Where Women Win: Closing the Gap in Congress

TECHNICAL REPORT

Political Parity

Author and Lead Researcher

Shauna Shames

Acknowledgements

We also extend thanks to
the following contributors:

Abbie Erler

Holly Boux

Tess Wise

Danielle Thomsen

Barbara Palmer and Dennis Simon

Johanna Ettin

Marni Allen

Malliga Och

Political Parity

Background and Methodology

After Political Parity released its groundbreaking "Twin States" research in 2013, revealing the phenomenon that states with one woman in a top office are more likely to elect other high-level women, we wondered if this was true at the congressional level as well. Beginning in 2014, our "Where Women Win" project was supervised by Dr. Shauna Shames of Rutgers University - Camden, with statistical expertise from Dr. Abbie Erler of Kenyon College and consultation from Political Parity staff and Parity Research Director Malliga Och.

Dataset Construction

Political Parity contracted with several independent researchers to produce a dataset of congressional candidates between 1980 and 2012. To enable statistical analysis and ensure adequate tests of this project's central questions, we collected political and demographic information for congressional districts (CD), as well as political and demographic information for each candidate (primary and general election) for that CD's House seat in each year of our time period. The unit of analysis in the dataset is the individual candidate, although for certain purposes we also used the CD as a unit of observation.

Dataset Composition

On the CD level, we collected the following variables for each election cycle in the time period:

- total population of CD for that year
- total white population
- total black population
- total Asian population
- total Hispanic population
- size of CD in square miles (varies with redistricting)
- district partisanship (percent identifying as Democrat in CD for that year)
- type of primary system
- open seat or incumbent
- population in CD over 65
- population unemployed
- population in primary/secondary school
- urban vs. rural in CD
- whether or not CD contained a large city (100,000 or more)
- median income in the CD for that year
- population of blue-collar workers in CD for that year
- percent women in state legislature in that year
- percent women in state legislature leadership in that year
- number of "top women" in state in that year (governor or senator(s))

To this district-level dataset, we added political and candidate-specific information for all primary and general election candidates, including:

- name
- gender
- party
- incumbent, challenger, or open-seat
- whether or not candidate won the primary
- whether or not candidate won the general election
- whether or not candidate won a special election
- whether or not candidate won a run-off election

For female candidates in 2012 only, we constructed a separate and smaller dataset to assess race and whether the district was majority-minority.

For most "district demographic" variables, the source of data was the U.S. Census, including the American Community Survey and other Census Bureau sub-units. For political variables, we relied on existing datasets, especially the DIME (Database on Ideology, Money in Politics, and Elections), constructed by Dr. Adam Bonica of Stanford University.[1] This dataset includes information on the average district partisanship votes in most of the election years in our time period. We supplemented this dataset with information on candidate gender, incumbency status, and whether the candidate won or lost the primary and the general election, which was provided by Danielle Thomsen, then a post-doctoral Fellow at Duke University, whose dissertation necessitated the collection of a similar dataset.

If the list of candidates differed between these two datasets for the same CD in the same election, we used the candidate names supplied by the Thomsen dataset, which is based on the "America Votes" dataset. We prioritized the Thomsen data as the DIME/Bonica dataset is based on information pulled directly from FEC files, which can introduce errors for two reasons: 1) candidates file papers but don't always run, and 2) candidates run but do not raise enough to file an FEC report. The resulting Parity dataset is a compilation of multiple sources and the most reliable data available. The final version of this merged dataset, completed in July 2015, will be available for free public access starting June 15, 2016 at politicalparity.org.

The resulting dataset contains more than 23,000 candidate observations across three decades for more than 12,000 district-year pairs. We attempted to correct any errors that appeared systematic, but multiple individual-level errors were likely in the previous datasets on which we based our dataset. There is also a large amount of missing data, particularly in the "incumbency status" variable. When possible, we corrected errors through spot-checking, but it was impossible to systematically reconstruct prior databases.

