



JOINT RESEARCH CENTRE
Directorate JRC.I - Competences (Brussels)
Unit I.2: Foresight, Modelling, Behavioural Insights and
Design for Policy

**Request for Service under Framework Contract CHAFEA/2019/CP/01 to provide services
for**

“Collection of data on attitude towards refugees”

Deliverable 8
Data delivery (wave 1)

Authors: Rocco Crimeni

Support: Alice De Berardis, Andrea Temporin

Version: 1

Date: 09/09/2022

History:

Version	Date	Authors	Comment
1	09/09/2022	Rocco Crimeni	

VVA contacts:

Name	Function	Email address
Francesco Rocciola	Project Manager	f.rocciola@vva.it
Rocco Crimeni	Research Manager	r.crimeni@vva.it

Contents

Introduction	4
First wave: fieldwork implementation and conclusion.....	4
Data report methodology	4
Metadata & paradata	5
Survey length	5
Timestamps for choice experiment	5
Response rate	6
Refusal rate	6
Items non-response	6
Opened quotas history	7
Track of changes	7
Weighting system	7
Annex: Final datasets	9

Introduction

This report documents the fieldwork phase for the 1st wave of data collection. Specifically, our focus is on the metadata, sample distribution and weights applied. The report is structured as following:

1. Introduction
2. Fieldwork implementation and conclusion for the first wave
3. Data report methodology
4. Survey feedback based on metadata and paradata collected
5. Opened quotas history
6. Weighting system
7. Annex: final datasets

First wave: fieldwork implementation and conclusion

The fieldwork started Tuesday, August 23 and finished Tuesday, September 6. The panel provider followed predefined quotas making sure that a variety of respondents in terms of age, gender, education, geographical distribution and employment status answered the questions.

Number of interviews per country for the first wave (2022):

Total	Poland	France	Germany	Hungary	Romania	Czech
14009	3500	2101	2101	2101	2106	2100

Initial target per country:

Total	Poland	France	Germany	Hungary	Romania	Czech
14000	3500	2100	2100	2100	2100	2100

As showed by the number of interviews, the fieldwork respected the initial target in terms of numerosity and completed all the interviews in exactly 2 weeks. The number of completed interviews is slightly higher for France, Germany, Hungary and Romania. However, the final distribution of interviews needed some adjustments to balance it with the composition of the above socio-demographic variables (see section “weighting system”).

Data report methodology

Within this report we have analysed the following 3 aspects of the survey: metadata gathered, quotas history per country and weights computation.

For the metadata we analysed the following:

- Survey length
- Timestamps for choice experiment
- Response rate
- Refusal rate
- Items Non-response

Furthermore, the section dedicated to the quotas concentrate on tracking the changes per each single variable included in the sample design.

Lastly, the focus on weights includes a detailed section to explain the computation related to the quotas reached and its connection with the original sample.

Metadata & paradata

Survey length

The LOI (length of interview) was assumed to an average of 15 minutes. During the first wave (2022), respondents for all countries guarantee the possibility to conclude the survey according to the initial forecast. As for the pilot report, also in this case the extreme outliers (duration higher than 45 minutes) have been excluded from the calculation. Here in detail the survey length for the whole sample:

Total average	Poland	France	Germany	Hungary	Romania	Czech
12 m 55 sec	13 m 36 sec	12 m 5 sec	11 m 59 sec	13 m 9 sec	12 m 58 sec	13 m 45 sec

Considering the different countries, the duration is slightly higher for Czechia, even if it is not so far from Poland. On the other hand, Germany and France are much lower than the total average. These differences in terms of length could be related to the answers the respondents gave to the open-ended questions and the level in which respondents engage with the refugees' crisis in each country.

Timestamps for choice experiment

Timestamps have been used to monitor different aspects:

- the entire questionnaire (from the beginning till the end)
- specific tasks (especially the choice experiment and perspective taking questions)
- any other questions

In detail, for the experiment (question Q7), the purpose has been to check for any potential difficulty in answering to this section. Please see below the average time to complete the choice experiment:

Total average	Poland	France	Germany	Hungary	Romania	Czech
2 m 14 sec	2 m 14 sec	1 m 55 sec	2 m 8 sec	2 m 18 sec	2 m 19 sec	2 m 31 sec

According to the duration per country, the experiment does not take more than 2 minutes and 30 seconds to be completed. The total average is slightly lower than the pilot tests (2 mins and 51 secs), the difference is much higher in France: little less than 2 minutes to complete this section.

Same as for the other reports, you can find as an annex the complete database with timestamps in a separate dataset.

Response rate

The response rate for the total sample and per each country:

Total average	Poland	France	Germany	Hungary	Romania	Czech
82%	82%	80%	79%	84%	82%	84%

Refusal rate

On the other hand, the refusal rate for the total sample and per each country:

Total average	Poland	France	Germany	Hungary	Romania	Czech
18%	18%	20%	21%	16%	18%	16%

The sum of the response and refusal rate equals to 1. The refusal rate includes all those panelists that have been invited to the survey but did not access it or accessed it but did not complete it.

The percentage related to the refusal rate for the respondents is really low. The reason is easily explainable: The survey participants are pre-recruited and existing panelists. In other words, the possibility to have pre-recruited panels allows to minimize the refusal rate.

Furthermore, the percentages both for the response and the refusal rate are in line with the pilot phase: this means that the initial sample considered to test the survey was reliable and well representative of the whole panelists.

Items non-response

The questions including items non-response were Q14, Q16, Q17 and Q20. Specifically, these questions include a “Don’t know” option. Confirming the data from the pilot phase, Q14, Q16, Q17 have a percentage of respondents answering “I don’t know” still acceptable.

