

CODEBOOK

Do Ukrainians Still Prefer Self-Defense Against Russia At Any Cost

Janina Dill*

Marnie Howlett[†]

Carl Müller-Crepon[‡]

October 24, 2025

Abstract

This file contains the description of the main analysis dataset contained in the replication dataverse.

*Department of Politics and International Relations, University of Oxford and Nuffield College.
Corresponding author: janina.dill@politics.ox.ac.uk

[†]Department of Politics and International Relations, University of Oxford

[‡]Department of Government, London School of Economics and Political Science.

1 Main dataset

The main dataset is `data_long_analysis.rds`. It records each profile examined by each respondent as a column. `data_wide_analysis.rds` contains the same data, but in a “wide” format, with one row per respondent. `acled_analysis.rds` contains ACLED conflict data, for plotting purposes and following the standard format documented by ACLED.¹ The following describes the variables in `data_long_analysis.rds`. The last part of this codebook contains the full questionnaire of the survey that produced the data, in English translation and the Ukrainian original.

1.1 Variable description

- `experiment` Experiment ID. 0 (for the data from DHM, 2022), and 1 and 2 for experiments I and II, respectively.
- `in_2022` PSU in DHM sample
- `in_2024` PSU in representative 2024 “main sample”
- `experiment` Experiment ID. 0 (for the data from DHM, 2022), and 1 and 2 for experiments I and II, respectively.
- `oblast` Region, string. Provided by interviewer.
- `chain` Sampling chain ID
- `id` Respondent ID
- `date_day` Day of interview
- `date_month` Month of interview
- `settlement_type` Type of settlement.
 - Village
 - Urban type village
 - Small town (up to 20K)
 - Town with population 20-49K
 - Town with population 50-99K
 - City with population 100-499K
 - Large city 500K+
- `n_chain` ID of respondent in chain. Provided by interviewer.

¹<https://acleddata.com/>

- `language` Language of interview. From question on page Q3 of questionnaire: Which language is most convenient for the respondent to talk to the interviewer.
- `livedFeb23` Lived in current location on 24 February 2022. Question: Until February 24, 2022, that is, before the full-scale Russian invasion, did you live in the same settlement or in another? See questionnaire, page Q3
 1. In the same
 2. In another
- `hh_adults` Number of adults in household. See questionnaire, page Q4. Question: How many adult citizens, over 18 years of age, have permanently resided and currently reside with you, not taking into account those who moved to you after February 23, 2022?
- `sex` Gender of selected respondent.
- `age_years` Age of selected respondent in years.
- `q1, q2, ... q8` Variables encode answers to questions 1 to 8 in the questionnaire below. Items with multiple responses are encoded with multiple variables, e.g., `q7_1, q7_2`, etc. Where applicable, `qX_text` encodes free text answers corresponding to “other” (e.g., `q6_text` encodes “other nationalities” in question 6).
- `q14, q15, ... q20` Variables encode answers to questions 14 to 20 in the questionnaire below. Items with multiple responses are encoded with multiple variables, e.g., `q14_1, q14_2`, etc. Where applicable, `qX_text` encodes free text answers corresponding to “other” (e.g., `q15_5_text` encodes “other type of war affectedness” in question 15).
- `attr_order` Order of attributes 2 to 4, string.
- `interviewer` Interviewer ID
- `age_5gr` Age in 5-year groups
- `pair` Pair ID, by respondents
- `strat` Strategy ID, by pair
- `attr1_level` Numeric level of Attribute 1. Ranges between 1 and 3, see Table in Questionnaire for labels of levels 1 to 3.
- `attr2_level` Numeric level of Attribute 2. Ranges between 1 and 3.

