The effect of survey design on nonresponse and nonresponse bias in cross-cultural surveys: findings from the sixth European Working Conditions Survey

Mathijn Wilkens (Eurofound), Aleksandra Wilczynska (Kantar Public) and Gijs van Houten (Eurofound)
mwi@eurofound.europa.eu

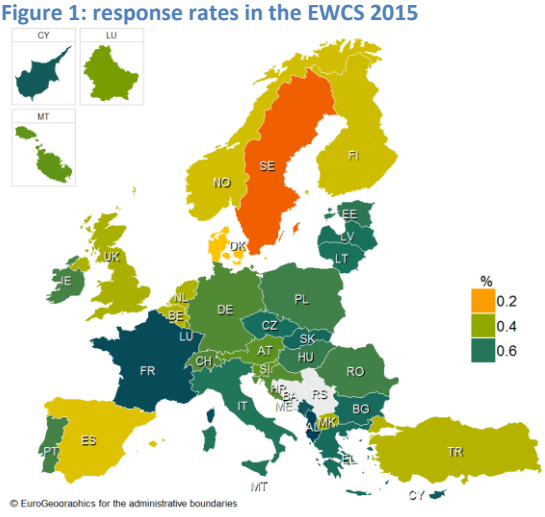
Introduction

Nonresponse bias is an important source of error in surveys (Groves and Couper, 1998). With response rates falling at different speeds in different countries (de Leeuw & de Heer, 2002) and with differences between respondents and non-respondents varying between countries (Stoop, Matsuo, Koch, & Billiet, 2010b) it is even more problematic in cross-cultural surveys.

Eurofound's European Working Conditions Survey (EWCS) is a face-to-face survey of working conditions and quality of work and employment in Europe conducted each five years. The main objectives of the survey are to assess and quantify working conditions across Europe on a harmonised basis and analyse relationships between different aspects of working conditions to contribute to European policy development in particular on quality of work and employment issues.

Response rates of the EWCS vary by wave and by country (Figure 1) and declining response rates are a concern for Eurofound because varying response rates may impede comparability across time and across countries. In addition, declining response rates might cause non-response bias to increase. Therefore, this paper aims to assess non-response bias in the EWCS). The goal is to uncover the extent to which non-response bias is present in the EWCS and to analyse which survey design features limit non-response bias. Lessons learnt are valuable for the design of future waves of the EWCS and other Eurofound surveys.

Figure 1: response rates in the EWCS 2015



Many of the best practices in survey design are geared towards maximising the likelihood that respondents are contacted and subsequently successfully recruited into the survey. The survey design of the most recent wave of the EWCS (6th wave, 2015) includes a variety of features aimed at limiting unit response, limiting non-response bias, and limiting the variability between countries in both. The paradata for the sixth edition of the EWCS conducted in 2015 include both contact sheet information (level-of-effort data) and interviewer observations. Extensive metadata is available on the specificities of the survey design throughout the survey cycle.[1]

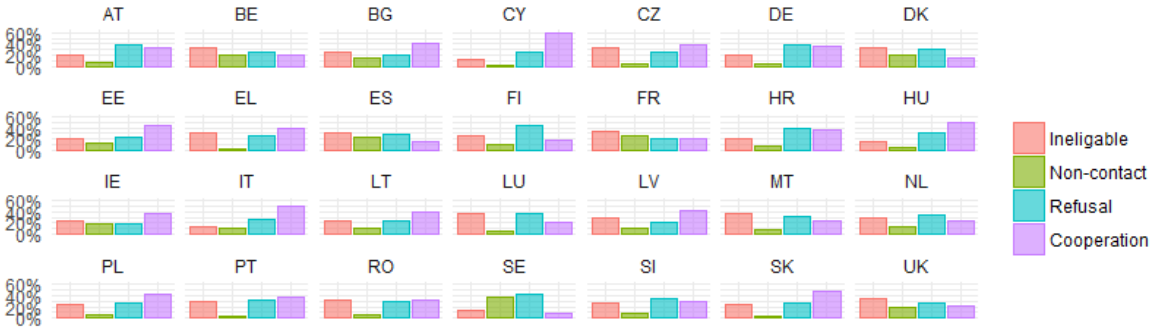
Method

In this paper we attempt to answer the following research questions:

1) *To what extent can noncontacts be predicted using information on the survey design?*
2) *To what extent can refusals be predicted using information on the survey design?*
3) *To what extent are predicted probabilities of being contacted and of cooperating associated with key substantive variables in the sixth EWCS?*

[1] See: https://www.eurofound.europa.eu/surveys/european-working-conditions-surveys

Figure 2: Final fieldwork outcomes by country



To do so we will show the nonresponse patterns in the sixth edition of the EWCS, distinguishing between nonresponse due to non-contacts and nonresponse due to non-cooperation. We will then estimate the probability to be contacted using a multi-level survival model, and within those that were contacted, we will estimate the probability to be interviewed. This will produce the estimated probability to be contacted and the estimated probability to be interviewed, for which we can then test associations with the variables of interest of the survey, among the survey respondents – indicating the nonresponse bias in the data as well as the effectiveness of various design features as a means of reducing nonresponse bias.

Design of the sixth EWCS

A stratified random sample of workers (employees and self-employed) was interviewed face-to-face in their homes. The geographical coverage of the survey has expanded from 12 countries in the first edition to 35 countries in the sixth edition that was carried out in 2015.[2]

The target sample size was 1,000 for most countries. To reflect the larger workforce in larger countries, the target was increased to 1,200 in Poland, 1,300 in Spain, 1,400 in Italy, 1,500 in France, 1,600 in the UK and 2,000 in Germany and Turkey. Eurofound also offered its stakeholders the opportunity to pay for a larger sample in their country, consequently larger target sample sizes were set for Belgium (2500), Slovenia (1600) and Spain (3300).

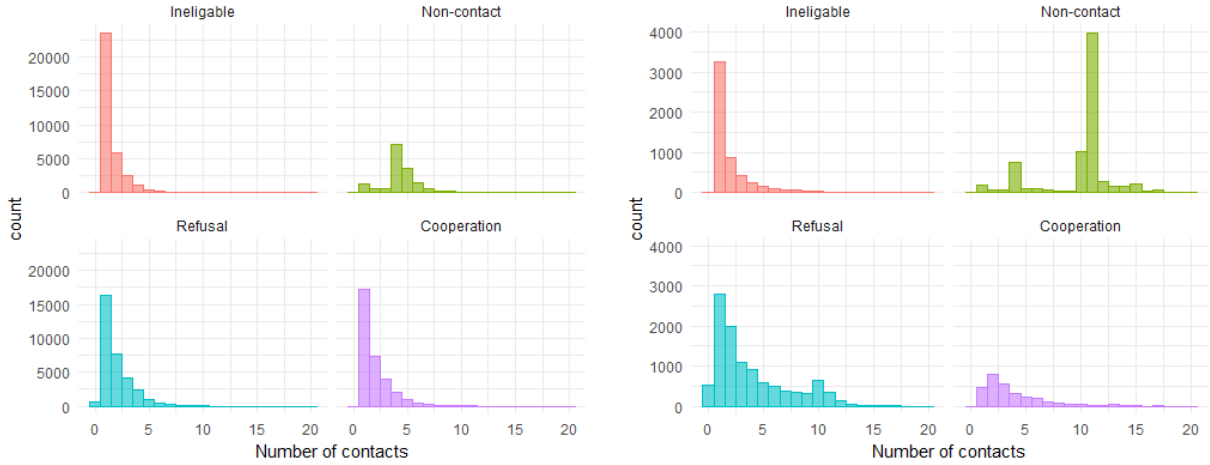
Contact with sample entries is made face to face, except for Denmark, Finland and Sweden, where contact was established over the telephone. In face to face contacting, a sample entry was required to be visited at least 4 times before it could be discarded as a noncontact (10 calls were required in case of phone contacting). Contact attempts were required to take place at different days of the week (including weekends) and at different times of the day. At least two weeks needed to pass between the first and the final contact attempt.

