Cowburn Congressional Primary Dataset Codebook

Variables in Dataset

Variable	Definition
unique_id	Seven-digit unique identifier made up of PartyYearYearStateStateDistrictDistrict. Democratic = 1, Republican = 2, state codes taken are state_fips, district codes are district_number
year	Year of primary contest
party	Political party D = Democratic R = Republican
state	State name
state_code	Two letter state abbreviation
state_fips	Two-digit state FIPS code
district_number	US Congressional District, Senate contests 99, special Senate 98
geoid	state_code * 10 + district_number
panel	geoid plus 0.1 for Democratic primary and 0.2 for Republican primary, used as the panel variable to set the data
district_pvi	Cook PVI rating, figures taken from: https://cookpolitical.com/pvi-0
relative_pvi	Rescaled version of district_pvi relative to the party, where a D+5 district would take the value 5 for the Democratic primary and -5 for the Republican primary.
white_pct	Percentage of white citizens in the district, data from US Census ACS website: https://www.census.gov/programs-surveys/acs
median_income	District median income, data from US Census ACS website: https://www.census.gov/programs-surveys/acs
presidential_vote	Party's most recent presidential vote share in the district
primary_type	Signifies whether the incumbent is standing in the primary or if no primary taking place. Incumbent = current office holder running in the contest, Incumbent^ = Incumbent defeated in the primary, Challenger = current office holder standing for alternative party, Open = current office holder not running, None = no primary contest taking place
factional_primary	Signifies whether the two highest placing candidates are proximate to different factions, 1 = different ideological factions against each other, 0 = non-factional contest, blank = no contest. See "Decision Rule – Faction" below.
ideology_primary	Signifies whether the reason for contest is coded as "Ideology", "Ideology-I" or "Ideology-C". Takes the value 1 if yes, 0 if no, blank if uncontested. See "Decision Rule – Reason for Contest" below.
polarized_primary	Signifies whether the two highest placing candidates are both proximate to the realigner faction, 1 = two realigner candidates, 0 = any other contest, blank = no contest. See "Decision Rule – Faction" below.
reason_for_contest	Main reason for primary contest taking place, metric of analysis originally developed by Robert Boatright (2013) independently applied to this data set by Mike Cowburn. See "Decision Rule – Reason for Contest" below.
primary_turnout	Number of voters in the primary contest, figures from FEC website: https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
voting_age_pop	Voting age population (VAP) of district, data from US Census ACS website: https://www.census.gov/programs-surveys/acs
primary_turnout_pct	primary_turnout / voting_age_pop
pres_consecutive	Presidential primary held on the same day? Takes the value 1 if yes, 0 if no. All midterm elections take the value 0.
fractionalization	Standard measure of primary competitiveness: $F = 1 - \Sigma[(C_1)^2 + (C_2)^2 + (C_3)^2 + (C_3)^2 + (C_4)^2 \ldots]$
primary_candidates	Number of primary candidates on the ballot.
winner_name	Name of candidate who wins the primary.
winner_icpsr	ICPSR code for candidate who wins the primary.
winner_gender	Gender of candidate who wins the primary.

winner_cfscore	CFscore of candidate who wins the primary. Data from Adam Bonica's Database on Ideology, Money in Politics, and Elections (DIME) https://data.stanford.edu/dime
winner_dwn	DW-NOMINATE score of the candidate who wins the primary. Data from www.voteview.com
winner_nokkenpoole	One-Congress-at-a-time-NOMINATE (Nokken Poole score) for the forthcoming Congress of candidate who wins the primary. Data from www.voteview.com
winner_p	Percentage primary vote share of the candidate who wins the primary, figures from FEC website: https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
winner_g	Percentage general election vote share of the candidate who wins the primary, figures from FEC website: https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
won_general	Dichotomous variable that takes the value 1 if the primary winner won the general election, 0 if the primary winner lost the general election, blank = no candidate. Data from https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
winner_ge_votes	Total number of votes for primary winner in the general election. Data from https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
general_turnout	Total turnout in general election. Data from https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
winner_fec_id	FEC ID for primary winner.
winner_receipts	Primary winner receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
winner_individuals	Primary winner individual receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
winner_pacs	Primary winner PAC receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
winner_self	Primary winner self-financed receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
winner_disbursement	Primary winner disbursement, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
winner_faction	Primary winner proximity to factional ideal types, assigned by Mike Cowburn. See "Decision Rule – Faction" below.
winner_quality	Primary winner 'quality' code of having held previous elected office. In line with Jacobson (1978), coded by Mike Cowburn using Ballotpedia, VoteSmart, plus candidate websites.
winner_extr1	Primary winner relative CFscore vs. party median for that year
winner_extr2	Primary winner relative CFscore vs. nearest opponent.
second_name	Name of candidate who placed second in the primary.
second_gender	Gender of candidate who placed second in the primary.
second_cfscore	CFscore of candidate who placed second in the primary. Data from Adam Bonica's Database on Ideology, Money in Politics, and Elections (DIME) https://data.stanford.edu/dime
second_dwn	DW-NOMINATE score of the candidate who placed second in the primary. Data from www.voteview.com
	One-Congress-at-a-time-NOMINATE (Nokken Poole score) for the forthcoming Congress of candidate who wins the primary. Data from www.voteview.com
second_p	Percentage primary vote share of the candidate who placed second in the primary, figures from FEC website: https://www.fec.gov/introduction-campaign-finance/election-and-voting-information/
second_fec_id	FEC ID for candidate who placed second in the primary.

