#0

Nonresponse Bias

Achim Koch
GESIS – Leibniz Institute for the Social Sciences

Michael Blohm
GESIS – Leibniz Institute for the Social Sciences

Publication date: December 6, 2016

Nonresponse is a source of error in surveys. In the present contribution, the concept of nonresponse bias is explained and the relationship between the response rate and nonresponse bias is discussed. Different methods to determine nonresponse bias are presented, and measures to deal with the problem of nonresponse bias during data collection are addressed.

"No issue in survey research is more misunderstood or controversial than nonresponse." (Dixon & Tucker, 2010)

1 Introduction

Nonresponse is one of the possible sources of error in surveys. It occurs when one does not succeed in collecting data for all units in a random sample. Usually, a distinction is made between two types of nonresponse: unit nonresponse, where all data for a sampled unit are missing, and item non response, where only part of the data could not be obtained. The present deliberations address the problem of unit nonresponse. Although they refer to the model of a face-to-face survey of the general population, they apply also – possibly in modified form – to other survey modes.

Nonresponse is a ubiquitous problem in survey research because it is almost impossible in any survey to conduct interviews with all selected target persons. In face-to-face surveys, some persons cannot be reached during the field period, while others are contacted but are unable or unwilling to take part in the survey. Over the past decades, nonresponse rates have increased in many Western countries. Nonresponse gives rise to problems in several respects. First, the sampling error increases when the realised sample size does not correspond to the originally targeted sample size. Second, the more elaborate the data collection procedures for achieving a certain response rate and sample size are, the higher are the costs. Third, surveys are affected by nonresponse bias if nonresponse is not randomly distributed across the target population.

In practice, the response rate, that is, the percentage of persons in the originally selected sample with whom interviews could be conducted, is often used as a proxy measure of nonresponse bias. The assumption is that the higher the response rate is, the lower is nonresponse bias. However, as we shall see in what follows, this is correct only when certain conditions are fulfilled. In Section 2, the relationship between the response rate and nonresponse bias will be explained. In Section 3, different methods to empirically determine nonresponse bias are presented. And finally, Section 4 deals with the question of how the problem of nonresponse bias can be addressed during data collection.

2 Response rate and nonresponse bias

The response rate is one of the most frequently used indicators of survey quality. The attractiveness of the response rate as a quality indicator is due, among other things, to the fact that it is (supposedly) easy to measure and that a single compact measure of survey quality is provided. However, this view ignores the fact that nonresponse is only one of several potential sources of survey error (Groves et al., 2009). Moreover, the response rate is only an imperfect indicator of nonresponse bias in surveys, as it represents only one of the two components that determine the magnitude of this bias.

In the case of linear statistics, such as means and proportions, nonresponse bias is a multiplicative function between the nonresponse rate and the difference between survey respondents and nonrespondents with respect to the variable of interest (Groves, 2006).

The nonresponse bias of a mean $\bar{y}$, for example, is calculated according to the following formula

$$NRB(\bar{y}) = NRR \times \left(\bar{y}R - \bar{y}NR\right)$$

where NRB is the nonresponse bias, NRR is the nonresponse rate, and $\bar{y}R$ is the mean of the respondents and $\bar{y}NR$ the mean of the nonrespondents on the survey variable in question.

This formula makes two things clear:

1. Only the *potential* for nonresponse bias can be inferred from the level of the nonresponse rate: The higher the nonresponse rate is, the greater the magnitude of the bias may be. In principle, however, nonresponse bias may be low even when the response rate is low – namely, when nonresponse is predominantly random, in other words, when the differences between the respondents and the nonrespondents on a particular variable are minor.
2. Nonresponse bias is a variable-specific phenomenon. A survey with a low response rate may be clearly biased with respect to one variable (because respondents and nonrespondents differ with respect to that variable) but may be largely unbiased with respect to another variable (because there is no difference, or only a slight difference, between respondents and nonrespondents with respect to that variable). A meta-analysis of nonresponse bias studies (Groves, 2006) revealed that much of the variation in nonresponse bias estimators lay within studies (and not across studies with different response rates).

