

Description of Variables

The data set SDR_W prepared for the International Training Event "Cross-national Survey Harmonization and Analysis: Weights, Data Quality and Multi-level Modeling", May 11-16, 2015, OSU Main Campus (Columbus, OH).

A unified schema of user-defined missing values was applied across target and control variables:

-1 = DON'T KNOW

-5 = VARIABLE NOT IDENTIFIED IN SURVEY

-8 = QUESTION NOT ASKED IN SURVEY

-9 = MISSING DATA

T_SOURCE_TABLE

Label: SOURCE TABLE NAME (CORRESPONDS TO SOURCE DATA FILE)

Description: According to applied schema all source data files are stored in corresponding source tables in a relational database. Variable holds the name of source table/file containing a given case.

T_CASE_ID

Label: UNIQUE CASE IDENTIFIER IN SOURCE TABLE (DATA FILE)

Description: Unique case identifier within a source table. Can be used for identification of the case in a corresponding source data file via SPSS \$CASENUM system variable.

T_SURVEY_NAME

Label: SURVEY PROJECT NAME

Description: Acronym of a survey project.

T_SURVEY_EDITION

Label: SURVEY PROJECT EDITION (WAVE/ROUND)

Description: Wave of a survey project. Empty value for standalone projects.

T_COUNTRY_YEAR

Label: YEAR IN WHICH COUNTRY SURVEY WAS CONDUCTED

Description: A year in which survey was conducted in a given country, preferably, stemming from original interview dates, otherwise, from survey documentation. If survey lasted more than a year, the one was chosen which saw most respondents or greater number of months of fieldwork.

T_COUNTRY

Label: COUNTRY (TERRITORY) NAME

Description: Harmonized ISO 3166-1 alpha-2 code of a country. For some territories or national populations a self-invented extension was used, e.g. IL-ARB meaning Arab population in Israel.

S_WEIGHT

Label: SOURCE VALUE OF WEIGHT

Description: Source value of a design or post-stratification weight.

T_WEIGHT

Label: WEIGHT

Description: Harmonized value of a design or post-stratification weight. In some cases its value may differ from a source value due to specific meaning of a given weight. For example, original weight in Americas Barometer reduces samples of all countries to the same size of 1500 respondents. Target value restores actual sample size of a country.

S_GENDER

Label: SOURCE VALUE OF R'S GENDER

Description: Source value of a respondent's gender.

T_GENDER

Label: R'S GENDER

Description: Harmonized respondent's gender.

Values: 0=male, 1=female.

S_AGE

Label: SOURCE VALUE OF R'S AGE

Description: Source value of a respondent's age.

T_AGE

Label: R'S AGE

Description: Harmonized value of a respondent's age.

T_BIRTH_YEAR

Label: R'S YEAR OF BIRTH

Description: Harmonized value of a respondent's year of birth (result of subtraction T_COUNTRY_YEAR-T_BIRTH_YEAR).

S_EDU

Label: SOURCE VALUE OF R'S EDUCATION LEVEL

Description: Source value of the highest education level achieved by respondent.

T_EDU

Label: R'S EDUCATION LEVEL

Description: Harmonized ISCED-2011 value of the highest education level achieved by respondent.

Values: 0=less than primary, 10=primary, 20=lower secondary, 30=upper secondary, 40=post-secondary non-tertiary, 50=short-cycle tertiary, 60=bachelor's or equivalent level, 70=master's or equivalent level, 80=doctoral or equivalent level, 90=not elsewhere classified.

C_EDU_AND_HIGHER

Label: IS R'S EDUCATION LEVEL INCLUDING HIGHER LEVELS?

Description: Dummy variable controlling if respondent's education level extends a given level (for example, MA and higher).

Values: 0=false, 1=true.

C_EDU_INCOMPLETE

Label: IS R'S EDUCATION LEVEL INCOMPLETE?

Description: Dummy variable controlling if respondent's education level is incomplete.

Values: 0=false, 1=true.

C_EDU_VOCATIONAL

Label: IS R'S EDUCATION LEVEL VOCATIONAL?

