Title: 2006 Congressional Election Study

Principal Investigator: Walter J. Stone

Funding Agency: Division of Social Sciences, University of California, Davis

Bibliographic Citation: Stone, Walter J. 2010. 2006 Congressional Election Study. University of California, Davis, CA. (http://electionstudy.ucdavis.edu/)

Summary: In October of 2006, we conducted a survey of political experts in 155 congressional districts. 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 2004 Democratic and Republican National Conventions and state legislators from both parties. The survey was conducted by mail and we received responses from 970 delegates and state legislators for a response rate of 21 percent. 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.

There are 3 data files and 3 corresponding codebooks. Each data file is in Stata format.

- **Expert Informant Data**: 2006_ucd_ces_expert_survey_data.dta
    - Unit of analysis: expert informant (N = 970)
    - Observations are the 970 delegates and state legislators surveyed in our sample
    - Data consist of delegates' and state legislators' responses to the informant survey
    - Informant survey questionnaire is included in the data file: 2006_ucd_ces_expert_survey_questionnaire.pdf
    - Codebook: 2006_ucd_ces_expert_survey_codebook.pdf

- **Registered Voter Data**: 2006_ucd_ces_voter_survey_data.dta
    - Unit of analysis: individual registered voter (N = 1,000)
    - Observations are the 1,000 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 questionnaire is included in the data file: 2006_ucd_ces_voter_survey_questionnaire.pdf
    - Codebook: 2006_ucd_ces_voter_survey_codebook.pdf

- **District Data**: 2006_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: 2006_ucd_ces_district_codebook.pdf

Notes:

- The "2006 District-Level Data" dataset contains two variables for every question asked of informants. The first, which ends in "_mn" (for "mean"), is the mean response of informants within the district after recoding all "Don't Knows" to missing. The second version of each question, ending in "_pmn" (for "purged mean"), represents 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 informantptyid 1/3=-1 4=0 5/7=1 8=., g(pid_3)

  foreach var of varlist districtptyid winnerpredict patriotemph-
  repptyemph distlc demcanlc repcanlc demprilc repprilc demcansc-votersc
  demcangm-votergm demcanip-voterip datten-rnpol {
       regress `var' pid_3
       g `var'_pmn=`var'-_b[pid_3]*pid_3
  }
  ```

- Individual respondents to the informant survey who are not in a sample district (N=40) are given missing data on the district variable. They are included in the file for anyone interested in the informant data for purposes not anticipated by this study, but they were dropped from aggregation analysis and are not included in the district-level data.

- We have included some district variables gathered for the project to make the data more useful to other scholars, including two variables collected by Gary Jacobson on challenger quality and whether the seat was open. We code the challenger-experience variable 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.

- The weight variable, "ptyweight" 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 "openorcompete" 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).

---

[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.