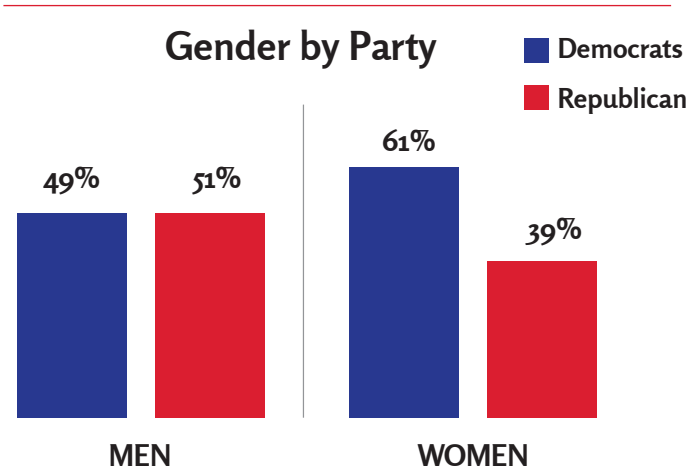
Statistical Analysis

Once the dataset was constructed, cleaned, and spot-checked, we commenced statistical analyses. The results in this report come from extensive tests on the dataset, moving from simple to complex statistics. We began with descriptive statistics of all variables and correlation tests between variables. For the major analysis phase, we conducted time-series and other types of regression analyses. This public version of the report does not report in full the results of all quantitative tests; if interested in greater numerical detail, please contact Political Parity.

1 More information on the DIME Dataset is available at data.stanford.edu/dime

Women as Candidates in Overall Dataset (1980-2012)

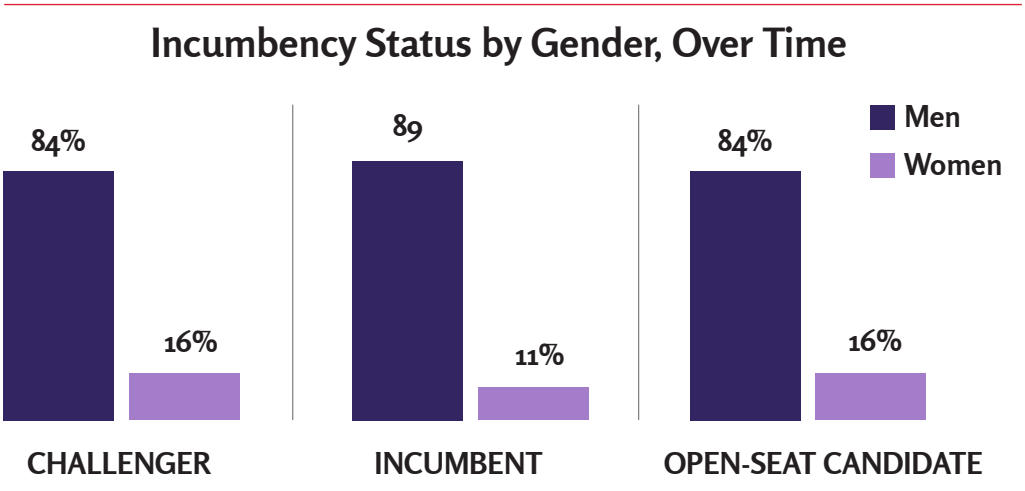
In total, we have data on 23,709 candidates who ran in congressional primaries and general elections between 1980 and 2012 across all 50 states (not including the territories or DC). While the number of female candidates dramatically increased after 1992, the "Year of the Woman" in American politics, in the overall time span, women constitute only 13% of congressional candidates. These female candidates were more likely to be Democrats than Republicans; 61% Democrat and 39% Republican. Male candidates were fairly evenly split across the two parties.



Across time, 70% of the districts in the sample did not have a female candidate, 23% had one female candidate, and 4.8% had two female candidates. About 2% had more than two female candidates. A female candidate ran in a primary election in 30% of the districts, in a general election in 21%, and won in 10.2% of these districts.

Looking at party differences, 20.5% of districts had a female Democratic candidate in the primary, 13.3% had a female Republican candidate, 14% had a Democratic woman in the general election, and 8.7% of districts had a GOP woman in the general. In 6.7% of the districts, a Democratic woman won the general, while a Republican woman won in 3.4%.

