure Reduce Bias?," *Quarterly Journal of Economics*, 2009, *124* (4), 1497–1540.

Berkman, M. B. and R. E. Oconnor, "Do Women Legislators Matter - Female Legislators and State Abortion Policy," *American Politics Quarterly*, 1993, *21* (1), 102–124.

Brown, R., "Further comment on the risky shift," *American Psycologist*, 1974, *29*, 468–70.

Burnstein, E., A. Vinokur, and Y. Trope, "interpersonal comparisons versus persuasive argument: a more direct test of alternative explanations for group-induced shifts in individual choices," *Journal of Experimental Social Psycology*, 1973, *9*, 236–245.

Carter, D.A., B.J. Simkins, and W.G. Simpson, "Corporate governance, board diversity, and firm value," *The Financial Review*, 2003, *38*, 33–53.

Carter, David A., Frank D'Souza, Betty J. Simkins, and W. Gary Simpson, "The Gender and Ethnic Diversity of US Board and Board Committees and Firm Financial Performance," *Corporate Governance: An International Review*, 2010, *18* (5), 396–414.

Chattopadhyay, R. and E. Duflo, "Women as policy makers: Evidence from a randomized policy experiment in India," *Econometrica*, 2004, *72* (5), 1409–1443.

Clots-Figueras, Irma, "Are Female Leaders Good for Education? Evidence from India," *Mimeo*, 2006.

Coate, S. and G. C. Loury, "Will Affirmative-Action Policies Eliminate Negative Stereotypes," *American Economic Review*, 1993, *83* (5), 1220–1240.

DiNardo, J. and D. S. Lee, "Economic impacts of new unionization on private sector employers: 1984-2001," *Quarterly Journal of Economics*, 2004, *119* (4), 1383–1441.

Ehrenberg, Ronald G., George H. Jakubson, Mirinda L. Martin, Joyce B. Main, and Thomas Eisenberg, "Do Trustees and Administrators Matter? Diversifying the Faculty Across Gender Lines," *NBER Working Paper 15606*, 2009.

Erhardt, Niclas L., James D. Werbel, and Charles B. Shrader, "Board of Director Diversity and Firm Financial Performance," *Corporate Governance: An International Perspective*, 2003, *11*, 102–111.

Ferreira, Fernando and Joseph Gyourko, "Do political parties matter? Evidence from U.S. cities," *Quarterly Journal of Economics*, 2009, *124* (1), 399–422.

Fox, R. L. and Z. M. Oxley, "Gender stereotyping in state executive elections: Candidate selection and success," *Journal of Politics*, 2003, *65* (3), 833–850.

GAO, "Women in Management: Female Managers' Representation, Charateristics, and Pay," 2010.

Gneezy, U., M. Niederle, and A. Rustichini, "Performance in Competitive Environments: Gender Differences," *Quarterly Journal of Economics*, 2003, *118* (3), 1049–1074.

Hambrick, D.C., T.S. Cho, and M.J. Chen, "The influence of top management team heterogeneity on firms' competitive moves," *Administrative Science Quarterly*, 1996, *41*, 659–84.

Kochan, T., K. Bezrukov, R. Ely, S. Jackson, A. Joshi, K. Jehn, J. Leonard, D. Levine, and D. Thomas, "The effects of diversity on business performance: report of the diversity network," *Human Resources Management*, 2003, *42* (1), 3–21.

Kurtulus, Fidan A. and Donald Tomaskovic-Devey, "Do Women Mangers Help Women Advance? A Panel Study Using EEo-1 Records," *Working Paper, Department of Economics, University of Massachusets - Amherst*, 2009.

Lee, D. S., "Randomized experiments from non-random selection in US House elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

——, **E. Moretti, and M. J. Butler**, "Do voters affect or elect policies? Evidence from the US house," *Quarterly Journal of Economics*, 2004, *119* (3), 807–859.

McDermott, M. L., "Voting cues in low-information elections: Candidate gender as a social information variable in contemporary United States elections," *American Journal of Political Science*, 1997, *41* (1), 270–283.

Niederle, M. and L. Vesterlund, "Do Women Shy away from Competition? Do Men Compete too Much?," *Quarterly Journal of Economics*, 2007, *122* (3), 1067–1101.

Page, Scott E., *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*, Princeton: Princeton University Press, 2007.

Paola, M. De, V. Scoppa, and R. Lombardo, "Can gender quotas break down negative stereotypes? Evidence from changes in electoral rules," *Journal of Public Economics*, 2010, *94* (5-6), 344–353.

Rehavi, M. Marit, "Sex and Politics: Do Female Legislators Affect State Spending?," *Mimeo*, 2007.

Rietz, A. Du and M. Henrekson, "Testing the female underperformance hypothesis," *Small Business Economics*, 2000, *14* (1), 1–10.

Rose, C., *The Composition of Boards and Financial Performance in Danish Listed firms - Are the Recommendations from the Norby Report Beneficial to Shareholders? (in Danish) WP 2004-2*, Handelshojskolen i Kobenhavn: Institut for Finansiering, 2004.

Shrader, C.B., V.B. Blackburn, and P. Iles, "Women in management and firm financial performance: an explorative study," *Journal of Mangerial Issues*, 1997, *9* (3), 355–72.

Smith, Nina, Smith Valdemar, and Mette Verner, "Do women in top managemnet affect firm performance? A panel study of 2,500 Danish firms," *International Journal of Productivity and Performance Management*, 2006, *55* (7), 569–593.

Thomas, Sue, "The Impact of Women on State Legislative Policies," *The Journal of Politics*, 1991, *53*, 958–76.

___ **and Susan Welch**, *The Impact of Women on State Legislatures: Numerical and Organizational Stength* The Impact of Women in Public Office, Bloomington: Indiana University Press, 2001.

Vega, A. and J. M. Firestone, "The Effects of Gender on Congressional Behavior and the Substantive Representation of Women," *Legislative Studies Quarterly*, 1995, *20* (2), 213–222.