Gender Composition and Group Behavior: Evidence from US City Councils.*

Emilia Brito, Jesse Bruhn,
Thea How Choon, and E. Anna Weber

July 2023

Abstract

How does gender composition influence individual and group behavior? To study this question empirically, we assembled a new, national sample of United States city council elections and digitized information from the minutes of over 40,000 city-council meetings. We find that replacing a male councilor with a female councilor results in a 25p.p. increase in the share of motions proposed by women. This is despite causing only a 20p.p. increase in the council female share. The discrepancy is driven, in part, by behavioral changes similar to those documented in laboratory-based studies of gender composition. When a lone woman is joined by a new female colleague, she participates more actively by proposing more motions. The apparent changes in behavior do not translate into clear differences in spending. The null finding on spending is not driven by strategic voting; however, preference alignment on local policy issues between men and women appears to play an important role.

*Emilia Brito: Department of Economics, Brown University. Email: ebritore@brown.edu. Jesse Bruhn: Department of Economics, Brown University. Email: jesse_bruhn@brown.edu. Thea How Choon: Department of Economics, St. Lawrence University. Email: thowchoon@stlawu.edu. E. Anna Weber: Department of Social Sciences, United States Military Academy. Email: anna.weber@westpoint.edu. All opinions expressed in this presentation are those of the authors and do not represent the opinions of the United States Military Academy (USMA), the United States Army, or the Department of Defense

Whether it is in politics, the labor force, or the boardroom, improving the representation of women is a global concern (Chattopadhyay and Duflo, 2004; Ahern and Dittmar, 2012). Ideally, this push for representation is meant to ensure that the preferences and expertise of women are given a "seat at the table", so that they can wield substantive influence over the outcomes of high-stakes decisions. However, classic theories of "critical-mass" suggest that small changes from a low baseline may not be enough, especially in political decision making bodies (Kanter, 1977; Bratton, 2005). Nominal representation will not necessarily translate into substantive representation if complicated group behavior inhibits female action when they are in the minority. This hypothesis also has some empirical support: experimental evidence finds that when women are in the minority of a group, and especially when they are a lone woman among a group of men, they withhold substantive participation (Coffman et al., 2021; Stoddard et al., 2020).

Yet, the finding that group gender composition causes consequential changes in substantive participation has never been replicated outside of a laboratory or "lab-in-the-field" type setting. This raises the possibility that this existing body of work, where the subjects largely consist of students performing tasks in laboratories or classrooms, may not translate to the types of high-stakes decision making processes in politics and the workforce that originally motivated this line of inquiry. Evidence on this question from outside laboratory or "lab-in-the-field" type settings is virtually non-existent because it is exceedingly rare to find data which contains the necessary ingredients for identification: random or quasi-random variation in the gender composition of a group paired with rich outcome data that captures the internal workings of a high-stakes decision making process.

We study how gender composition influences individual and group behavior using novel data on the internal workings of US city councils. We assembled the data by hand collecting and coding over 40,000 PDFs from hundreds of municipal government websites containing the minutes of specific city council meetings. We pair this with information on electoral outcomes assembled from both publicly available files in California and newly collected national data we built from over 3,500 FOIA requests. The resulting data contains over 500 distinct councils spanning nearly 200 cities across 34 states over the course of eight years. It includes numerous measures of the behavior of individual city-councilors and the outcomes of city council decision making. Thus by leveraging close elections between a man and a woman, we can isolate exogenous variation in the gender composition of the councils and estimate causal effects on group behavior, individual behavior, and the workings of local government.

We find that nominal representation of women on city councils does translate into their substantive participation. On average, replacing a male city-councilor with a female city-councilor causes the share of motions offered by females to increase by 25pp. These effects are largest at low levels of baseline female representation, particularly for councils that would otherwise contain a single, "token" woman. Yet, replacing a man with a woman only causes the female share of the council to increase by 20p.p. The overall effects on substantive representation are therefore larger than what would be expected if the newly added women simply made their, "fair share," of motions. What drives this discrepancy?

At the individual level, we find that adding an additional woman to the council can cause changes to the behavior of other female councilors. When a second woman is added to a council that contains a single token woman, the share of motions offered by the lone woman increases by 14.8p.p. This increase in the share of motions comes from an additional 0.865 motions offered by the existing female councilor per meeting. There is also a comparably sized *drop* in participation by the other male councilors (-0.717 motions per meeting); however, it is imprecisely estimated. We do not find evidence that adding an additional woman to a council that already has multiple women engenders further changes in their behavior.

Despite these changes in substantive representation, we find little evidence that the addition of a new female councilor translates into consequential changes in the operation of city government. For example, we find precise null effects for the impact that an additional female councilor has on the level or composition of public expenditure. Importantly, these patterns hold for both the average council *and* the subset of councils that would otherwise contain a lone, token female councilor.

Why does the apparent increase in women's substantive representation fail to translate into changes in policy? We find little evidence that coalitional or strategic behavior pertaining to the way councilors vote on motions is responsible. For example, we find precise null effects on the share of motions passed unanimously, the share of motions rejected, and the average vote margin.

We find suggestive evidence that close alignment of policy preferences between men and women plays an important role. To probe the role of preferences, we conduct two supplementary analyses that rely on the topical content of motions offered as a proxy for male/female preferences. First, we show descriptively that there is little difference in the topical content of the motions offered by the average male and the average female councilor in our sample. Second, we explore the causal effect of an additional female councilor on the types of motions offered. We find no evidence that the gender composition of the group changes the topics of discussion.

We argue that these empirical facts can be rationalize by a stylized model of "critical-mass" in the spirit of Kanter (1977). The model features two key parameters: psychic costs to participation while in the minority, and preference heterogeneity between men and women. The model predicts that substantive female participation can vary discontinuously with group gender composition in response to gains in nominal representation, especially at low levels of baseline representation. However, gains in substantive representation only translate into changes in voting patterns and policy outcomes when the average preferences of men and women are sufficiently different. This is because, when men and women's policy preferences agree, changes in gender composition have no effect on the topics of motions or the frequency of non-unanimous votes. To our knowledge, this is the first formal mathematical model of critical mass,[1] and it suggests that the preferences of male / female policy makers is likely a key source of heterogeneity that could explain the diverse, and sometimes contradictory, findings in the literature regarding the policy impact of increases in women's nominal representation.

Our paper relates to a rich literature in behavioral and experimental economics by being the first to explore connections between gender composition and group behavior outside of a laboratory or lab-in-the-field type setting. Of particular importance is Stoddard et al. (2020), who conduct a field experiment among randomly-assigned groups of college students. The authors find that when a group has only one woman, she is less likely to be chosen as spokesperson or rated as influential by her peers, as compared to women in groups with multiple women. In related work, Coffman (2014) shows that in mixed-sex groups, women are less likely than men to contribute to a group when the topic of discussion is traditionally male-associated; this harms overall group performance. Bordalo et al. (2019) find that when working in pairs to answer trivia questions, women have less confidence in their relative ability in male-dominated domains when their partner is revealed to be male, again leading to worse group performance. Women are also less willing to lead mixed-sexed groups than all-female groups (Chen and Houser, 2019; Born

[1] In related work, Hughes et al. (2023) develop a model of committee decision making among groups with diverse signals and preferences; however, their model focuses on truth-telling rather than participation. They predict that, when groups differ in their informational structures but not in their preferences, welfare is increasing in diversity.

et al., 2022). Coffman et al. (2021) finds that within mixed-sexed groups, women rank themselves less favorably when they are in the minority in the group.

Our ability to study the effect of adding an additional woman on the behavior of women already on the council speaks to a broad interdisciplinary literature proposing the need for a "critical mass" of women to be represented to effect a change in outcomes in politics and other group settings, usually around 15-30% (Kanter, 1977; Bratton, 2005).[2] Using a similar regression discontinuity design as we implement here, Ferreira and Gyourko (2014) find little impact of the gender of U.S. mayors on the size of the municipal government or spending composition. In contrast, Chattopadhyay and Duflo (2004)'s results indicate that in India, the gender of local leaders does matter, with villages led by women spending more on issues which tend to be important for women. Gagliarducci and Paserman (2012) find that Italian municipalities headed by female mayors are more likely to be dissolved early and that this problem is more severe when the entire rest of the council is male. Women in California are also less likely than men to re-run for local office if they lose (Wasserman), and this gap is the largest for offices with low levels of female representation.[3] To date, the existing empirical literature on women's representation in politics has focused solely on the outcome of the decision making process and has not specifically explored thresholding or "critical mass" effects. Our paper is the first to explore behavioral changes, and in particular the implications of critical mass theory, that could underpin these findings.

1 Data and Background

1.1 City Councils in the United States

Collectively, local government in the US is responsible for managing over $1.9 trillion in revenues (Urban Institute, Accessed 2023). City councils are not only responsible for determining how the majority of these funds are spent; they are also responsible for other

[2] There is also some empirical support for the notion of a critical mass of women from the study of groups outside of politics. For example, among publicly traded companies in Germany, Joecks et al. (2013) find a U-shaped relationship between firm performance and women's representation on the board. When a token woman is present, performance is lower, but performance returns to the level of all-male boards when the share of women rises above 30%.

