Respondent's Housing Conditions as a Factor Influencing the Probability of Contact with the Respondent in an F2F Interview (on the Example of ESS Paradata Analysis)

Yeliena Kovalska [*] [0000- 0003-1981-2255], Maria Dzhevaga [0000- 0001-5773-7300]

Taras Shevchenko National University of Kyiv, Ukraine
[*] *e_kovalskaya@knu.ua*

ABSTRACT

Paradata is used to improve the quality of analysis and data collection in F2F interviews. Its collection allows to increase the response rate by analyzing non-responses. Based on the binary logistic regression constructed for datasets including ESS 2010 (ESS-5) and 2018 (ESS-9) data in Germany, was found that a respondent's life circumstances, including the characteristics of their household, influence the consent of the ESS survey: the variables "house type" and "house condition" only partially predict the possibility of obtaining interview consent from the respondent. The chances of a successful interview were also assessed depending on the type of household and its physical condition.

Keywords: paradata, ESS, response rate, quality of sociological information, respondent.

1. INTRODUCTION

Predicting the success of contact with the respondent is currently one of the key tasks of empirical sociology. The ability to use a correct prediction allows researchers to avoid or reduce the impact of system errors and improves the quality of the study itself. One way to improve a response rate is to use paradata. Paradata is data about the process of collecting information received during the survey. This includes the respondent's behavior, characteristics of his housing conditions, reactions to questions, time of answer and more. [1]. They are used to correct response rate and the data collection process itself. This additional information helps to improve future surveys, predicts the likelihood of next contacts, points out errors to the researcher and allows to look at the survey through the eyes of the respondent, gathering external information about him, as well as interviewers' and respondents' comments on the survey process. Having collected all the factors that affect the level of consent to the interview, we can predict possible contact and thus predict the response rate and post factum sampling.

The role of letters with paradata, which are filled in by the interviewers, is significant [2]: researchers collect information on all typical units, including refusals (contact form includes information about the interviewer's visual observations, assessment [3] of the environment) for each of the attempts to contact the respondent. This makes it possible to describe potential participants who refuse to take part in the research and those who are willing to pass it, to assess the likelihood of participation for each category of respondents. The results of the ESS study tell about the importance of assessing the household conditions of the respondent.

European Social Survey (ESS) interviewers use an adaptive method of engagement, which means that the more information they have about the respondent before the first contact is made, (with the right approach and interviewer training) the greater the likelihood of the respondent's commitment and even how sincerely the respondent will answer. The adaptive method used by researchers integrates paradata into design decisions when planning the stage of fieldwork, which can reduce the costs and errors associated with survey data collection [4]. Although the interview plan and the interviewer's instructions include all the respondent's response options and subsequent interview scenarios, there are still no clear instructions for the interviewers depending on the type of house where the respondent lives, the area of residence.

Another issue is the problem of distinguishing the influence [5] of different reasons for refusal to participate

in the research, the importance and impact on the results of reasons for refusal explores DeMaio [6], but the question of connection between refusal for various reasons with other factors remains open.

1.1. Analysis of current research and publications

The study of the paradata is presented in the works of K. Beullens, K. Vandenplas, G. Loosveldt, R. Groves, J. Shinn, J. Durrant, P. Smith, F. Kreuter, C. Casas-Cordero, Y. Dmitriev, P. Tarassenko, Y. Ustinov. K. Beullens, K. Vandenplas and G. Loosveldt [7, 8] pointed out in their work the importance of taking into account the interviewer's influence on the response rate in the results of the ESS study. In the ESS study, interviewers are sent (to respondents) depending on their area of residence (to reduce travel costs), this factor and other characteristics of the interviewer (such as experience, personality traits, extraversion) has a decisive influence on the consent of the respondent. The results of research conducted by R. Groves [9] showed that in addition to the characteristics of the interviewer (the probability of agreement increases with increasing experience of the interviewer) such factors as the region of residence of the respondent, population density, socio-economic situation, ethnic group, religion and characteristics of housing have a significant influence, optimization of the process of recruiting respondents is especially relevant in the context of the study of weighing the opinions of respondents in accordance with their professional status, education experience [10]. J. Shinn [11] concludes that the characteristics of the area (number of apartments and population density) are also predictors of lack of response. P. Smith. J. Durrant with colleagues and F. Kreuter, C. Casas-Cordero provide information on the effectiveness of using paradata to predict and correct non-response [12, 13]. Alexander Raichenko and colleagues [14] talk about the peculiarities of conducting personal interviews.