The division of a random sample into the two groups – respondents and nonrespondents – is a simple deterministic model. In recent years, an alternative, stochastic model has attracted increasing attention (Groves, 2006). It assumes that people cannot simply be characterised either as survey respondents or as nonrespondents. Different survey designs lead to different divisions of the same population into respondents and nonrespondents. This suggests that, depending on the topic of the survey, the length of the questionnaire, the use of incentives, the behaviour of the interviewers, etc., every sampled unit has a specific, nonzero response propensity, or likelihood of participation. This propensity cannot be observed directly but rather only estimated. Under this model, the nonresponse bias of the mean of a survey variable (y) can be represented by the ratio between the covariance of the survey variable and the response propensity (p) and the mean response propensity (which corresponds to the response rate):

$$NRB\left(\bar{y}\right) = \frac{\sigma_{y,p}}{\bar{p}}$$

According to this formula, nonresponse bias increases as the covariance between the response propensity and the survey variable in question increases. It makes clear that nonresponse bias occurs only if a relationship exists between the response propensity and the substantive survey variable of interest. This relationship may be due to the fact that the survey variable directly influences the response propensity (i.e., y influences p). For example, in a survey on literacy using a self-administered questionnaire, literacy itself is a determinant of the propensity to respond (p). Alternatively, the relationship between the response propensity and the variable of interest may be due to the fact that they are both influenced by a third variable (z), as the common cause of y and p. For example, in an election survey, an interest in politics (z) influences both the response propensity (p) and voting behaviour (y).

The decisive question with respect to the relationship between the response rate and nonresponse bias is: How does the difference between respondents and nonrespondents vary as the response rate changes? In the past, it was often (implicitly) assumed that the difference between respondents and nonrespondents was more or less fixed. In that case, every increase in the response rate would be accompanied by a decrease in nonresponse bias. However, empirical investigations of the effects of changes in the response rate on survey results have shown that this is not the case. For example, U.S. researchers found only minor differences in the results of two telephone surveys whose response rates differed by 25 percentage points due to the different levels of effort expended to obtain interviews (Keeter et al., 2000; 2006). A comparably designed study on the face-to-face German General Social Survey (ALLBUS) 2008 obtained a similar result (Blohm & Koch, 2009). Despite the fact that the supplementary study had a much higher response rate than the main ALLBUS study (63% compared to 40%), the results of the supplementary study differed from those of the main study only with respect to a few variables.

Several conclusions can be drawn from the above deliberations:

1. It is not useful to specify a certain (high) response rate limit that, if reached or surpassed, is deemed to solve the nonresponse problem because no nonresponse bias is to be expected (which would be the case only if an almost 100% response rate was achieved). Conversely, there is also no minimum response rate level below which the survey results should be generally deemed to be problematic (because they are biased). Specifications, for instance, that link the acceptance of a data set solely to the achievement of a certain response rate target are therefore devoid of scientific justification.
2. Nonresponse bias is a variable-specific phenomenon and should therefore also be determined in a variable-specific way. Hence, it is hardly possible to speak of the nonresponse bias of a survey, except, perhaps, in the case of surveys that focus on just one topic. The situation is

much more complicated in multi-topical surveys in which nonresponse bias can, in principle, be determined for dozens, or even hundreds, of variables.

3. The nonresponse bias that can be observed for different variables in a particular survey is not a direct consequence of the nonresponse rate per se but rather of the survey design and the procedures and methods employed (the so-called survey protocol). For example, the two main sources of nonresponse (non-contact and refusal) usually have different causes and therefore frequently correlate with different survey variables. Hence, a survey will exhibit different nonresponse bias patterns if nonresponse is primarily due to the fact that target persons could not be contacted than when it was primarily the result of target persons' unwillingness to cooperate.

