User's Guide and Codebook for the CAWP Statewide and Congressional Women Candidates Databases

Center for American Women and Politics

Rutgers University- New Brunswick

This document is a guide and codebook for the Center for American Women and Politics' (CAWP) Women Statewide and Congressional Candidates Databases. If you find any errors or have comments and/or questions about the dataset, please email Chelsea Hill (chelsea.hill@eagleton.rutgers.edu) or Claire Gothreau (claire.gothreau@eagleton.rutgers.edu).

Use of this data for non-commercial purposes is encouraged, provided that clear and visible credit is given to Center for American Women and Politics, Eagleton Institute of Politics, Rutgers University-New Brunswick. Any use of this data for publication or public presentation must credit CAWP. Commercial reproduction requires prior permission in writing from the Center for American Women and Politics.

Suggested Citations:

Center for American Women and Politics. 2020. *Center for American Women and Politics (CAWP) Congressional Women Candidates Database*. Dataset. New Brunswick, NJ: Center for American Women and Politics, Eagleton Institute of Politics, Rutgers University.

Center for American Women and Politics. 2020. *Center for American Women and Politics (CAWP) Statewide Women Candidates Database*. Dataset. New Brunswick, NJ: Center for American Women and Politics, Eagleton Institute of Politics, Rutgers University.

Contents

Data Collection

This data was collected by the staff at the Center for American Women and Politics at Rutgers University. CAWP relies on individual state's election lists and results (via state election agencies) and conducts gender and race verification on all candidates. We utilized CAWP historical candidate lists to assemble the dataset, as well as data from KnowWho data services for more recent years. We verified those lists in 2019 against individual state's election reports (filed candidates and election returns). When election returns were not available, we contacted the state's election officials for more information. Gender was confirmed by a CAWP staff member through Google News searches or through Lexis Nexus Academic. Additionally, we used the U.S. House Primary Election Results (1956-2010) dataset assembled by Stephen Pettigrew and colleagues as an additional crosscheck.[*]

---

[*] Pettigrew, Stephen; Owen, Karen; Wanless, Emily, 2014, "U.S. House Primary Election Results (1956-2010)", https://doi.org/10.7910/DVN/26448, Harvard Dataverse

Codebook

Candidate-Level Variables

1. **FName-** Candidate's first name

2. **MName-** Candidate's middle name

3. **LName-** Candidate's last name

4. **Party-** Candidate's party affiliation

    D = Democratic Party

    DFL = Democratic Farm Labor Party

    I = Independent

    IR = Independent-Republican Party

    R = Republican Party

    NP = Non-Partisan

5. **PrimarySeat-** Candidate's seat status during the primary election

    CP = Challenger in the Primary

    C = Challenger

    I = Incumbent

    IvI = Incumbent vs. Incumbent (redistricting)

    IU = Incumbent Unopposed

    O = Open

    WD = Withdrawn

6. **GeneralSeat-** Candidate's seat status during the general election

    C = Challenger

    I = Incumbent

    IvI = Incumbent vs. Incumbent (redistricting)

    IU = Incumbent Unopposed

    O = Open

    WD = Withdrawn

    NA = Not Applicable

7. **PrimaryWin-** Result of the primary election

L = Lost

W = Won

WD = Withdrawn before Election Day but appeared on ballot

NA = Not Applicable (see below)

Died = Candidate died before Election Day but appeared on ballot

8. **GeneralWin-** Result of the general election

L = Lost

W = Won

WD = Withdrawn before Election Day but appeared on ballot

NA = Not Applicable (see below)

Died = Candidate died before Election Day but appeared on ballot

9. **Race-** Race or Ethnicity of Candidate

AP = Asian American/Pacific Islander

B = Black

MENA = Middle Eastern/North African

H = Hispanic/Latina

W = White

MU = Multi-Racial

NA = Native American/Native Alaskan

Codebook

District-Level Variables

1. **State-** Name of the State

2. **Year -** Year of election

3. **Office-** Office up for election

   Agriculture
   Attorney General
   Auditor
   Chief Financial Officer
   Commissioner of Insurance
   Commissioner of Labor
   Commissioner of Public Lands
   Comptroller
   Corporation Commissioner
   Governor
   Lieutenant Governor
   Public Regulatory Commissioner
   Public Service Commissioner
   Public Utilities Commissioner
   Railroad Commissioner
   Secretary of State
   State Controller
   State School Superintendent
   Superintendent of Public Instruction
   Superintendent of Education
   Tax Commissioner
   Treasurer
   U.S. Representative
   U.S. Senator