The problem of research is the low level of analysis of the impact of the characteristics of the respondent's household and place of residence on the level of non-response. The aim of the article is to study the influence of the characteristics of the respondent's household and place of residence on the response rate in the ESS study. The task is determining the influence of the respondent's household conditions on the probability of their participation in the research.

1.2. Research methodology

The study was conducted based on the analysis of ESS 5 and 9 wave datasets in Germany. The methodological documents presented on the ESS website were also taken in the analysis. Logistic regression was used to build a model for assessing the respondent's participation in the research based on the characteristics

of their households. For analysis was used the SPSS Statistics program.

2. RESEARCH RESULTS

For consideration of internal and external processes of ESS research the methodical documents which are in free access of the ESS site were collected and processed [15]. The paradata in the contact letters contains information about the conditions in which the respondent lives, whether the area is safe, the type of accommodation and, if the respondent refuses to participate in the research, the reason for refusal. Such information is logged for each contact attempt. Methodical documents have their own features, depending on the country, but retain the main provisions of the research for the similarity of the survey procedure. Also, before conducting the survey, each country determines the rate for the time of the interview and the reach of the respondents, changing the instructions for the interviewer.

In order to analyze the relationship between the respondent's housing conditions and consent to participate in the survey, we analyzed the main block questions present in each wave and the alternating questions for different waves and countries (different for most participating countries). Due to the main questions, we can get information about the objective status of the respondent, the decile distribution of income, the type of employment (or unemployment of the respondent) and the level of education.

2.1. Analysis of methodical documents ESS

We analyzed paradata and instructions for the interviewers of ESS 2010 (ESS5) and 2018 (ESS9) in Germany. This country was chosen because it is a regular participant in the study. Also, the difference of 8 years between the waves makes it possible to analyze how the errors of previous years were corrected and analyze the evolution of the ESS methodology.

Several methodological documents noted the problem of non-response of respondents living in certain types of households, but as the fact of the existence of this problem was described in several sentences, it can be concluded that it is not given much attention to this problem. [15]. The topic of the influence of subjective social status and income on how likely the respondent will agree to participate in the study was almost not covered. Despite this, ESS is moving towards an individualized approach to each respondent.

Earlier [16] we considered the difference between the methodological aspects of the research in 2010 and 2018, and analyzed changes in the operationalization of housing conditions of respondents. In particular, the typology of the respondents' apartments appears, the list of possible variants of the interview results changes: additional alternatives appear, which reduces the risk of

the interviewer misunderstanding the further ways to continue the interaction with the respondent. The appearance in the data set of contact letters of 27 options of refusal of the respondent, indicates how accurately the interviewer should identify the reason for refusal during the survey, which allows to analyze the relationship between the reasons for refusal and living conditions of respondents.

In three waves (6, 7 and 8) the research methodology changed, the interviewer needs to describe the respondents' attitude to the survey, the characteristics of their homes and living conditions. We analyzed 9 wave, where these changes are most noticeable, to see if there is a relationship between the level of refusal, non-response and housing conditions of the respondent, characteristics of the house where the survey takes place, employment of the respondent, level of education, income level of the household, personal income and objective social status of the respondent [17, 18].

2.2. Highlighting the reasons for respondents' refusal to participate in the survey

In [16] we explored the influence of respondents' housing characteristics on the level of non-responses. The analysis of the data set showed that in 2018 (ESS9) there were the most refusals by the residences of apartment buildings, residents of individual buildings refused much less often, so there is reason to investigate these factors in more depth. The result confirms the findings [19] on the increased likelihood of successful interviews with respondents living in individual houses or duplexes. This is consistent with previous studies that have found a low level of cooperation of people living in apartments [4, 20] and may indicate the impact of wealth on the ability to provide certain housing.

