www.eurocohort.eu

EuroCohort Working Paper No. 1

05/2019

Handling Non-Response in Longitudinal Surveys: A Review of Minimization and Adjustment Strategies and Implications for Eurocohort

Daniela Cocchi[1], Francesco Giovinazzi[1] and Peter Lynn[2]

[1]University of Bologna; [2]University of Essex

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

EuroCohort Working Papers

EuroCohort Working Papers are peer-reviewed outputs from European Cohort Development Project (ECDP) (http://www.eurocohort.eu/). The European Cohort Development Project (ECDP) is a Design Study which will create the specification and business case for a European Research Infrastructure that will provide, over the next 25 years, comparative longitudinal survey data on child and young adult well-being. The infrastructure developed by ECDP will subsequently coordinate the first Europe wide cohort survey, named EuroCohort.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777449. The content of the working papers does not reflect the the official opinion of the European Union. Responsibility for the information and views expressed in the working paper lies entirely with the authors.

The ECDP Consortium: Manchester Metropolitan University (coordinator), Ivo Pilar Institute of Social Science, Tallinn University, University of Bremen, Catalan Youth Agency, Panteion University of Social and Political Sciences, University of Debrecen, University of Essex, University of Saints Cyril and Methodius, Daugavpils University, University Institute of Lisbon, University of Jyväskylä, University of Bologna, University College London, Generations and Gender Programme and European Social Survey European Research Infrastructure Consortium.

Suggested reference to this paper:

Cocchi, D., Giovinazzi, F. and Lynn, P. (2019). Handling non-response in longitudinal surveys: A review of minimization and adjustment strategies and implications for EuroCohort. *EuroCohort Working Paper* No. 1. Daugavpils: Daugavpils University. Available at: http://www.eurocohort.eu/

Corresponding author: Daniela Cocchi - daniela.cocchi@unibo.it

ISBN 978-9984-14-875-5

ISSN 2661-5177

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

TABLE OF CONTENTS

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

ABSTRACT

This report aims to portray the state-of-the-art in minimization and adjustment strategies for both unit and item non-response in the framework of longitudinal surveys. First, we consider the aspects of survey design and implementation procedures that may reduce the occurrence of non-response in a survey. Secondly, we focus on modern weighting methods and imputation techniques that can be used to reduce the impact of non-response on estimation and data analysis. Drawing on these reviews we then make some recommendations for minimising the extent and impact of nonresponse and attrition on EuroCohort

1 INTRODUCTION

Survey studies are affected by nonresponse whenever a portion of the sampled population does not provide the requested information, or provides information that is not complete and usable. We distinguish two types of nonresponse (Bethlehem et al., 2011):

- unit nonresponse, when the respondent does not provide any information;
- item nonresponse, when the respondent only provides partial information.

A direct consequence of unit nonresponse is a realized sample size that is smaller than the planned one (assuming that no substitution strategy is used). Such a smaller sample size induces an increase in the variance of the estimates, and thus a lower precision (Särndal and Lundström, 2005). Moreover, if data is not missing at random (MAR, Graham, 2012), both types of nonresponse can lead to a severe bias of the estimates. This situation is typical in the case of selective nonresponse, when some subgroups in the sampled population behave differently with respect to the object of investigation, resulting in their over- or under-representation in the final sample. Understanding the nature and the causes of nonresponse is crucial in order to reduce it or to correct its effects.

These considerations apply to any kind of survey, including longitudinal panel surveys, where we repeatedly collect data from the same respondents over time (Lynn, 2009). In this case, we add a new dimension to the response process, the time dimension, increasing the complexity of possible nonresponse patterns. A unit may respond in a time occasion and fail to respond in another, members may leave the sample over time (dropouts), causing an erosion of the sample itself. This inevitable process of “erosion” is called attrition. Panel attrition (Lynn, 2018), if not handled, can lead to both low precision and high bias, for this reason a great amount of literature has flourished around this topic in the past two decades.

This project has received funding from the European Union’s Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

In this report, we aim to present some of the most recent developments in the field. We can broadly separate two different (but complementary) streams of research that deal with the implications of nonresponse in longitudinal survey studies:

- the minimization strategies, concerning the prevention or avoidance of nonresponse before it occurs, for example through a targeted survey design, with dedicated implementation techniques and a focused fieldwork;
- the adjustment strategies, concerning data correction, for example through weighting and missing data imputation, in order to properly account for the nonresponse that has already occurred.

Singer (2006) points out that statisticians have been concerned mainly with adjustment, while social scientists and survey experts have tended to focus on minimization.

2 STRATEGIES FOR MINIMIZING UNIT NONRESPONSE

If we wish to minimize the occurrence of a phenomenon like attrition, or nonresponse in general, it is important to understand the reasons behind it (Lynn et al., 2005; Massey and Tourangeau, 2013). In this section, we refer mainly to unit nonresponse. In the case of longitudinal surveys, nonresponse and sample attrition can depend on both:

- endogenous characteristics of the survey design, survey protocols and modes of interview (web surveys, CATI, CAPI, …), the type of interaction between interviewers and respondents;
- exogenous characteristics of the context in which the survey is implemented, such as demographic variables and cultural differences among groups in the population.