4. **District-** District number

   AL = At Large

Coding Guidelines: Rules for Inclusion in the Database

Drop Outs and Withdrawals

1. If the candidate appeared on the primary ballot, they are counted as a filed candidate and included in the database. Appearance on the primary ballot was confirmed for each candidate by a CAWP staff member.
2. If a candidate appeared on the primary ballot but withdrew before the election, they are counted as a filed candidate but their result for the string variable "PrimaryWin" will be "WD" for withdrawn. If a candidate withdrew from a primary election contest and did not appear on the ballot, they are removed from the CAWP database. If a candidate dies and does not appear on the primary ballot, they are removed from the CAWP database.
3. If the candidate won the primary despite withdrawing from the election, they are still counted as a primary winner, but their result for the string variable "GeneralWin" will be "WD" for withdrawn. If a nominee withdrew from a general election contest and did not appear on the ballot, their general election result is "NA."
4. If a candidate dies after filing but appears on the primary ballot, they are counted as a filed candidate. However, their primary result for the string variable "PrimaryWin" will be "Died."
5. If a candidate dies after winning the nomination and they are subsequently replaced on the ballot, then they are counted as a primary winner or a "W" for the string variable "PrimaryWin" and "NA" for "GeneralWin." If a nominee dies before the general election, but still appears on the ballot, they will be coded "Died" for "GeneralWin."
6. If a candidate did not appear on a primary election ballot but became a general election nominee, their "PrimaryWin" code will be "NA." In most cases, these candidates were selected to replace a primary nominee on the general election ballot.

Special Elections

1. Special election candidates for the are counted as filed candidates and nominees if the general election date coincides with the regular general election. The only exception is when there is both a special and regular election being held on the same day for the same district (seat), which is more likely in U.S. House contests. In this case, we only count the regular election candidate as filed candidates or nominees to avoid double counting candidates. In that case, the primary and general election results refer to the general election race. Individuals who ran for *only* the special election – and thus a limited term from November to January – are not included in our dataset, but those candidates who ran for both are included once and their results reflect the result for the full-term primary and general elections.

2. Our candidate database only includes women candidates who ran in regularly-scheduled elections, even if that election was deemed a special election (where the general election date for the special election coincides with that of the regular election). The result listed is for the full

term (for House) or the completion of a term that goes beyond that session (for Senate). For House completion of term, we do not distinguish results apart from footnotes that indicate if a woman won a special completion and did NOT win the full House term (i.e. Brenda Jones in 2018). For special elections held outside of normal election schedules (not simultaneous/coinciding with regular general election dates), women candidates are not included.

Runoffs

1. If a state holds runoff elections in a primary and/or general election outside of a jungle primary system, the runoff result for each phase is what count for the string variables "PrimaryWin" and "GeneralWin."
2. We do not count the pre-runoff election separately, unless it's the terminal contest (i.e. doesn't have to go to a runoff).
3. If a runoff occurs, anyone who fails to make the runoff as well as the loser(s) of the runoff are categorized as losers for that phase of the election.

Jungle Primaries

1. Jungle primaries are non-partisan primaries that coincide with General Election days. If no candidate gets more than 50% of the vote, the race goes to a runoff. In these cases, we count the jungle primary as the primary result for string variable "PrimaryWin" and the runoff election as the general election result for string variable "GeneralWin."
2. Nominees include either a) the top two vote getters, when no single candidate receives more than 50% of the vote; or b) the top vote getter who receives 50% or more of the vote. In the latter case, there is only one general election nominee and everyone else is coded as a primary loser and has no result for the general election field. In the former case, the two top vote getters are coded as primary winners and their result in the runoff election determines their general election result code for string variable "GeneralWin."

Top Two Primary

1. If a candidate is challenging a member of their own party in a top-two primary, they are coded "CP" for challenger in the primary for the string variable "PrimarySeat."
2. If a candidate is challenging a member of the opposing party in a top-two primary, they are coded as "C" for challenger for the string variable "PrimarySeat."
3. If the candidate was a challenger in the primary and moves on to the general election, then their seat status for the general election will be "C," but only if the incumbent also moves on to the general election. If the candidate was a challenger in the primary and moves on to the general election, but the incumbent does not, then their seat status for the general election will be "O."

Multi-winner Races

1. There are a small number of statewide executive contests that are conducted for multiple seats (e.g. on a corporation commission) on one body based on one set of votes. In these cases, if the number of seats being contested matches the number of incumbents running, then all non-incumbent candidates will be coded "C" for challenger for seat status. If the number of incumbents running is less than the number of seats being contested, then all non-incumbent candidates will be coded as "O" for open seat for seat status.

Nominees Chosen by Convention

1. If the state counts convention candidates as filed candidates (i.e. they are listed on the state website as a filed candidate), then they are included in the dataset. In states where these candidates do not file with the state, we only count the candidates who win their party's convention. Historical comparisons of primary candidates in states like Virginia should be made with caution because a district can alternate in how it selects its nominees (whether by convention or traditional primary).

Lieutenant Governors

1. Candidates for lieutenant governor who are selected by the gubernatorial nominee after nomination are counted as primary winners and are coded as such for the variable "PrimaryWin."
2. Candidates for lieutenant governor who are selected by the gubernatorial nominee before nomination are given the same code as the gubernatorial candidate for "PrimaryWin" even if their name is not on the primary ballot.

Rules on Third Party Candidates

1. Third party candidates are not included in the dataset unless they are incumbents or they win the general election. In those cases, they are coded as missing for "PrimaryWin" and "PrimarySeat." For the "GeneralWin" and "GeneralSeat" variables, they are coded normally.

Candidate Race Verification

1. Candidate race is coded by a team of CAWP researchers in two ways. First, we rely on candidate self-identification. Where self-identification was not provided to us, we rely on a multiple source verification process for coding. If verification sources were unavailable or unclear, we leave the candidate as uncoded for race identification in the database.