Codebook of the Data for
"Revealing Issue Salience via Costly Protest:
How Legislative Behavior following Protest Advantages Low-Resource Groups"

LaGina Gause
University of California San Diego
Lgause@ucsd.edu

This file describes the data in "bjps_costlyprotest_replication_data.csv". The analyses in the article rely on several different datasets. Primarily, the analyses in the article rely on the Dynamics of Collective Action (DCA). Other data are from govtrack.us; the 1992, 1994, and 1996 American National Election Study (ANES) surveys; the US Census Bureau; Lublin (1997); Voteview: Congressional Roll Call Votes, and the CQ Press Voting and Elections Collection. The data are available at the following links:

- The DCA data and codebook are available at https://web.stanford.edu/group/collectiveaction/cgi-bin/drupal/node/17
- The govtrack.us data and codebook is available at https://www.govtrack.us/about-our-data
- Data and codebooks for the 1992, 1994, and 1996 ANES are available at https://electionstudies.org/data-center/
- Population estimates are obtained from the 1990 Decennial Summary File 3 and accessible via the United States Census Bureau's DataFerret tool https://dataferrett.census.gov/
- Lublin (1997) data are available at https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/OUX1BA
- Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet (2019). Voteview: Congressional Roll-Call Votes Database. https://voteview.com/.
- CQ Press Voting and Elections Collection data are available at http://library.cqpress.com/elections/

Table 1. Overview of the data

Variable	Description	Source
eventid	Unique identifier for protest events	DCA
statecd (string)	State code and congressional district number	DCA, ANES, Voteview, Census, Lublin (1997), CQ Press
statecd1 (numeric)	State code and congressional district number	DCA, ANES, Voteview, Census, Lublin (1997), CQ Press
cnrptyy	Congressional district corresponding to the variable year	DCA, ANES, Voteview, Census, Lublin (1997), CQ Press
majortopic	Policy Agendas Project classification code categorizing the relevant issue area for a protest claim	DCA, Voteview
lib	Indicates whether protesters were acting in favor or support of Civil Rights, Minority Issue, or Civil Liberties	DCA
cons	Indicates whether protesters were acting against or in opposition to Civil Rights, Minority Issue, or Civil Liberties	DCA
rccount	First, final passage roll call vote on issue relating to protest claim	DCA, govtrack, Voteview
support	Indicates whether a legislator's vote on a protest related bill aligns with protesters' preferences	DCA, govtrack, Voteview
supportall2	Indicates whether a legislator's vote on a protest related bill aligns with protesters' preferences (in districts where a protest occurred) or constituents' preferences (in districts where a protest did not occur).	DCA, ANES, govtrack, Voteview, Census, Lublin (1997)
protest	Indicates whether the observation describes a protest	DCA
protestincd	Indicates whether there was ever a protest in the congressional district before the roll call vote in a congressional session	DCA, Voteview
poorh	Indicates whether any protesters are described as being poor or homeless	DCA
nonwhite	Indicates whether protesters were from a racial or ethnic minority group	DCA

Table 1. Overview of the data (cont'd)

Variable	Description	Source
noorgig	Indicates whether a formal interest group participated in the collective action event	DCA
poorhc	Describes whether there was a protest and whether any protesters were poor or homeless	DCA
nonwhitec	Describes whether there was a protest and whether protesters were from a racial or ethnic minority group	DCA
noorgigc	Describes whether there was a protest and whether a formal interest group participated in the collective action event	DCA
educ	Percent of people in a congressional district 25 or older who have completed 4 years of college or more	Lublin (1997)
educm	Mean education within a congressional district over time	Lublin (1997)
dem	Whether the legislator is a Democrat	Lublin (1997)
lengthserve	Legislator's total length of service (in years)	Lublin (1997)
lenm	Mean length of service within a congressional district over time	Lublin (1997)
votemargin	Legislator's margin of victory in the previous election	CQ Press
margm	Mean margin of victory over time for a legislator	CQ Press
blkrep	Black Representative	Lublin (1997)
hisprep	Latino Representative	Lublin (1997)
womanrep	Woman Representative	Lublin (1997)
wnom10to1	W-NOMINATE scores for each legislator for each congress	Voteview
wnommin10to1	W-NOMINATE scores for each legislator for each congress based only on minority roll call votes	Voteview
osize	Number of participants reported as taking part in the event	DCA