second_receipts	Primary second receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
second_individuals	Primary second individual receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
second_pacs	Primary second PAC receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
second_self	Primary second self-financed receipts, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
second_disbursement	Primary second disbursement, as per FEC 12P pre-primary filing: https://www.fec.gov/data/reports/house-senate/?data_type=processed&amendment_indicator=N&report_type=12P
second_faction	Primary second proximity to factional ideal types, assigned by Mike Cowburn. See "Decision Rule – Faction" below.
second_quality	Primary second 'quality' code of having held previous elected office. In line with Jacobson (1978), coded by Mike Cowburn using Ballotpedia, VoteSmart, plus candidate websites.
third_pct	Third placed candidate primary vote share, used only to calculate fractionalization
fourth_pct	Fourth placed candidate primary vote share, used only to calculate fractionalization
fifth_pct	Fifth placed candidate primary vote share, used only to calculate fractionalization
sixth_pct	Sixth placed candidate primary vote share, used only to calculate fractionalization
seventh_pct	Seventh placed candidate primary vote share, used only to calculate fractionalization
eighth_pct	Eighth placed candidate primary vote share, used only to calculate fractionalization
ninth_pct	Ninth placed candidate primary vote share, used only to calculate fractionalization
tenth_pct	Tenth placed candidate primary vote share, used only to calculate fractionalization
eleventh_pct	Eleventh placed candidate primary vote share, used only to calculate fractionalization
twelfth_pct	Twelfth placed candidate primary vote share, used only to calculate fractionalization
thirteenth_pct	Thirteenth placed candidate primary vote share, used only to calculate fractionalization
fourteenth_pct	Fourteenth placed candidate primary vote share, used only to calculate fractionalization
fifteenth_pct	Fifteenth placed candidate primary vote share, used only to calculate fractionalization
sixteenth_pct	Sixteenth placed candidate primary vote share, used only to calculate fractionalization
seventeenth_pct	Seventeenth placed candidate primary vote share, used only to calculate fractionalization
eighteenth_pct	Eighteenth placed candidate primary vote share, used only to calculate fractionalization
republican	Takes value 1 if Republican, 0 if Democratic.
senate	Take value 1 if Senate, 0 if House.
notes	Brief note and sources on each primary competition.

Decision Rule – Faction

The following table is an overview of metrics used to code candidate faction. For all candidates a minimum of two datapoints was used to code to a faction, for most candidates significantly more than two datapoints were used. Candidates were analyzed across multiple fields and coded accordingly.

Progressive Democrats	**Establishment Democrats**	**Establishment Republicans**	**Reactionary Republicans**
Caucus Membership			
Congressional Progressive Caucus	New Democrat Coalition	Tuesday Group	Freedom Caucus
Medicare for All Caucus	Blue Dog Caucus	Main Street Partnership Caucus	Liberty Caucus
Blue Collar Caucus	Moderate Dems Working Group		Tea Party Caucus
Endorsements, Associations & Funding (Groups)			
Democracy for America	Blue Dog PAC	Chamber of Commerce groups	Americans for Prosperity
OurRevolution	NewDemPAC	Main Street Partnership (PAC)	Club for Growth
Justice Democrats	Democratic Leadership Council		Tea Party Groups
Progressive Change Campaign Committee	Third Way		FreedomWorks
Brand New Congress	New Democrat Network		House Freedom Fund
Progressive Democrats of America			
Swing Left			
Endorsements & Associations (Individuals)			
Bernie Sanders	Bill Clinton	Ronald Reagan	Donald Trump
Howard Dean	Dianne Feinstein	Mitt Romney	Ron Paul
Elizabeth Warren	Joe Biden	John McCain	Ted Cruz
Alexandria Ocasio-Cortez	Joe Lieberman	Susan Collins	Mark Meadows
Barbara Lee	Al Gore	John Kasich	Michele Bachman
Ro Khanna	Amy Klobuchar	Dick Cheney	Sarah Palin
Policy Positions & Campaign Themes			
Free College for All	Increase minimum wage with Inflation	Limit government using existing structures	Flat Taxation
Medicare for All	Tax reduction	"Pragmatic conservatism"	Leave the UN
$15 Federal Minimum Wage	ACA as healthcare reform	"Practical" solutions to illegal immigration	0% foreign aid
Abolish ICE	End gridlock	End gridlock	Status threat narratives
Healthcare/education as Right	Bipartisanship	Bipartisanship	Climate change denial
No corporate PAC donations	Work across the aisle	Work across the aisle	No amnesty for any illegal immigrants
Critical of the richest 1%/ billionaires	Problem-solver	Problem-solver	Construction of a wall at the southern border
Evoking European social democracies, esp. Scandinavia	Work within current systems to make	Commitment to internationalist foreign policy	Eliminating entire govt departs (esp.