In order to determine the influence of the respondent's living conditions on the probability of their participation in the survey, we analyzed the objective and subjective reasons for refusing the interview. In the ESS9 study, there are 26 options for refusal and one option for consent to an interview. Among the refusals are conscious categorical refusals due to unwillingness to participate in the study, refusals due to other circumstances with the possibility of further cooperation, options for not conducting an interview due to the relocation of the respondent, or respondent's illness or death. Existing reasons for refusal can be divided into objective, which combine reasons, independent of the respondent's opinion and subjective, which reflect the motivation for refusing to participate in the research. In the vast majority of studies that do not use paradata, all refusals are considered as one result - failure to interview the respondent. However, it should be noted that different reasons for refusal may be related to different characteristics of the respondent, and combining all options for refusal may primitivize and distort the results.

Therefore, for the analysis we identified a subjective refusal (due to reluctance, interrupted interviews due to the fault of the respondent, refusal of the household) and refusal for objective (other) reasons (the respondent fell ill, moved, unable to answer for various reasons). Invalid interviews due to the interviewer's fault were also deleted. Based on these characteristics, two binary variables "consent / conscious refusal", "consent / refusal for other reasons" were created.

2.3. Isolation of factors influencing the level of non-responses among the characteristics of the respondent and his household characteristics (based on binary logistic regression)

In order to explore the influence of the respondent's household characteristics on the response rate, we constructed binary logistic regressions for the dichotomous variable "Was the interview attempt successful" (for two newly created variables) based on ESS 9 wave data. In the first case, to determine the impact of household characteristics, the independent variables were: type of house, access to it (whether there is an intercom or other means of communication), the presence of garbage nearby, vandalism and physical condition of the house, its walls, roof.

Binary logistic regression for the dependent variable "agreement / subjective refusal" with independent variables, "physa"(physical condition of the house), "type" (type of household), "access" (access to the house), "littera"(presence of garbage near the apartment) and "vandaa" (presence of vandalism near the apartment). In this model, the influence of the variables "access to the house" and "presence of garbage" and "presence of vandalism" turned out to be insignificant variables - they improved the model by less than 0.1%, so they were excluded from further analysis of the model.

The quality of the first model for the variables "consent / subjective refusal" is low (Nagel-Kerke - 1%), namely the subjective refusals of the respondent cannot be explained by household characteristics. However, the quality of the model for the variables "consent / objective rejection" is higher (Nagel-Kerke - 8%) compared to the first, but also not sufficient. Thus, the living conditions of the respondent can to some extent be used to predict the success of the contact, but do not have a strong impact. This is in line with previous studies [20, 21], which indicate a link between the type of home and the likelihood of participating in the survey, in particular less cooperation of respondents living in apartment buildings may indicate the effect of social status [17, 18] as a latent variable. In addition, the low quality of the model can be explained by the nonlinearity of communication - respondents with low and high social status are less likely to cooperate [19].

2.4. Dependence of a valid questionnaire depending on the condition and type of household of the respondent (depending on the reasons for refusals)

In order to predict the success of the contact by the condition and type of housing for different reasons of refuse, we analyzed the conjugation tables. Based on the analysis of the distributions, it can be concluded that the worse the condition of the house, the lower the chance of agreement for the interview from respondent ($xi2$ is significant at $<.005$). We can see such results for both subjective and objective refusals to conduct interviews. Regarding the significance of the type of house, the results of the analysis showed that for a subjective reason for failure, this variable does not have a strong impact. But for objective conditions, there is the connection with the type of house (42.6% of refuses in apartment buildings, against 22.2% - in individual houses). In general, this indicates the influence of the condition and type of residence of the respondent on the probability of successful contact with the researcher.

3. CONCLUSIONS

Taking into account the results of recent ESS research and analyzing existing data from contact forms about the respondent's life circumstances, researchers have the opportunity to develop approaches to engage groups of respondents based on their characteristics, thus increasing the success of interviews, including respondents with different types of housing for which could be made a larger margin in quantity to reduce weight in further analysis.