In detail, this is the distribution per country and considering the total sample:

Items non-response: final results

Questions	Total	Poland	France	Germany	Hungary	Romania	Czech
Q14	6%	5%	8%	6%	5%	6%	4%
Q16	19%	21%	19%	13%	21%	20%	17%
Q17	17%	10%	22%	16%	22%	21%	17%
Q20	31%	26%	48%	28%	32%	23%	30%

Comparing this information with the pilot phase (table below), Q19 changed its logic including different price ranges and removing the “Don’t know” option. Another question that changed its answer options is Q20:

Q20. If a person coming from Ukraine is religious, what would be your best guess about the denomination?

Previously it included 4 different options for “Roman Catholic”, “Eastern Catholic”, “Eastern Orthodox” and “Protestant”. Following the pilot results, the above-mentioned options have been aggregated into “Christian”.

Items non-response pilot results

Questions	Total	Poland	France	Germany	Hungary	Romania	Czech
Q14	5%	5%	10%	4%	6%	3%	3%
Q16	17%	16%	19%	10%	29%	13%	13%
Q17	16%	8%	26%	14%	24%	10%	16%
Q19	56%	50%	77%	56%	57%	46%	49%
Q20	32%	12%	51%	36%	41%	17%	34%

Opened quotas history

Opening quotas is a standard practice when the fieldwork gets stuck and number of responses slow down. This is done step by step, constantly checking the percentage of completion, to stay as close as possible with the sample design.

Track of changes

The track of changes, consider our detailed experience in this specific project.

Friday, 2 September (3 working days for the end of the fieldwork), with the number of missing interviews lower than 200 for all countries, except for Hungary, Romania and Poland, we decided to open the first quotas for these 3 countries. In detail we slightly opened age groups 46+ (both male and female) and we set the new quota limit to 20 interviews higher for each variable.

Monday, 5 September, during our daily monitoring, the fieldwork showed that we are still missing a bit less than 100 interviews per country. However, due to the slow rate of new responses, we decided to relax the youngest age ranges and the education levels. The expression of "relaxing a quota", means to fix a limit of interviews higher than the original sample design. This amendment was done for all countries, except for Czechia. For the latter we opened the quota for education levels afterward. During the same day (Monday) in the late afternoon, the quota for regions was fully opened: the target was to conclude the fieldwork by the upcoming day.

Tuesday, 6 September, the only missing interviews were related to Romania: during the last day we decided to open the last quota too, employment status. At the end of the day, we conclude the fieldwork with 70 more interviews for "active" respondents in Romania.

Weighting system

Our weighting methodology follows a three-step procedure:

- **Step 1:** identify the variables in which there is a deviation between sample and population
- **Step 2:** calculate the weight to be applied where necessary

- **Step 3:** apply weights for every single variable if required

Different methods can be used for weighting. Among these, the best known are Cell weighting and Rim weighting. Cell weighting aligns sample and population proportions in every cell in the matrix constructed by fully cross classifying all the weighting variables, whereas Rim weighting seeks to align sample and population proportions only on the marginal distributions of this matrix.

Cell weighting works well with a reduced number of variables, on the other side Rim weighting runs better for big number of variables. For this survey we used the **Rim (random iterative method) weighting**.

An example of this process can be the hypothetical case where in a country the gender distribution is 45% male and 55% female. If the gender distribution of the completed responses is 50% male and 50% female it can be apply a weight, with the followed procedure:

Example: weights computation

Gender	Sample	% Sample	Final responses	Computation	Weight
Male	450	45%	500	450/500	0.9
Female	550	55%	500	550/500	1.1

Considering our specific case, to derive weights for our final dataset, we used the SPSS RAKE procedure. In detail, the statistical software includes a command, **“Rake weights” function**, that is the most common way to weight data in the market research. The SPSS command follows these steps:

- Setting as input all the original quotas per each country
- SPSS reads the quotas using a Rim and General Log-Linear Analysis to compute the weighting
- The procedure identifies the optimal weights, trying to satisfy the initial quotas as much as possible

Here below an example of SPSS syntax per Czechia.

Example: syntax for weights computation (Czechia)

SPSS Syntax

```
SPSSINC RAKE
dim1 = dim_gender
1 48.839152839626
2 51.160847160374
3 0.238095238095238
dim2 = dim_age
1 4.69312487368446
2 8.42138179825218
3 10.2701106718678
4 9.16832080574044
5 7.48567401971478
6 9.96138783074036
7 4.22985400556062
8 7.54916196715407
9 9.2081518252119
10 8.37259536915094
```



```
11 7.3909431668524
12 13.2492936660701
13 0.238095238095238
dim3 = dim_edu
1 8.73560857901965
2 68.5602803198521
3 22.7041111011282
dim4 = dim_empl
1 62.9648419065597
2 37.0351580934403
dim5 = dim_regio
1 12.5670053970579
2 12.749663886036
3 11.5784796303938
4 10.3798948737567
5 14.1699404640901
6 15.9430373645394
7 11.384833298646
8 11.22714508548
FINALWEIGHT=weight.
des weight.
```

The procedure is approximately the same per each country, the only changes are related to the quota proportions.

Annex: Final datasets

Please find the attachments including all datasets that will accompany our report.

In detail, the attachments include:

- Final weighted dataset (SPSS) – completed interviews
- Final weighted dataset with values (excel) – completed interviews
- Final weighted dataset with labels (excel) – completed interviews
- Final dataset (SPSS) – completed and incomplete interviews
- Final dataset with values (excel) – completed and incomplete interviews
- Final dataset with labels (excel) – completed and incomplete interviews
- Codebook & datamap
- Metadata & paradata
- Rim weighting with IBM SPSS