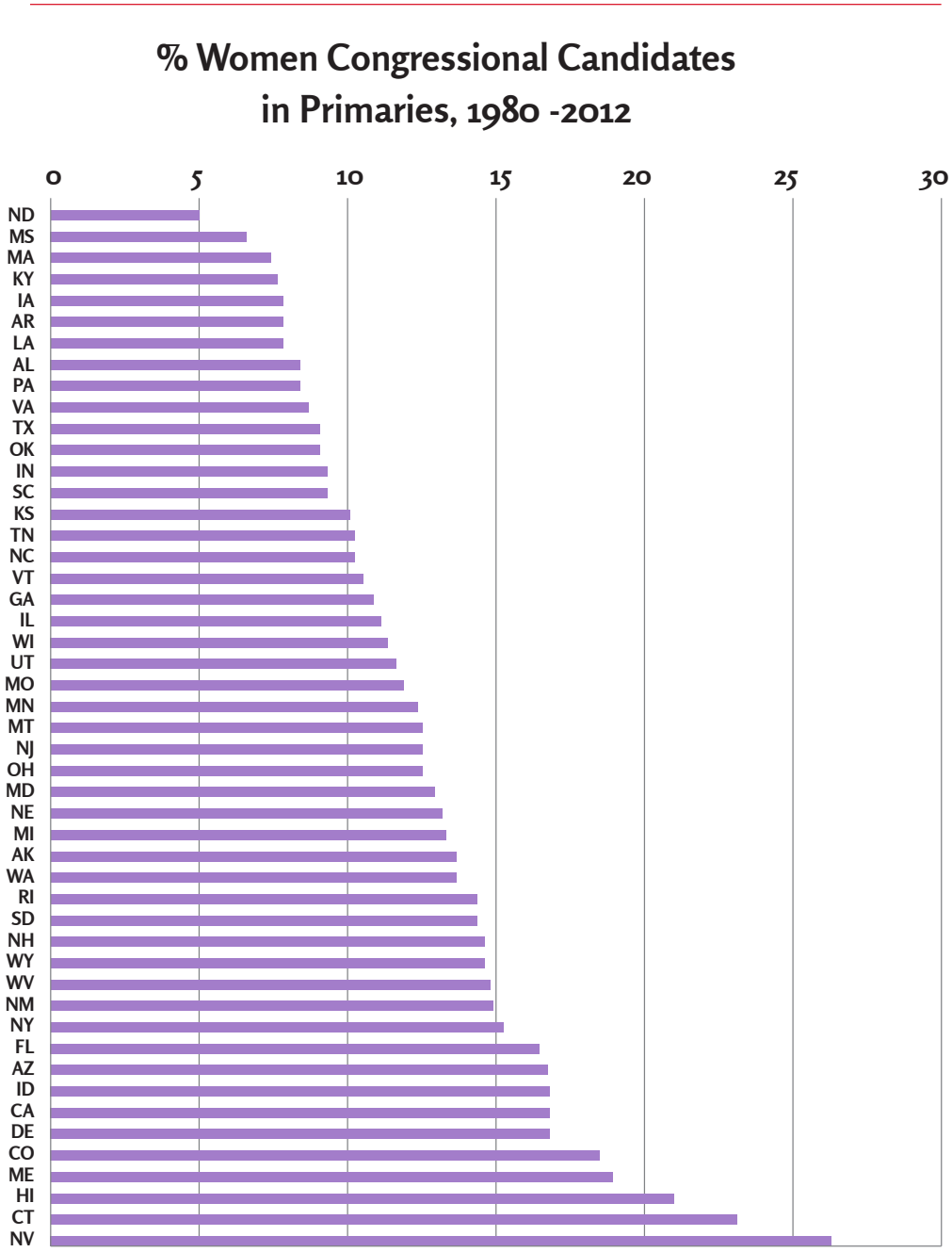
In terms of win-rates, we found that male and female candidates are equally likely to win their primaries, but men are far more likely to win in general election races (probably because they're more likely to be incumbents, and incumbents are far more likely to win than either challengers or open-seat candidates, as previous research has documented[2]). When women are incumbents, they're just as likely as men to win both their primaries and general elections. But women are far more likely to be open-seat candidates and/or challengers than incumbents.