Table 1. Overview of the data (cont'd)

Variable	Description	Source
numprotest	The number of protests in a congressional district prior to the roll call vote.	DCA
numportestpoor	The number of protests in a congressional district prior to the roll call vote featuring low-income protesters.	DCA
numprotestnp	The number of protests in a congressional district prior to the roll call vote featuring higher income protesters.	DCA
numprotestp	The number of protests in a congressional district prior to the roll call vote given the protesters' income.	DCA
numprotestnonw	The number of protests in a congressional district prior to the roll call vote featuring non-white protesters.	DCA
numprotestwhite	The number of protests in a congressional district prior to the roll call vote featuring white protesters.	DCA
numprotestnw	The number of protests in a congressional district prior to the roll call vote given the protesters' race.	DCA
numprotestnoorg	The number of protests in a congressional district prior to the roll call vote not featuring a formal interest group.	DCA
numprotestig	The number of protests in a congressional district prior to the roll call vote featuring a formal interest group.	DCA
numprotestnoig	The number of protests in a congressional district prior to the roll call vote based on whether a formal interest group participated in the protest.	DCA
disruptive0	Indicates whether an event had disruptive characteristics	DCA
stories0	Indicates whether more than one article was written on the collective action event	DCA

eventid
Unique identifier for collective action events extracted from the DCA.

statecd
String variable used to merge many of the datasets. In many cases, the state was identified using the state's full name or it's abbreviation. The corresponding state code was combined with the congressional district number to create the statecd variable. For example, Alabama's (AL) state code is 1. The statecd for Alabama's 1st congressional district is "1 1".

statecd1
Numeric version on the statecd string variable.

cnrptyy
For each dataset, cnrptyy was created using the year the variable was recorded to identify the corresponding congressional district. For example, rptyy is a variable extracted from the DCA for the year in which a protest event occurred. cnrptyy identifies the corresponding congressional district. cnprtyy==102 if a collective action event occurs in 1992.

majortopic
Policy Agendas Project classification code categorizing the relevant issue area for a protest claim or roll call vote.

lib
Created based on the valence and major topic classification of each collective action event's claim. Civil Rights, Civil Liberties, and Minority Issue area protests with valence ==1 (i.e., Protesters are acting for or in favor of the issue represented by the claim) are coded as lib==1, otherwise lib==0.

cons
Created based on the valence and major topic classification of each collective action event's claim. Civil Rights, Civil Liberties, and Minority Issue area protests with valence ==2 (i.e., Protesters are acting against or in opposition to the issue represented by the claim) are coded as cons==1, otherwise cons==0.

rccount
For each protest claim, I searched govtrack.us and recorded the first, final passage roll call vote relevant to the protest.

support
Indicates whether a legislator for a congressional district in which a protest occurs supports protesters' issue preference based on the legislator's vote on the rccount and the protesters' claim as stated in the DCA.

supportall2
If there is a protest in a congressional district, supportall2==support. If there is not protest in a congressional district, supportall2 indicates whether a legislator's vote on a rccount aligns with the preferences of constituents in districts in which protest does not occur. A description of the ANES questions used to determine protest related constituency opinion is available in the online appendix. The article describes the MRP process that produced constituency preference estimates for districts without a protest.

protest
If an observation has a nonmissing eventid, then protest==1, otherwise protest==0.

protestincd
For each rccount, statecd, and cnrptyy, if protest==1 for any observation within a congressional district and prior to a roll call vote in the congressional term then protestincd==1, otherwise protestincd==0.

poorh
If any of the initiating group characteristics (igrp1c1-igrp3c1) from the DCA record poor or homeless participants, then poorh==1. Otherwise, poorh==0.

nonwhite
If any of the initiating group characteristics (igrp1c1-igrp3c1) from the DCA record any of the following descriptions:

- 0410 Asian, SE Asian, Vietnamese, Chinese, Japanese, Korean, Indonesian, Malaysian, Other Asian
- 0415 Pacific Islanders, Filipino, Hawaiian, Samoan, Tongan, Melanesian, Fijian, Micronesian, Chamorro, Guam
- 0416 Asian Indian
- 0424 Caribbean, Haitian, Dominican Republic, Bahamian, Jamaican, Trinidadians, Other Caribbean
- 0434 Arab/Other Middle Eastern
- 0435 Native American/American Indian
- 0460 Foreign Born, Immigrant, Other Ethnic
- 0495 People of Mixed Race

then nonwhite==1. Otherwise, nonwhite==0.