Under a different perspective, the practical causes of nonresponse can be sorted into three main categories (Lepkowski and Couper, 2002): non-location, non-contact, and refusal to co-operate. In details:

- non-location refers to a failure to locate a sample unit at a particular wave of a panel survey;
- non-contact refers to a failure to reach a located sample unit;
- refusal to co-operate refers to the case when a sample unit, that has been successfully located and contacted, does not agree to participate to the survey.

Problems with non-location, non-contact, and refusal can be particularly challenging in cross-national longitudinal studies due to cultural, organizational, technical and financial barriers (Stoop et al., 2010).

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

2.1 Accounting for national variation

The existence of national variation in both the extent and the correlates of nonresponse requires various considerations regarding the design of cross-national surveys. Behr et al. (2005) analyse panel attrition in the European Community Household Panel (ECHP), revealing strong differences in terms of respondent participation patterns across countries. Panel attrition has high variability across countries as well as for different waves within one country. Between-country differences in response rates were found to be strongest in particular subpopulations, for example households that moved in the sample period or the cases in which the interviewer changed. This cross-national variability can be the result of demographic and cultural differences between nationalities of the respondents (considered as subgroups of the sampled population). The impact of the characteristics of respondents and interviewers across countries has been investigated also in Billiet et al. (2007) for the European Social Survey (ESS), which found that the relationship between the type of respondent (cooperative or reluctant) and some attitudinal and background variables was not characterized by the same direction in all countries. On the other hand, analysing the ESS, Blom et al. (2011) show systematic differences between countries are affecting not only respondents, but also interviewers. Blom et al. (2010) gives a complete overview on cross-national differences in survey participation processes. Interviewer training, contacting and cooperation strategies as well as survey climates differ across countries, thus influencing differential nonresponse processes and possibly nonresponse biases. A recent article from Bianchi and Biffignandi (2018) shows the importance of community attachment indicator on the likelihood of contacting with members of the panel. Indicators of social participation can be significant in explaining cooperation differences across countries, and are found to be more significant than personality factors and well-being related variables.

2.2 Subgrouping

The phenomenon of cross-national differences in nonresponse can be explained also by interpreting countries or geographic areas as subgroups or strata of the population to sample. Subgroups based on this or any other criteria can be associated with response rates. A body of research is devoted to the idea of targeting respondents on the basis of subgroups, whether they are defined by observed characteristics or estimated propensities. As an example of probabilistic criterion, Lugtig (2014) shows how to study attrition in a latent class framework. This method allows to distinguish different latent groups of respondents, each following a different and distinct process of attrition. The study shows that attriters have different types of personality, and that they value survey participation differently from loyal stayers: attriters have less commitment and higher levels of panel fatigue. Similarly, Earp et al. (2018) discuss a method for analysing nonresponse in a longitudinal survey using regression trees. Regression tree models are used to identify subgroups of units that present vulnerabilities during the data collection process. This information can be used to direct additional resources to these groups in order to improve the response rate. In the context of Understanding Society, Bianchi and Biffignandi (2017b) address the problem of under- and over-represented groups of individuals. Once identified, these groups should be targeted, in order to correct the representativeness of the panel, adapting the survey design. Fumagalli et al. (2013) suggest that a tailored approach for tracking

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

strategies and respondent reports (addressing for example the needs of young people and busy people) increases the rate of cooperation.

2.3 Improving co-operation

In order to minimize attrition in a longitudinal panel survey study it is necessary to implement strategies aimed at improving cooperation and securing respondents loyalty. In this regard, Aughinbaugh et al. (2017) showed how attrition in the U.S. "National Longitudinal Survey of Youth 1979" remained remarkably low, thanks to the policy of proposing the interview to all sampled individuals in each round, regardless of their participation in previous rounds. McGinley et al. (2015) investigates the use a social network to reduce attrition in longitudinal studies; the results of their experiment suggest that Facebook might be an effective mean of minimizing attrition in longitudinal studies, reducing risks on non-location and non-contact.

2.3.1 Improving cooperation by incentives

Cooperation may be improved by means of material incentives. The effect of incentives over nonresponse is based on the social mechanism of altruistic reciprocity: an unconditionally given gift endows trust in the social exchange between researchers and respondents. According to current findings in survey research, unconditional gifts, in particular prepaid monetary incentives, are the most efficient and effective strategy for reducing unit nonresponse and strengthening the cooperation of respondents in social scientific surveys. Olsen (2005) suggests the use of targeted incentive payments as a cost effective way of holding attrition in check. Ryu et al. (2005) give a very good literature review on the impact of incentives; they compared cash incentive vs. an in-kind incentive in both a face-to-face and mail survey, showing that cash incentive yielded higher response rates, while the in-kind incentive was less effective. Laurie and Lynn (2009) give a review on the use of respondent incentives on longitudinal surveys, producing experimental evidence that attrition rates would be higher in the absence of incentives. Evidence of the beneficial effect of incentives in a mixed-mode longitudinal survey is provided by Jäckle and Lynn (2008). Becker and Mehlkop (2011) carried out an experiment to investigate whether mail survey response rates can be influenced by monetary incentives. They showed that prepaid monetary incentives elicit higher response rates than incentives promised once the questionnaire is returned, strengthening the respondents' trust towards the researcher. Lynn (2001) suggests that interviewer attitudes can be a barrier to the effectiveness of respondent incentives in interviewer-administered surveys.