[3] There is also a related strand of literature that looks at the impact of political representation among minorities. For example, Beach and Jones (2017) show that, in California, the election of an ethnic-minority councilor results in a decrease in spending on public goods, which they attribute to increased intra-council disagreement.

important regulatory functions related to zoning, local ordinances, and the administration of local programs (National League of Cities, Accessed 2023). However, it is important to note that the political context for city council members is quite different than for other elected officials. For example, the typical city council is small, with an average council containing only 6 councilors (National League of Cities, Accessed 2023). Most city council elections are also explicitly non-partisan (National League of Cities, Accessed 2023). The actual process of policy making on councils usually involves close collaboration in committee settings with other members in order to develop the details of projects and proposals.[4] While national, hot-button issues do sometimes emerge in this process, much of the consequential tasks therefore involve the straightforward administrative decisions necessary for city government to function.

1.2 Electoral and Spending Data

Our data on the outcomes of municipal elections come from two sources: the California Elections Data Archive (CEDA), and a series of FOIA requests sent to cities across the United States. The CEDA data contains vote totals and election results since 1995 for nearly every city in California. We supplement this with data we obtained via a series of FOIA requests sent to over 3,500 additional US cities. The cities where we sent FOIA requests were chosen from the set of cities represented in the Annual Survey of State and Local Government.[5] These FOIA requests produced nearly 200 additional cities with usable election data spanning 38 states.

As a result of the FOIA requests, our data has more representative geographic coverage than prior work.[6] Figure 1 plots the national geographic distribution of the municipalities in our data. Panel (a) plots the "candidate set" of cities for FOIA requests that are contained in the Annual Survey of City Government. Panel (b) shows the set of cities where we were able to obtain data on electoral outcomes either via FOIA request or via the publicly available data in California. Within the set of cities where we have electoral

[4]According to a survey conducted by the National League of Cities, 81% of city councils rely on committees in order to "Provide groups of council members the opportunity to thoroughly consider particular items of business then recommend action on those items to the full council," National League of Cities (Accessed 2023).

[5]For this data product, the Census surveys every city in the United States ever five years, with a rotating sample of cities surveyed every year for a period of 5 years. We sent FOIA requests to the cities represented in the rotating sample.

[6]Prior work looking at city council elections in the US has focused exclusively on California (e.g. Beach and Jones, 2017).

data, panel (b) also further distinguishes between cities where we were (and were not) able to collect data on city council meeting minutes (see section 1.3 for more discussion of the minutes data).

Our research design relies on comparing close elections between men and women. Since the raw election data does not contain gender, we proxy for candidates' genders using first names. Using information from the Social Security Administration, we calculate for each first name the share of babies born between 1950 and 2000 given that name which are male. If more than 99% are male, we assign any candidates with that name as male. If less than 1% are male, we assign any candidates as female. For candidates in contested elections whose names fell within an intermediate range, we conducted internet searches to establish candidate gender from news articles or photographs. For a handful of cities, the gender of councilors was provided in the response to the FOIA request, in which case we rely on this information.

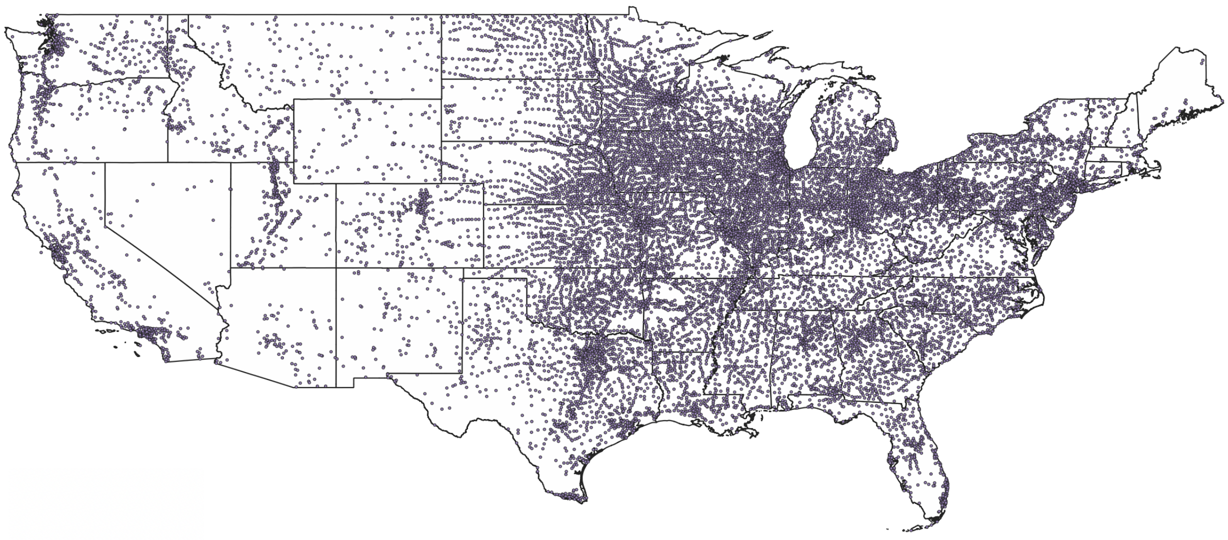
Our outcomes on local government spending come from two sources. The Annual Survey of State and Local Government provides expenditure data for cities outside of California. For cities in California, we rely on municipal spending information that is publicly available from the California State Controller's office. Most of the national data covered the period from 2007-2015; thus, we restrict our attention to 2007-2015 in both data sets to ensure a common overlap.

Table 1 provides summary statistics for the cities that enter our sample and for the population of cities from which they are drawn. The "All" column corresponds to every city represented in the Annual Survey of City Government. The "Election" column corresponds to the subsample of cities where we were able to collect election data. The "Gendered Election" column corresponds to the subsample of cities that have at least one opposite gender election necessary to work with our regression discontinuity research design. The "Minutes Sample" column corresponds to the subset of cities with a gendered election where we were able to collect data on the internal workings of city council meetings via their minutes (see section 1.3 for more detail on the minutes data). Panel A gives averages of census characteristics drawn from the ACS as measured in 2012. Panel B gives averages of municipal spending as measured in 2012. Panel C gives averages of council characteristics compiled from the election data.[7] Relative to the average city in the United States, the gendered election sample has a larger population and higher

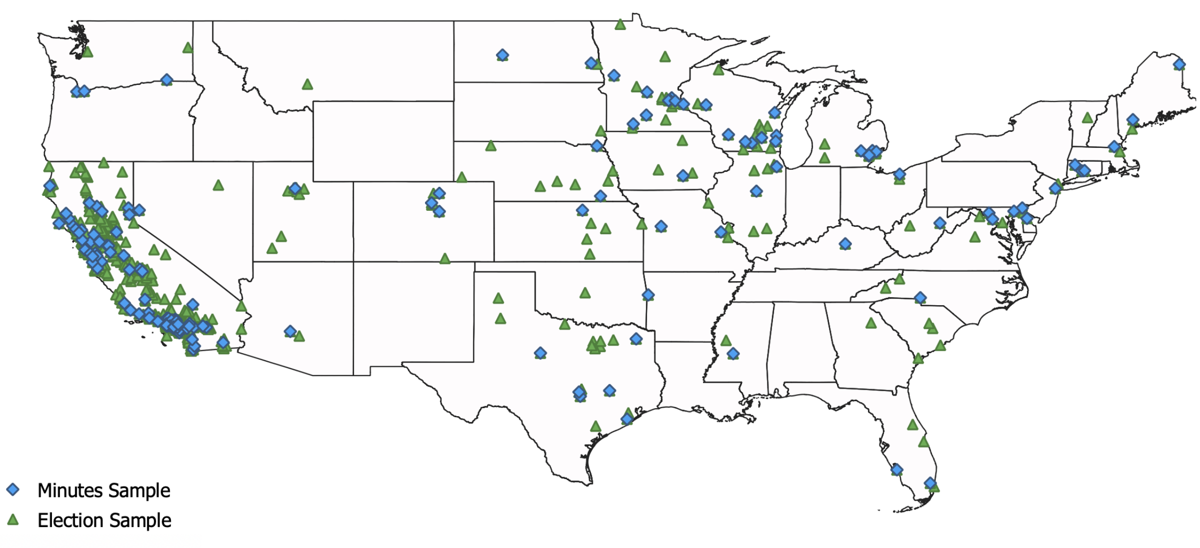
[7]Since there is substantial variation both in the time periods covered by our election data and in the years different municipalities actually hold elections, it is impossible to fix a common reference year for panel C. For that reason, we give the averages in panel C over all available time periods.

Figure 1: Geographic Distribution of City Councils

(a) Councils Represented in Annual Survey of City Government



(b) Councils Contained in Election and Minutes Data



Note: Panel (a) shows the geographic distribution of the initial set of cities and towns we used to build our analysis data. Panel (b) shows the distribution of cities and towns where we were able to collecting data on electoral outcomes and the internal workings of city council meetings via their minutes. Green triangles denote cities where we were only able to obtain electoral data. Blue diamonds denote cities where we were able to obtain both electoral data and data from meeting minutes.

annual expenditures. This is primarily due to the fact that most of the cities where we could obtain election data were large relative to the average city, as evident in the election column of table 1.

