TRUST IN GOVERNMENT - General Target Variable Report (GVR)

1. General Information

We provide two harmonized measures of respondents' trust in the country's government: T_TRGOV_11 and T_TRGOV_DISTRIB. Section 3.2 outlines the harmonization rules for each target variable.

T_TRGOV_11 measures respondents' reported trust on an 11-point scale. Target values range from 0 (lowest degree/intensity of trust) to 10 (highest degree/intensity of trust).

T_TRGOV_DISTRIB measures the relative position of a respondent in the distribution of trust in the country's government in a given national survey. The scores of this target variable are percentiles within the national sample that indicate what share of respondents reports the same or lower trust than the individual.

Both target variables are accompanied by harmonization control variables that capture special features of the source variables. Three of them refer to the response scales: length (C_TRGOV_SCALE_LENGTH), direction (C_TRGOV_ASCEND) and polarity (C_TRGOV_POLARITY). One control variable captures special features of the source variables with regard to wording (C_TRGOV_NONSTANDARD) (see Table 1.1 and Section 3.3).

The target variable report for TRUST IN GOVERNMENT is accompanied by the following Excel documents:

- The Detailed Variable Report (DVR) T_TRUSTS_DVR_SDR2.xlsx. DVR Excel files in SDR2 systemize all information about source variables that were used for harmonization into a given target variable of the SDR2 database;
- A Crosswalk Table (CWT): T_TRGOV_CWT_SDR2.xlsx. CWT Excel files in SDR2 contain details about mapping of source values to target values.

Table 1.1. TRUST IN GOVERNMENT: Description of the target, source, and control variables

	Variable description	**Variable name**	**Variable values**[a]
Target variables	Trust in government (11-point scale)	T_TRGOV_11	0 = lowest degree 10 = highest degree
	Trust in government (distribution-preserving scale)	T_TRGOV_DISTRIB	0 = lowest percentile point in distribution 100 = highest percentile point in distribution
Source variable			See:T_TRUSTS_DVR_SDR2.xlsx and T_TRGOV_CWT_SDR2.xlsx
Control variables	The length of the rating scale (i.e. number of	C_TRGOV_SCALE_LENGTH	2 = 2-point scale 4 = 4-point scale 5 = 5-point scale

	answer options) that the source variable uses		7 = 7-point scale 10 = 10-point scale
	Source values: scale direction	C_TRGOV_ASCEND	0 = descending 1 = ascending
	Polarity of source rating scale	C_TRGOV_POLARITY	1 = UNIPOLAR (TRUST) 2 = BIPOLAR (TRUST/DISTRUST) 3 = BIPOLAR (AGREE/DISAGREE) 4 = UNIPOLAR (FREQUENCY)
	Nonstandard wording of source question on 'trust in government'	C_TRGOV_NONSTANDARD	0 = standard question wording, and answer options refer to intensity 1 = question includes "do what is right" 2 = question includes "do what is right," and answer options concern frequency

[a] Missing values are assigned according to the SDR2 missing codes schema, provided in the Appendix.

2. Survey projects

Source variables that we used for T_TRGOV appear in 19 international survey projects: ABS, AMB, ARB, ASES, CB, CDCEE, EB, EQLS, EVS, ISJP, ISSP, LB, LITS, NBB, NEB, PA1, PA2, VPCPCE, WVS, 80 waves and 1517 national surveys. The data cover 125 countries and years from 1973 to 2017.

3. General Rules and Procedures

3.1. Source data description

The question about trust in government is among the most frequently included political attitudes items in the surveys we selected for harmonization. Yet, the question wording and, especially, the type of response scales and coding of answers, vary across survey projects. Table 3.1 illustrates variation in properties of response scales.

Table 3.1. TRUST IN GOVERNMENT: Types of scales

Length of scale	Direction of scale	
	Descending	**Ascending**
2	EB (55.1, 57.1, 60.1, 63.4, 66.1, 68.1, 73.4, 77.3, 79.3, 81.4, 83.3, 85.2, 87.3)	

4	ABS/1-4, ARB/2-4, ASES, CDCEE/1-2, EB/24, EVS/4, LB/1995-1996 and 2001-2016, NBB/3, PA1, PA2, WVS/2-6	
5	ISJP/1-2, ISSP (2004,2010,2014), VPCPCE	CB/2008-2015, LITS/1-3
7		AMB/2004-2016, NEB/2-5
10		EB (72.1, 74.1), EQLS/2-3

3.2. Rules of transformation of the source variable into target variables

To construct the 11-point scale and distribution-preserving target variables, we first create preparatory scales. This involves recording the values of the source scales using the consecutive numbers *k*, where *k* ranges from 1 to n. The value 1 of the preparatory scale corresponds to the lowest trust, and higher scores correspond to higher trust (ascending direction). Each preparatory scale is of the same length as the source scale it was derived from.

