

## Replication Data Codebook 2

for

Heather Stoll, 2013, *Changing Societies, Changing Party Systems*, New York: Cambridge University Press:

Data for the presidential elections analysis (Chapter 4), including the original aggregate index measuring social heterogeneity for democracies with popularly elected presidents (Chapter 3)

Variable	Description and Source
country	Country
year	Election year
enpres	Effective number of presidential candidates (Golder 2005); see p. 103 for details
fused	Dummy variable coded "1" for a fused electoral system and "0" otherwise; fused electoral systems were eliminated from the analysis, so all cases included in the replication data set are coded "0" (Golder 2005)
presplurality	Dummy variable coded "1" if the presidential electoral is simple plurality and "0" otherwise (Golder 2005); see p. 103 for details
prespluralityus	An alternative version of variable "presplurality" that codes the United States as using simple plurality for its presidential elections (because of the United States's electoral college, variable "presplurality" codes the United States as <i>not</i> using simple plurality)
index	My index of social heterogeneity; see p. 107 for an overview of its construction and p. 77-87 for details; sum of variables "gini"; "rural"; "highincome"; "ethroe"; "war"; and "relf"
gini	Gini coefficient, earliest year available, except for Eastern European countries, for which the earliest data available since independence or democratization is used (World Bank Group DDG 2009); see p. 82-83 for details
gini1	Alternative to variable "gini", where the earliest data available for all countries is employed
gini2	Alternative to variable "gini", where the average of the available data is taken over the period of analysis (1980-2000)
gini3	Alternative to variable "gini", where the earliest data available is used, except for either newly independent or newly democratic countries (e.g., Malawi and Russia), for which data is taken from the year closest to their first appearance in the data set (privileging prior years over following years)
rural	Percentage of the population that is rural, earliest year available (World Bank Group DDG 2009); see p. 81-82 for details
rurf	Rurality fractionalization index, an alternative to variable "rural", calculated in the usual way for the two groups of urban and rural individuals $(1 - (\text{rural}^2) - (1 - \text{rural})^2)$
rural1	Alternative to variable "rural", for which data is taken from 1960 or the earliest year available, except for either newly independent or newly democratic states (e.g. Malawi and Russia), for which data is taken for the year preceding their first appearance in the data set
rural2	Alternative to variable "rural", where the average of the available data is taken over the period of analysis (generally, 1960-2000)
highincome	Dummy variable coded "1" for high income countries (those with a 2008 gross national income per capita of \$11,906 or more) and "0" otherwise (World Bank Group DDG 2009); see p. 83 for details

war	Percentage of the post-World War II period (1946-2005) in which a state existed and was engaged in international armed conflicts (World Bank Group HSRP 2008); see p. 83-84
warnum	Alternative to variable "war"; the number of international armed conflicts in which a state engaged in the post-World War II period (1946-2005) when it existed
ethroe	Ethnolinguistic fractionalization (ELF) index created by Soviet geographers in 1961, as modified by Roeder (2001) to make use of data on sub-groups (taking into account racial divisions within ethnolinguistic groups); for countries not in existence in the early 1960s (e.g., Russia), the Roeder variant of ELF calculated using Soviet measurements from 1985; see p. 80-81 for details
ethale	Alternative to variable "ethroe": ethnic fractionalization index of Alesina et al. (2003)
elf	Alternative to variable "ethroe": the original ethnolinguistic fractionalization index (ELF), with data taken from Fearon and Laitin (2003)
ethpolale	Alternative to variable "ethroe": ethnic polarization index of Montalvo and Reynal-Querol (2001), calculated using data from Alesina et al. (2003)
culfr	Alternative to variable "ethroe": Cultural fractionalization index of Fearon (2003)
relf	Index of religious fractionalization of Alesina et al. (2003); see p. 80-81 for details
relann	Alternative to variable "relf": religious fractionalization index of Annett (2001)
relfealat	Alternative to variable "relf": religious fractionalization index of Fearon and Laitin (2003)
relpolale	Alternative to variable "relf": religious polarization index of Montalvo and Reynal-Querol (2001), calculated using data from Alesina et al. (2003)
prespowers	De jure presidential powers, primarily coded by Hicken and Stoll (Hicken and Stoll 2008, variable "prespowers2" from their replication data set); see p. 104-106 for more details about this measure's construction and variable "prespowers2" below for more information about the source of extant codings.
coding1a	Dummy variable coded "1" for country-years for which Hicken and Stoll (2008) did not themselves code constitutions (i.e., cases for which they either drew upon extant codings or extrapolated codings from other years) and "0" for the cases they themselves coded
coding1b	Dummy variable coded "1" for country-years for which Hicken and Stoll (2008) extrapolated the coding of presidential powers and "0" otherwise
prespowers2	Alternative measure of presidential powers, which replaces Hicken and Stoll's codings with extant codings where there is disagreement (Hicken and Stoll 2008; variable "prespowers" from their replication data set). Source is as given in variable "coding2"
coding2	Source of presidential powers data of variable "prespowers2": 0 denotes extrapolation of existing coding by other scholars, in absence of Hicken and Stoll's knowledge of any constitutional changes in that year; 1 denotes Shugart and Carey (1992); 2 denotes Frye, Hellman and Tucker (2000); and 3 denotes Hicken and Stoll's own coding of the case where they were aware of recent (post-1992 in the case of Shugart and Carey and post-1995 in the case of Frye, Hellman and Tucker) constitutional changes, or when the case was not included

	in either of the original (Shugart and Carey or Frye, Hellman and Tucker) data sets (Hicken and Stoll 2008, variable “coding” in their replication data set)
regime2	An ordinal scale measuring the type of regime: 2 = true presidential; 1 = mixed; 0 = parliamentary (Golder 2005); see p. 105-106 for details
regf	A nominal version of the ordinal scale variable “regime2,” with regime types coded “Presidential”, “Mixed”, or “Parliamentary”
region	Region of the world to which the country belongs, one of six categories: “Advanced Industrial”; “Asia”; “Latin America”; “Eastern Europe”; “Africa”; and “Other” (Pacific and Caribbean islands) (Hicken and Stoll 2008); see p. 106 for details
newdem	Dummy variable coded “1” for elections that are the first since either independence or a transition to democracy, and “0” otherwise (Hicken and Stoll 2008); see p. 106 for details
oneelect	Dummy variable coded “1” for countries with only one election in the pre-list-wise deleted data set and “0” otherwise

Notes:

- Missing data is coded “NA”. Cases with missing data that otherwise satisfy criteria for inclusion are included in the replication data set. For Models 1 and 2 reported in Chapter 4, these cases are listwise deleted.
- Variables appearing in Model 1 are: enpres; index; prespowers; presplurality; and region (omitted category is advanced industrial democracies).
- Variables appearing in Model 2 are: enpres; index; regime; presplurality; and region (omitted category is advanced industrial democracies).
- All other variables included in the data set are used as part of the sensitivity analysis reported in the text (i.e., as alternative measures and for alternative model specifications).
- See Stoll (2013) for the list of references cited here.