- `attr3_level` Numeric level of Attribute 3. Ranges between 1 and 3.
- `attr4_level` Numeric level of Attribute 4. Ranges between 1 and 3.
- `attr5_level` Numeric level of Attribute 5. Ranges between 1 and 3. Note that the scale is reversed to make it consistent with the other attributes (low values indicate “good” attribute values and high values indicate “bad” attribute values): level 1 = full autonomy; level 2 = negotiated neutrality; level 3 = Russian-controlled government.
- `score` Scoring outcome given to a strategy, 6 values rescaled to 0-1.
- `choice` Forced choice outcome, 0 (strategy is not chosen) or 1 (strategy is chosen by respondent)
- `attr1_amb` Dummy variable that takes 1 if there is no variation on attribute 1 in a strategy pair and takes the value of 0 if there is variation on attribute 1 in a pair.
- `attr1_comp` Dummy variable that takes 1 if a strategy has a better (i.e. lower) attribute 1 than the other strategy in the pair, 0 if the values are the same, and -1 if the value is worse (i.e., higher) than in the other strategy.
- `attr2_amb` Dummy variable that takes 1 if there is no variation on attribute 2 in a strategy pair and takes the value of 0 if there is variation on attribute 2 in a pair.
- `attr2_comp` Dummy variable that takes 1 if a strategy has a better (i.e. lower) attribute 2 than the other strategy in the pair, 0 if the values are the same, and -1 if the value is worse (i.e., higher) than in the other strategy.
- `attr3_amb` Dummy variable that takes 1 if there is no variation on attribute 3 in a strategy pair and takes the value of 0 if there is variation on attribute 3 in a pair.
- `attr3_comp` Dummy variable that takes 1 if a strategy has a better (i.e. lower) attribute 3 than the other strategy in the pair, 0 if the values are the same, and -1 if the value is worse (i.e., higher) than in the other strategy.
- `attr4_amb` Dummy variable that takes 1 if there is no variation on attribute 4 in a strategy pair and takes the value of 0 if there is variation on attribute 4 in a pair.
- `attr4_comp` Dummy variable that takes 1 if a strategy has a better (i.e. lower) attribute 4 than the other strategy in the pair, 0 if the values are the same, and -1 if the value is worse (i.e., higher) than in the other strategy.

- `attr5_amb` Dummy variable that takes 1 if there is no variation on attribute 5 in a strategy pair and takes the value of 0 if there is variation on attribute 5 in a pair.
- `attr5_comp` Dummy variable that takes 1 if a strategy has a lower attribute 5 than the other strategy in the pair, 0 if the values are the same, and -1 if the value is higher than in the other strategy.
- `PSU` PSU ID
- `x_noise` x-coordinate with random noise
- `y_noise` y-coordinate with random noise
- `battle_all_d` Were there any ACLED battles within 10km (0: no, 1: yes)
- `shelling_all_d` Were there any ACLED shelling events within 10km (0: no, 1: yes)
- `onesided_all_d` Were there any ACLED one-sided violence events within 10km (0: no, 1: yes)
- `attr1_fac` Attribute 1 (territorial integrity) as factor, see Table in Questionnaire, page Q2 for levels 1 to 3.
- `attr1_d1` Attribute 1, level 1 (full integrity) as dummy variable (0: no, strategy is not characterized by attribute 1, level 1; 1: yes, strategy is characterized by attribute 1, level 1)
- `attr1_d2` Attribute 1, level 2 (full integrity minus Crimea) as dummy variable (0: no, strategy is not characterized by attribute 1, level 2; 1: yes, strategy is characterized by attribute 1, level 3)
- `attr1_d3` Attribute 1, level 3 (full integrity minus Donbas and Crimea) as dummy variable (0: no, strategy is not characterized by attribute 1, level 3; 1: yes, strategy is characterized by attribute 1, level 3)
- `attr2_fac` Attribute 2 (civilian fatalities) as factor, see Table in Questionnaire, page Q2 for levels 1 to 3.
- `attr2_d1` Attribute 2, level 1 (6'000 civilian fatalities) as dummy variable (0: no, strategy is not characterized by attribute 2, level 1; 1: yes, strategy is characterized by attribute 2, level 1)
- `attr2_d2` Attribute 2, level 2 (12'000 civilian fatalities) as dummy variable (0: no, strategy is not characterized by attribute 2, level 2; 1: yes, strategy is characterized by attribute 2, level 2)