1.3 City Council Meeting Minutes

Our meeting minutes data was hand-collected by the research team over the course of 2 years by visiting individual municipal web-pages and manually downloading them.[8] The research team prioritized cities that had close gendered elections within the relevant time period, since these are the cities that contribute the most identifying power using the close gendered election research design. In total, we collected 42,610 PDFs. Figure 2 provides an example of the minutes from a city council meeting.

We then sent these PDFs to a data entry firm and had them extract basic outcomes such as meeting dates, start times, stop times, and councilors present/absent. For a subsample of these PDFs, we asked the firm to collect more detailed data on each motion offered at the particular city council meeting.[9] These motion-level outcomes included the names of the councilors who made and seconded the motion, as well as which councilors voted for or against the motion, and a categorization of the motion topics.

We will sometimes wish to distinguish ”behavioral” effects of gender composition on individual councilors from the contribution of the newly elected council member who was involved in the close election. For that reason, we linked councilors from the minutes data directly to their corresponding entry in the municipal elections data. This allows us to create variables that correspond to the behavior of individual councilors, including those not involved in the election, by gender. We will refer to councilors elected in close elections as a "focal" councilor, and we will refer to members of the council who were not part of the close election as "non-focal" councilors. In general, non-focal councilors may be those who were elected in prior elections, those who were elected in the same election but in separate races, or non-marginally elected candidates within the same race.[10]

The final unit of observation in our analysis data is a "council," which is defined at the

[8]There was enough heterogeneity across municipal government websites such that scraping them proved impossible.

[9]More precisely, we asked them to extract detailed data from 3 randomly chosen meetings per council-year.

[10]Some councils in our data elect representative by holding open elections for multiple city council seats, in which case the "top-N" vote-getters are elected to the council. In these cases, we define a gendered election as one where the "worst winner" and the "best loser" are of different genders, which implies that the "better" or "non-marginal" winners are exogenous from the perspective of the close gendered election.

Table 1: Summary Statistics: Municipalities

	All	Election	Gendered Election	Minutes
Panel A: Demographics				
Total Population	9,949	63,601	71,455	68,317
Black (share)	7.8%	5%	5.3%	5.5%
Hispanic (share)	7.9%	29.5%	30.1%	24.2%
Income (per-capita)	24,321	31,224	30,937	30,702
Panel B: Spending (Millions)				
Total	26	122.7	140.4	126.5
Police and Fire	4.1	25.9	29.5	26.8
Utilities	4.1	25.6	29.3	27.1
Waste and Sewer	2.3	10.7	12.1	13.2
Panel C: Councils				
Council size		5.5	5.5	6.1
Races		1.7	2.1	2.3
Candidates		5.8	6.1	7
Female candidates		1.7	2.1	2.3
Observations (cities)	19,276	630	523	153

Note: Table 1 provides descriptive statistics for four samples of US cities. "All" corresponds to every city represented in the Annual Survey of City Government. "Election" corresponds to the subsample of cities where we were able to collect election data. "Gendered Election" corresponds to the subsample of cities where we were able to identify at least one election where the marginal winner and the marginal loser in the election were of different genders. "Minutes Sample" corresponds to the subset of cities where we were able to collect data on the internal workings of city council meetings. Panel A gives averages of census characteristics drawn from the ACS as measured 2012. Panel B gives averages of municipal spending from the Annual Survey of City Government as measured in 2012. Panel C gives averages of council characteristics compiled from the election data. Since there is substantial variation both in the time periods covered by our election data and in the years different municipalities actually hold elections, it is impossible to fix a common reference year for panel C. For that reason, we give the averages in panel C over all available time periods.

Figure 2: Example of Meeting Minutes

Minutes
City Council/Redevelopment Agency/Public Financing Authority
City of Huntington Beach

Monday, November 15, 2010
4:00 PM - Room B-8
6:00 PM - Council Chambers
Civic Center, 2000 Main Street
Huntington Beach, California 92648

An audio recording of the 4:00 PM portion of this meeting and a video recording of the 6:00 PM portion of this meeting are on file in the Office of the City Clerk and are archived at www.surfcity-hb.org/government/agendas/

4:00 PM - ROOM B-8

The City Clerk recessed until 4:30 PM due to a lack of quorum.

CALL TO ORDER - 4:35 PM

ROLL CALL

Present: Carchio, Coerper (arrived at 5:15 PM for Closed Session), Hardy, Green, Bohr, Dwyer, and Hansen

ANNOUNCEMENT OF COMMUNICATIONS RECEIVED AFTER AGENDA DISTRIBUTION

Pursuant to the Brown "Open Meetings" Act, City Clerk Joan Flynn announced that the following communication was received by her office after distribution of the City Council agenda packet:

Communication received from Shari L. Freidenrich, CPA, City Treasurer, dated November 15, 2010 entitled City of Huntington Beach Investment Advisory Board (IAB) Annual Report to the City Council for the Period October 1, 2009 to September 30, 2010.

28. Approved for introduction Ordinance No. 3905 amending Chapter 14.12 of the Huntington Beach Municipal Code (HBMC) relating to fees, rates and deposits for water billing.

A motion was made by Coerper, second Hardy to after the City Clerk reads by title, approve for introduction Ordinance No. 3905, "An Ordinance of the City of Huntington Beach Amending Chapter 14.12 of the Huntington Beach Municipal Code Relating to Water Billing." The motion carried by the following roll call vote:

AYES: Carchio, Coerper, Hardy, Green, Bohr, Dwyer, and Hansen
NOES: None

ORDINANCES FOR ADOPTION

29. Adopted Ordinance No. 3902 amending Huntington Beach Municipal Code (HBMC) section 8.43 relating to emergency response costs Approved for introduction November 1, 2010.

A motion was made by Hansen, second Carchio to after the City Clerk reads by title, approve for introduction Ordinance No. 3905, "An Ordinance of the City of Huntington Beach Amending Chapter 14.12 of the Huntington Beach Municipal Code Relating to Water Billing." The motion carried by the following roll call vote:

AYES: Carchio, Coerper, Green, Dwyer, and Hansen
NOES: Hardy, and Bohr

30. Adopted Ordinance No. 3903 accepting modifications to Local Coastal Program Amendment (LCPA) No. 2-10 approved by the California Coastal Commission and amend the Local Coastal Program (LCP) accordingly Approved for introduction November 1, 2010.

A motion was made by Hansen, second Coerper to after the City Clerk reads by title, adopt Ordinance No. 3903, "An Ordinance of the City Council of the City of Huntington Beach amending sections 203.06, 216.04, 216.18, 221.10 and 230.82 of the Huntington Beach Zoning and Subdivision Ordinance thereof to conform LCP Amendment No. 2-10 made by the California Coastal Commission." The motion carried by the following roll call vote:

AYES: Carchio, Coerper, Hardy, Green, Bohr, Dwyer, and Hansen
NOES: None

Note: This figure displays two pages taken from the November 15, 2010, meeting of the Huntington Beach city council. The left panel provides information on attendance and the meeting date / time. The right panel shows examples of motions which were offered for a vote and the outcomes of votes on those motions.

city-by-electoral-term level. This unit of observation captures the notion of a consistent set of individuals who make decisions pertaining to city government in-between elections. This is the natural unit of observation in our setting, since it is the economic unit which is "treated" as a result of the close gendered election. For that reason, we create outcomes of interest from the raw minutes data by creating "per-meeting" averages at the council (i.e. city-by-term) level. So if a council met two times over the course of a term, and in the first meeting there were 5 motions offered, and in the second meeting there were 15 motions offered, the corresponding outcome variable would indicate that the council made an average of 10 motions per meeting.

Table 2 provides summary statistics for the meeting minutes data. The "All" column refers to all councils where we collected minutes data. The "1 Non-Focal" column refers to the subset of councils that, were it not for the focal woman involved in the close election, would otherwise have contained a single, token woman. The "> 1 Non-Focal" column contains the subset of councils that had multiple non-focal women. Panels A and B contain variables related to the process of meeting and the motions offered. Panel C contains the 4 most common motion topics. Note that motion topics are not mutually exclusive and hence shares do not need to add up to one. Numerical values in this table represent averages across councils.

The typical council in our data met 45 times at an average length of 146.7 minutes over course of a term. 93% of motions offered during a typical term are passed, and 90% are passed unanimously. While this may seem high, nearly 61% of all motions have a purely "administrative," function such as calling the meeting to order or putting the minutes into record. Other common topics of motions include those related to finance (19%), regulation (10%), and public utility management (13%). However, there is still a non-trivial amount of disagreement. 10% of motions do not pass unanimously, and 7% ultimately fail.