noorgig
If a social movement organization participated in the collective action event, then noorgig==0. Otherwise noorgig==1.

poorhc
If there was no protest in the district, poorhc==0. If poorh==0, then poorhc==1. If poorh==1, then poorhc==2.

nonwhitec
If there was no protest in the district, nonwhitec==0. If nonwhite==0, then nonwhitec==1. If nonwhite==1, then nonwhitec==2.

noorgigc
If there was no protest in the district, noorgigc==0. If noorgig==0, then noorgigc==1. If noorgig==1, then noorgigc==2.

educ
Percent of people in a congressional district who are 25 or older and who have completed 4 years of college or more. Coded 0 to 1 by dividing each congressional district's educ value by the most educated congressional district's educ value.

educm
Mean education within a congressional district over time, created using the egen mean command.

dem
dem==1 if partycode==100, dem==0 if partycode==200, dem==., and if partycode==300.

lengthserve
Legislator's total length of service (in years). Coded 0 to 1 by dividing each legislator's length of service by the longest serving legislator's length of service.

lenm
Mean length of service within a congressional district over time, created using the egen mean command.

votemargin
Legislator's margin of victory in the previous election.

margm
Mean margin of victory over time for a legislator, created using the egen mean command.

blkrep
Black Representative (1 = Black, 0 = Other).

hisprep
Latino Representative (1 = Mexican, Puerto Rican, or Cuban; 0 = Other)

womanrep
Women Representative (1 = Woman, 0 = Man)

wnom10to1
W-NOMINATE scores for each legislator for each Congress. Created using the wnominate package in R.

wnommin10to1
W-NOMINATE scores for each legislator for each congress based only on roll call votes with the Civil Rights, Minority Issues, and Civil Liberties major topic classification and roll call votes on social welfare and inequality issues based on the definition in Gillion (2012). Created using the wnominate package in R.

osize
Number of participants reported as taking part in the event. Exact numbers of participants are grouped into the following categories:

1 – Small, few handful (1-9 people)
2 – Group, committee (10-49 people)
3 – Large, gathering (50-99 people)
4 – Hundreds, mass, mob (100-999 people)
5 – Thousands (1,000-9,999 people)
6 – Tens of thousands (10,000 or more people)

numprotest
The number of protests in a congressional district prior to the roll call vote. Created using the egen command.

numportestpoor
The number of protests in a congressional district prior to the roll call vote featuring low-income protesters. Created using the egen command.

numprotestnp
The number of protests in a congressional district prior to the roll call vote featuring higher income protesters. Created using the egen command.

numprotestp
If poorh==1, then numprotestp==numprotestpoor. If poorh==0, then numprotestp==numprotestnp.

numprotestnonw
The number of protests in a congressional district prior to the roll call vote featuring non-white protesters. Created using the egen command.

numprotestwhite
The number of protests in a congressional district prior to the roll call vote featuring white protesters. Created using the egen command.

numprotestnw
If nonwhite==1, then numprotestnw==numprotestwhite. If nonwhite==0, then numprotestnw==numprotestnonw.

numprotestnoorg
The number of protests in a congressional district prior to the roll call vote not featuring a formal interest group. Created using the egen command.

numprotestig
The number of protests in a congressional district prior to the roll call vote featuring a formal interest group. Created using the egen command.

numprotestnoig
If noorgig==1, then numprotestnoig==numprotestnoorg. If noorgig ==0, then numprotestnoig==numprotestig.

disruptive0
disruptive0==1 if the article mentions that a collective action event contained property damage, police presence, arrests, weapons, injury, or death. Otherwise, disruptive0==0.

stories0
Indicates whether more than one article was written on the collective action event.

References

Gillion, Daniel Q. 2012. "Protest and Congressional Behavior: Assessing Racial and Ethnic Minority Protests in the District." Journal of Politics 74, no. 4 (October): 950–962.

Lewis, Jeffrey B., Keith Poole, Howard Rosenthal, Adam Boche, Aaron Rudkin, and Luke Sonnet (2019). Voteview: Congressional Roll-Call Votes Database. https://voteview.com/

Lublin, David. 1997. The Paradox of Representation. Princeton: Princeton University Press.