3.2.1 Constructing the 11-point target scale

To construct the 11-point target scale, we use the preparatory scales and assign scores to them in the interval from 0 to 10, according to the following linear transformation:

$$l(k) = \frac{10}{n*2} + (k - 1) * \frac{10}{n}$$

where: $l(k)$ is a target score corresponding to the preparatory score k, and n is the number of k-values.

This process involves "stretching" preparatory (and thus, source) scales that have fewer than 10 points and keeping 11-point scales as such. Figure 3.2.1 and Table 3.2.1 depict transformations using this type of rescaling.

Figure 3.2.1. Transformation of source values into the target 11-point scale with 0 to 10 range

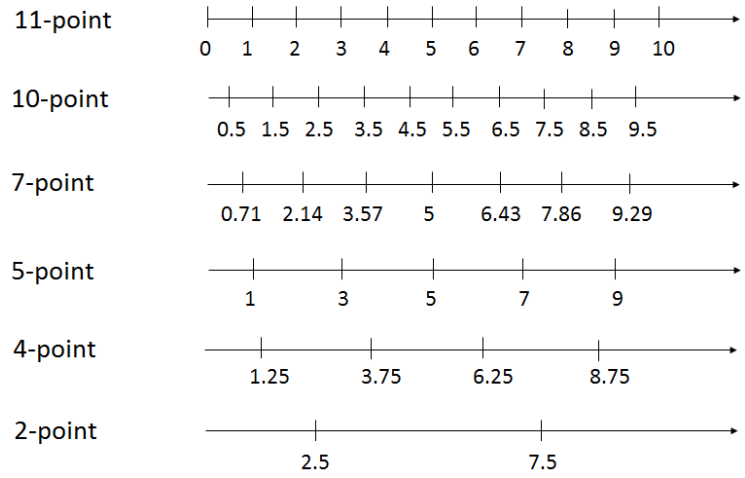

Table 3.2.1. Creating the 11-point scale (from 0 to 10), with median and mean values 5, and minimized inter-scale differences in the variability

Source scale length	Recodes	Median Mean	Average of absolute deviations	Variance	Standard deviation
11-point	0,1.0, 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0	5.0	2.72	10.00	3.16
10-point	0.5, 1.5, 2.5, 3.5, 4.5, 5.5, 6.5, 7.5, 8.5, 9.5	5.0	2.50	8.25	2.87
7-point	0.7, 2.1, 3,6, 5.0. 6,4, 7.9, 9,3	5.0	2.46	8.25	2.87
5-point	1.0, 3.0, 5.0, 7.0, 9.0	5.0	2.40	8.00	2.83
4-point	1.25, 3.75, 6.25, 8.75	5.0	2.50	7.81	2.80
3-point	1.7, 5.0, 8.3	5.0	2.20	7.26	2.69
2-point	2.5, 7.5	5.0	2.50	6.25	2.50

3.2.2 Constructing the distribution-preserving target scale

To construct the distribution-preserving target scale, we take into account respondents' position in the distribution of reported trust values in a given national sample. For an *n*-point preparatory scale, for values *k* that range from 1 to *n,* where X_i is the percent distribution of the variable in sample *s*, *k* is recoded to:

$$k = \sum_{i=1}^{k-1} X_i + \frac{X_k}{2}$$

The distributional score for the answer option *k* is the sum of percentiles of all previous answer options up to *k-1* plus half of the percentile of the answer option *k*.

For a given sample, each scale point of the distribution target scale corresponds to the midpoint of the cumulative distribution of scores *k* (see Table 3.2.2). Put differently, the scores of the distributional target scale are percentiles that indicate what share of respondents within a national sample reports the same or lower value than the individual. The target variable is computed using unweighted samples.

Note that we do not include 2-point source scales to create the distribution-preserving target scale.

Table 3.2 illustrates how we transform **preparatory** variables (which recode **source** variables' values in ascending direction) with 5 response options into the distribution-based target variable.

Table 3.2. Example of the distribution-based transformation of 5-point preparatory variables into T_TRGOV_DISTRIB.

Preparatory variable values, based on source values k	**Percentage distribution** X_k	**Cumulative percentage distribution** $\sum_{i=1}^{k} X_i$	**Interval** $\sum_{i=1}^{k-1} X_i$	**Interval lower bound plus interval midpoint** $\sum_{i=1}^{k-1} X_i + \frac{X_k}{2}$	**Target value (rounded to integer)**
1 = lowest trust	10.68	10.68	0	= 10.68/2 = 5.34	5
2	32.75	43.44	10.68	= (10.68 + 32.75)/2 = 27.05	27
3	32.11	75.55	43.44	= (43.44 + 32.11)/2 = 59.49	59
4	21.69	97.23	75.55	= (75.55 + 21.69)/2 = 86.39	86
5 = highest trust	2.77	100	97.23	= (97.23 + 2.77)/2 = 98.61	99

Missing values and different situations that warrant to be treated as missing data are coded according to the SDR2 missing codes schema, provided in Table A.1 in the Appendix.