	healthcare more affordable		Dept for Education and EPA)
Electoral reform: advocation of structural change to the Democratic Party and democratic system	Critical of ideological battles	Invoking JFK, frequent quotation "Let us not seek the Republican answer or the Democratic answer but the right answer"	Positioning establishment conservatives as RINOs
Business regulation to ensure benefits of economic growth received by all	Incremental additions to healthcare cover	Not mentioning or associating their campaign with Donald Trump	Whites as discriminated against, narratives of white nationalism
End the death penalty	Promotion of growth narrative	Pro-business, economic growth	Openly embracing Donald Trump
End war on drugs, legalization	Minimal regulation of business	Free market economics, critical of tariffs	America First
Prioritization of climate change & renewable energy, Green New Deal	Concern about leftward movement of the party	Invoking party as anti-slavery, referencing Abraham Lincoln	Rhetoric about immigrants bringing drugs and crime into America
Obama presidency as not progressive enough	Support free trade, especially TPP	Support for UN as part of American military might	Framing cultural issues as threats to (white) American identity

Decision Rule – Reason for Contest

Reason	Description
Competence/ Age	In incumbent primaries, these were criticism of the ability of the current officeholder, criticisms included doing a poor job of representing the district, failing to deliver resources/projects to the district, spending too long in Washington, or being too old. In non-incumbent races these were usually framed in positive terms with experience in other public office or experience in business most frequently given as the reason why they are the most competent candidate for the district. In challenger races this was often positioned as being most likely to defeat the incumbent in the general election.
Ideology	Non-incumbent primaries which focused on ideological difference between the candidates. These were sometimes framed in positive ways such as being the "true conservative" in the race or negatively by branding an opponent as ideologically out of step with the district, too-centrist or not sufficiently committed to the party platform e.g., using RINO as a slur.
Ideology - C	Primary where an incumbent is challenged by a more centrist alternative on ideological grounds, often featuring criticism of the incumbent for being too extreme for the district.
Ideology - I	Primary where an incumbent is challenged by a less centrist alternative on ideological grounds with the incumbent criticized for being too moderate, "primarying".
Local Issue	Focused on a single local issue that was important to the district, a frequent example being water supply in districts in southern California.
National Issue	Focused on either a single national issue or where candidates focused their campaigns on different issues with no ideological context, often these races included candidates promoting specific plans in a national policy area without clashing directly with their opponent. Frequent examples included teachers who focused on education policy or medical professionals who focused on healthcare. Non-incumbent contest frequently featured candidates focused solely on their issue (or couple of issues) without referencing their opponent.
Other/ No Reason	Either the contest was taking place for an unspecified reason e.g. "I just felt like doing it"/ "all races should be contested" or for a reason which fell outside of the other categories. This was often for tactical reasons, to disrupt the competition, or to help the alternative party.
Race	Race was specifically mentioned as the reason for the contest, where the main reason was the race of one of the candidates. Many involve claims that only a minority should represent a minority-majority district or challenges between candidates from different minority groups.
Scandal	The dominant theme of the campaign was a specific perceived wrongdoing by one of the candidates. Most of these involved allegations of either sexual or financial misconduct.
Unknown	Unable to find enough information about one candidate or the nature of the contest to categorize.
Challenger Solo	No primary contest: one candidate running for office in a challenger seat
Incumbent Solo	No primary contest: incumbent only candidate running for office
No Candidate	No primary contest: no candidates running for seat, no candidate stood in general election
Open Solo	No primary contest: one candidate running for office in an open seat