2 Research Design

Our goal is to estimate the causal effect of replacing a male city councilor with a female city councilor on group behavior, individual behavior, and the outcomes of city council decision making. For identification, we rely on a close election regression discontinuity design similar to prior work that has explored the impact of gender representation on public policy (Ferreira and Gyourko, 2014; Gagliarducci and Paserman, 2012; Beach and

Table 2: Summary Statistics: Meeting Minutes

	All	1 Non-Focal	> 1 Non-Focal
Panel A: Process			
Attendance (share)	0.81	0.82	0.79
Number of meetings	45.3	45.2	43.4
Meeting length	146.7	138	160.7
Panel B: Motions			
Total	10.5	10.3	11.5
Moved by women (share)	0.3	0.26	0.43
Seconded by women (share)	0.35	0.3	0.5
Passed	0.93	0.92	0.93
Passed Unanimously	0.9	0.9	0.9
Failed	0.07	0.08	0.07
Panel C: Topics			
Admin (share)	0.61	0.6	0.6
Finance (share)	0.19	0.19	0.2
Regulation (share)	0.10	0.11	0.09
Public Utility (share)	0.13	0.14	0.13
Observations (councils)	325	136	124

Note: This table presents summary statistics on the internal workings of the city council for the meeting minutes sample. "All" refers to all councils in this sample. "1 Non-Focal" refers to the subset of councils that, were it not for the focal woman involved in the close election, would otherwise have contained a single, token woman. "> 1 Non-Focal" contains the subset of councils that had multiple non-focal women. Numerical values in table represent council averages. The variables in panel A are observable for every meeting of the council in the relevant term. So "Number of meetings" equal to 45.3 in the "All" column indicates that the average council met 45 times during an average term (typically 2 years). "Attendance" is the average share of meetings attended by an individual councilor during a term. Panels B and C contains variables that were more costly to extract, and hence are only observable for 3 (randomly chosen) meetings per year. So for example, the variable "Total" equal to 10.5 indicates that the average number of motions made during these three meetings, and "Moved by women (share)" indicates that and average of 30% of these motions were made by women.

Jones, 2017). This strategy relies on an assumption that the conditional expectation function mapping the running variable into outcomes would, in the absence of treatment, be smooth through the cutoff (Lee and Lemieux, 2010). We discuss the testable implications of this assumption further when we develop our formal econometric model in section 2.1.

One unique feature of the city council setting is that there are many electoral races in which multiple seats are at stake in a single contest. For example, a city may have two "at large" seats for which a large pool of candidates are eligible. Voters may vote for up to two candidates, and the two highest vote-getters win seats. Therefore, to determine a close election within each race, we focus attention on the vote differential between the winner with the lowest vote total ("worst winner") and the loser with the highest vote total ("best loser").

For races where the worst winner and the best loser are different genders, we construct the female vote differential by taking the number of votes received by the female best loser/worst winner, minus the number of votes received by the male best loser/worst winner. Consistent with prior work we normalize the vote differential by the total number of votes cast in the race and hence our running variable is the vote-share (as in Lee, 2008, for example). Our treatment variable is an indicator for whether the female candidate was the winner between this pair of candidates.

We acknowledge here that recent work has criticized the close election regression discontinuity design on the grounds that the characteristics that define "treatment" in this setting are likely correlated with other councilor characteristics in a potentially complicated way. For example, if women are more likely to be democrats than men, then replacing a man with a woman is likely to change the ideological composition of the council, not just the gender.

Our view is that the combined influence of all of these characteristics on group behavior is precisely the treatment of interest in our setting. Put simply, if the causal channel that changes group behavior runs through population differences in ideology across genders, then that still implies a real impact on the lived experience of the women involved in these high stakes decisions.[11] Our view is that this lived experience is a worthy estimand to study. For that reason, we think that these critiques, while quite valuable for

[11] We also struggle conceptually to understand what it means to measure the treatment effect of varying group gender composition while holding other population characteristics across gender constant, since this would mechanically generate a form of selection on the dependent variable and generate estimates of a treatment effect that does not correspond to the population level ATE of replacing a man with a woman (see appendix **??** for a short proof of this fact).

contextualizing the treatment effects for close-election research designs in many settings, are less important here.

2.1 Econometric Model

Our estimating equation takes the form:

$$Y_c = \beta D_c + F\left(W_c\right) + \Gamma X_c + \varepsilon_c$$

where Y_c is the outcome for council c.[12] D_c is an indicator that takes a value of 1 if the worst winner is female. W_c is a running variable that captures the margin of victory between the worst winner and the best loser.[13] Our baseline model specifies F as piece-wise linear within a bandwidth around the threshold. In order to ensure that our point estimates are generated using a consistent sample and hence mutually comparable within a sample across outcomes, our baseline model uses a common bandwidth for all outcomes within each sample. We choose the common bandwidth within each sample as the MSE-optimal bandwidth for a specification where Y_c is the share of motions moved by non-focal women (Calonico et al., 2014). This implies that the estimates from our preferred model for the share of motions by non-focal women are identical to those found using the RDrobust package implementation of Calonico et al. (2014). However, we find qualitatively similar results for other outcomes when we allow the bandwidth to vary outcome-by-outcome, when we fix the bandwidth across samples, and when we vary the functional form of F by changing the order of the polynomial. The vector X_c contains council level control variables. For precision, our baseline model includes controls for council size, term length and baseline expenditures; however, point estimates are similar with alternative sets of controls and with no controls at all. The parameter of interest in this model is β, which measures the expected difference in the outcome variable at the RD threshold when a woman wins an election against a man.

The key assumption we rely on for causal identification is that the conditional expectation functioning mapping the running variable into outcomes is continuous across the threshold in the absence of treatment (Lee and Lemieux, 2010). Under this assumption, we can interpret β as the causal effect of a female win when vote shares between the worst winner and best loser are equally split. We provide evidence in support of this as-

[12] As discuss in section 1, we define a "council" at the level of a city-by-electoral term

[13] See section 1 for see for a precise definition of these terms.

sumption in section 2.2. For inference, we account for within-city serial correlation in the outcome over time by clustering our standard errors at the city level.

In addition to estimating the average effect of an additional woman, we will also present estimates for two important sub-samples. The first sub-sample, which we call the "1 Non-Focal Woman," sample, corresponds to councils where the gendered election has the potential to add a second female to a council that would otherwise have only one female member. This cut of the data is motivated both by classic work on critical mass theory (e.g. Kanter, 1977; Bratton, 2005) and by existing experimental work which finds that isolated, "token" women behave differently in group settings (Stoddard et al., 2020).

The second sub-sample, which we call the "> 1 Non-Focal Women" sample, corresponds to councils where the gendered election has the potential to add a woman to an environment where there are already at least two non-focal women. Critical mass theory predicts that, once the threshold has been breached, there should be no further behavioral responses over and above the mechanical effect of replacing a man with a woman for this sample (see model in section 5 for a more formal version of the argument). Therefore, this sub-sample is interesting because it allows us to explore this implication of the classic theory.

2.2 Validity of the Regression Discontinuity Design

Our key identifying assumption, which is that the conditional expectation functioning mapping the running variable into outcomes would be continuous across the threshold in the absence of treatment, has testable implications (Lee and Lemieux, 2010). First, it implies that the sample should be balanced on baseline and otherwise exogenous characteristics at the threshold. Table 3 provides evidence of balance for the full sample, and the two sub-samples of interest, by putting baseline and otherwise exogenous council characteristics on the left-hand side of our preferred model (2.1). We do not find evidence of sample imbalance, individually or jointly, for the full sample or the 1 Non-Focal sample. For the >1 Non-Focal sample, we find evidence that two characteristics may exhibit discontinuities across the threshold which may raise some concern for this sub-group; however, the joint test cannot reject that all coefficients are equal to zero for this sample which is consistent with two significant coefficients representing false positives in light of the many coefficients checked for balance in table (2.1).

