Reducing the differences between respondents and nonrespondents or increasing the response rate? What is the best order to reduce nonresponse bias? Example from the European Social Survey

Koen Beullens* [1] and Geert Loosveldt[1]

[1]Centre for Sociological Research, Leuven (Belgium)

Nonresponse Workshop, Oslo, August 31-September 2, 2015

1 Introduction

A rather simple expression unifies three key quantities related to nonresponse:

$$B = K(1 - R), \tag{1}$$

where B expresses the bias of the mean of a variable ($B = \bar{y}_{full} - \bar{y}_{resp}$), or the difference between the mean of the full sample and the mean among the respondents only[1]. K indicates the contrast between the respondents and nonrespondents regarding the mean of y ($K = \bar{y}_{nonresp} - \bar{y}_{resp}$) and R is the response rate. The expression suggests that bias can be reduced by (1) reducing the contrast between respondents and nonrespondents and (2) increase the response rate.

The way K and R influence B is not strictly linear. In fact, the impact of the contrast K depends on the response rate R, as illustrated by Figure 1. The figure seems to suggest that whenever the response rate is relatively low (e.g. $<$ 70-80%) the contrast needs to be low too in order to keep the bias under a certain threshold (e.g. $<$0.05 standard deviation between respondents and full sample). With high response rates (e.g. $>$ 70-80%), the contrast becomes rather unimportant to combat bias. For example, the contrast (K) can be 1 (one full standard deviation between respondents and nonrespondents) and still the bias is 0.05, only if the response rate is 95%. Therefore, above a certain response rate benchmark, the risk of nonresponse bias is fairly minimal. This may explain why some survey researchers have proposed such benchmarks. Bailey (1987) proposes a minimal acceptable response rate of 75%. The Office of Management and Budget of the American federal government (OMB, 2006) requests that survey procedures are generally designed to yield an 80% response rate. The European Social Survey requires response rates of 70% or more from all participating countries.

As response rates of 70% or more have become rather exceptional, the concern for nonresponse bias has shifted the attention towards the composition of the sample as well. There seems to be an increasing awareness amongst survey researchers that a single-minded focus on response rates alone is not advisable. As Groves (2006) finds, high response rates do not necessarily imply low bias in the eventual survey estimates, more refined strategies may more appropriate. Instead of the blind pursuit of high response rates, an informed pursuit of high response rates is wiser.

Setting their fieldwork objective, survey organizations have a choice between:

1. Aim for a very high response rate (>70-80%), which practically excludes the risk of bias

*koen.beullens@soc.kuleuven.be

[1]we do not consider the population mean $\bar{Y}$ to compare the respondent-only mean with since more than nonresponse error might occur (e.g. sampling error, coverage error).