McGonagle et al. (2013) examine the basic utility of the between-wave mailing and investigates how incentives affect cooperation to the update requests. Boys et al. (2003) describe the methods employed to maximize retention in a longitudinal study of alcohol use in 15–17 year olds with two follow-up occasions. In this age group, rather than prepaid money incentives, the use of music vouchers and prize-draws played an important role in encouraging participation of young people in the study. On the other hand, the price of the vouchers comprised a significant proportion of the study

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

costs. In addition, Becker and Glauser (2018) evaluate short- and long-term effects of prepaid monetary incentives and reminders on young people's cooperation and response rate. The authors argue that a monetary incentive has a direct and positive effect on the response rate in the shorter field phase, while it is an insufficient instrument against panel attrition in the long-term if it is the only strategy applied.

Sánchez-Fernández et al. (2012) show that the use of post-incentives and the use of frequent reminders were not found to have a significant effect on retention rate in web-based survey. The authors argue that personalization and targeting are the most effective strategies in this kind of surveys to control attrition.

2.4 Survey modes and targeting strategies

The modes in which a survey is administered may directly affect the response rates. In the last decade, web-surveys are becoming more and more popular: there are considerable cost and timeliness advantages associated with web interviewing, compared to face-to-face administration and to CAPI and CATI. However, web surveys do not perform well in terms of coverage and participation. Jäckle et al. (2015) demonstrated that transitioning from a face-to-face longitudinal survey to a mixed mode web and face-to-face survey is not straightforward. It can produce considerable cost savings, increasing in attrition and item nonresponse. In addition, Lynn (2013) showed that a sequential mixed-mode data collection, from telephone to face-to-face interviewing, resulted in a lower response rate than a single-mode face-to-face protocol. However, Bianchi et al. (2017) showed that the negative effect of a protocol transition on panel attrition was limited to the first wave after the transition, and that it disappeared or even reversed after a few waves. Furthermore, Cernat and Lynn (2018) showed that longitudinal surveys provide additional opportunities to boost web participation by collecting and using email addresses for participants. Prenotification letters can effectively boost participation in both self-completion (Taylor and Lynn, 1998) and interviewer-administered modes (Lynn et al., 1998), but communicating additional information to sample members does not always have a beneficial effort on response (Lynn, 1991).

In general, face-to-face protocols base their strength on the relationship that is created between respondent and interviewer. The skills, attitudes and training of interviewers can be an important determinant of co-operation rates in interviewer-administered surveys (Jäckle et al., 2013), as can the extent of effort made obtain participation (Lynn and Clarke, 2002; Hall et al., 2013. Refusal conversion can be particularly effective in a longitudinal survey context (Burton et al., 2006). Hill and Willis (2001) illustrate how assigning the same interviewer wave after wave has a strong positive effect on response rates. The stability of the relationship between respondent and interviewer in face-to-face or telephone surveys is particularly important, independently from the length of the interview. Lynn (2014a) has found no effect of the length of a first wave survey instrument on subsequent participation propensity, however another experimental study described in Lynn et al. (2014) provided evidence of heterogeneous effects of interviewer continuity on cooperation by panel survey members, depending on the age and other characteristics of the interviewer. This supports the notion

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

that interviewer continuity may be beneficial in some situations, but not necessarily in others. Whether interviewer continuity is beneficial may depend on the characteristics of the previous interviewer, the available alternative interviewers, and the respondent.

Moreover, the leverage-salience theory suggests that the "leverages" of single survey design features, on the cooperation decision of a respondent, depended on the saliency of the attribute. Following this approach, Groves et al. (2000) suggested a "tailored" approach, largely dependent on the communication skills of the interviewer (on whether the attribute was made salient to during the interview). The theme of tailoring and targeting is always present. Given that survey researchers tend often to treat all members of a survey sample the same way, Lynn (2014b) argues that treating sample subgroups differently, depending on what we know about them, improves participation in the subgroups. This is what can be referred to as a targeted response inducement strategy, aimed at improving cooperation. This concept is enforced in Lynn (2016) showing that, as an example, a targeted initial letter can increase response rates, but these effects are uneven across sample subgroups. Several other examples of successful targeting have been reported in the literature (Lynn, 2017).

2.4.1 Special populations and adaptive survey design

The above mentioned causes of nonresponse may be even more problematic when referring to special populations. Flores et al. (2017) address attrition minimization in minority, low-income children populations. The authors illustrate a successful retention strategic framework, based on optimizing cultural/linguistic competency of interviewers, staff training on participant relationships and trust, comprehensive participant contact information, electronic tracking database, reminders for upcoming outcomes assessment appointments, frequent and sustained contact attempts for non-respondents, financial incentives, individualized rapid-cycle quality-improvement approaches to non-respondents, reinforcing study importance, and home assessment visits.

The idea of using a tailored approach and a targeted survey design fostered the literature on Adaptive Survey Design (ASD). A very good review on this topic can be found in Tourangeau et al. (2017). ASD implies prioritizing worst performing subgroups and attempting to equalize or balance response propensities (Schouten et al., 2016). Within ASD, a particular issue is whether longitudinal surveys should target hard-to-get sample members with disproportionate resources. On this topic, Watson and Wooden (2018) show that hard-to-get cases are distinctly different from easy-to-get cases, and that simply removing hard-to-get cases may result in severe biases.