Figure 3 plots a histogram of the running variable for our full sample. As first noted by McCrary (2008), the RD identifying assumption implies that councils should be un-

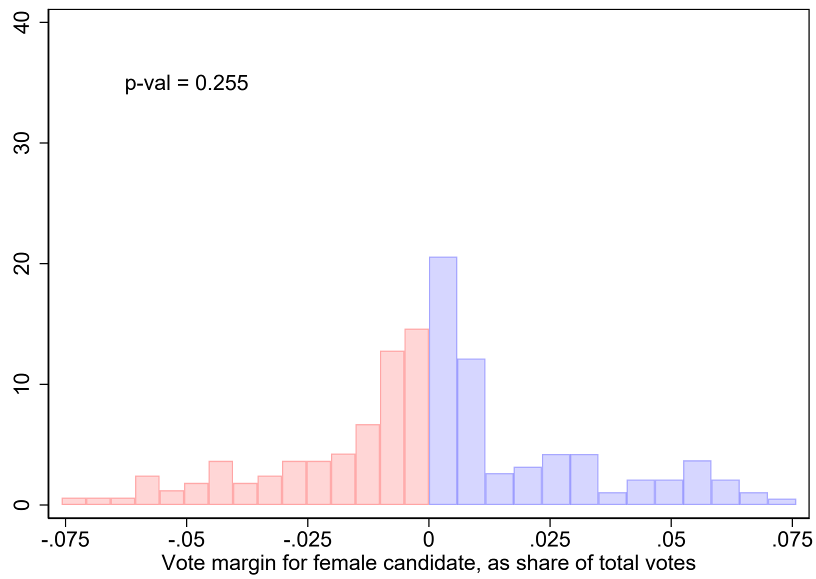
Table 3: Balance

	All	1 Non-Focal	> 1 Non-Focal
Council size	0.136	-0.329	0.512
	(0.470)	(0.373)	(0.997)
Term duration	0.093	0.084	0.209
	(0.106)	(0.182)	(0.181)
Turnover	-0.128*	-0.127	-0.272**
	(0.065)	(0.082)	(0.137)
Candidates	0.215	-1.047	2.317
	(0.895)	(1.277)	(1.557)
Female candidates	0.145	-0.765*	1.240**
	(0.316)	(0.435)	(0.543)
Races	0.002	-0.083	0.377
	(0.252)	(0.280)	(0.496)
Total population	8.734	-16.809	40.595
	(26.410)	(20.257)	(61.470)
Black (share)	0.021	-0.002	0.041
	(0.019)	(0.014)	(0.043)
Hispanic (share)	-0.027	0.005	-0.088
	(0.050)	(0.097)	(0.078)
Per capita income	-3.027	-3.401	4.114
	(4.887)	(9.546)	(5.491)
Joint P-value	0.209	0.332	0.179
Observations	325	136	124

Note: This table presents evidence in support of our identifying assumption for the entire sample (All), for the sub-sample of councils with only one non-focal member (1 Non-Focal), and for the sub-sample of councils with more than one non-focal member (>1 Non-Focal). Each number in this table is estimated using a separate regression and our preferred model (equation 2.1). Rows denote baseline and other exogenous council characteristics which were placed on the left hand side of model 2.1. Standard errors, clustered at the city level, are reported in parentheses. P-values from a joint test across outcomes are reported in the second to last row. "Observations" provides the number of councils contained in each column. Stars denote statistical significance as follows: *** p<0.01, ** p<0.05, * p<0.1.

able to manipulate the running variable to determine treatment status. This implies that the density of the running variable should be continuous around the RD threshold and is confirmed visually in figure 3. We test this formally using the method described in McCrary (2008) and find no evidence of manipulation ($P = 0.255$).

Figure 3: Density of Councils around RD Threshold, Full Sample

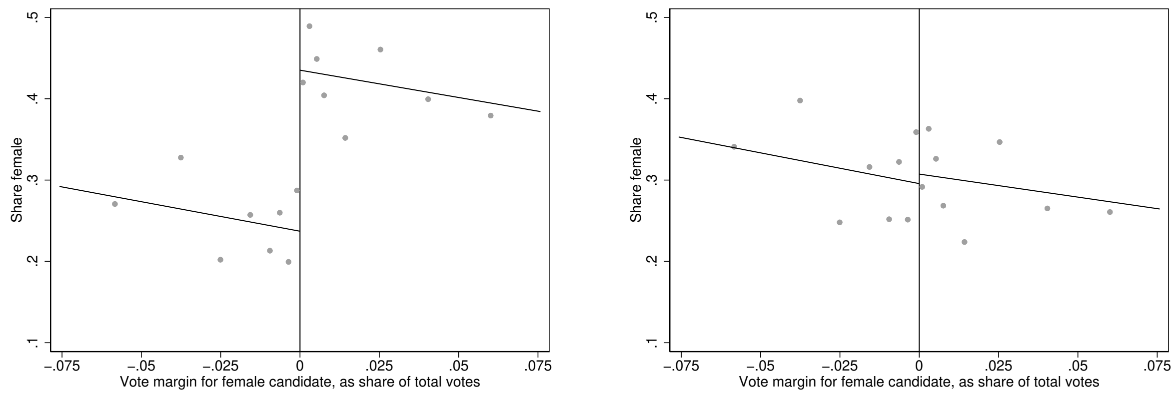


Note: This figure plots the density of cities represented in our election sample by the vote margin of the female candidate in the cross-gender election. The p-value represents the result of a formal test for continuity of density across the threshold as described in McCrary (2008).

2.3 "First Stage" Effects on Nominal Representation

Figure 4 provides evidence that winning a close election does, in fact, generate consequential changes in the gender composition of the council. Panel (a) plots the overall share of councilors that are female as a function of the vote-differential running variable. In theory, the addition of a woman in a close election could cause other women (or men) to leave the council or retire, which would complicate our interpretation of the treatment effects identified by model 2.1; however, we find that crossing the threshold causes a 20 percentage point increase in the overall female share. Given an average council size of ≈ 5.5 members (see table 1), this effect is consistent with a change in the gender of one member from male to female. Consistent with this finding, in panel (b) of figure 4, we show that a female winning a close election has no impact on the gender composition of the other members. This also serves as a natural "placebo" test for our preferred interpretation of the estimates as representing the impact of randomly adding and additional woman to the council, since we panel (b) suggests that there are no *additional* downstream changes in council gender composition that are caused by a female close election win.

Figure 4: Discontinuities in Nominal Representation



(a) Female Share: All Councilors **(b)** Female Share: Non-Focal Councilors

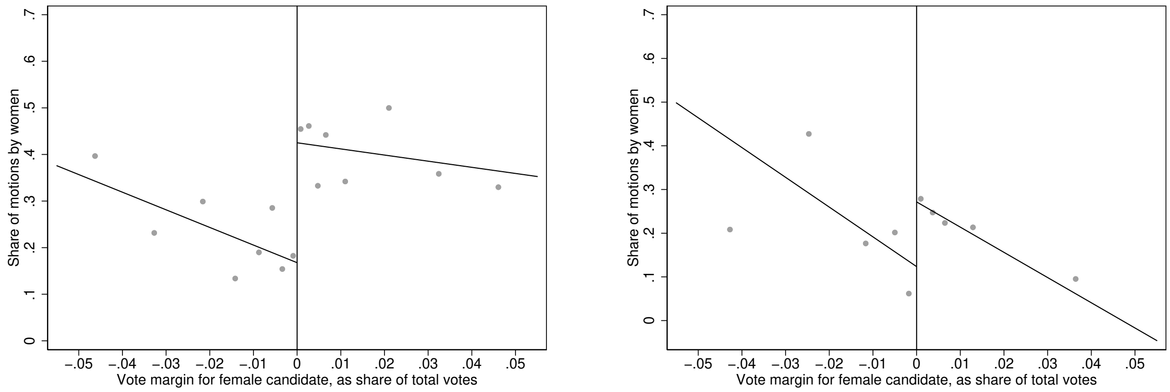
Note: This figure plots the running variable (vote-share normalized to be zero at the cutoff) against measures of council gender composition within the RD bandwidth. Panel (a) plots the overall share of the council that is female. Panel (b) is identical to panel (a) except that it excludes the councilor involved in the close election when calculating the female share. It therefore serves as a natural "placebo test" that any effects we see must operate through the addition of a new female councilor as opposed to other downstream changes in the composition of the other councilors.

3 Findings

3.1 Overall Effects on Substantive participation

We find evidence that nominal representation does translate into substantive participation. Figure 5 shows visual RD evidence of effects on our key outcomes related to women's substantive participation in decision making. Table 4 provides corresponding point estimate from model 2.1 for these key outcomes and others across three samples: the full sample of councils where we observe meeting data, the sub-sample of these councils with 1 non-focal woman, and the sub-sample of councils with multiple non-focal women.

Figure 5: Discontinuities in Substantive Participation



(a) Share of Motions made by Women **(b)** Share by Non-focal Women ($N_w = 1$)

Note: This figure plots the running variable (vote-share normalized to be zero at the cutoff) against key outcomes from our meeting-minutes data within the RD bandwidth. Panel (a) displays results for the "Share of motions made by women" and is calculated using all council members. Panel (b) contains results for the "Share of motions made by Non-focal women" for the sample of councils where there is exactly one pre-existing female councilor other than the "focal" councilor involved in the close election.

Panel (a) of figure 5 plots the share of motions made by women in the council against the female vote margin running variable. We see a large jump in the number of motions made by women precisely at the female victory threshold. Table 4 reveals that this jump corresponds to an increase of 25p.p., which is large relative to the baseline average female share in this sample of 30p.p. (hence over 80% of the mean). Perhaps more surprisingly, this increase is also 25% larger than the overall impact that a female victory has on overall female representation on the council, a point we return to in section 3.2.

We find similar patterns when examining the number of motions seconded by women, and joint outcomes that describe both the number of motions made and seconded by women. We find broadly similar patterns when examining sub-samples of our data that do (and do not) have low levels of baseline female representation. For both the 1 non-focal woman sample, and the > 1 non-focal women sample, we find that nominal representation does translate into substantive representation.

