Codebook for "U.S. House Primary Election Results (2012-2018)."

This dataset compiles primary election results for Democratic and Republican U.S. House candidates running in elections from 2012 to 2018. The data include results from all states except those conducting "top two primaries" (California and Washington for the entirety of this time period) as well as from Louisiana, which does not hold traditional primary elections.

These data were collected by Michael G. Miller (mgmiller@barnard.edu) and Nicki Camberg, both of Barnard College. If you publish with these data, please cite as:

Miller, Michael G., and Nicki Camberg. 2020. "U.S. House Primary Election Results (2012-2018)." Harvard Dataverse, Version 1. Available at https://doi.org/10.7910/DVN/CXVMSY

Citation information is also available on the Harvard Dataverse page.

This file employs the same architecture as the following dataset:

Pettigrew, Stephen; Owen, Karen; Wanless, Emily, 2014, "U.S. House Primary Election Results (1956-2010)," available at https://doi.org/10.7910/DVN/26448. Current version information can be found at the preceding link.

This dataset can therefore be used as a stand-alone file, or can be appended to Pettigrew et al. to yield primary election data for years 1956 to 2018. If users append this dataset to Pettigrew et al., both datasets should be cited.

Data collection procedures follow Pettigrew et al.: Election returns were compiled from the *America Votes* book series. Candidate-level data are derived from Gary Jacobson's dataset of House general elections. Data on candidate quality are incomplete. Where we have such data, they are obtained using similar methods as Pettigrew et al.: Searching news databases for biographical information.

Note: Run-offs and pre-runoffs are denoted with the *runoff* variable (see below). Run-offs are also denoted with "_r" appended to the *raceid.* Pre-runoffs (denoted with "_p"), are defined as elections in run-off states in which no candidate received a majority of the vote, and thus triggered run-off elections. If users are employing these data to denote candidates who advanced to the general election (as in say, a regression discontinuity design), pre-runoffs should be discarded.

If you discover errors in these data, please contact Michael G. Miller at mgmiller@barnard.edu.

District level variables:

raceid: A nine or eleven digit code. The first four digits denote the year, digits 5 and 6 denote the state, digits 7 and 8 denote the congressional district, and digit 9 denotes the party (0=Republican, 1=Democrat).
An appended "_r" denotes runoff election
An appended "_p" denotes a primary that resulted in a runoff election

year: The year the election was held

stcd: A four-digit code (the last four digits of *raceid*). The first two digits denote the state; the second two digits denote the congressional district.

state: State name

seat: Denotes the party holding the seat.

0=Republican incumbent
1=Democratic incumbent
2=Open seat held by Republicans
3=Open seat held by Democrats
4=Open seat, new (reapportionment)
5=Two incumbents (redistricting)

party: The party holding the primary.
0=republican
1=democratic

redist: Denotes whether redistricting has occurred since the last election.
1=Redistricting occurred.
0=Redistricting did not occur.

fr: Denotes whether the incumbent is a freshman.
0 = Incumbent is not freshman
1 = Freshman elected previous general election
2 = Freshman elected in special election more than 1 year earlier
3 = Freshmen elected in a special election during this election year
9 = Seat not defended by major party incumbent

law: Denotes the laws used to run the primary
0=Open primary system
1= Closed primary system
2=Semi-open system; registered voters vote in their party's primary; unregistered voters can vote in either.
3= Multiparty ballot. The candidate with the most votes by party won the party's nomination.
4=Modified system. Democrats/Republicans could vote with only their party, but unaffiliated voters could register with the Democrats or Republicans to vote in that party's primary.

candnumber: The total number of candidates running in the primary.

prez: The presidential two-party vote for the relevant party in the district in the previous election cycle.

votep: The congressional two-party vote for this party in the district in the previous election cycle.

incname: The name of the seat's incumbent at the time of the primary.

type: Primary Competitiveness
0=unopposed primary
1=opposed primary
2=caucus/convention
3=no candidate

runoff: Run-off Status
0=Primary is in a non-runoff state, or the primary did not result in a run-off
1=No candidate received a majority in a run-off state; this primary triggered a run-off. No candidate has won the nomination as a result of this election.
2=This is a two-candidate run-off election. The winner advances to the general election.

Candidate-level variables:

candidate: The candidate's name.

candvotes: The number of votes for the candidate in the primary.

tvotes: The total number of votes cast in this party's primary.

candpct: The percentage of votes received by the candidate (*candvotes*/*tvotes*).

winner: 0 = candidate lost election; 1 = candidate won election

inc: 0 = candidate is a challenger; 1 = candidate is the incumbent

gender: 0 = candidate is male; 1 = candidate is female

qual: Candidate Quality
0=has not held elected office previously
1=has held elected office
2=incumbent
3=challenger who is former US Representative

office:
30 current incumbent
29 state wide elected office (governor, lieutenant governor)
28 state senator
27 state representative
26 mayor
25 judge (elected)
24 district or prosecuting attorney
23 city council member
22 other elected office
21 previous incumbent in most recent Congress who lost
20 former House member (not most recent Congress)
19 former senator
18 attorney general
17 state's attorney
16 state commissioner
15 alderman
14 sheriff
13 treasurer
12 businessm/business owner
11 journalist
10 lawyer
9 judge (non-elected)
8 minister
7 farmer/rancher
6 military
5 local party leader/activist
4 doctor/dentist/vet
3 teacher/professor/educator
2 other government-non-elected/former congressional aide or advisor
0 indicates no electoral experience