World Values Survey Wave 7 in Czechia: Sample Design

Survey organization: DEKK Institute oversees the whole process, ranging from designing methodology to translation of WVS questionnaire from English to Czech. The fieldwork itself was conducted by FOCUS polling agency, that is specialized in collecting social scientific data on a large scale in Czech Republic, since it has a large, trained network of pollsters distributed all around the country.

Fieldwork in Czech Republic was planned to start in the second half of December and take whole January and February 2022. Depending on the pandemic situation (e.g. in case of a lockdown or strong spread of Covid-19 among the population), data collection may be extended into the following month.

Target population: People aged 18 years and more, no upper age limit. Citizens, residents, and foreigners alike. Usual place of residence is the address where a person usually resides. It may also be his/her legal residence. This may be different from the place where he/she is at the time of the survey. For the purposes of the survey, a person's usual residence is the place where they normally sleep. We are not surveying people living in the institutionalized population such as students living in dormitories, prisoners, hospital patients, etc.

Target sample size is 1200 individuals, which corresponds with WVS guidelines for countries of Czech size (approx. 10.7 million inhabitants).

Whole country was divided into 120 PSUs, proportional to the population density. On average, 65% of PSUs are in cities, and 35% are in rural areas.

Sampling model: Random route sampling - a classic method of face-to-face interviewing in social sciences, ensuring the equal probability of selection. It is mostly used in places where no proper listing of residents is available for research purposes, as is the case of Czechia. The sample design is a multi-stage random (probability) one. The sampling points (PSUs) are drawn with probability proportional to population size, for a total coverage of the country. The sampling points are drawn after stratification by equivalents to regions (NUTS III level) and by degree of urbanization (settlement hierarchy) within these. This means that sample has to be proportionate by these two criteria.

Sampling plan:

We have 120 PSUs, so in average in one PSU (from one starting point) 10 interviews will be made.

Selection of households: The general idea is the right-hand rule and the every-third dwelling selection. **Case 1) A street is provided as starting point**. In this case interviewer goes and finds number 1 on the street; this will be her/his starting point. She/ he stands on the street at this starting point facing towards the end of the street. Interviewer will choose than the side of the street, which is to her/his right, and he/she walk towards the end of the street.

Case 2 A building provided. In this case interviewer first has to find out the address of the building. If this is not possible for some reason (it's not written on the door, or nobody is there), he/she needs to find the main entrance, which will be his/her starting point anyway.

Then he/she stands at the main entrance facing towards the end of the street. Interviewer will choose then the side of the street, which is to his/her right, and will walk towards the end of the street. Now, interviewer knows where he/she has to be, which side of the street he/she

has to be, and which direction he/she has to go. The task is to find out at which door the interviewer must ring. For this purpose, we need a number: and this number is three (3). Now interviewer goes ahead and walks in the direction he/she has to and finds the third door opening to the street on the right-hand side. What interviewer looks for are dwelling units. By dwelling unit, we mean living quarters for one household, whether it is a single house, half a duplex, a basement or attic apartment in a multiple family house, an apartment over a garage or store, or an apartment in a high-rise building.

Selection of starting points: The starting points have two significant characteristics: they belong to a settlement and are undoubtedly specified in geographic terms. 1) Selection of settlements: We have a list about each settlement in Czechia, indicating the region this given settlement belongs to. We distribute our PSUs among these locations. From towns and villages, we choose the appropriate number by random method. Our sampling program assigns random numbers to settlements, then rank orders it by this random number and it chooses the uppermost towns/cities and villages from the complete listing. Completing this procedure, we have a list of settlements with the number of PSUs assigned to them – now we have to translate these PSUs into starting points. 2) Selection of starting points: In this step we specify a point for each PSU, where the interviewer must start her/his route. There is no complete listing of all residential addresses in Czech Republic. So, we must use the maps in this step or in cases of smaller settlements – a public building or another structure may be chosen as a starting point. 1) If a map of the settlement is available, usually there is also a complete list of streets. We select the starting point in this case by assigning an ordinal number ranging 1 to "n" to the individual streets. We select the respective street using a chart of random numbers. We select the beginning of the street and specify that as starting point. 2) Unfortunately, there are many small settlements, which are not mapped. In these settlements we rotatedly assign the church, the train station, the city hall, the first house the interviewer saw when entered the village, the bus station and general store as starting point.

Respondent selection:

When somebody opens the door at the selected addresses, the interviewer will make a list of adults living in the household and the software will randomly select the member of the household with whom the interview is to be conducted. This is how we secure that each member has equal chance to be included in the survey. The person interviewed must be: - 18 years of age or over - a permanent resident of the household contacted (s/he only must live there on a permanent basis, we do not care if s/he administratively registered as living elsewhere) - the only household member interviewed - interviewed individually without disturbances or suggestions from anyone else.

Interviews are typically conducted in respondent's place of actual residence (not necessarily place of official residence).

Data collection method:

Face to face interview. Tablets with questionnaires are used by pollsters who ask questions and fill in the answers. Respondents are interviewed individually without disturbances or suggestions from anyone else. If the epidemiological situation requires it or the respondent requests it, the interview can take place outside his/her household (e.g. in a nearby café/restaurant, in the yard, in the garden, etc.).

Quality control:

Trained pollsters ensure that the questionnaire is filled by a respondent who understands the questions. They also code in concrete location into tracking sheets, so we can be sure the pollster followed the recommended route and stopped in a certain household an appropriate time to conduct an interview. Software used by the FOCUS agency screens the questionnaires – it does allow to code in unfilled questionnaires, but the unfulfilled ones are not counted towards the final number of polled individuals – it serves as an information for the agency and as a proof of work for the pollster. Typically, very few people that agree to be polled later refuse to finish the questionnaire (it occasionally happen when a situation arises that does not allow for the finish of an interview, for example a mother of a child is polled and something happens to the child). The software also notifies the agency if someone would want to fake the completion, like a questionnaire filled too quickly (using the time stamp control method) without allowing for an actual thought process to go between reading the question and answering or filling the same answer all the time. Post survey, DEKK staff checks the logic of all the surveys, and our AI-assisted analysis identifies the outliers and anomalies. If there are anomalies far above the expected Gauss distribution, we can investigate. We may either discover a new interesting social reality social scientific literature was not aware of, or we can identify a methodological mistake, or a fraud.

We plan to check 20% randomly chosen respondents to verify if the interviews actually took place. Based on the pandemic situation, we will either use combined method of personal visits and calls or rely on calls completely. On average 60% of respondents agree to be called for the sake of control – that is why the combined method is preferred whenever possible.

Polling agency does record the place where interview was conducted, though the data are stored separately from personal data, according to Czech and EU legislation. DEKK Institute is not given the personal data of the respondents, except those for phone numbers or addresses of those randomly selected for secondary control.