2.5 Refreshment samples

Minimization and adjustment strategies are not to be considered alternative approaches, they often are combined to produce better results, in particular in longitudinal panel surveys; some adjustment strategies require auxiliary information to be collected and planned since the design phase. This is the

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

case of refreshment samples, where panel data sets are augmented by replacing dropouts with new respondents sampled from the original population. On this topic the works of Hirano et al. (2001), Tunali and Ekinci (2007), Deng et al. (2013), Si et al. (2015) and Bianchi and Biffignandi (2017a) give a good methodological and applied coverage. However, while refreshment samples offer the potential to reduce cross-sectional bias, they are rarely able to provide retrospectively the data missing from earlier waves of data collection and consequently cannot compensate for bias in (most) longitudinal analyses.

3 ADJUSTMENT STRATEGIES

In the previous section, we have seen how minimization strategies try to prevent attrition, acting directly on response rates during data collection. Once data collection is over, if we are still in the presence of selective nonresponse – unit nonresponse and/or item nonresponse – we can try to reduce its negative effects on estimation by means of adjustment strategies. These strategies have the potential to correct for possible bias, and can be divided into two broad families of procedures: weighting techniques and imputation methods.

Weighting and imputation share the common technical objective to reduce the potential bias of the estimates, using auxiliary information to rebalance the responses or to complete the missing cases.

3.1 Weighting

Weighting adjustment is typically applied to adjust for the effects of unit nonresponse. Kalton and Flores-Cervantes (2003) and Hofler et al. (2005) give a general overview on many classic weighting techniques, such as poststratification, raking or multiplicative weighting, linear weighting, GREG weighting and logistic regression weighting, with a review of the major ideas and principles regarding the computation of adjustment weights and the analysis of weighted data. A critical review can be found in Brick (2013) with comments on the current state-of-the-art and possible future developments.

Adjustment weights are computed on the basis of auxiliary variables, whose distribution in the population has to be known and that have to be measured again in the survey. Weighting adjustment is effective in bias reduction only if there is a strong population relationship between these auxiliary variables and response behaviour. In the special case of propensity weighting, instead of adjusting for whole set of observed auxiliary variables, the univariate response propensity score is used. Response propensity is the conditional probability that a person responds to the survey, given her/his specific available background characteristics. In order to compute response propensities, auxiliary information for all sample elements is needed. Also in the case of longitudinal surveys, the problem of attrition can be dealt with weighting adjustment. One way to do this is by inverse probability weighting; for

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

this purpose, we compute the probability of remaining in the sample instead of the probability of response. This can be done again using auxiliary information, same as in a cross-sectional survey, exploiting as well the observed target variables from previous waves of the survey.

Propensity score weighting is widely used in research and survey practice; for example, Matsuo et al. (2010) show its good performance referring to the European Social Survey (ESS), with propensity scores computed on the basis of key sociodemographic and attitudinal variables. In the longitudinal framework, Mostafa and Wiggins (2015) construct inverse probability weights to adjust for sample loss on data coming from "1970 British Cohort Study" (BCS70). Corry et al. (2017) illustrate the good results of a comprehensive survey weighting strategy in the Millennium Cohort Family Study. Vandecasteele and Debels (2006) examine the effectiveness of weighting in the European Community Household Panel (ECHP). Cai and Wang (2018) compare the performance of different weighting strategies on data collected from the Hong Kong Panel Study of Social Dynamics Survey (HKPSSD) and the Beijing College Students Panel Survey (BCSPS). Bishop et al. (2018) use a quasi-experimental weighting methodology based on propensity scores on National Early Intervention Longitudinal Study. The usual nonresponse analysis and nonresponse adjustment methods do not distinguish among the different causes of nonresponse; Iannacchione (2003) proposes the sequential weight adjustment method for a two-stage response process consisting of contact and participation conditional on contact. The method sequentially fits a number of logistic regression models, obtaining estimates for the contact probability and for the participation probability, then combined in a single final weight.

3.2 Imputation

Imputation methods are a tool to deal with item nonresponse. These methods consist of replacing each missing value by a value that is estimated on the basis of auxiliary information. Various estimation methods are possible (Durrant, 2009; Nakai and Ke, 2011). Little and Rubin (2014) give a complete treatment of the topic and the many variants of these methods. Single imputation is performed when a missing value is replaced by a single synthetic value; the replacement can be performed with either deterministic of probabilistic techniques. Multiple imputation is performed when a missing value is replaced by a set of synthetic values. This leads to a corresponding set of complete data sets, from which we obtain estimates of population parameters, that are then combined to produce the final estimate and confidence intervals that incorporate missing-data uncertainty. A milestone methodological book entirely dedicated to the extensive treatment of multiple imputation is Rubin (2004).

Longitudinal imputation differs from its cross-sectional counterpart since it is possible, and typically desirable, to use the observations from previous waves as auxiliary information. Fitzmaurice et al. (2012) discuss a number of methods that can be used for longitudinal imputation. When there is dropout in a longitudinal survey, it is profitable to use the longitudinal character of the survey instead of cross-sectional information: previous observations on the same subject are highly correlated with the target variables in the wave in which the dropout occurs, making good predictors for the missing value. The quality of the imputation using information from previous waves is improved with respect

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

to the case of using only cross-sectional information. In this context, Little and Su (1989) proposed a longitudinal imputation method. It is a nearest neighbour technique, where imputation is based on the observations that are the closest to the missing one in some defined sense. The definition of nearest neighbours takes into account both cross-sectional and longitudinal information. The method takes into account variation over time (previous waves) in the target variable for the unit under study as well as variation between units for the target variable. Beaumont (2005) explores the possibility of using paradata[1] for nonresponse bias adjustment, showing that paradata can be used for nonresponse adjustment without introducing additional bias or an additional variance component in the estimates of population totals. In general, multiple imputation has good estimation properties, but it may introduce issues for secondary analysts if they are not aware of the model underlying the imputation strategy.