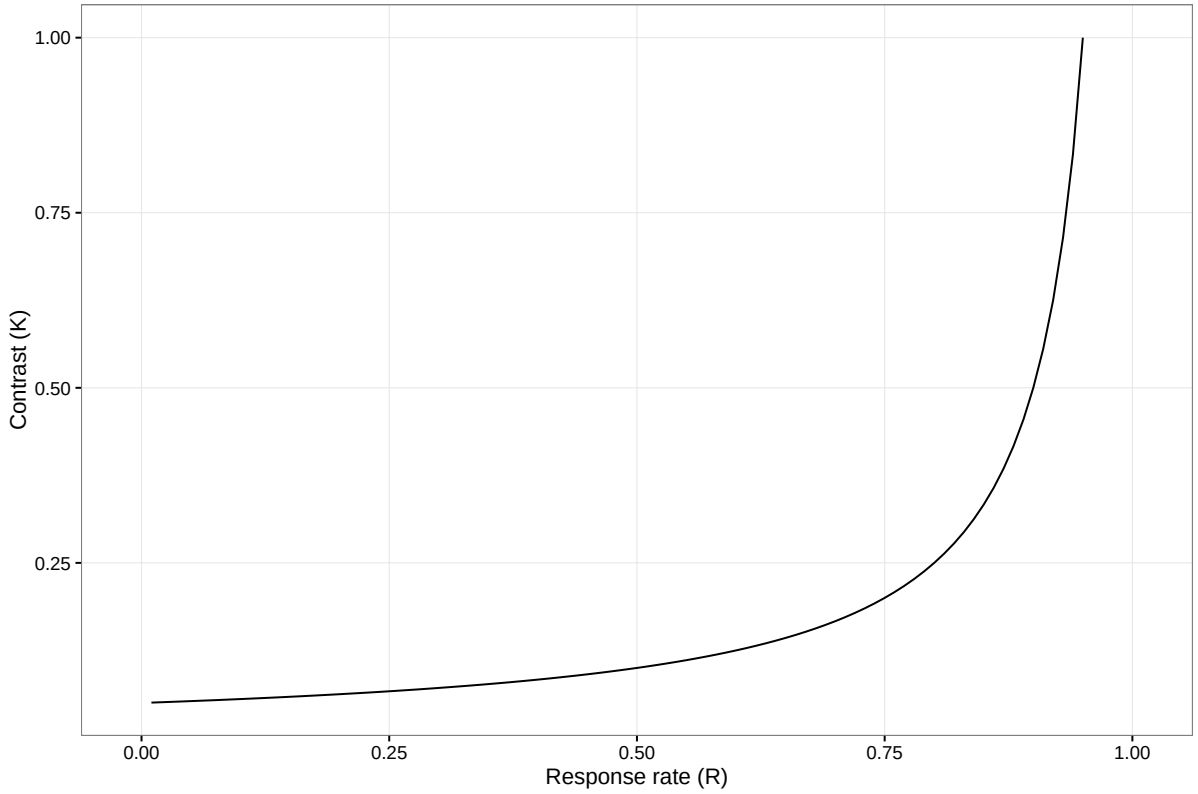

Figure 1: Combinations of Response rates (R) and Contrast (K) for equal level of bias

2. Relax the response rate objective, but instead aim for a fair sample composition (reducing the contrast)

In fieldwork practice, there is a risk that the fieldwork result combines the worst of the two options: having a low response and a high contrast. This will be particularly the case when the fieldwork prioritizes the so-called low hanging cases (easy, high response propensity cases). This path of least resistance strategy maximizes the response rates, but may completely ignore the composition of the sample, resulting in a high bias, since the response is still not high enough (Beullens & Loosveldt, 2012).

In this paper, we seek to monitor how bias and contrast evolve as a function of the increasing attempts during the fieldwork. Multiple attempts per cases may be done during the course of the fieldwork in order to advance the response. At each of these stages (attempts), the bias, contrast and response rate can be monitored. Furthermore, some cases may not be selected for new contact attempts and are therefore considered final nonresponse at attempt t. This paper will also try to find an indication as to whether higher propensity cases (low hanging fruit) have been prioritized or not. It can be expected, given that the response is the main objective for the ESS), that high propensity cases will be given priority, which may be in turn be disadvantageous regarding contrast and bias.

2 Measuring nonresponse bias in the ESS

In order to evaluate the extent to which nonresponse affects survey estimates, the ESS fieldwork protocol has chosen to collect data recorded by the interviewers at their first visit. The following questions were asked:

- TYPE: What type of house does the (target) respondent live in? The interviewer could choose between: farm, detached house, semi-detached house, terraced house, the only housing unit in a building with another purpose (commercial property), flat, student apartment, retirement

house, House-trailer or boat, or other. With regard to this variable, we will rather focus on the distinction between apartment dwellers and non-apartment dwellers (1 binary variable).

- GATE/DOOR: Before reaching the (target) respondent's individual door, is there an entry phone system or locked gate/door? (1: Entry phone system; 2: locked gate/door; 3: both entry phone system and locked gate/door; 4: neither of these). property), flat, student apartment, retirement house, House-trailer or boat, or other. With regard to this variable, we will rather focus on the distinction gate / no gate and entry phone / no entry phone (two binary variables)
- PHYSA: What is your assessment of the overall physical condition of this building/house? (Very good, good, satisfactory, bad, very bad)
- LITTER: In the immediate vicinity, how much litter and rubbish is there? (very large amount, large amount, small amount, none or almost none)
- VANDAA: In the immediate vicinity, how much vandalism and graffiti is there? (very large amount, large amount, small amount, none or almost none)

Apart from these observable data, many countries (particularly individual sample based frame countries) could also provide gender and age information about the sampled individuals. In sum, eight variables are available to make a nonresponse bias assessment.

