TRUST IN PARLIAMENT - General Target Variable Report (GVR)

1. General Information

We provide two harmonized measures of respondents' trust in the national parliament: T_TRPARL_11 and T_TRPARL_DISTRIB. Section 3.2 outlines the harmonization rules for each target variable.

T_TRPARL_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_TRPARL_DISTRIB measures the relative position of a respondent in the distribution of trust in the national parliament 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 response scales of the source variables: length (C_TRPARL_LENGTH), direction, (C_TRPARL_ASCEND) and polarity (C_TRPARL_POLARITY) (see Table 1.1 and Section 3.3).

The target variable report for TRUST IN PARLIAMENT 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;
- The Crosswalk Table (CWT): T_TRPARL_CWT_SDR2.xlsx. CWT Excel files in SDR2 contain details about mapping of source values to target values.

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

	Variable description	**Variable name**	**Variable values**[a]
Target variable	Trust in parliament (11-point scale)	T_TRPARL_11	0 = Lowest degree 10 = Highest degree
	Trust in parliament (distribution-preserving scale)	T_TRPARL_DISTRIB	0 = Lowest percentile point in distribution 100 = Highest percentile point in distribution
Source variables			See: T_TRUSTS_DVR_SDR2.xlsx and T_TRPARL_CWT_SDR2.xlsx

Control variables	The length of the rating scale (i.e. number of answer options) that the source variable uses	C_TRPARL_LENGTH	2 = 2-point scale 4 = 4-point scale 5 = 5-point scale 7 = 7-point scale 10 = 10-point scale 11 = 11-point scale
	Source values: scale direction	C_TRPARL_ASCEND	0 = Descending 1 = Ascending
	Polarity of source rating scale	C_TRPARL_POLARITY	1 = Unipolar (Trust) 2 = Bipolar (Trust/Distrust)
	Quality control: Processing error types (realized values)	QP_TRPARL	0 = No processing errors 3 = Contradictory variable values 4 = Discrepant variable values 5 = Lack of variable value labels 10 = Contradictory variable values & lack of variable value labels 11 = Discrepant & lack of variable value labels

[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_TRPARL appear in 19 international survey projects: ABS, AFB, AMB, ARB, ASES, CB, CDCEE, CNEP, EB, EQLS, ESS, EVS, ISSP, LB, LITS, NBB, NEB, VPCPCE, WVS, 103 waves and 2045 national surveys. The data cover 153 countries and years from 1981 to 2017.

3. General Rules and Procedures

3.1. Source data description

The question about trust in the national parliament 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 PARLIAMENT: Types of scales

Length of scale	Direction of scale	
	Descending	**Ascending**

2	EB/54.1, 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/1-4, ASES, CDCEE/1-2, EVS/1-4, ISSP/1991(SI),2008(VE), LB/1995-2016, NBB/1,3, WVS/1-6	AFB/2-6
5	ISSP/1991,1998,2008, VPCPCE	CB/2008-2015, LITS/1-3
7		AMB/2004-2016, NBB/5,6, NEB 2-5,7
10		EB/72.1, 74.1, EQLS/2,3
11		CNEP/3(ES,UY), CNEP/4(ES), ESS 1-8

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

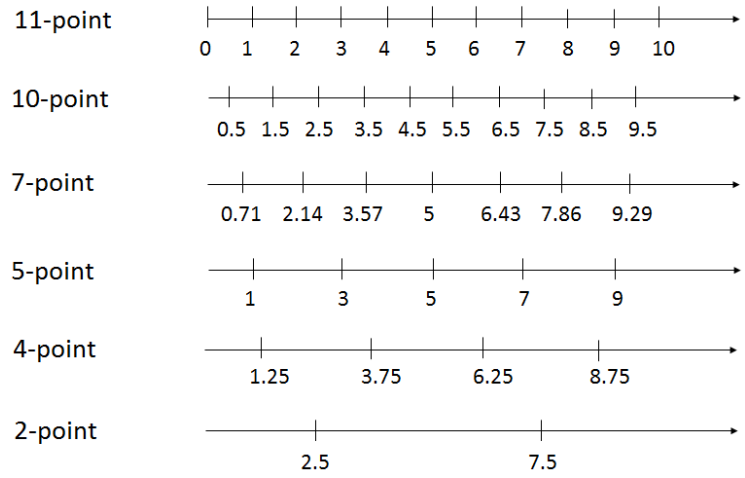

Table 3.2.1. Creating the 11-point scale (from 0 to 10), with median and mean values 5, and minimized inter-scale differences in the variability

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_TRPARL_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 PARLIAMENT

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

1. C_TRPARL_LENGTH is a nominal variable that identifies the length (i.e. number of answer options) of source scales used to construct T_TRPARL_11 or T_TRPARL_DISTRIB. It can take the following values: 2, 4, 5, 7, 10, 11.

   Since we do not harmonize 2-point scales into T_TRPEOPLE_DISTRIB, for this target, C_TRPEOPLE_LENGTH takes the values 4, 5, 7, 10, 11.

2. C_TRPARL_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_TRPARL_POLARITY is coded 1 if the source scale is unipolar (e.g. from no trust to a lot of trust), and 2 if the scale is bipolar (e.g. from a lot of distrust to a lot of trust).

3.4. Quality control variable for processing error types

QP_TRPARL, stored in the SDR2 PLUG_SURVEY file, measures processing error types that we identified when comparing information about the source variables available in the survey documentation (codebook, questionnaire) with information about the same source variables in data records in the source data files (data dictionaries). QP_TRPARL accounts for inconsistencies between different metadata elements, such as data records on the one hand (i.e., variable values, variable labels, value labels - in the source data files), and information in the codebook or questionnaires, on the other hand.