3 Methods to determine nonresponse bias

If one wishes to empirically investigate nonresponse bias, one is faced with the problem that the two components that determine its magnitude are differentially accessible to measurement. The response, or nonresponse, rate can, at least in theory, be easily and unequivocally determined. However, information about the second component of nonresponse bias – the difference between respondents and nonrespondents with respect to a certain survey variable or the covariance between the survey variable and the response propensity – is not normally available. However, there are a number of different methods to approach the problem. These methods have their own specific strengths and weaknesses. Although not every method can be applied in every survey, it is advisable to use different methods to approach the problem in order to achieve a better understanding of the situation in the survey in question. The most important methods are (Montaquila & Olson, 2012):

1. Comparison of survey results with other data sources/aggregate statistics
2. Analysis of (individual) data for respondents and nonrespondents
3. Analysis of variations within the group of respondents
4. Analysis of the effects of different weighting procedures

3.1 Comparison of survey results with other data sources/aggregate statistics

The most frequently used method for studying nonresponse bias is to compare the survey data with information from another, more accurate, data source. One example of this is the comparison of survey results with data from Germany's Microcensus, which has a low nonresponse rate because participation is compulsory (Koch, 1998). As a rule, such a comparison must be restricted to a few sociodemographic characteristics. The prerequisite for the comparison is that these characteristics were collected in a comparable way and that there are no (marked) differences between the target population and the time of data collection of the two surveys. In practice, the measurements are often not completely identical. Differences may then stem both from nonresponse error and from different measurements. Because both the survey in question and the Microcensus are based on samples, sampling error must also be taken into account in the comparison. Apart from that, it should be noted that it is not possible to infer from the absence of differences between the demographic characteristics in question that the actual survey variables of interest are unbiased. Bias, or the absence of bias, in demographic characteristics determines bias in other variables only to the extent that these variables are correlated with the respective demographic characteristics. In a meta-analysis, Groves & Peytcheva (2008) found that bias in substantive survey variables could not be predicted by bias in demographic characteristics.

3.2 Analysis of (individual) data for respondents and nonrespondents

Another method for investigating nonresponse bias uses information that is available for the entire survey sample (respondents and nonrespondents). If such information is available, the data for the respondents can be compared to those for the nonrespondents in order to obtain estimates for the nonresponse bias. Relevant information may already be included in the sampling frame itself. Examples include information on age and gender in a population register sample or information on the duration of membership in the case of a survey of members of an association. Further information can be gained if individual data from other sources can be matched to the sample (for example, health or employment data from administrative records). Sometimes, aggregate information may also be used. The MICROM data (MICROM, 2011), for example, which were originally collected for the purpose of direct marketing, contain information for aggregates of eight households or for street sections. This information can be matched to a sample on the basis of household addresses (Goebel et al., 2007). Observations by interviewers are a further source of information about the entire sample. In the European Social Survey, for example, interviewers are instructed to classify for all sampled cases the type of house in which the target person lives and to make and record certain observations about the immediate vicinity (Stoop et al., 2010).

The advantage of this method is that comparable measurements are then available for survey respondents and nonrespondents. On this basis, estimates of nonresponse bias can be gained for the observed (auxiliary) variables. In addition, profiles can be created for the different categories of nonresponse (e.g., target persons who could not be contacted or who were unwilling to cooperate) thereby providing clues about the sources of the nonresponse bias. However, the available characteristics do not usually include the actual survey variables of interest (if they did, the survey would not have been necessary). In general, it is important that the auxiliary variables used are closely correlated with the survey variables of interest (this can be empirically investigated for the group of survey respondents) and that further analyses reveal that they are also linked to participation behaviour. If this is the case, information about the presumed nonresponse bias of the survey variables of interest can be derived from the auxiliary variables. One potential disadvantage of this method is that the information in question is frequently not available for all sample units. Moreover, the quality of the measurement of the auxiliary variables is sometimes questionable. For example, observations frequently vary across interviewers (Olson, 2012).