3.2 Behavioral Effects on Non-focal Councilors

We find evidence that increasing female representation causes consequential changes in the tendency for other women to participate. As noted in section 3.1, the impact of adding

Table 4: The Impact of Representation on Substantive Participation

	Full sample (1)	1 Non-Focal Woman (2)	> 1 Non-Focal Women (3)
No. of Motions	-0.091	-0.977	1.987
	(0.902)	(1.451)	(1.254)
	[10.095]	[10.374]	[9.877]
Share moved by women	0.247***	0.301***	0.156*
	(0.048)	(0.054)	(0.080)
	[0.226]	[0.182]	[0.379]
Share seconded by women	0.235***	0.267***	0.214***
	(0.055)	(0.068)	(0.074)
	[0.258]	[0.222]	[0.417]
Share moved or seconded by women	0.335***	0.394***	0.217***
	(0.060)	(0.072)	(0.083)
	[0.371]	[0.338]	[0.574]
Share moved and seconded by women	0.077**	0.084***	0.100
	(0.035)	(0.025)	(0.073)
	[0.050]	[0.000]	[0.134]
Observations	325	136	124

Note: This table presents estimates from our preferred model (equation 2.1) for key outcomes across our three samples of interest. Standard errors clustered by city are presented in parentheses. Corresponding outcome means within each sample are denoted in square brackets. Stars denote statistical significance as follows: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

an additional woman on substantive representation (25p.p.) is considerably larger than the corresponding impact on nominal representation (20p.p.). This pattern becomes more extreme at low levels of baseline representation. A female victory causes a 30p.p. increase in the female vote share in councils with 1 non-focal woman. The corresponding change in nominal representation for this sample is also 20p.p.

What drives this discrepancy? Panel (b) of figure 5 plots the share of motions made by non-focal women for councils where they are isolated. Consistent with a large laboratory literature documenting the behavioral impacts of increased gender representation on lone, "token" women, we find visual evidence that adding an additional female councilor causes the non-focal woman to offer more motions. While less visually stark than the full council figure, there is a clear jump in the female motion share upon crossing the female victory threshold.

Table 5 presents estimates for the impact of a narrow female victory on the behavior of non-focal council members. Importantly, all outcomes in this table are calculated *excluding* the man or woman involved in the close election and therefore identify a pure "behavioral" effect that is not directly connected to the change in nominal representation. The evidence in table 5 confirms that the apparent visual jump in panel (b) of figure 5 is, indeed, statistically significant.

Table 5: Meeting Minutes Results: Non-Focal Councilor Outcomes

	1 Non-Focal Woman		> 1 Non-Focal Women	
	Men (1)	Women (2)	Men (3)	Women (4)
No. of Motions	-0.717	0.856**	0.687	0.467
	(0.906)	(0.362)	(0.881)	(0.765)
	[5.560]	[1.469]	[4.806]	[2.826]
Share moved by women		0.148**		-0.045
		(0.058)		(0.095)
		[0.219]		[0.452]
Share seconded by women		0.088		0.032
		(0.076)		(0.089)
		[0.282]		[0.471]
Share moved or seconded by women		0.219***		0.142
		(0.064)		(0.087)
		[0.270]		[0.481]
Bandwidth	0.055	0.055	0.070	0.070
Observations	136	136	124	124

Note: This table presents estimates from our preferred model (equation 2.1) for key outcomes as measured over the men and women not involved in the close election. Thus these estimates can be interpreted as the pure behavioral effect of a female victory on the other councilors. Standard errors clustered by city are presented in parentheses. Corresponding outcome means within each sample are denoted in square brackets. Stars denote statistical significance as follows: *** p<0.01, ** p<0.05, * p<0.1.

We find that addition of a second woman to a council with a lone token woman causes the share of motions made by the female councilor to increase by 14.8p.p. This is large, constituting a 37% increase relative to the dependent variable mean in this sample (39.4p.p.). It is also large relative to the overall council level effect of 30p.p., suggesting that roughly "half" of the change in substantive representation that accrues from a female victory in this sample comes from the behavioral changes of isolated women rather than the mechanical effect of replacing a male councilor with a female councilor. This change in the share comes about via an increase in the number of motions offered by the non-focal woman of 0.86 per meeting. Point estimates for men are negative and imprecisely estimated.

By way of contrast, we do not find statistically precise evidence that female victories cause additional behavioral changes for non-focal woman when baseline representation is larger than one. In fact, the point estimates suggest the overall share of motions made by females among non-focal members actually *declines* by 4.5p.p. This suggests some possible substitution between focal and non-focal councilors when baseline representation is large. However, we note that the standard errors for non-focal females in councils with high levels of baseline representation are large and hence it is difficult to make precise statistical claims on this point.

3.3 Policy Outcomes

We do not find evidence that the large changes in nominal representation and substantive participation translate into consequential changes in policy. Table 6 plots point estimates for spending per-capita overall and within specific sub-categories.[14]

Table 6 reveals a pattern of consistent and, in most cases reasonably precise, null results. For example, the point estimates for total per-capita spending suggest that replacing a man with a woman causes reductions of 38 million which is just 2% relative to the mean. Even taking the upper and lower bounds of the 95% confidence intervals suggest that the close election changes spending by anywhere from 10 to 14% relative to the mean. Across all of the outcomes considered, only spending on Airports and Water approaches conventional levels of statistical significance. However, even that may be a statistical artifact of the many hypothesis tested in table 6: the P-values from a joint-F test across outcomes cannot reject the hypothesis that all coefficients are zero.

[14]In this case, we augment our preferred model with a control for lagged spending, which improves our precision considerably. However, we obtain similar conclusions (albeit with larger standard errors) if we omit this control.

Table 6: Effects on Per Capita Municipal Spending

	Total (1)	Public Utility (2)	Health and Hospital (3)	Parks and Recreation (4)	Library (5)	Housing and Com Dev (6)	Airports and Water Ports (7)	Police and Fire (8)	Sewerage and Waste (9)	Roads and Parking (10)
Panel A: All councils										
RD estimate	-38.452	-2.012	6.023	-14.833	-5.115	19.289	16.717**	11.885	-41.357	-44.890
	(123.392)	(44.313)	(9.443)	(13.942)	(5.177)	(40.067)	(8.363)	(20.837)	(55.610)	(38.663)
	[1943.420]	[346.339]	[43.541]	[96.842]	[30.700]	[82.654]	[16.707]	[412.077]	[236.724]	[231.033]
Panel B: 1 non-focal women										
RD estimate	101.065	-19.421	-5.829	-21.668	-12.070**	103.363	17.964*	40.074	-0.296	-4.017
	(115.167)	(54.343)	(15.453)	(23.146)	(4.690)	(92.616)	(9.697)	(41.725)	(32.759)	(28.196)
	[2015.504]	[320.570]	[78.754]	[105.677]	[33.643]	[94.113]	[32.201]	[402.336]	[268.547]	[245.980]
Panel C: >1 **non-focal woman**										
RD estimate	-147.084	-82.800	1.860	8.677	-8.021	-16.694	2.894	-2.103	-4.974	-16.891
	(128.751)	(66.792)	(1.807)	(13.891)	(10.932)	(19.533)	(2.997)	(14.655)	(38.811)	(25.149)
	[1944.416]	[347.455]	[17.268]	[88.327]	[33.561]	[70.150]	[2.730]	[431.536]	[196.802]	[206.931]

Note: This table presents estimates from our preferred model (equation 2.1) for spending outcomes. Columns denote the type of expenditure. Rows denote the sample used for estimation. Standard errors clustered by city are presented in parentheses. Corresponding outcome means within each sample are denoted in square brackets. Stars denote statistical significance as follows: *** p<0.01, ** p<0.05, * p<0.1.

4 Mechanisms

Why does nominal representation lead to large changes in substantive participation that do not translate into shifts in the composition of public policy? In this section, we consider several candidate explanations: strategic voting behavior, party affiliation, and preferences. We find little evidence that voting behavior or party affiliation play an important role. However, we find evidence that close alignment of preferences between men and women could rationalize these patterns.

4.1 Strategic Voting Behavior

Strategic voting is one candidate explanation that could rationalize these patterns. For example, it could be that change in group composition also creates important changes in the alignment of coalitions within the council such that the new motions brought by women ultimately are voted down and fail to pass.

We explore this possibility in table 7, where we show the causal effect of a close female victory on various outcomes meant to capture voting patterns with respect to motions. We do not find evidence in favor of this explanation. For virtually all outcomes considered, in table 7 we find precisely estimate null effects. For example, our null result on the share of unanimous motions is precise enough to rule out changes larger than 6p.p. (approximately 6% of the dependent variable mean). We find broadly similar patterns when we explore other outcomes like number of votes for or against the motion and the difference between votes for and against (i.e. the vote margin).