3.3. Methodological variables that accompany TRUST IN GOVERNMENT

The main variation in source questions about trust in government stems from differences in response scales and wording. We provide four harmonization control variables that store specific features of the source variables (see Table 1.1):

1. C_TRGOV_LENGTH is a nominal variable that identifies the length (i.e. number of answer options) of source scales used to construct T_TRGOV_11 or T_TRGOV_DISTRIB. It can take the following values: 2, 4, 5, 7, 10.

   Since we do not harmonize 2-point scales into T_TRGOV_DISTRIB, for this target, C_TRGOV_LENGTH takes the values 4, 5, 7, 10.

2. C_TRGOV_ASCEND is dichotomous. It takes the value 1 for source scales whose values are in ascending order (i.e. response options are ordered from least to most trust); it takes the value 0 when source scale values are in descending order (i.e. response options are ordered from most to least trust).
3. C_TRGOV_POLARITY is coded 1 if the source scale is unipolar, e.g. from no trust to a lot of trust), 2 if the scale is bipolar in item-specific way, e.g. from a lot of distrust to a lot of trust, 3 if the scale is bipolar using agree - disagree scale, and 4 if the source scale is unipolar specifying how frequently respondent trusts government's actions.
4. C_TRGOV_NONSTANDARD is coded 1 if the source question asks whether the government can be trusted to do what is right or best for the country; 2 if the source question asks whether the government can be trusted to do what is right and answers concerns frequency and not intensity; 0 if the source question wording is standard (i.e. the most frequent type of wording in the source data) and answer options refer to intensity of trust.

4. Special cases

- ABS waves 1 and 2 feature a discrepancy between the scale direction presented in source questionnaires on the one hand, and codebooks and data dictionaries on the other. Specifically, for wave 1 the questionnaire provides the following scale:

  1 = A great deal of trust; 2 = Quite a lot of trust; 3 = Not very much trust; 4 = None at all;

  <u>while the codebook and data dictionary provide:</u>

  0 = Not sure; 1= None at all; 2 = Not very much trust; 3 = Quite a lot of trust; 4 = A great deal of trust; 98 = Don't know; 99 = No Answer.

  For wave 2, the scale according to the questionnaire is:

  1 = A great deal of trust; 2 = Quite a lot of trust; 3 = Not very much trust; 4 = None at all; 7 =DU; 8 = CC; 9 = DA

  <u>while the scale from codebook and data dictionary reads:</u>

  1 = None at all; 2 = Not Very Much Trust; 3 = Quite a Lot of Trust; 4 = A Great Deal of Trust; 7 = Do not understand the question; 8 = Can't choose; 9 = Decline to answer).

We use the scales as provided in the codebook and data dictionary, whose values are in ascending order, to construct the target variable. Yet, for the control variable C_TRGOV_ASCEND it is important what scale direction the respondent was presented with; hence, we assign C_TRGOV_ASCEND the code 0 (descending), based on the direction of the scale as available in the questionnaire.

- A difference in question wording beyond the phrase “to do what is right” appears in LB/2001 (*I am going read a statement. Please, tell me if you strongly agree, somewhat agree, somewhat disagree or strongly disagree. ´You can generally trust the people who run our government to do what is right´.*)
- In LB/2008-2013 value labels in the questionnaire consist of one letter non-unique (and somewhat misleading) abbreviations: 1 = L, 2 = S, 3 = L, 4 = N; in SPSS dictionary the values are labeled with full words: 1 = A lot of confidence, 2 = Some confidence, 3 = Little confidence, 4 = No confidence at all.

Appendix: Codes for missing values in SDR2

In the SDR database v.2 we identify different situations that warrant to be treated as missing data. Table A.1 lists all SDR2 missing value codes:

Table A.1. Codes for missing values in SDR2

SDR tag [a]	SPSS (STATA) codes	Label
		Standardized source codes for missing values
DK	-1 (.a)	Don't know
NA	-2 (.b)	No answer
REF	-3 (.c)	Refusal
DU	-4 (.d)	Don't understand the question
DNR	-5 (.e)	Any combination of DK, NA, REF, DU
INAP	-6 (.f)	Inapplicable
NEC	-7 (.g)	Not elsewhere classified
		SDR created codes for missing values
UNFIT	-8 (.h)	Source value does not fit to target
ERR	-9 (.i)	Errors in source data and undocumented source values
COMBI	-10 (.j)	Different missing codes on multiple sources taken for a target
CINAP	-11 (.k)	For control variables only: inapplicable
INSUF	-12 (.l)	For survey: Insufficiently defined response categories
QNA	-20 (.t)	For survey: Question not available

[a] Abbreviations for the labels corresponding to the SDR2 codes for missing values. These tags are used in the Crosswalk Table (CWT) files (Excel) that accompany documentation of SDR2 target variables.

In exceptional situations when codes for missing data listed in Table A.1 cannot be used, we apply a system missing <null> value.