For each of these variables and for each new contacts attempt per case, an estimate of the bias for that variable can be obtained by calculating the difference between the means among respondents only and the full sample. That bias is then standardized by dividing it by the standard deviation of that variable. Finally, an average is obtained for all available auxiliary variables for that country. It is plotted in Figure 2 (red line). The blue lines represent the equivalent contrasts. It is the difference between the respondents and nonrespondents[2].

Not for every country the full set of auxiliary variables is available. In Czech Republic, France, Ireland, Israel and Portugal the gender and age information is not available, whereas in Germany, Estonia, Norway, Poland and Sweden, no auxiliary variables but age and gender are available.

3 Measuring prioritization during the fieldwork

We will use previous contact outcomes to predict the success probability of a case at the current contact attempts. The underlying idea is that profiles that deemed to be converted more easily ('just away', non-contact, ...) are expected to be more frequently given a second (or third, fourth, ...) chance.

In order to explain the method by which the hypotheses will be tested, consider the fictitious sample as shown in Table 1. The sample consists of six individuals for whom the contact history is presented. Three individuals (1, 3, and 5) eventually participate, while the other three will be considered as final nonrespondents after respectively 1, 3, and 3 unsuccessful contact attempts. In fact, the non-selection of an individual for a renewed attempt is the reason the case is considered as a final nonrespondent. In this particular example, the last outcomes of the contact sequence are refusals or a language barrier; nonresponding profiles that can probably be expected to have a low follow-up success rate. Further, other nonresponse outcome codes that relate to relatively high follow-up success probabilities, such as noncontact or a broken appointment, seem to be consistently re-selected. Therefore, this extract seems to strongly reflect the idea that fieldwork decisions follow the path of least resistance.

Using a much larger dataset, before each contact attempt, the probability of a successful follow-up can be modelled, conditional on the contact history or

[2]Formally, the average bias can be denoted $\frac{1}{8}\sum_{i=1}^{8}\left|\frac{\bar{y}_{r,i}-\bar{y}_{f,i}}{\sigma_{y_i}}\right|$, where for the i^{th} variable y, the difference is determined between the respondent mean $\bar{y}_r$ and the full sample mean $\bar{y}_f$, and standardized by dividing by σ_y. Equivalently, the average contrast is determined by $\frac{1}{8}\sum_{i=1}^{8}\left|\frac{\bar{y}_{r,i}-\bar{y}_{nr,i}}{\sigma_{y_i}}\right|$, where $\bar{y}_{nr}$ is the nonrespondent mean

Table 1: Fictitious extract of contact process data. Strong evidence for the line of least resistance

ID	attempt number	outcome	response?	re-selection?	ID	attempt number	outcome	response?	re-selection?
1	1	noncontact	0	1	4	1	moved	0	1
1	2	noncontact	0	1	4	2	noncontact	0	1
1	3	refusal by proxy	0	1	4	3	refusal by target	0	0
1	4	interview	1	n.a.					
					5	1	refusal by proxy	0	1
2	1	refusal by target	0	0	5	2	noncontact	0	1
					5	3	language barrier	0	0
3	1	appointment	0	1					
3	2	noncontact	0	1	6	1	just away	0	1
3	3	interview	1	n.a.	6	2	interview	1	n.a.