QP_TRPARL is a nominal variable that can take values from 0 (no processing errors) to 13. Values larger than 0 identify what type, or combination of types, of processing errors occur in the source variables associated with T_TRPARL.

Generally, in the SDR2 project we check for five types of processing errors that we define as follows:

- **Illegitimate Variable Values** captures values of the source variable that are outside of the range that SDR proposes as acceptable, e.g., respondents' number of education years exceeding 25, and age when full‑time education was completed below 6.

- **Misleading Variable Values** shows that variable values, as coded in the questionnaires/codebooks, are not congruent with the data records in the source data files.

- **Contradictory Variable Values** captures inconsistencies in how the same variable value is labeled in a codebook vs. a questionnaire or a data dictionary, such as when

"don't know" value is recorded as 88 in the codebook vs. 888888 in the data dictionary.

- **Variable Values Discrepancy** concerns situations when several or all values in the same variable are inconsistently labeled in codebook, vs. questionnaire or a data dictionary (i.e., there are reversed scales or contradictory value labels for more than one variable value), e.g., age of responds when finished education is recorded as year ranges in the questionnaire vs. as years in the data dictionary.

- **Lack of Variable Value Labels** concerns undefined "nulls" and codes (variable values) in the source data file that are not explained in any source of documentation defining variables and their values (i.e., "wild codes: for a given variable, such as value 7 in data records while the documentation describes the variable as a 1 to 5 scale).

Correspondingly, the full list of SDR2 codes for Processing Error Types is:

0 = No processing errors
1 = Illegitimate variable values
2 = Misleading variable values
3 = Contradictory variable values
4 = Variable values discrepancies
5 = Lack of variable value labels
6 = Illegitimate & misleading variable values
7 = Illegitimate & contradictory variable values
8 = Illegitimate & discrepancy variable values
9 = Illegitimate & lack of variable value labels
10 = Contradictory variable values & lack of variable value labels
11 = Discrepant & lack of variable value labels
12 = Illegitimate & contradictory & lack of variable value labels
13 = Illegitimate & discrepant & lack of variable value labels

Values 6 to 13 indicate that a combination of different processing error types occurs within one source variable.

The source variables that we used to create a target T_TRPARL had contradictory variable values (QP_TRPARL = 3), discrepant variable values (QP_TRPARL = 4) and lack of variable value labels (QP_TRPARL = 5) or a combination of these processing error types (QP_TRPARL = 10 and 11). Other types of processing errors did not occur in source data used for this target variable, hence the realized values in the QP_TRPARL variable are 0, 3, 4, 5, 10 and 11.

4. Special Cases

- In the VPCPCE survey there are two questions about trust in parliament, one is country-specific (only for Russia) and another one for all countries participating in the survey

except Russia. The question for Russia (RQ6X1) reads: *I'd like to ask how much you feel you can trust some people and other things. Please could you choose the answer that best represents your opinion? The old parliament suspended by Yeltsin* (Response categories: *1 = Completely trust; 2 = Mostly trust; 3 = Neither; 4 = Mostly distrust; 5 = Completely distrust; 9 = DK*). The general question (ERQ6X): *I'd like to ask how much you feel you can trust some people and other things. Please could you choose the answer that best represents your opinion? Parliament* (Response categories: *1 = Completely trust; 2 = Mostly trust; 3 = Neither; 4 = Mostly distrust; 5 = Completely distrust; 9 = DK*). We combined information from both questions to construct the target variables. A second variable for Russia only (RQ6X2) was excluded from harmonization, because it was asked only as a part of a panel study within the project.

- In EB 66.1, 68.1, 73.4, 77.3 and 79.3 the question was not asked in CY/TCC.
- ABS waves 1 and 2 feature a discrepancy between variable values in questionnaires and values in codebooks and datasets; for wave 1, the question (q010) reads *I'm going to name a number of institutions. For each one, please tell me how much trust you have in them. Is it a great deal of trust, quite a lot of trust, not very much trust, or none at all? Parliament* (scale from questionnaire: *1 = A great deal of trust; 2 = Quite a lot of trust; 3 = Not very much trust; 4 = None at all;* vs. scale from codebook and SPSS dictionary: *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 (q010): *I'm going to name a number of institutions. For each one, please tell me how much trust do you have in them? Parliament* (scale from questionnaire: *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;* vs. scale from codebook and SPSS dictionary: *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*). For the sake of the scale transformation, we treat the source scale as if it were ascending (i.e. as in the data) but the control C_TRPARL_ASCEND is set to 0 (i.e. as it was presumably presented to a respondent).
- CNEP/3/ES, CNEP/3/UY – unipolar question, bipolar scale. The control variable for this survey is coded as bipolar because respondents were presented a show card, which showed a bipolar response scale (0 = Absolutely distrust, 10 = Absolutely trust).
- Some questionnaires use different words to describe the institutional body equivalent to the Parliament, e.g. National Legislature (AMB/2012, AMB/2016/AG codebook), National Congress (AMB/2004-2010, 2014-2016; National Congress/Parliament (LB 2000-2016). In ISSP and EB different institutional bodies were named on country level (e.g. in ISSP: Philippines: Senat, Russia: Supreme Soviet of the USSR, USA: U.S. Congress; in EB: France: Assemblée Nationale, Germany: Bundestag, Great Britain: House of Commons). Sometimes, English language questionnaires and codebooks for the same survey wave contain different words for the Parliament (e.g. AMB/2012 National Legislature in the questionnaire vs. Parliament in the codebook).
- 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.