From the results of the analysis it can be concluded that the living conditions of the respondent do not explain the level of non-responses, the dynamics of dependence on the type of house and its physical condition remains the same. Respondents living in apartments and dormitories are more likely to refuse interviews than respondents living in private homes. In addition to the factor of the type of housing where the potential respondent lives (which has the greatest impact), the factor of the physical condition of the apartment should also be taken into account, because the analysis revealed that the worse the condition of the apartment, the more

REFERENCES

[1] Sydorov, M.V.-S. (2011), "Using of paradata in social researches", *Sociology: theory, methods, marketing*, vol. 2011/4, October-December, pp. 198–208, available at: https://www.researchga te.net/publication/312974578_Vikoristanna_parada nih_u_sociologicnih_doslidzennah_Ispolzovanie_p aradannyh_v_sociologiceskih_issledovaniah_Usin g_of_paradata_in_social_researches

likely the respondent will refuse the interview. Putting this question on consideration for ESS member countries, which have problems involving respondents in the survey, it is possible to improve the work of interviewers and reduce the level of non-response.

RESEARCH PROSPECTS

If we take into account the latest research of the ESS, we can see that there are methods to determine the level of prestige of the area where the respondent lives, based on available data. Having created such a methodology, we will be able to see whether there is a connection between how prestigious the respondent's area of residence is and the chance that respondent will agree to participate in the study.

Also, if ESS representatives provide at least minimal information about the interviewers, such as gender and number of years of interview experience, we will be able to follow Groves and Couper's path [21] and see if is true for the ESS research that the woman interviewers are more likely to obtain respondent's consent, and to what extent the experience gained by the interviewer affects the success of the involvement of respondents in the study.

Knowing the information about which attempt the respondent agreed to participate in the study and knowing whether the interviewers who visited the same respondent changed (and how often), we can see how this affected the response rate.

The question of the influence of the characteristics of housing in Ukrainian studies on the level of responses remains open, in particular, there is the problem of studying the peculiarities of living in high-rise buildings [22] and comparing them with the European context.

Last but not least, the impact of timing in conducting an interview, depending on where the respondent lives. Since the contact letter contains information about the day of the week and the time of engagement of the respondent, there will be no problems with finding the necessary data (a study of the impact of time and day of the week on the final result was conducted in 2011 by Z. Jill [23]).

[2] Social survey ESS, Response Based Quality, available at: https://www.europeansocialsurvey.org/

[3] Akinsanmi, Akinbolati and Olufemi J., Agunbiade (2020), "Assessment of Error Bounds for Path Loss Prediction Models for TV White Space Usage in Ekiti State, Nigeria", *International Journal of Information Engineering and Electronic Business (IJIEEB),* vol. 12, no. 3, pp. 28-39. DOI: 10.5815/ijieeb.2020.03.04

[4] Couper, M.P. (1998), "A Measuring Survey Quality in a CASIC Environment", *Invited paper presented at the Joint Statistical Meetings of the American Statistical Association, Proceedings of the Survey Research Methods Section*, ASA, Achieving Quality in Surveys, Dallas, August, pp. 41-49, available at: http://www.amstat.org/sections/srms/Proceedings/y1998f.html

[5] Wei, Wang Likun, Cui and Zhuo, Li (2011), "Theoretical Design and Computational Fluid Dynamic Analysis of Projectile Intake", *International Journal of Intelligent Systems and Applications (IJISA)*, vol. 3, no. 5, pp.56-63. DOI: 10.5815/ijisa.2011.05.08

[6] DeMaio, T.J. (1980), "Refusals: who, where and why", *Public Opin, Quart*, vol. 44 (2), pp. 223–233, available at: https://academic.oup.com/poq/ article-abstract/44/2/223/1863870?redirectedFrom =fulltext

[7] Beullens, K. and Loosveldt, G. (2016), "Interviewer Effects in the European Social Survey Centre for Sociological Research", *Leuven*, Belgium, vol. 10, no. 2, pp. 103-118. DOI:10.18148/srm/2016.v10i2.6261

[8] Vandenplas, C. Loosveldt, G. and Beullens, K. (2017), "Fieldwork Monitoring for the European Social Survey: An illustration with Belgium and the Czech Republic in Round 7", *Journal of Official Statistics*, vol. 33 (3), pp. 659-686. DOI: 10.1515/JOS-2017-0031

[9] Groves, R.M. (2006), "Nonresponse Rates and Nonresponse Bias in Household Surveys", *The Public Opinion Quarterly*, vol. 70 (5), pp. 646–675. DOI: 10.1093/POQ/NFL033