Many studies compare in the longitudinal framework the use of imputation methods with weighting strategies, particularly concerning epidemiological applications. Sterne et al. (2009) emphasize the potential for multiple imputation to improve the validity of medical research results and to reduce the waste of resources caused by missing data. Biering et al. (2015) illustrate the use of multiple imputation in longitudinal studies with repeated measures of patient-reported outcomes. Seaman and White (2013) compare multiple imputation and inverse probability weighting on the British Birth Cohort data. Examples of novel applications of multiple imputation can be found in Doidge et al. (2017) and Nooraee et al. (2018). Multiple imputation can be performed through statistical matching, which is a very fruitful technique to treat non response in the longitudinal case, as explored by *Ucar* et al. (2016), Donatello et al. (2016) and D'Orazio et al. (2017).

4 CHALLENGES FOR THE EUROCOHORT

All of the above cited strategies and methods are highly valuable in handling nonresponse, each one having strengths and weaknesses. In order to address this issue, Eurocohort must take into account the peculiar nature of the population under study. Children and young adults are a special population characterized by unique needs and problems.

[1] Data about the survey data collection process. In the case of nonresponse adjustment, indicators of the difficulty of making contact (number of interviewer visits or phone calls needed) and the reluctance of sample members to participate (whether an initial refusal needed to be 'converted'; interviewer assessment of co-operativeness) can be particularly helpful.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

We stress some of these peculiar factors:

- children live and grow in very different environments across Europe and within countries; they are exposed more than adults to the influence of the geographical and socio-cultural group structure;
- the family composition and the role of its members vary consistently among European countries, influencing the mentality of children;
- children and their families are exposed to economic fragility, particularly in poorer countries;
- young families with children can be mobile, difficult to track and reach;
- the education and the job placement systems are also different.

All these factors increase the risk of attrition and each of them can be contrasted with an appropriate minimization or adjustment strategic proposal.

As regards minimization of nonresponse and attrition:

- the extent and nature of support for the survey from institutions responsible for the sampling frame, associated information, and access to sample members (such as government ministries, registries and schools – depending on national sampling frames and information systems) is likely to influence survey response rates. Gaining the full co-operation of such institutions will be an important step towards the success of EuroCohort;
- the group structure, latent or deterministic, in the population of children and young adults in Europe suggests the use of a targeted/tailored approach to survey design. We note that the extent to which that can be done is likely to vary between countries at the first wave, depending on the sampling frame, but that approaches to targeting can be more standardised at all subsequent waves, drawing upon earlier survey responses to define the targeting groups;
- Some resources should be devoted to the development of targeted strategies and the identification of appropriate groups to target;
- the economic fragility and the needs of families with young children may make them reactive to monetary or material incentives;
- taking into account a possible digital divide between countries and families, we suggest a mixed-mode for survey administration, with face-to-face, CAPI and web-surveys (adapting the mode to each wave);

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

- in addition to the use of static ASD at each wave to motivate sample members to participate, dynamic ASD is a promising possibility for addressing unexpected field results, but relies on responsive and flexible field management and may require additional resources;
- the high mobility of young families points to the need for special measures to maximise the chances of locating sample members at subsequent waves, we suggest collecting a range of contact information at the end of the first interview, including phone numbers and email addresses for both parents and "stable" contact details, to the extent that parents are willing to provide these;
- the contact information should be updated at subsequent waves and managed within a sample management database designed to support the survey throughout is lifetime;
- between-wave contacts may be a cost-effective way of maintaining respondent interest and keeping in contact with mobile families, particularly if email can be used for a majority of such contacts;
- at each wave, an appropriate programme of reminders (if self-completion) or follow-ups (if interviewer-led) should be planned and implemented, following best practice regarding timing and quantity.

As regards adjustment:

- an obstacle is the availability of auxiliary information at a population level. Auxiliary information is different according to countries due to cultural and legal issues. This affects the adjustment methods than can be used to deal with non-response at wave 1. A choice will present itself between (limited) standardised adjustment across countries and the use of nationally optimal procedures that may vary greatly between countries;
- from wave 2 onwards, survey data from previous wave(s) can form the auxiliary data for adjustment for attrition;
- given that users of the EuroCohort data are likely to be a large and heterogenous group, provision of general purpose adjustment weights will be important;
- single imputation may be appropriate if important measures suffer relatively high item nonresponse rates;
- neither multiple imputation nor replicate weights are recommended for the majority of users, due to the complexity of analysis using these methods. Nevertheless, some specialist users may prefer to use these methods.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

5 REFERENCES

Aughinbaugh, A., Pierret, Ch.R. and Rothstein, D.S. (2017). *Attrition and its Implications in the National Longitudinal Survey of Youth 1979. JSM Proceedings.* Alexandria, VA: American Statistical Association.