One possible interpretation of the null results for our voting outcomes is that there are real changes to the content of the bills being debated and passed by the council, but that this change is masked in the voting process due to some change in the number or composition of the motions actually brought to the floor for a vote. We provide three pieces of evidence against this interpretation. First, table 4 explores the impact of a female victory on the total number of motions brought to the floor. The point estimate suggest that the change in the number of motions is not large, roughly -0.091 motions per meeting relative to a mean of 10.1; however, we acknowledge that the lower bound of the corresponding 95% confidence interval is not small enough to rule out changes which are large in magnitude. Second, we show in the section 3.3 that female victories also do not result in consequential changes in public spending, which is consistent with female victories having no substantive change in the public policy contained in the motions that

Table 7: Effects on Voting Patterns

	Full sample (1)	1 Non-Focal Woman (2)	> 1 Non-Focal Women (3)
Share unanimous	0.005	-0.015	0.011
	(0.026)	(0.052)	(0.033)
	[0.899]	[0.894]	[0.905]
Share rejected	0.010	-0.001	0.022
	(0.022)	(0.046)	(0.025)
	[0.928]	[0.921]	[0.938]
Vote margin	0.012	-0.084	0.061
	(0.138)	(0.252)	(0.215)
	[5.688]	[5.350]	[6.265]
Votes in favor	0.007	-0.062	0.036
	(0.072)	(0.140)	(0.111)
	[5.974]	[5.633]	[6.551]
Votes against	0.003	0.010	0.027
	(0.044)	(0.068)	(0.076)
	[0.160]	[0.154]	[0.173]
Observations	325	136	124

Note: This table presents estimates from our preferred model (equation 2.1) for spending outcomes. Columns denote the type of expenditure. Rows denote the sample used for estimation. Standard errors clustered by city are presented in parentheses. Corresponding outcome means within each sample are denoted in square brackets. Stars denote statistical significance as follows: *** p<0.01, ** p<0.05, * p<0.1.

are brought to the floor and passed. Third and finally, in section 4.2, we will explore the impact of female victories on the topics of motions brought to a vote before the council and find no evidence of an effect. This is also consistent with female victory having no real impact on the substance of the public policy debated on or passed by the council.

4.2 Preference Alignment

In this section, we present evidence that our null results on spending and voting may be driven by an alignment of preferences between men and women. Table 8 shows the share of motions offered by male and female councilors in our sample broken down by the topical category of the motion. Note that motions can be categories as having multiple topics, thus the numerical values in this table need not sum to one.

As is evident in the table, men and women in our sample appear to offer motions pertaining to specific issues at similar rates. In fact, across the range of topics in our data, we find no statistically significant difference in the share offered by men and women, individually or jointly. In addition, in ongoing analysis (available upon request), we find that a woman winning a close election against a man also does not impact the types of motions that are brought to the floor. Thus, it appears as if men and women have broadly similar policy preferences, at least regarding the types of motions they tend to bring to the floor for a vote.

5 A Stylized Model of Critical Mass

In this section, we present a stylized model of critical mass to help interpret our results. Our model is motivated by the institutional details discussed in section 1.1.

Consider a city council engaged in a collaborative decision that determines public policy. The process begins with a potential project arriving before the council. Concretely, this project could emerge via a mandate from a higher level of government, from a petition filed by a constituent, or from the routine business regularly preformed within the council.

When a project arrives before the council, individual councilors sequentially choose whether to participate in a committee (denoted by I) that will determine the project's overall quality. Denote each councilor's idiosyncratic valuation of the project by u_i. If no one volunteers to join the committee, the project fails and councilors receive the status quo utility u_0. If a project is taken up by the committee, then each council member on the committee contributes an additional value of δ_i to the project, at personal cost c_i.

Once a project is finished in the committee process, a councilor (chosen at random) proposes a motion to approve the project. If the motion is approved by a majority, each councilor receives the following utility

$$U_i = u_i + \sum_{j \in I} D_j \delta_j$$

where D_j is an indicator for councilor j's participation.

If the motion to approve the project fails, council members receive u_0.

Table 8: Gender Differences in Motion Topics

	Moved by men	Moved by women	Difference
Admin	0.581	0.580	-0.002
			(0.008)
Finance	0.206	0.219	0.014
			(0.035)
Regulation	0.100	0.092	-0.009
			(0.005)
Property	0.183	0.175	-0.008
			(0.006)
Public Utilities	0.133	0.129	-0.004
			(0.005)
Health	0.009	0.012	0.002
			(0.002)
Parks	0.030	0.031	0.001
			(0.003)
Libraries	0.005	0.005	0.000
			(0.001)
Police	0.033	0.033	0.000
			(0.003)
Airports	0.005	0.006	0.001
			(0.001)

Note: This table shows the average share of motions offered by men and women broken down by the category of motion. Note that motions can fall into multiple categories; hence, the coefficients in this table need not sum to 1. Standard errors for the gender difference are displayed in parentheses.

We consider the unique subgame-perfect equilibrium of the game, assuming members vote truthfully and that whenever a member is indifferent between participating or not, they break the tie in favor of participating.

Preferences and Costs for Men and Women

Suppose that the council can be partitioned into two identity groups: a majority and a minority group. For simplicity and to match our setting and data, let the majority group be men ($M \subseteq I$), and we shall refer to the minority group as women ($W \subset I$)[15], although any underrepresented identity may fit the model. For simplicity, assume that for each project, all men have identical policy preferences given by $u_i = u_m$ and that all women have identical policy preferences given by $u_i = u_w$.

Motivated by the large experimental literature which documents that women are less likely to contribute to group decision making processes when they are in the minority (Stoddard et al., 2020; Coffman, 2014; Coffman et al., 2021), we assume that councilor i faces a cost of participating $c(N_i)$, which is strictly decreasing in N_i, the number of councilors that share i's gender. For simplicity, let $c(N_i) = \frac{C}{N_i}$. Also assume that all councilors contribute equal values to the committees they join so that $\delta_i = \delta$ for all i. We show that women's participation disproportionately increases once the number of women reaches a critical mass N^*, given by $\frac{C}{N^*+1} \leq \delta < \frac{C}{N^*}$. Our empirical results and the tokenism literature (Stoddard et al., 2020) are consistent with $N^* = 2$, however our model allows for any $N^* \geq 2$.

Finally, assume that women choose first.[16] We now present simple testable predictions, focusing on the on-the-equilibrium-path outcomes.[17]

5.1 Comparative Statics and Testable Predictions

The model presented above yields testable predictions as the number of women crosses the critical mass threshold. Compared to the council with a token woman, a more gender-balanced council sees an increase in women's participation (and as a result the number of motions she offers) that is larger than the expected "mechanical" effect of replacing a man

[15]In our data, less than XXX% of city councils is majority female, which limits us from investigating female majority effects.

[16]This does not affect our qualitative predictions, but rules out some borderline cases that depend on the sequential order

[17]A complete description of councilors' equilibrium strategies is provided in Appendix XX.

with a woman. However, the model's predictions regarding the types of motions that are brought to a vote and the outcomes of these votes depends critically on the alignment of male / female preferences with respect to the projects brought before the council.

Our results show a differential effect of achieving critical mass on two types of projects: gender-neutral, defined as $|u_m - u_w| \leq \varepsilon$, and gendered, where $|u_m - u_w| > \varepsilon$, for some $0 < \varepsilon < \delta$. Gendered projects can be female-preferred ($u_w > u_m + \varepsilon$) or male-preferred ($u_m > u_w + \varepsilon$).

PROPOSITION 1. Participation. *When critical mass is achieved, female participation increases disproportionately in both gender-neutral and gendered projects.*

If $N_w = 1$, the woman councilor participates only in female-preferred projects where her participation is pivotal, withholding participation in male-preferred projects. If $N_w \geq 2$, then women participate at the same rate as men in all projects.

PROPOSITION 2. Non-unanimous votes. *Whether or not critical mass is achieved, gender-neutral and female-preferred motions are passed unanimously conditional on being proposed. When critical mass is achieved, there are fewer non-unanimous male-preferred motions but more dissenting votes per non-unanimous motion.*

Any motion proposed with $d(u_m, u_w) \leq \varepsilon$ (FIND ε) is passed unanimously for all $0 \leq N_w \leq N$.