[10] Dmitriev, Y. Tarassenko, P. and Ustinov, Y. (2017), "Determinacy Analysis of Weights as Mathematical Basis of the Future Sociology", *Proceedings of the IV International research conference "Information technologies in Science, Management, Social sphere and Medicine" (ITSMSSM 2017)*. DOI: 10.2991/itsmssm-17.2017.50

[11] Shinn, Glen Baker, Matt and Briers, Gary (2007), "Response Patterns: Effect of Day of Receipt of an E-Mailed Survey Instrument on Response Rate, Response Time, and Response Quality", *Journal of Extension*, vol. 45, no. 2, available at: https://www.researchgate.net/publication/2916519 06_Response_patterns_Effect_of_day_of_receipt_ of_an_E-mailed_survey_instrument_on_response_ rate_response_time_and_response_quality

[12] Durrant, G. Maslovskaya, O. and Smith, P. (2017), "Using Prior Wave Information and Paradata: Can They Help to Predict Response Outcomes and Call Sequence Length in a Longitudinal Study?", *Journal of Official Statistics*, vol. 33 (3), pp. 801-833. DOI: 10.1515/JOS-2017-0037

[13] Kreuter, F. and Casas-Cordero, C. (2010), "Paradata", German Council for Social and Economic Data (RatSWD), *Working Paper Series of the German Council for Social and Economic Data*, available at: https://www.researchgate.net /publication/46455102_Paradata

[14] Raichenko, A. Avetisyan, G. Gromova, O. Latfullin, G. Orlova, L. and Ibyatov, F. (2019), "Interviewer and public organizations", *Proceedings of the International Conference on Man-Power-Law-Governance: Interdisciplinary Approaches (MPLG-IA 2019)*. DOI: 10.2991/MPLG-IA-19.2019.79

[15] ESS methodological documents, available at: https://www.europeansocialsurvey.org/data/method ological_data.html

[16] Kovalska, E. and Dzhevaga, M. (2020), "Relationship Between the Level of Non-Responses in Personal Interviews and the Type of Housing of the Respondent (on the Example of Analysis of Paradata ESS)", *Sociological studios*, no. 2 (17), pp. 30-40. DOI: 10.29038/2306-3971-2020-02-30-40

[17] Xie Z., (2021) Selective Attention Ability's Relationship with Socioeconomic Status. Proceedings of the 2021 6th International Conference on Modern Management and Education Technology（MMET 2021）Access mode: https://doi.org/10.2991/assehr.k.211011.104

[18] Ye, N. Zhang, X. and Yu, H. (2018), "Social class, happiness and environmentally friendly behavior", *Proceedings of the 2018 3rd International Conference on Humanities Science, Management and Education Technology (HSMET 2018)*. DOI: 10.2991/HSMET-18.2018.143

[19] Bristle, J. Celidoni, M. Bianco, C.D. and Weber, G. (2019), "The contributions of paradata and features of respondents, interviewers and survey agencies to panel co-operation in the Survey of Health, Ageing and Retirement in Europe", *Journal of the Royal Statistical Society. Series A, (Statistics in Society)*, vol. 182 (1), pp. 3–35. DOI: 10.1111/rssa.12391

[20] Goyder, J. (1987), "The Silent Minority: Nonrespondents on Sample Surveys (1st ed.)", *Routledge*. DOI: 10.4324/9780429314520

[21] Groves, R. and Couper, M. (1998), *Nonresponse in Household Interview Surveys*, Wiley Interscience, New York, available at: https://www.semanticschol ar.org/paper/Nonresponse-In-Household-Interview -Surveys-Ziegel-Groves/03813889a193b4882837 ba7ead6c9b073e0545da

[22] Litvintsev, D.B. Osmuk, L.A. and Litvintseva, G.P. (2019), "Condos in Russia: Socio-Economic Aspect", *Proceedings of the "New Silk Road: Business Cooperation and Prospective of Economic Development" (NSRBCPED 2019)*. DOI: 10.2991/AEBMR.K.200324.039

[23] Jill, Z. (2011), "What Day of the Week Should You Send Your Survey?", *SurveyMonkey Blog*, available at: https://www.surveymonkey.com/curiosity/day-of-the-week/