Beaumont, J.-F. (2005). On the use of data collection process information for the treatment of unit nonresponse through weight Adjustment. *Survey Methodology*, 31(2), pp. 227-231.

Becker, R. and Glauser. (2018). Are prepaid monetary incentives sufficient for reducing panel attrition and optimizing the response rate? An Experiment in the Context of a Multi-wave Panel with a Sequential Mixed-mode Design. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 139(1), pp. 74-95.

Becker, R. and Mehlkop, G. (2011). Effects of prepaid monetary incentives on mail survey response rates and on self-reporting about delinquency-empirical findings. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, 111(1), pp. 5-25.

Behr, A., Bellgardt, E. and Rendtel, U. (2005). Extent and determinants of panel attrition in the European community household panel. *European Sociological Review*, 21(5), pp. 489-512.

Bethlehem, J., Cobben, F. and Schouten, B. (2011). *Handbook of Nonresponse in Household Surveys,* vol. 568. John Wiley & Sons.

Bianchi, A. and Biffignandi, S. (2017a). Integrating a refreshment sample in a longitudinal survey: effects of different weighting methods over four waves. In: *New Techniques and Technologies for Statistics 2017*. NTTS. DOI: 10.2901/EUROSTAT.C2017.001

Bianchi, A., and Biffignandi, S. (2017b). Representativeness in panel surveys. *Mathematical Population Studies*, 24(2), pp. 126-143.

Bianchi, A., Biffignandi, S. and Lynn, P. (2017). Web-face-to-face mixed-mode design in a longitudinal survey: Effects on participation rates, sample composition, and costs. *Journal of Official Statistics*, 33(2), pp. 385-408.

Bianchi, A. and Biffignandi, S. (2018). Social indicators to explain response in longitudinal studies. *Social Indicators Research*, pp. 1-27.

Biering, K., Hjollund, N.H. and Frydenberg, M. (2015). Using multiple imputation to deal with missing data and attrition in longitudinal studies with repeated measures of patient-reported outcomes. *Clinical Epidemiology*, 7, pp. 91-106.

Billiet, J., Philippens, M., Fitzgerald, R. and Stoop, I. (2007). Estimation of nonresponse bias in the European Social Survey: Using information from reluctant respondents. *Journal of Official Statistics*, 23(2), p. 135-162.

Bishop, C.D., Leite, W. L. and Snyder, P.A. (2018). Using propensity score weighting to reduce selection bias in large-scale data sets. *Journal of Early Intervention*, 40(4), pp. 347-362.

Blom, A.G., Jäckle, A. and Lynn, P. (2010). The Use of Contact Data in Understanding Cross-National Differences in Unit Nonresponse. *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*. Wiley Online Library, pp. 333-354.

Blom, A.G., Leeuw, E. de, and Hox, J. (2011). Interviewer effects on nonresponse in the European Social Survey. *Journal of Official Statistics,* 27(2), pp. 359-377.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

Boys, A., Marsden, J., Stillwell, G., Hatchings, K., Griffiths, P. and Farrell, M. (2003). Minimizing respondent attrition in longitudinal research: Practical implications from a cohort study of adolescent drinking. *Journal of Adolescence*, 26(3), pp. 363-373.

Brick, J.M. (2013). Unit Nonresponse and Weighting Adjustments: A Critical Review. *Journal of Official Statistics*, 29(3), pp. 329-353.

Burton, J., Laurie, H. and Lynn, P. (2006). The long-term effectiveness of refusal conversion procedures on longitudinal surveys. *Journal of the Royal Statistical Society Series A (Statistics in Society)*, 169(3), pp. 459-478.

Cai, T. and Wang, H. (2018). Comparing weighting methods for nonresponse in the HKPSSD survey and the BCSPS. *Chinese Sociological Review*, 50(1), pp. 1-26.

Cernat, A. and Lynn, P. (2018). The role of email communications in determining response rates and mode of participation in a mixed-mode design. *Field Methods*, 30(1), pp. 70-87.

Corry, N.H., Williams, C.S., Battaglia, M., McMaster, H.S. and Stander, V.A. (2017). Assessing and adjusting for non-response in the millennium cohort family study. *BMC Medical Research Methodology*, 17(1), p. 16.

Deng, Y., Hillygus, D.S., Reiter, J.P., Si, Y. and Zheng, S. (2013). Handling attrition in longitudinal studies: The case for refreshment samples. *Statistical Science*, 28(2), pp. 238-256.

Doidge, J.C., Edwards, B., Higgins, D.J. and Segal, L. (2017). Adverse childhood experiences, non-response and loss to follow-up: Findings from a prospective birth cohort and recommendations for addressing missing data. *Longitudinal and Life Course Studies*, 8(4), pp. 382-400.

Durrant, G.B. (2009). Imputation methods for handling item nonresponse in practice: methodological issues and recent debates. *International Journal of Social Research Methodology*, 12(4), pp. 293-304.

Earp, M., Toth, D., Phipps, P. and Oslund, Ch. (2018). Assessing nonresponse in a longitudinal establishment survey using regression trees. *Journal of Official Statistics*, 34(2), pp. 463-481.

Fitzmaurice, G.M., Laird, N.M. and Ware, J.H. (2012). *Applied Longitudinal Analysis*, vol. 998. John Wiley & Sons.