Table 2: Response propensity model, ESS6, 16 countries

Constant[4]	-1.43*	Hungary	0.49*	Refusal by target	-0.37*
Switzerland	-0.29	Ireland	0.46*	Refusal by proxy	-0.48*
Czech Rep.	0.83*	Israel	2.10*	Noncontact	-0.11*
Germany	-0.68*	Lithuania	1.77*	'Just away'	0.12*
Denmark	-0.07	Netherlands	-0.22	Short term illness	-0.07
Estonia	0.47*	Norway	0.29	Long term illness	-0.52
Spain	-0.14	Poland	0.64*	Language barrier	0.06
Finland	-0.45*	Portugal	0.82*	Partial interview	1.30*
France	-0.10	Sweden	-1.04*	Other	0.06
UK	-0.44*	Slovenia	0.15		

*: significant at p<0.05

$$
\begin{aligned}
ln\left(\frac{p_{success,it}}{1-p_{success,it}}\right) &= \#noncontacts_{i(t-1)}\beta_1 \\
&+ \#refusal.by.target_{i(t-1)}\beta_2 \\
&+ \#refusals.by.proxy_{i(t-1)}\beta_3 \\
&+ \ldots + \#language.barrier_{i(t-1)}\beta_C \qquad (2)
\end{aligned}
$$

where C is the number of possible nonresponse outcome categories such as refusals or noncontacts, i is the index for the individual and t represents the contact number (chronological rank order in the contact attempt sequence within individual i). The variables $\#noncontacts$ $\#soft.refusals$, ... $\#language$ count the number of previous contact outcomes of that particular nonresponse code within the sequence of the call records of an individual. For example, ID 1 in table 1 will obtain $\#noncontacts = 2$ and $\#soft.refusals = 1$ at $t = 4$.

We have used ESS6 to estimate the propensity model, based on which the propensities per cases and per contact attempt could be calculated for ESS7. Table 2 provides the parameter estimates of that logistic model[3].

It could be expected that countries would be different regarding the success probability of their contact attempts (for example Lithuania or Israel have much higher success probabilities than Germany or Sweden[5]). More important are the estimates related to previous outcomes. Cases that

[3]Only countries that participated in both round 6 and 7 have been used in this model. Only countries that have less than 5% missing values in at least one of the eight auxiliary variables of ESS7 have been kept in the model.

[5]Although it can not be ruled out that these differences are due to under- or over-reporting in the contact forms.

refused in the past seem to have much lower propensities during the current or future contact attempts as compared to cases that have recorded partial interviews or 'just away' in the past.

With this model, the expected propensities can be projected on every case/attempt combination of ESS7. At every occasion for a new contact attempts, the mean of the projected propensities of the selected cases is calculated and compared to the propensity mean of the cases that were not selected. If the difference in positive, the fieldwork tends to prioritize high propensity cases.

4 Results

In Figure 2 the x-axis in each country represents the response rate which progresses as a function of renewed contact attempts (these attempts are represented in by the dots in the curves). For example, in Belgium the response rate after the first contact attempt was 8%, after the second attempt 26%, ..., and 57% after the last contact attempt. At each of these stages (contact attempts) the bias (red) and contrast (blue) are provided. The prioritization indicator is only provided from the second contact onwards (since all cases have been given a first contact attempt).

- Of all 16 countries, only one (Israel) has attained the required response rate of 70%. It was already mentioned that below this threshold, the contrast between respondents and nonrespondents becomes increasing important to reduce bias.
- In 10 countries, there are indications that the contrast between respondents and nonrespondents is increasing during at least some stages of the fieldwork. This might suggest missed opportunities to even further minimize the bias. In Norway, this has even led to a bias increase.
- Most countries (except in Finland and Israel) seem to generally prioritize the high propensity cases.
- The prioritization indicator correlates with contrast (+0.26) and bias (+0.13). This implies that when countries would give more priority to low propensity cases, the contrast and eventually the bias might further reduce.

This analysis suggests that fieldwork objectives should be re-considered on order to combat nonresponse bias. Since response rates $>$ 70-80% are no longer realistic, fieldworks' attention should shift towards a more appropriate case prioritization, even if this might lead to further response rate erosion.

References

Bailey, K. (1987). *Methods of Social Research* (3rd ed.). New York: Free Press.

Beullens, K., & Loosveldt, G. (2012). Should High Response Rates Really Be a Primary Objective. *Survey Practice*, *5*(3), 1–5.

Groves, R. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, *70*(5), 646–675.