3.3 Analysis of variations within the group of respondents

The aim of this design is to gain information about nonresponse bias by comparing different subgroups of survey participants. To this end, respondents are divided into two groups – easy and difficult cases – according to the level of effort expended to obtain an interview. The division of the cases may be based on different criteria. For example, (a), early responders can be distinguished from those who were recruited later in the survey period. Or (b), respondents can be distinguished according to the number of contact attempts that were necessary. Also conceivable, (c), is a distinction according to whether respondents were immediately willing to participate or whether they initially refused and could be convinced to take part in the interview only after further attempts were made to encourage them to do so. These paradata are compiled during the fieldwork period – usually in contact protocols. The greatest advantage of this method is that it allows one to investigate for every survey variable whether there are differences between the different subgroups. The most serious disadvantage is that no information is available about the actual nonrespondents. The – usually untested – assumption is that the nonrespondents are more similar to the "difficult" cases than to the "easy" cases. Empirically, however, this is by no means inevitably the case, as demonstrated, for example, in a study conducted by Lin & Schaeffer (1995). A further limitation of this approach is that the indicators of difficulty cannot be unequivocally assigned to processes of reachability or willingness to be interviewed. It is advisable, for example, not to distinguish respondents according to the total number of contact attempts, but rather to consider only the contact attempts prior to the first contact, as this distinction is more directly related to the process of reachability.

A special case of the analysis of variations within the group of survey respondents is to conduct a special nonresponse follow-up study. Such a study is usually carried out after the actual survey, and an attempt is made to persuade the nonrespondents (or a subset of the nonrespondents) to participate in the survey after all by increasing the effort expended (e.g., by offering monetary incentives). To increase the chances of success, the length of the questionnaire is sometimes reduced and/or a different survey mode is used (Stoop, Billiet, Koch, & Fitzgerald, 2010). However, this renders it more difficult to compare the participants in the original survey to those in the follow-up study. The different field periods may also raise problems in this regard, especially when it cannot be ruled out that the variables measured change over time (as in the case of the measurement of attitudes, for example). However, the fundamental problem with nonresponse follow-up studies is that, despite all efforts, interviews cannot usually be conducted with all the target persons. In other words, the follow-up study also faces a (sometimes considerable) nonresponse problem.

3.4 Analysis of the effects of different weighting procedures

For the sake of completeness, it should be mentioned here that comparing unweighted (or merely design-weighted) data with nonresponse-weighted data may yield information about the extent to which nonresponse bias can be compensated for by weighting (Gabler et al., 2016). If different weighting procedures are available, their results can be compared with each other and with the unweighted results. The main problem with this method is that no binding standard is available that might help one to decide which results reflect reality better.

4 Data collection governed by the precept of optimising response rates and minimising nonresponse bias

Anyone who is planning a survey should actively engage with the nonresponse problem. Normally, one will design the survey in such a way that nonresponse is reduced and the highest possible response rate is achieved. There are several different tried-and-tested measures available for this purpose (Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009; Koch et al., 2012). In face-to-face surveys (just as in other survey modes) one will not limit oneself to individual measures but rather employ a whole bundle of measures in order to try to minimise the problems of reachability, ability to respond, and willingness to cooperate. While some measures improve the chances for success in all sub-processes (especially contact and motivation), others are targeted more specifically towards successfully contacting or successfully motivating target persons. For example, the deployment of experienced, well-trained, and adequately remunerated interviewers is generally a good basis for successfully minimising nonresponse. Measures such as determining a sufficiently long field period and specifying the number and scheduling of contact attempts (day of the week, time, interval between attempts, etc.) are geared primarily towards successfully contacting target persons, whereas training interviewers in refusal avoidance strategies and using advance letters or incentives is primarily aimed at motivating target persons to participate in the survey.

The obvious question is whether the said measures are also a suitable means of avoiding or reducing possible nonresponse bias. It is hardly possible to give a general answer to this question. Previous research has primarily investigated the effects of these measures on the response rate and has hardly studied their effects on nonresponse bias at all. Moreover, one should bear in mind that the effect of individual measures depends also on the respective circumstances of the specific survey – that is, on the population studied, the topic and length of the questionnaire, and the other planned measures to increase the response rate, etc. Technically speaking, when assessing the effectiveness of individual measures to reduce nonresponse bias, the focus is likely to be on interaction effects rather than main effects (Groves, 2006; Groves & Peytcheva, 2008). For example, if nonresponse bias is expected in a survey on a specific topic because persons who are very interested in the topic will exhibit a higher response propensity, the use of respondent incentives may be expedient. As has been demonstrated, these incentives can help to motivate, in particular, target persons with little interest in the survey topic to participate in an interview (Groves et al., 2004). However, it cannot be ruled out that respondent incentives may be counter-productive in some situations – namely, when they increase the response propensity of persons who are already overrepresented in the sample.