2 Palmer, Barbara and Simon, Dennis: 2008. *Breaking the Political Glass Ceiling: Women and Congressional Elections.* NY: Routledge; see also Palmer, Barbara and Simon, Dennis: 2012. *Women and Congressional Elections: A Century of Change.* CO: Lynne Rienner.

Female candidates are more likely to come from districts that are smaller geographically, measured in square miles, and contain a large city (100,000+ residents) and that are racially diverse (in a simple bivariate correlation, candidate gender is strongly correlated with percent black, percent Hispanic, and percent Asian in the population). Women are less likely to run in districts with a high proportion of whites.

Certain states (particularly those with more women in the state legislature) are far more likely to encourage women to run in congressional primaries. States with the highest proportion of women as primary candidates, in order, include: Nevada, Connecticut, Hawaii, Maine, Colorado, Delaware, California, Idaho, Arizona, and Florida. States with the lowest percentages of women as candidates, in order, include: North Dakota, Mississippi, Massachusetts, Kentucky, Iowa, Arkansas, Louisiana, Alabama, Pennsylvania, and Virginia. (This pertains to female primary candidates, not general election.)



SUCCESSION DATA FOR ELECTED WOMEN 1980-2012

Of the 147 elected women in the dataset, 70 left office. Of those who left: 39 (55.7%) retired, 25 were defeated (35.7%), and six died in office (8.6%). Of those who retired, 14 ran for U.S. Senate, nine lost their bids for higher office, and five won. Of those who ran, nine were Republican. Female Republican members of Congress fared worse in their bids for Senate than Democratic women. Of the nine Republicans who ran for Senate, only one was successful (Olympia Snowe of Maine). Of the five Democratic women members of Congress who ran for Senate, four won and one lost (Denise Majette of Georgia).

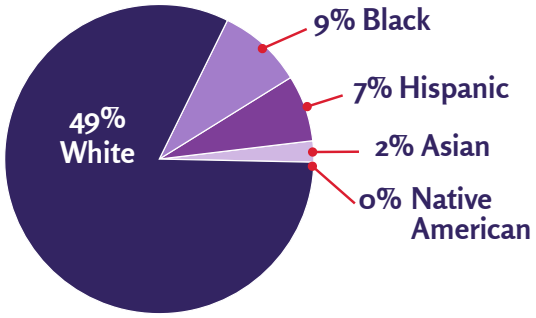
Of the 70 women representatives who left office during this time, only 15 were succeeded by another woman. Of these 15, five retired, two died in office, and eight were defeated. In the cases of retirement and death, the seat was passed on to another woman in their party. In the cases of defeat, two lost in the primary. Both involved black Democratic women (Cynthia McKinney of Georgia and Barbara Rose-Collins of Michigan) who were replaced by black Democratic women. In the other cases of defeat, the seats changed parties – two from Republican to Democrat (Colorado and Ohio) and four seats from Democrat to Republican (Kansas, Florida, Utah, and Washington).

RACIAL DESCRIPTIVE DATA FOR FEMALE CANDIDATES IN 2012

We collected a small and separate dataset on race and ethnicity with 297 observations of female candidates in House primaries in 2012. For each of these candidates, we collected information from CAWP's files on the 2012 election.

Non-white women were very likely to run in majority-minority districts: 61.5% of black female candidates, 29% of Hispanic female candidates, and 86% of Asian American female candidates ran in majority-minority districts. The large majority of white female candidates (87%) ran in white-majority districts. In line with these findings, districts that were majority-minority, in addition to being more female-friendly for candidates, were also very likely to contain a large city (100,000+ residents); 80% of majority-minority districts have a large city, compared with 54% of districts that are not majority-minority.

Female 2012 Primary Candidates, by Race/Ethnicity



If nominated in the primary, non-white women were much more likely than white women to win the general election. This differs by party, with Democratic women winning at a rate of 45% and Republican women at 41%. Women of color, who are mostly Democrats, win general elections at a very high rate probably because they're concentrated in majority-minority districts. Once through the primaries, black women won their general elections at a rate of 87%, Hispanic women at 61.5%, Asian American women at 86%, and white women at a rate of 42.3%.