PROPOSITION 3. Motion projects. *When critical mass is achieved, women increase their participation in all gender-neutral projects proportionally. In gendered projects, women switch from participating only in female-preferred projects to participating in both female- and male-preferred projects.*

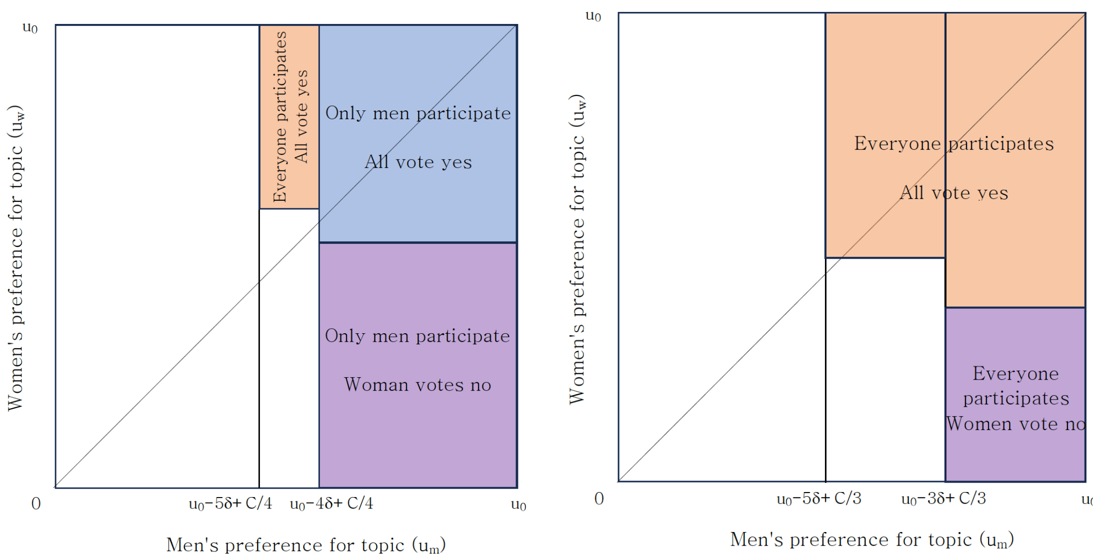
For projects where preferences are not *too* dissimilar between men and women,[18] the model predicts a small decrease in the likelihood of a motion reaching the floor, and no change in voting patterns among motions brought to a vote. In contrast, if preferences differ significantly across genders,[19] then as more women join the council, there may be: (1) a decrease in non-unanimous motions, but more dissenting votes per non-unanimous

[18] Projects that fall along the 45-degree line in figure 6

[19] Projects fall outside the 45-degree line in figure 6

motion; or (2) Women switching from selective participation in a narrow range of women-oriented projects to participating widely in all projects.

Figure 6: Preferences, Participation and Votes in a 5-Member Council



(a) $N_w = 1$ **(b)** $N_w = 2$

6 Conclusion

We examine the causal effect of gender representation on decision-making and municipal spending by assembling an ambitious dataset on city council elections, meetings and expenditures in the United States. By exploiting close gendered elections as an exogenous source of gender variation, we estimate causal effects of gender at different margins of representation. Our results provide evidence to support theories of tokenism and critical mass, specifically an effect on other women's propensity to propose motions when a male-majority council switches from one to two women. At that margin of representation, the mechanical effect of increasing women's representation (by around 20p.p.) is accompanied by an almost equal behavioral effect increasing the share of the non-focal woman's motions by 14.8 p.p. This is consistent with the predictions from our stylized model, as well as theories of tokenism, which posit that adding more women can relieve the effects of isolation and change the "forms of relationships and peer culture" within the group (Kanter (1977)).

However, we find that municipal spending patterns are generally not affected by the increased representation of women, consistent with prior evidence about female mayors in the U.S. (Ferreira and Gyourko, 2014). Our ongoing analysis seeks to reconcile these two facts by exploring the topics of the motions proposed by male and female councillors. Although not causal, Table 8 shows the distribution of topics proposed by men and women. There are notably few if any differences in the overall topics proposed by gender. We view this as suggestive evidence that men and women may not have substantially defenses in political preferences in the local U.S. context. Women's participatory respose to increased diversity might then be a response to social or cultural factors rather than the political content of the discussion.

There are important implications for policies aiming to increase women's representation in policy-making settings. Our findings support the well-established notion that the effectiveness of such policies is context-dependent. In particular, policies that only add one woman to an otherwise all-male group may work very differently from policies that lead to a more substantial level of representation. Our research highlights additional benefits of increased representation, in the form of positive behavioral effects on other women.

References

Kenneth R. Ahern and Amy K. Dittmar. The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, 127(1):137–197, 2012. ISSN 00335533, 15314650. URL http://www.jstor.org/stable/41337208.

Brian Beach and Daniel B. Jones. Gridlock: Ethnic diversity in government and the provision of public goods. *American Economic Journal: Economic Policy*, 9(1):112–36, February 2017. doi: 10.1257/pol.20150394. URL http://www.aeaweb.org/articles?id=10.1257/pol.20150394.

Pedro Bordalo, Katherine Coffman, Nicola Gennaioli, and Andrei Shleifer. Beliefs about gender. *American Economic Review*, 109(3):739–73, March 2019. doi: 10.1257/aer.20170007. URL https://www.aeaweb.org/articles?id=10.1257/aer.20170007.

Andreas Born, Eva Ranehill, and Anna Sandberg. Gender and Willingness to Lead: Does the Gender Composition of Teams Matter? *The Review of Economics and Statistics*, 104 (2):259–275, 03 2022. ISSN 0034-6535. doi: 10.1162/rest_a_00955. URL https://doi.org/10.1162/rest_a_00955.

Kathleen A. Bratton. Critical mass theory revisited: The behavior and success of token women in state legislatures. *Politics Gender*, 1(1):97–125, 03 2005. URL https://login.usmalibrary.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/critical-mass-theory-revisited-behavior-success/docview/194681330/se-2. Copyright - Copyright Cambridge University Press Mar 2005; Last updated - 2018-10-06.

California State Controller's Office. Cities financial data, 2019. URL https://cities.bythenumbers.sco.ca.gov/#!/year/default.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326, 2014.

Raghabendra Chattopadhyay and Esther Duflo. Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443, 2004. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/3598894.

Jingnan Chen and Daniel Houser. When are women willing to lead? the effect of team gender composition and gendered tasks. *The Leadership Quarterly*, 30(6):101340, 2019. ISSN 1048-9843. doi: https://doi.org/10.1016/j.leaqua.2019.101340. URL https://www.sciencedirect.com/science/article/pii/S1048984318304260.

Katherine Coffman, Clio Bryant Flikkema, and Olga Shurchkov. Gender stereotypes in deliberation and team decisions. *Games and Economic Behavior*, 129:329–349, 2021. ISSN 0899-8256. doi: https://doi.org/10.1016/j.geb.2021.06.004. URL https://www.sciencedirect.com/science/article/pii/S0899825621000798.

Katherine Baldiga Coffman. Evidence on self-stereotyping and the contribution of ideas *. *The Quarterly Journal of Economics*, 129(4):1625–1660, 2014. doi: 10.1093/qje/qju023. URL http://dx.doi.org/10.1093/qje/qju023.

Fernando Ferreira and Joseph Gyourko. Does gender matter for political leadership? the case of u.s. mayors. *Journal of Public Economics*, 112:24 – 39, 2014. ISSN

0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2014.01.006. URL http://www.sciencedirect.com/science/article/pii/S0047272714000073.

Stefano Gagliarducci and M. Daniele Paserman. Gender interactions within hierarchies: Evidence from the political arena. *The Review of Economic Studies*, 79(3):1021–1052, 2012. ISSN 00346527, 1467937X. URL http://www.jstor.org/stable/23261378.

Niall Hughes, Zia-Ul-Hassan Khan, and Friederike Mengel. Diversity in committees. Working Paper, 2023. URL https://www.dropbox.com/s/5nbf4haflvfco4f/Diversity_28_June_2023.pdf?dl=0.

Institute for Social Research at California State University, Sacramento. California elections data archive (ceda), 2007-2015. URL https://csu-csus.esploro.exlibrisgroup.com/esploro/outputs/dataset/California-Elections-Data-Archive-CEDA/99257830890201671?institution=01CALS_USL.

Jasmin Joecks, Kerstin Pull, and Karin Vetter. Gender diversity in the boardroom and firm performance: What exactly constitutes a ”critical mass?”. *Journal of Business Ethics*, 118(1):61–72, 2013. ISSN 01674544, 15730697. URL http://www.jstor.org/stable/42921212.

Rosabeth Moss Kanter. Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, 82(5):965–990, 1977. ISSN 00029602, 15375390. URL http://www.jstor.org/stable/2777808.

David S Lee. Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697, 2008.

David S. Lee and Thomas Lemieux. Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355, June 2010. doi: 10.1257/jel.48.2.281. URL https://www.aeaweb.org/articles?id=10.1257/jel.48.2.281.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008. ISSN 0304-4076. doi: https://doi.org/10.1016/j.jeconom.2007.05.005. URL https://www.sciencedirect.com/science/article/pii/S0304407607001133. The regression discontinuity design: Theory and applications.

National League of Cities. National League of Cities - City Councils. https://www.nlc.org/city-councils/, Accessed 2023.

Social Security Administration. National data on the relative frequency of given snames in the population of u.s. births where the individual has a social security number, 2017. URL https://www.ssa.gov/oact/babynames/limits.html.

Olga B. Stoddard, Christopher F. Karpowitz, and Jessica Preece. Strength in numbers: A field experiment in gender, influence, and group dynamics. Iza discussion paper no. 13741, September 2020.

United States Census Bureau. Annual survey of state and local government finances: Historical data, 2017.

United States Census Bureau. Annual survey of state and local government finances: Individual unit files, 2017-2019.

Urban Institute. Urban Institute - State and Local Revenues. https://www.urban.org/policy-centers/cross-center-initiatives/state-and-local-finance-initiative/state-and-local-backgrounders/state-and-local-revenues, Accessed 2023.

Melanie Wasserman. Gender differences in politician persistence. Working paper. URL https://static1.squarespace.com/static/57bdef2ecd0f687f9fbfec30/t/6112ee007a11e63e81048ab1/1628630534669/politicians-2021-06-30-formatted.pdf.