Groves (2006, p. 668) gave the following general advice to survey practitioners: "Blind pursuit of high response rates in probability samples is unwise; informed pursuit of high response rates is wise." From this perspective, it cannot be a question of maximising the response rate at all costs. Such maximisation often follows the path of least resistance by targeting the easy cases. If a correlation exists between the survey variables of interest and the response propensity, then such a strategy might even increase an existing nonresponse bias (Beullens & Loosveldt, 2012).

Formulated in general terms, the objective of minimising nonresponse bias during data collection might be to achieve the same response propensity for all sampled units by the end of the survey period. This necessitates identifying target persons with a low response propensity and making intensive efforts to enlist their cooperation (Peytchev et al., 2010). Here, auxiliary variables are needed that influence the response propensity and are closely correlated with the survey variables. The identification and collection of such variables is no easy undertaking. Often, only a few demographic data are available (e.g., the age and sex of the target person from the sampling frame, or interviewer observations regarding the target person's dwelling and living environment). If these auxiliary variables reveal that certain groups exhibit a response rate that is below a defined level, targeted interventions should be initiated in order to increase their response propensity. This may involve increasing contact efforts, deploying particularly competent interviewers, or offering a (higher) monetary incentive.

The implementation of a corresponding system of targeted interventions calls for the intensive observation and analysis of fieldwork processes and the capacity to actually intervene in data collection (on so-called "responsive designs," see Groves & Heeringa, 2006). If the fielding of a face-to-face population survey involves collaboration with a commercial survey institute, there are considerable obstacles to the implementation of such an approach. The prevailing philosophy and the organisational procedures (e.g., interviewer training and remuneration) in survey practice are primarily aimed at achieving a certain number of interviews and the highest possible response rate. Interviewers are freelancers who often work for several institutes and are hardly bound by instructions from the survey institute. In such circumstances, the orientation of the fieldwork towards target persons who are disproportionately hard to reach or hard to motivate can be realised only to a limited extent. Moreover, leaving aside these obstacles to implementation, it should be stressed once again that the success of such an approach depends crucially on the extent to which one succeeds in finding good auxiliary variables. If the available auxiliary variables influence response propensity but are only weakly correlated with the key survey variables, the orientation of the fieldwork towards these auxiliary variables means (sometimes considerable) extra effort without nonresponse bias being reduced significantly as a result.

5 Concluding remarks

In the quotation that prefaces this contribution, nonresponse is described as the most misunderstood and controversial issue in survey research. That nonresponse will continue to be one of the key challenges in survey research is unlikely to be a subject of great controversy. Survey response rates tend to be low in Germany and in many other Western countries, and it will hardly be possible to significantly increase them. Nor will the methods to minimise nonresponse bias at the fieldwork stage yield a comprehensive solution to the problem. Against this background, dealing with the nonresponse problem by means of statistical correction procedures at the data analysis stage will gain in importance (Brick, 2013). In these approaches, too, the availability or collection of meaningful auxiliary variables is crucial. The scientific penetration of survey participation behaviour and the systematic design of the entire data collection process are thus on the agenda.

5.1 References

5.1.1 Further references:

Groves, R. M., & Couper, M. P. (1998). Nonresponse in Household Interview Surveys. New York: John Wiley & Sons.

