Title: 2010 Congressional Election Study

Principal Investigator: Walter J. Stone

Funding Agency: National Science Foundation (SES-0852387)

Bibliographic Citation: Stone, Walter J. 2013. 2010 Congressional Election Study. University of California, Davis, CA. (http://electionstudy.ucdavis.edu/)

Summary: In early summer 2009 and in October 2010, we conducted a survey of political experts in 155 congressional districts, the same districts used in our 2006 Congressional Election Study (see http://electionstudy.ucdavis.edu/ for more information on the 2006 study). The sample is composed of 100 randomly sampled districts from the contiguous 48 states (sample06=1) and a supplemental purposive sample of 55 competitive and/or open seats (sample06=0). In June of 2006 we consulted Congressional Quarterly, Cook Report, Sabato Crystal Ball, and National Journal for districts anticipated to be competitive. Districts rated as "tossup" or "leaning competitive" by any source were included in the competitive supplemental sample. We identified 72 districts in this manner, 17 of which were included in the random cross section.

The expert informant survey was of delegates from the 2008 Democratic and Republican National Conventions, state legislators from both parties, and individuals identified from Polimetrix's panel of respondents as being politically knowledgeable based on an expertise battery. The survey was conducted by mail and online. For the first wave in early summer 2009, we received responses from 5,466 delegates, state legislators, and Polimetrix experts (1,294 delegates and state legislators; 4,172 Polimetrix experts). For the second wave in October 2010, we received responses from 4,871 delegates, state legislators, and Polimetrix experts (726 delegates and state legislators; 4,145 Polimetrix experts). Respondents were asked questions over a variety of topics related to the congressional district, the candidates, and their campaigns. We aggregate informant perceptions to the district level to obtain measures of district- and candidate-level characteristics.

The expertise battery used to identify individuals as politically knowledgeable in the Polimetrix panel consisted of questions related to news consumption, House incumbent name recognition, how informed the respondent said they were about government and politics in their state, and how confident they were answering questions about their congressional district. The questions and screening rules are included at the end of this document.

There are 5 data files and 5 corresponding codebooks (3 data files corresponding to 2010 and 2 data files corresponding to 2009). Each data file is in Stata format.

- **2010 Expert Informant Data**: 2010_ucd_ces_expert_survey_data.dta
  - Unit of analysis: expert informant (N = 4,871)
  - Observations are the 4,871 delegates, state legislators, and Polimetrix experts surveyed in our sample
  - Data consist of delegates', state legislators', and Polimetrix experts' responses to the informant survey
  - Informant survey questionnaires (online and mail) are included in the data file: 2010_ucd_ces_expert_survey_questionnaire.pdf
  - Codebook: 2010_ucd_ces_expert_survey_codebook.pdf

- **2010 Registered Voter Data**: 2010_ucd_ces_voter_survey_data.dta
  - Unit of analysis: individual registered voter (N = 2,000)
  - Observations are the 2,000 registered voters surveyed in our sample of districts
  - Data consist of voters' responses to the Cooperative Congressional Election Study (CCES) UC Davis Module survey. Data also includes responses to the CCES common content survey questions (see http://projects.iq.harvard.edu/cces/ for more information on the CCES common content questions).
  - The CCES UC Davis Module survey questionnaires (pre- and post-election) are included in the data file: 2010_ucd_ces_voter_survey_questionnaire.pdf
  - Codebook: 2010_ucd_ces_voter_survey_codebook.pdf

- **2010 District Data**: 2010_ucd_ces_district_data.dta
  - Unit of analysis: congressional district (N = 155)
  - Observations are the 155 congressional districts in our sample
  - Data consist of the aggregated informant perceptions
  - Codebook: 2010_ucd_ces_district_codebook.pdf

- **2009 Expert Informant Data**: 2009_ucd_ces_expert_survey_data.dta
  - Unit of analysis: expert informant (N = 5,466)
  - Observations are the 5,466 delegates, state legislators, and Polimetrix experts surveyed in our sample
  - Data consist of delegates', state legislators', and Polimetrix experts' responses to the informant survey
  - Informant survey questionnaires (online and mail) are included in the data file: 2009_ucd_ces_expert_survey_questionnaire.pdf
  - Codebook: 2009_ucd_ces_expert_survey_codebook.pdf

- **2009 District Data**: 2009_ucd_ces_district_data.dta
  - Unit of analysis: congressional district (N = 155)
  - Observations are the 155 congressional districts in our sample
  - Data consist of the aggregated informant perceptions
  - Codebook: 2009_ucd_ces_district_codebook.pdf

Notes:

- The "2010 District-Level Data" dataset contains three variable versions for every question asked of informants, each representing our attempt to adjust for partisan bias in informant perceptions, since the majority of informants included in our sample strongly identify with either the Republican or Democratic parties. We correct for partisan bias by regressing informant responses to each item on a party dummy (coded -1 for Democrats, 0 for independents, and +1 for Republicans).[1] The resulting coefficient estimates reflect the average partisan bias associated with each question. We subtract these estimates from informant responses, leaving us with values that approximate the answers independent experts would have given. We use the following Stata code:

```
recode ipid 1/3=-1 4=0 5/7=1, g(pid_3)

foreach var of varlist  winnerpredict-icertain {
      quietly regress `var' pid_3 if inftype==1
      quietly g `var'_pp10=`var'-_b[pid_3]*pid_3 if inftype==1
      quietly regress `var' pid_3 if inftype==2
      quietly g `var'_pd10=`var'-_b[pid_3]*pid_3 if inftype==2
      quietly regress `var' pid_3
      quietly g `var'_pc10=`var'-_b[pid_3]*pid_3
}
```

  where `inftype==1` indicates a Polimetrix expert and `inftype==2` indicates a delegate or state legislator. The three versions of these variables represent which type of informants we aggregated to create the variable. The first version, which ends in "_pc10" for "purged combined," includes all expert informants (delegates, legislators, and Polimetrix experts) in our attempt to adjust for partisan bias in informant perceptions. The second version, "_pd10" for "purged delegates," includes only delegates and legislators. The third, "_pp10" for "purged Polimetrix," includes only Polimetrix experts.

  We use the same method to adjust for partisan bias in informant perceptions in 2009. The "2009 District Data" dataset contains three variable versions for every question asked of informants. The first version, which ends in "_pc09" for "purged combined," includes all expert informants (delegates, legislators, and Polimetrix experts) in our attempt to adjust for partisan bias in informant perceptions. The second version, "_pd09" for "purged delegates," includes only delegates and legislators. The third, "_pp09" for "purged Polimetrix," includes only Polimetrix experts.

- Since the first wave was administered in early summer 2009, before the 2010 House candidates in the district emerged, the 2009 informant survey questions were related to the incumbent currently serving in the district. We transformed the data to be in terms of democratic and republican candidates before implementing the method for correcting for partisan bias described above.

[1] For questions related to the performance or ideology of incumbents, we use a three-point variable coded -1 when the informant and incumbent were in the opposite party, 0 for independents, and +1 when the informant and incumbent were in the same party.

- We have included some district variables gathered for the project to make the data more useful to other scholars. The source of these variables is noted in the variable label.

    - These district variables include variables collected by Gary Jacobson. Two of these variables are on challenger quality and whether the seat was open. We code the challenger-experience variable ("chexp10") to indicate the experience of the candidate in the party opposite that of the incumbent's party (whether or not the incumbent actually ran). In this coding "challenger" equates to "out party" and is not restricted to candidates running against incumbents seeking reelection.

    - Other sources include: CQ's Politics in America (CQ)[2]; Cooperative Congressional Election Study (CCES); 2000 and 2010 U.S. Census; Federal Election Commission (FEC); State Secretary of State; New York Times; The Almanac of American Politics (Almanac)[3].

- The weight variable, "ptyweight06," in the 2010 district dataset may be used to correct for the proportion of Democrats and Republicans in the random cross-section (sample06=1). In the unweighted cross-section sample, Democrats are over-represented.

- The "openorcompete06" variable is coded 1 if Congressional Quarterly, Cook Report, Sabato Crystal Ball, or National Journal rated the district as "tossup" or "leaning competitive" in June 2006, and 0 otherwise. 72 districts are considered open or competitive (55 in the supplemental purposive sample and 17 in the random cross section).

- The variable "weight" in the informant datasets was automatically generated by Polimetrix. The weight can be used to approximate a representative population with respect to age, gender, race, and education.

- In the 2010 informant dataset, the "caseid" variable is the identification number generated by Polimetrix to identify individuals from the Polimetrix sample. The "id" variable is a unique identification number for *all* informants, including the delegates and Polimetrix informants.

[2] McCutcheon, Chuck and Christina L. Lyons. 2009. *CQ's Politics in America 2010: The 111th Congress.* CQ Press.
[3] Barone, Michael and Chuck McCutcheon. 2011. *The Almanac of American Politics 2012*. University of Chicago Press. Barone, Michael, Richard E. Cohen, and Jackie Koszczuk. 2009. *The Alamanac of American Politics 2010.* National Journal Group.

Polimetrix Expert Informants Screening Questions and Rules:

Informed
How well informed would you say that you are about government and politics in your state?
<1> Extremely well informed
<2> Well informed
<3> About average
<4> Not so well informed
<5> Not well informed

Dropped if > 1

Confidence
How would you rate your confidence in your ability to answer questions about your congressional district?
<1> Extremely high confidence
<2> High confidence
<3> So-so confidence
<4> Low confidence
<5> Extremely low confidence

Dropped if >3

Incumbent Name Recognition
What is the name of the current U.S. House incumbent in your district?
<1> $Senator1
<2> $Senator2
<3> $House
<4> $Governor
<5> $Neighbor
<6 > Not sure

Dropped if !=3

News Consumption
How often do you watch TV news about politics and public affairs?
<1> Less than once a week
<2> About once a week
<3> Several times a week
<4> Every day

How often do you read a newspaper about politics and public affairs?
<1> Less than once a week
<2> About once a week
<3> Several times a week
<4> Every day

How often do you consult Internet sources about politics and public affairs?
<1> Less than once a week
<2> About once a week
<3> Several times a week
<4> Every day

Dropped if all three were < 4

Approximately 5% of the Polimetrix sample was allowed to fail one or more of these conditions to make up for a low number of expert informants in 13 districts (`pass==0`).