Flores, G., Portillo, A., Lin, H., Walker, C., Fierro, M., Henry, M. and Massey, K. (2017). A successful approach to minimizing attrition in racial/ethnic minority, low-income populations. *Contemporary Clinical Trials Communications*, 5, pp. 168-174.

Fumagalli, L., Laurie, H. and Lynn, P. (2013). Experiments with methods to reduce attrition in longitudinal surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(2), pp. 499-519.

Graham, J.W. (2012). *Missing Data: Analysis and Design*. Springer Science & Business Media.

Groves, R.M., Singer, E. and Corning, A. (2000). Leverage-saliency theory of survey participation: description and an illustration. *Public Opinion Quarterly*, 64(3), pp. 299-308.

Hall, J., Brown, V., Nicolaas, G. and Lynn, P. (2013). Extended field efforts to reduce the risk of non-response bias: Have the effects changed over time? Can weighting achieve the same effects? *Bulletin of Sociological Methodology*, 117, pp. 5-25.

Hill, D.H. and Willis, R.J. (2001). Reducing panel attrition: A search for effective policy instruments. *Journal of Human Resources,* pp. 416-438.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

Hirano, K., Imbens, G.W, Ridder, G. and Rubin, D.B. (2001). Combining panel data sets with attrition and refreshment samples. *Econometrica*, 69(6), pp. 1645-1659.

Höfler, M., Pfister, H., Lieb, R. and Wittchen, H.-U. (2005). The use of weights to account for non-response and drop-out. *Social Psychiatry and Psychiatric Epidemiology*, 40(4), pp. 291-299.

Iannacchione, V.G. (2003). Sequential weight adjustments for location and cooperation propensity for the 1995 national survey of family growth. *Journal of Official Statistics*, 19(1), pp. 31-43.

Jäckle, A. and Lynn, P. (2008). Respondent incentives in a multi-mode panel survey: cumulative effects on non-response and bias. *Survey Methodology*, 34(1), pp. 105-117.

Jäckle, A., Lynn, P., Sinibaldi, J., and Tipping, S. (2013). The effect of interviewer experience, attitudes, personality and skills on respondent co-operation with face-to-face surveys. *Survey Research Methods*, 7(1), pp. 1-15.

Jäckle, A., Lynn, P. and Burton, J. (2015). Going online with a face-to-face household panel: Effects of a mixed mode design on item and unit non-response. *Survey Research Methods*, 9, pp. 57-70.

Kalton, G. and Flores-Cervantes, I. (2003). Weighting methods. *Journal of Official Statistics*, 19(2), pp. 81-97.

Laurie, H. and Lynn, P. (2009). The use of respondent incentives on longitudinal surveys. *Methodology of Longitudinal Surveys*, pp. 205-233.

Lepkowski, J.M. and Couper, M.P. (2002). *Nonresponse in the Second Wave of Longitudinal Household Surveys.* Survey Nonresponse, Wiley, pp. 259-272.

Little, R.J.A. and Rubin, D.B. (2014). *Statistical Analysis with Missing Data*, vol. 333. John Wiley & Sons.

Little, R.J. and Su, H.L. (1989). Item non-response in panel surveys. In: D. Kasprzyk, G. Duncan and M.P. Singh, eds., *Panel Surveys*. New York: Wiley, pp. 400-425.

Lugtig, P. (2014). Panel attrition: separating stayers, fast attriters, gradual attriters, and lurkers. *Sociological Methods & Research*, 43(4), pp. 699-723.

Lynn, P. (1991). A split-run experiment in a postal survey. *The Statistician*, 40, pp. 61-66.

Lynn, P. (2001). The impact of incentives on response rates to personal interview surveys: role and perceptions of interviewers. *International Journal of Public Opinion Research*, 13, pp. 326-337.

Lynn, P. (2009). Methods for longitudinal surveys. In: P. Lynn, ed., *Methodology of Longitudinal Surveys*. John Wiley & Sons, Ltd, pp. 1-19.

Lynn, P. (2013). Alternative sequential mixed-mode designs: effects on attrition rates, attrition bias, and costs. *Journal of Survey Statistics and Methodology*, 1(2), pp. 183-205.

Lynn, P. (2014a). Longer interviews may not affect subsequent survey participation propensity. *Public Opinion Quarterly*, 78(2), pp. 500-509.

Lynn, P. (2014b). Targeted response inducement strategies on longitudinal surveys. In: U. Engel, B. Jann, P. Lynn, A. Scherpenzeel and P. Sturgis, eds., *Improving Survey Methods: Lessons from Recent* Research, 1st ed. Routledge, pp. 322-338.

Lynn, P. (2016). Targeted appeals for participation in letters to panel survey members. *Public Opinion Quarterly*, 80(3), pp. 771-782.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

Lynn, P. (2017). From standardised to targeted survey procedures for tackling non-response and attrition. *Survey Research Methods*, 11(1), pp. 93-103.

Lynn, P. (2018). Tackling panel attrition. In: D.L. Vannette and J.A. Krosnick, eds., *Palgrave Handbook of Survey Research.* Palgrave Macmillan, pp. 143-153.

Lynn, P. and Clarke, P. (2002). Separating refusal bias and non-contact bias: evidence from UK national surveys, *Journal of the Royal Statistical Society Series D (The Statistician)*, 51(3), pp. 319-333.