- `attr2_d3` Attribute 2, level 3 (24'000 civilian fatalities) as dummy variable (0: no, strategy is not characterized by attribute 2, level 3; 1: yes, strategy is characterized by attribute 2, level 3)
- `attr3_fac` Attribute 3 (military fatalities) as factor, see Table in Questionnaire, page Q2 for levels 1 to 3.
- `attr3_d1` Attribute 3, level 1 (6'000 military fatalities) as dummy variable (0: no, strategy is not characterized by attribute 3, level 1; 1: yes, strategy is characterized by attribute 3, level 1)
- `attr3_d2` Attribute 3, level 2 (12'000 military fatalities) as dummy variable (0: no, strategy is not characterized by attribute 3, level 2; 1: yes, strategy is characterized by attribute 3, level 2)
- `attr3_d3` Attribute 3, level 3 (24'000 military fatalities) as dummy variable (0: no, strategy is not characterized by attribute 3, level 3; 1: yes, strategy is characterized by attribute 3, level 3)
- `attr4_fac` Attribute 4 (nuclear risk) as factor, see Table in Questionnaire, page Q2 for levels 1 to 3.
- `attr4_d1` Attribute 4, level 1 (no nuclear risk) as dummy variable (0: no, strategy is not characterized by attribute 4, level 1; 1: yes, strategy is characterized by attribute 4, level 1)
- `attr4_d2` Attribute 4, level 2 (low nuclear risk) as dummy variable (0: no, strategy is not characterized by attribute 4, level 2; 1: yes, strategy is characterized by attribute 4, level 2)
- `attr4_d3` Attribute 4, level 3 (high nuclear risk) as dummy variable (0: no, strategy is not characterized by attribute 4, level 3; 1: yes, strategy is characterized by attribute 4, level 3)
- `attr5_fac` Attribute 5 (political autonomy) as factor, see Table in Questionnaire, page Q2 for levels 1 to 3. Note that the scale of levels is reversed to make it consistent with the other attributes: level 1 = full autonomy; level 2 = negotiated neutrality; level 3 = Russian-controlled government.
- `attr5_d1` Attribute 5, level 1 (full autonomy) as dummy variable (0: no, strategy is not characterized by attribute 5, level 1; 1: yes, strategy is characterized by attribute 5, level 1)
- `attr5_d2` Attribute 5, level 2 (negotiated neutrality) as dummy variable (0: no, strategy is not characterized by attribute 5, level 2; 1: yes, strategy is characterized by attribute 5, level 2)

- `attr5_d3` Attribute 5, level 3 (Russian-controlled government) as dummy variable (0: no, strategy is not characterized by attribute 5, level 3; 1: yes, strategy is characterized by attribute 5, level 3)
- `resp_pair` Unique ID per respondent-strategy pair.
- `resp_pair_s` Unique ID per respondent-pair-strategy.
- `survnat` Belief that survival of Ukrainian nation is at stake in the war, binary factor. Based on q18. Values: “no”: answer level 1; “yes”: answer level 2, NA: answer level 3.
- `trustpres` Trust in president, binary factor. Based on q19. Values: “low”: answer levels 1 and 2; “high”: answer levels 3 and 4, NA: answer level 5.
- `victimp` Victory in war is extremely important, binary factor. Based on q17. Values: “Extremely important”: answer level 6; “all others”: answer levels 1 to 5.
- `famaffected` Family affected, binary factor (no, yes, NA: missing / response 13). Based on q16. Response 12 coded as not affected by war, all others except 13 (missing) coded as affected.
- `selfaffected` Respondent affected, binary factor (no, yes, NA: missing / response 8). Based on q15. Response 7 coded as not affected by war, all others except 8 (missing) coded as affected.
- `econdepr` Respondent is economically deprived, binary factor (no, yes). Based on q4. Responses 1 and 2 coded as deprivation.
- `region` Variable encoding eastern (“East??”) vs western (“West”) regions.
- `diraffected` Variable encoding regions first affected by Russian invasion, binary factor (no, yes). Affected regions (yes) are: Chernihiv, Zhytomyr, Kyiv region and city, Dnipropetrovsk, Zaporizhzhia, and Odesa, others are coded as no.
- `lang` language of interview, Ukrainian vs. Russian. Based on language.
- `urbrur` Urban vs rural, based on `settlement_type`
- `children` Any children, binary factor (no, yes). Based on q2. Response 1 coded as no children, all others as children.
- `educ` Level of education, five level ordered factor, based on items in q3.
- `natlang` Native language: Ukrainian, Russian, and other. Based on q5.

- `ethnic` Respondents' ethnicity, Ukrainian, Russian, and other, based on q6.

2 Survey questionnaire (English/Ukrainian)