Description: Dummy variable controlling if respondent's education level is vocational.

Values: 0=false, 1=true.

S_TR_PARLI

Label: SOURCE VALUE OF TRUST IN PARLIAMENT

Description: Source value of respondent's trust in parliament.

T_TR_PARLI_11

Label: TRUST IN PARLIAMENT (11-POINT SCALE)

Description: Harmonized 11-point scale of respondent's trust in parliament. A linear transformation of source scale was applied.

Values: 0 (=fully distrust) thru 10 (=fully trust).

T_TR_PARLI_DISTRIB

Label: TRUST IN PARLIAMENT (DISTRIBUTION-PRESERVING SCALE)

Description: Distribution-preserving scale of trust in parliament. Mid-points of cumulative frequency percentages corresponding to source values were taken.

Values: 0 thru 100.

C_TR_PARLI_ORIG_SCALE_SIZE

Label: ORIGINAL TRUST IN PARLIAMENT SCALE SIZE

Description: Control variable measuring original size of trust in parliament source scale.

Values: 2,4,5,7,10,11-point scales.

C_TR_PARLI_CONFIDENCE

Label: WAS IN QUESTION ABOUT TRUST IN PARLIAMENT USED WORD CONFIDENCE?

Description: Dummy variable controlling if question about trust in parliament used the word "confidence".

Values: 0=false, 1=true.

S_PR_DEMONST

Label: SOURCE VALUE OF TAKING PART IN PROTEST/DEMONSTRATION

Description: Source value of taking part in protest/demonstration.

T_PR_DEMONST_FACT

Label: RESPONDENT'S REPORTED PARTICIPATION IN PROTEST/DEMONSTRATION (AS FACT)

Description: Dummy variable about respondent's taking part in protest/demonstration.

Values: 0=false, 1=true.

T_PR_DEMONST_OPINION

Label: RESPONDENT'S OPINION ABOUT PROTEST/DEMONSTRATION

Description: Dummy variable about positive opinion regarding participation in protest/demonstration.

Values: 0=false, 1=true.

C_PR_DEMONST_YEARS

Label: WHAT PERIOD IN YEARS WAS QUESTION ABOUT TAKING PART IN PROTEST/DEMONSTRATION ASKING?

Description: Control variable measuring period of time given in years that question about taking part in protest/demonstration refers to.

Values: Last 1,2,3,4,5,8,10 and more (=11) years.

C_PR_DEMONST_EXTENDED

Label: DID QUESTION ABOUT TAKING PART IN PROTEST/DEMONSTRATION DEAL WITH OTHER FORMS OF PROTEST THAN DEMONSTRATION?

Description: Dummy variable controlling if question about taking part in protest/demonstration pertains to other forms of protest than demonstration.

Values: 0=false, 1=true.

C_PR_DEMONST_ILLEGAL

Label: WAS THE PROTEST ILLEGAL?

Description: Dummy variable controlling if question about taking part in protest/demonstration pertains to illegal activities.

Values: 0=false, 1=true.

C_PR_DEMONST_SET_OF_QUESTIONS

Label: WAS INQUIRYING ABOUT PROTEST/DEMONSTRATION A SET OF QUESTIONS?

Description: Dummy variable controlling if inquiring about taking part in protest/demonstration is based on the set of questions.

Values: 0=false, 1=true.

QW_IV

Label: NUMBER OF ILLEGITIMATE VALUES IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; ILLEGITIMATE VALUE is when the value in data looks suspicious and was not explained in codebook or SPSS dictionary of a source data file.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_VVD

Label: NUMBER OF VARIABLE VALUE DISCREPANCIES IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; VARIABLE VALUE DISCREPANCY is when a different value label in the codebook and SPSS dictionary of a source data file.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_CVL

Label: NUMBER OF CONTRADICTORY VALUE LABELINGS IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; CONTRADICTORY VALUE LABELING is when a completely different labeling in codebook and SPSS dictionary of a source data file was used.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_LL