Lynn, P., Turner, R. and Smith, P. (1998). Assessing the effects of an advance letter for a personal interview survey, *Journal of the Market Research Society* 40(3), pp. 265-272.

Lynn, P., Buck, N., Burton, J., Jäckle, A. and Laurie, H. (2005). *A Review of Methodological Research Pertinent to Longitudinal Survey Design and data collection.* Tech. rept. ISER Working Paper Series.

Lynn, P., Kaminska, O. and Goldstein, H. (2014). Panel attrition: How important is interviewer continuity? *Journal of Official Statistics*, 30(3), pp. 443-457.

Massey, D.S. and Tourangeau, R. (2013). Where do we go from here? Nonresponse and social measurement. *The Annals of the American Academy of Political and Social Science*, 645(1), pp. 222-236.

Matsuo, H., Billiet, J., Loosveldt, G., Berglund, F. and Kleven, Ø. (2010). Measurement and adjustment of non-response bias based on non-response surveys: the case of Belgium and Norway in the European Social Survey Round 3. *Survey Research Methods*, 4(3), pp. 165-178.

McGinley, S.P., Zhang, L., Hanks, L.E. and O'Neill, J.W. (2015). Reducing longitudinal attrition through Facebook. *Journal of Hospitality Marketing & Management*, 24(8), pp. 894-900.

McGonagle, K.A., Schoeni, R.F. and Couper, M.P. (2013). The effects of a between-wave incentive experiment on contact update and production outcomes in a panel study. *Journal of Official Statistics*, 29(2), pp. 261-276.

Mostafa, T. and Wiggins, R. (2015). The impact of attrition and non-response in birth cohort studies: a need to incorporate missingness strategies. *Longitudinal and Life Course Studies,* 6(2), pp. 131-146.

Nakai, M. and Ke, W. (2011). Review of the methods for handling missing data in longitudinal data analysis. *International Journal of Mathematical Analysis*, 5(1), pp. 1-13.

Nooraee, N., Molenberghs, G., Ormel, J. and Van den Heuvel, E.R. (2018). Strategies for handling missing data in longitudinal studies with questionnaires. *Journal of Statistical Computation and Simulation*, pp. 1-22.

Olsen, R.J. (2005). The problem of respondent attrition: Survey methodology is key. *Monthly Labor Review*, February, pp. 63-70.

Pollock, G., Ozan, J., Goswami, H., Rees, G. and Stasulane, A., eds. (2018). *Measuring Youth Well-being: How a Pan-European Longitudinal Survey Can Improve Policy*, vol. 19. Springer International Publishing.

Rubin, D.B. (2004). *Multiple Imputation for Nonresponse in Surveys*, vol. 81. John Wiley & Sons.

Ryu, E., Couper, M.P. and Marans, R.W. (2005). Survey incentives: Cash vs. in-kind; face-to-face vs. mail; response rate vs. nonresponse error. *International Journal of Public Opinion Research*, 18(1), pp. 89-106.

Sánchez-Fernández, J., Muñoz-Leiva, F. and Montoro-Ríos, F.J. (2012). Improving retention rate and response quality in web-based surveys. *Computers in human behaviour*, 28(2), pp. 507-514.

Särndal, C.-E. and Lundström, S. (2005). *Estimation in Surveys with Nonresponse*. John Wiley & Sons.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.

Schouten, B., Cobben, F., Lundquist, P. and Wagner, J. (2016). Does more balanced survey response imply less non-response bias? *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 179(3), pp. 727-748.

Seaman, S.R, and White, I.R. (2013). Review of inverse probability weighting for dealing with missing data. *Statistical Methods in Medical Research*, 22(3), pp. 278-295.

Si, Y., Reiter, J.P. and Hillygus, D.S. (2015). Semiparametric selection models for potentially non-ignorable attrition in panel studies with refreshment samples. *Political Analysis*, 23(1), pp. 92-112.

Singer, E. (2006). Introduction: Nonresponse bias in household surveys. *International Journal of Public Opinion Quarterly*, 70(5), pp. 637-645.

Sterne, J.A., White, I.R, Carlin, J.B., Spratt, M.P., Royston, P., Kenward, M.G, Wood, A.M. and Carpenter, J.R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, 338, pp. 157-160.

Stoop, I.A., Billiet, J., Koch, A. and Fitzgerald, R. (2010). *Improving Survey Response: Lessons Learned from the European Social Survey*. John Wiley & Sons.

Taylor, S. and Lynn, P. (1998). The effect of a preliminary notification letter on response to a postal survey of young people. *Journal of the Market Research Society* 40(2), pp. 165-174.

Tourangeau, R., Michael B.J., Lohr, Sh. and Li, J. (2017). Adaptive and responsive survey designs: A review and assessment. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 180(1), pp. 203-223.

Tunali, I. and Ekinci, E. (2007). *Dealing with Attrition when Refreshment Samples are Available: An Application to the Turkish Household Labour Force Survey.* Tech. rept. Mimeo.

Vandecasteele, L. and Debels, A. (2006). Attrition in panel data: the effectiveness of weighting. *European Sociological Review*, 23(1), pp. 81-97.

Watson, N. and Wooden, M. (2018). Chasing hard-to-get cases in panel surveys: Is it worth it? *Methods, Data, Analyses*, October, pp. 1-24.

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under the Grant Agreement no 777449.