STATISTICAL ANALYSES | MAJOR FINDINGS

Findings on Women as Candidates Over Time Using Regression Modeling with Controls (1980-2012)

We created statistical models to test several key questions at the House district level, specifically:

1. Where and when women run in primaries
2. Where women *win* in primaries (and thus run in generals)
3. Where women *win* in general elections

We treated each of these as separate dependent variables (DVs), necessitating separate statistical models. We tried several models for each, testing what was and was not significant in the proper model of regression analyses. The correct model changed based on the format of the DV. For a binary DV, we used a logistic regression model, while for a scale-level dependent variable, we used ordinary least squares regression. We also used a "rare events" model to construct the equations testing for when Republican or Democrat women won, as in both cases these were "rare events" happening in only about 10% or less of districts. In addition, we used an interrupted time series model to test for increases in women's candidacies.

These are mostly logistical models with robust standard errors clustered on the congressional districts. We ran the models with dummy variables for each cycle (year) to control for time. We provide initial findings for each DV separately, but some of the same findings appear in several areas, giving us greater confidence in these factors as causal elements in women's political success.

I. Where/When Women *Run* in Primaries in Congressional Districts (DV #1)

To find which factors strongly correlate with women running in a primary in a congressional district, we tested several types of regression models. One was a logistic regression model, using a simple binary dependent variable of whether or not any woman ran in that district's primary election in that year. Separately, we also calculated the percentage of women as primary candidates in that district in that year, as a different way of constructing the dependent variable and testing whether many women ran or only one. In general, because the number of women running varied widely, the second model wasn't as effective at predicting results as the first, simple model. Some of the same results, however, came up in both (see below).

Overall, certain demographic factors in a congressional district (CD) are strongly related to female candidates running in the primary. These factors are similar to those found by Palmer and Simon[3] and include:

- Women are more likely to run in districts with more women in the state legislature
- Women are more likely to run and win an open seat
- Women are more likely to run in districts with fewer blue-collar workers (we also used median income in the district, which worked the same as the blue-collar variable in almost all testing)

3 See Palmer and Simon 2008; see also Palmer and Simon 2012

- Women are more likely to run in districts that are more Democratic (even female Republican candidates)
- Women are more likely to run in districts with a higher percent of people living in urban areas (significant in some, but not all of the models)
- Women are more likely to run in more "compact" districts, based on square mileage (significant in some, but not all models)
- Women are more likely to run in districts with a lower percent of unemployed (significant in some, but not all models)
- Women are more likely to run in districts with a higher percentage of racial minorities (significant in some, but not all of the models)

Looking specifically at the percent of female primary candidates and open seats, the strongest predictor of having more women run in primaries is the percent of women in that state's legislature. (This is even stronger than district partisanship.)

With regards to female primary candidates, our evidence strongly suggests that there is a "trickle up" model of clustering, not a "trickle down" effect. Having one or more women in a top, state leadership position (senator or governor) does *not* seem to make a difference in encouraging women to run in district primaries.

We did find a strong correlation between having a top woman in state office and the percentage of women in the state's legislature, which could mean that high-level women have an indirect effect on women running in district primaries and general elections. This may also imply that having more women in the state legislature leads to more top women. The direction of causality of that correlation is not clear, but the relationship is strong.

When considering female candidates from the two major parties, the significant predictors for whether or not Democratic women run in primaries are similar to those for all women, including:

- A higher percentage of women in the state's legislature
- Open seat (not controlled by an incumbent)
- Fewer blue-collar workers
- More Democratic districts

For Republican women, the same factors are significant, although the "district partisanship" variable, while still showing some consequence, is not as important. This does suggest that Republican women are not as likely to run in more conservative districts and probably don't run in heavily liberal districts either, reducing the scope of where they can run.

II. Predicting Where/When Women *Win* in Primaries in Congressional Districts (and Where Women *Run* in General Elections) (DV #2)

From our analysis of each CD in each year, women tend to *win* primary elections in districts that have:

- A higher percentage of women in the state's legislature (very strongly predictive)
- An open seat (very strongly predictive)
- More Democratic in terms of partisanship
- A large African American population
- Fewer blue-collar workers (or a higher median income)
- Fewer senior citizens

Again, having a woman at the top (senator or governor) did not make a difference in predicting where female candidates win the primary and go on to the general election.