Label: NUMBER OF LACK OF LABELS FOR MISSING DATA IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; LACK OF LABELS FOR MISSING DATA is when no information about how "Don't know", "Refused to Answer", "Missing Data" was coded, while there is no description in either codebook or SPSS dictionary of a source data file.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_MV

Label: NUMBER OF MISLEADING VALUES IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; MISLEADING VALUE is when in the codebook and/or questionnaire there is not a complete information about what a user can expect to have in data, like giving only age categories, although in data there is the information about the exact age.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_ML

Label: NUMBER OF MISLEADING VARIABLE LABELS IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; MISLEADING VARIABLE LABEL is when differences between the question and the label are not clear and logical label.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_II

Label: NUMBER OF OCCURRENCES OF INSUFFICIENT INFORMATION IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; INSUFFICIENT INFORMATION is when the general codebook does not include important information, like country specific differences in asking/not asking/coding some question.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QW_TI

Label: NUMBER OF TRANSLATION ISSUES IN DOC/DATA (WAVE LEVEL)

Description: Wave-level quality variable; TRANSLATION ISSUE is when not all documentation is translated into English, e.g. English codebook, but Spanish data with data labels and values.

Values: 0 (=no problems identified) thru HIGHEST (=number of problems identified).

QS_RSP_RATE_VALUE

Label: RESPONSE RATE VALUE FOUND IN DOCUMENTATION (SURVEY LEVEL)

Description: Survey-level quality variable; value of response rate for a national survey as given in documentation.

QS_RSP_RATE

Label: RESPONSE RATE DUMMY (SURVEY LEVEL)

Description: Survey-level quality variable; dummy for existence of response rate information in national survey documentation.

Values: 0=false, 1=true.

QS_TRANSL

Label: DOES DOCUMENTATION INCLUDE INFORMATION ABOUT ANY TRANSLATION METHOD MORE SOPHISTICATED THAN TRANSLATION BY THE SURVEY TEAM ITSELF? (SURVEY LEVEL)

Description: Survey-level quality variable; dummy for documentation including information about any translation method more sophisticated than translation by the survey team itself.

Values: 0=false, 1=true.

QS_PRETEST

Label: DOES DOCUMENTATION INCLUDE INFORMATION ABOUT PRETESTING OR PILOTING? (SURVEY LEVEL)

Description: Survey-level quality variable; dummy for documentation including information about pretesting or piloting.

Values: 0=false, 1=true.

QS_FIELDW_CTRL

Label: DOES DOCUMENTATION INCLUDE INFORMATION ABOUT FIELDWORK CONTROL/BACK-CHECKING? (SURVEY LEVEL)

Description: Survey-level quality variable; dummy for documentation including information about fieldwork control/back-checking.

Values: 0=false, 1=true.

QR_DUPLICATE

Label: IS THIS CASE DUPLICATED? (RECORD LEVEL)

Description: Record-level quality variable; dummy marking the case being a duplicate of another case.

Values: 0=false, 1=true.

M_GDP

Label: GDP BASED ON PPP PER CAPITA GDP

Description: Macro variable; GDP based on PPP per capita GDP; source: EconStats, <http://econstats.com/weo/V011.htm>, accessed in April 2015.

M_POPULATION

Label: COUNTRY/TERRITORY POPULATION SIZE

Description: Macro variable; population size of a given country/territory/national population; source: World Bank, <http://data.worldbank.org/indicator/SP.POP.TOTL>, accessed in April 2015.

M_FREEDOM_CL

Label: FREEDOM HOUSE INDEX (CIVIL LIBERTIES)

Description: Macro variable; Civil Liberties index; source: Freedom in the World - Freedom House, <https://freedomhouse.org>, accessed in January 2013.

Values: 1 (=free) thru 7 (=not free).

M_FREEDOM_PR

Label: FREEDOM HOUSE INDEX (POLITICAL RIGHTS)

Description: Macro variable; Political Rights index; source: Freedom in the World - Freedom House, <https://freedomhouse.org>, accessed in January 2013.

Values: 1 (=free) thru 7 (=not free).