Schnell, R. (1997). Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske & Budrich. Retrieved from http://kops.ub.uni-konstanz.de/xmlui/bitstream/handle/urn:nbn:de:bsz:352-opus56148/Nonresponse_in_Bevoelkerungsumfragen.pdf?sequence=1

5.1.2 Journal special issues on the subject of nonresponse

The Annals of the American Academy of Political and Social Science (2013), 645(1): The Nonresponse Challenge to Surveys and Statistics

Journal of Official Statistics (2011), 27(2)

Journal of the Royal Statistical Society: Series A (2013), 176(1): The use of paradata in social survey research

Public Opinion Quarterly (2006), 70(5): Special Issue: Nonresponse bias in household surveys

Bibliography

Beullens, K., & Loosveldt, G. (2012). Should high response rates really be a primary objective?. *Survey Practice*, *5*(3), 1–5. https://doi.org/10.29115/sp-2012-0019

Blohm, M., & Koch, A. (2009). Ausschöpfungsquoten und Stichprobenqualität am Beispiel des ALLBUS 2008: Führt eine höhere Ausschöpfung zu anderen/besseren Umfrageergebnissen?. *Quality of Large-Scale Surveys.*

Brick, J. M. (2013). Unit Nonresponse and Weighting Adjustments: A Critical Review. *Journal of Official Statistics*, *29*(3), 329–353. https://doi.org/10.2478/jos-2013-0026

Dixon, J., & Tucker, C. (2010). Survey nonresponse. In P. V. Marsden & J. D. Wright (Eds.), *Handbook of Survey Research (2nd ed.)* (pp. 593–630). Emerald.

Gabler, S., Kolb, J.-P., Sand, M., & Zins, S. (2016). *Gewichtung (GESIS Survey Guidelines).* GESIS - Leibniz Institute for the Social Sciences. https://doi.org/10.15465/GESIS-SG_EN_007

Goebel, J., Spieß, C. K., Witte, N. R. J., & Gerstenberg, S. (2007). Die Verknüpfung des SOEP mit MICROM-Indikatoren: Der MICROM-SOEP Datensatz. *DIW Data Documentation 26.* http://www.diw.de/documents/publikationen/73/diw_01.c.78103.de/diw_datadoc_2007-026.pdf

Groves, R. M., Presser, S., & Dipko, S. (2004). The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly*, *68*(1), 2–31. https://doi.org/10.1093/poq/nfh002

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. Wiley. https://books.google.de/books?id=HXoSpXvo3s4C

Groves, R. M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, *70*(5), 646–675. https://doi.org/10.1093/poq/nfl033

Groves, R. M., & Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias: A Meta-Analysis. *Public Opinion Quarterly*, *72*(2), 167–189. https://doi.org/10.1093/poq/nfn011

Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the Impact of Growing Nonresponse on Estimates from a National RDD Telephone Survey. *Public Opinion Quarterly*, *70*(5), 759–779. https://doi.org/10.1093/poq/nfl035

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of Reducing Nonresponse in a National Telephone Survey. *Public Opinion Quarterly*, *64*(2), 125–148. https://doi.org/10.1086/317759

Koch, A., Fitzgerald, R., Stoop, I., Widdop, S., & Halbherr, V. (2012). *Field procedures in the European Social Survey round 6: enhancing response rates.* http://www.europeansocialsurvey.org/docs/round6/methods/ESS6_response_enhancement_guidelines.pdf

Koch, A. (1998). Wenn "mehr" nicht gleichbedeutend mit "besser" ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. *ZUMA Nachrichten*, *22*(42), 66–90.

Lin, I.-F., & Schaeffer, N. C. (1995). Using Survey Participants to Estimate the Impact of Nonparticipation. *Public Opinion Quarterly*, *59*(2), 236–237. https://doi.org/10.1086/269471

MICROM. (2011). *Microm Datenhandbuch. Arbeitsunterlagen für microm MARKET & GEO.*

Montaquila, J., & Olson, K. (2012). *Practical tools for nonresponse bias studies.* http://www.amstat.org/sections/srms/webinarfiles/NRBiasWebinarApril2012.pdf

Olson, K. (2012). Paradata for Nonresponse Adjustment. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 142–170. https://doi.org/10.1177/0002716212459475

Peytchev, A., Riley, S., Rosen, J., Murphy, J., & Lindblad, M. (2010). Reduction of Nonresponse Bias through Case Prioritization. *Survey Research Methods*, No1(2010). https://doi.org/10.18148/SRM/2010.V4I1.3037

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

Publisher