OMB. (2006). *Standards and Guidelines for Statistical Surveys.* Retrieved from `http://www.whitehouse.gov/omb/inforeg_statpolicy`

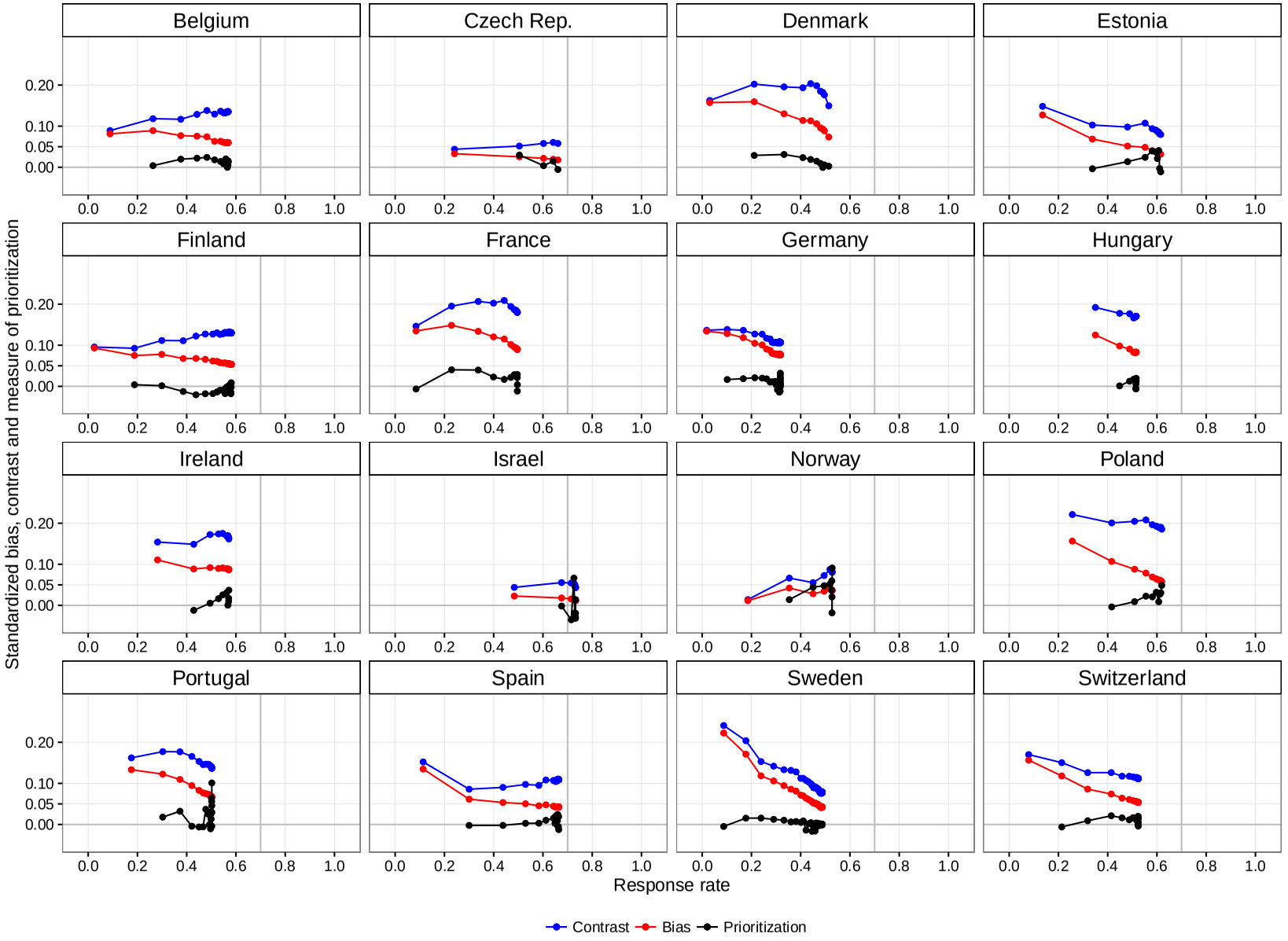

Figure 2: Bias, contrast and case prioritization during ESS7 fieldwork, 2014, 16 countries

a

[a]Swedens' first contact attempts has not been plotted, since the bias (0.945) and contrast (0.956) fell outside the y-axis range