Party-wise, while having more women in the state legislature is strongly significant for Democratic women to win the primary, it's not significant at all for Republican women. And having a top woman in the state as senator or governor negatively affected female Republican candidates in primaries. In several of the models, the percent of African Americans in a district was not statistically significant for women of either party.

III. Predicting Where/When Women *Win* in General Elections in Congressional Districts (DV #3)

From our analysis of each CD in each year, female candidates *win* general elections in districts that have:

- A higher percentage of women in the state's legislature (very strongly predictive)
- An open seat rather than one held by an incumbent (very strongly predictive)
- More Democratic in partisanship

When we break this down by party, some differences emerge. For Democratic women, the percentage of women in the state's legislature is very strongly significant in predicting whether a woman won the general election in that CD in that year. For Republican women, this is still somewhat significant, but not nearly as strong. Democratic women are more likely to win generals in CDs with a major city (100,000+ residents), while this is not true for Republican women.

Many of our "district demographic" predictor variables such as the number of senior citizens and blue-collar workers are not significant for either party when it comes to winning a general election. The big factor is whether or not there is an incumbent. If there is an incumbent holding the seat, almost always a male, he usually wins 98% of the time. As Palmer and Simon have concluded in previous work, incumbency is the "glass ceiling" for women in politics.[4]

4 Palmer & Simon 2008

The racial variables also reverse for Democratic versus Republican women and in primary versus general elections. While a higher percent of minorities appears to be a strong factor in Democratic women running and winning, this is not true for Republican women. In fact, when looking at predictors for Republican women winning general elections, percent black has a negative value, likely because African Americans tend to vote Democrat.

Although Republican women are more likely to run and win their primaries in more politically liberal districts, this reverses for the general election. Female Republican candidates are more likely to win general elections in more conservative districts.

STATISTICAL EVIDENCE FOR "BURST" EFFECT

In our "Twin States" report, we proposed three hypotheses to explain "clustering" effects at the state level, where certain states with women in top posts were more likely to elect additional women, while states with one or no top women were less likely to support female candidates. Our research showed that this clustering was not random, but related to a state's demographic and/or political factors, whether or not that state had a history of electing women and the current percent of women in elected leadership.

At the congressional district (CD) level, the same three factors also play a role in determining the distribution of women as candidates and elected officials. We found evidence of a "burst" effect, where multiple women run for and are elected to office only after the first woman breaks through certain barriers. These barriers can be psychological, in the sense that voters and funders can't support women because of preconceived notions about politics being a "man's job," or party leaders blocking women from running in certain districts. But when one female candidate breaks through, more women follow in a "burst" effect.

Statistical evidence supports this theory. We ran an interrupted time series regression model, using percent women as primary candidates as the dependent variable, with a key independent variable (IV) of whether or not a woman had been elected to Congress from that district in the past two election cycles (yes or no). The regression also included demographic controls, such as percent unemployed, percent blue-collar, percent senior citizens, percent women in the state legislature, number of top women in that CD's state, whether the district included a city or not, and the CD's partisanship. The key IV was very strongly significant despite all these controls, implying that a CD that has elected a woman in the previous two election cycles is strongly and significantly likely to have a higher percentage of female candidates.

We further tested this idea by creating a variable totaling all of the female candidates in a district between 1992 and 2012. We then checked that average for districts that elected a woman before 1992 (between 1980 and 1992) and the average for districts that did not elect a woman in that time period. Districts that elected a woman in this time period were labeled "early adopters," and were *FAR* more likely to have more women as candidates after 1992 than CDs that did not have a woman in office before 1992. The average number of female candidates in the two decades between 1992 and 2012 for the non-early-adopting CDs was only 4.1; the average for early-adopting-CDs was 7.5.

Lastly, we looked at the number of non-incumbent women as candidates in each CD, creating a simple "yes" or "no" variable: did a woman win in the CD in the time period 1980-2012? When that variable was a "yes," that CD was significantly more likely to have more women running, even as non-incumbent candidates.