Although the responsibility for call scheduling lies with the interviewers, the process was made as transparent as possible. In most countries call attempts were recorded in CAPI and available to the fieldwork coordinators almost in real-time. Where available, (only) registers were used for sampling. This was the case in 14 countries. In two countries the register was supplemented with address enumeration in some regions. In the other 19 countries all addresses had to be enumerated using a random walk approach. The enumeration of addresses was separated from the interview stage. So regardless of whether the sample was based on a register or on enumeration, interviewers would be given lists of addresses based on which they could schedule their visits.

In principle, interviewers were not allowed to carry out more than 40 interviews. In some cases this requirement had to be relaxed.

[2] The 28 EU Member States, Norway, Switzerland, Albania, the Former Yugoslav Republic of Macedonia (FYROM), Montenegro, Serbia and Turkey.

Figure 3: Final outcome by number of contacts for countries contacted face-to-face (left) and by telephone (right)



Contact data

Detailed contact outcomes were collected for each contact attempt. Figure 2 shows the non-contacts, refusals, interviews and ineligibles by country. Figure 3 shows the final fieldwork outcome by number of contact attempts, differentiating between countries where contact was made face-o-face and by telephone. For both contact strategies, the bulk of contacts were made with the first contact attempt. For face-to-face contacting, contacts seem to be more concentrated at the first contact attempt then for the countries that use phone contacting.

For the face-to-face contacting countries, cooperation seems to be more likely than refusal for contact that required fewer contact attempts (figure 4). The more contact attempts required, the more likely a refusal has occurred. This seems to indicate the importance of multiple contact attempts. For telephone countries, however, cooperation is very unlikely at the first contact attempt and becomes more likely for the second and third contact attempts, but decreases afterwards. The first contact attempt is strongly overrepresented by Finland.

To what extent is the number of contact attempts related to the post-stratification weights? A possible hypothesis is that underrepresented respondents require more contact attempts. Figure 5, however, does not confirm this hypothesis: the mean of the post-stratification weight is hardly related to the number of contact attempts required to contact the respondent. The mean seems to be decreasing for the face-to-face contacting data, but note that the minimum required number of contract attempts for that contract strategy was 4 (it was 10 for telephone countries), so these might be particular cases not representative for the overall face-to-face contacting.

Figure 4: Share of cooperation (cooperation / refusal) by the number of contact attempts for countries contacted face-to-face (left) and by telephone (right)

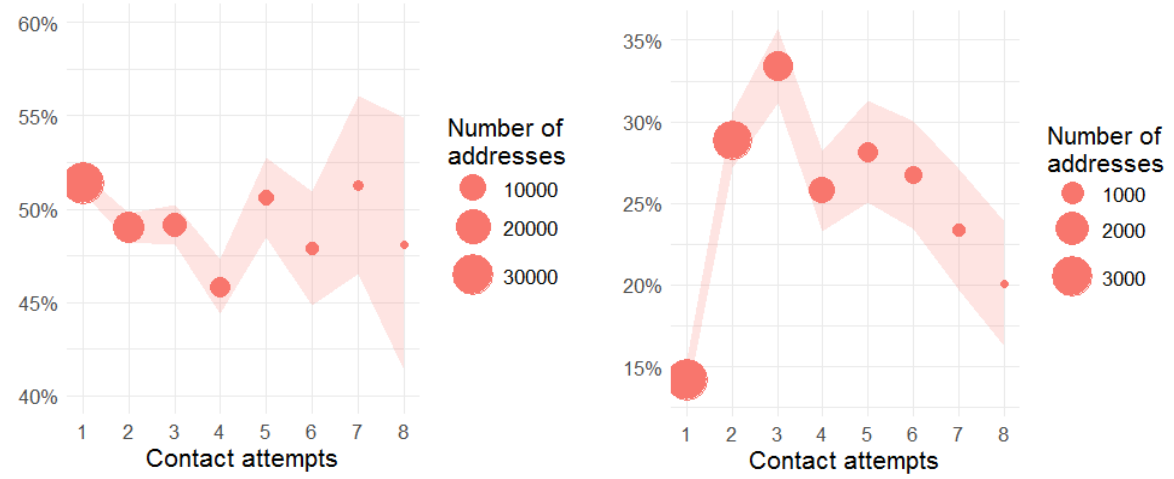


Figure 5: Mean of the post-stratification weight by the number of contact attempts required. Left: face-to-face contacting, right: telephone contacting

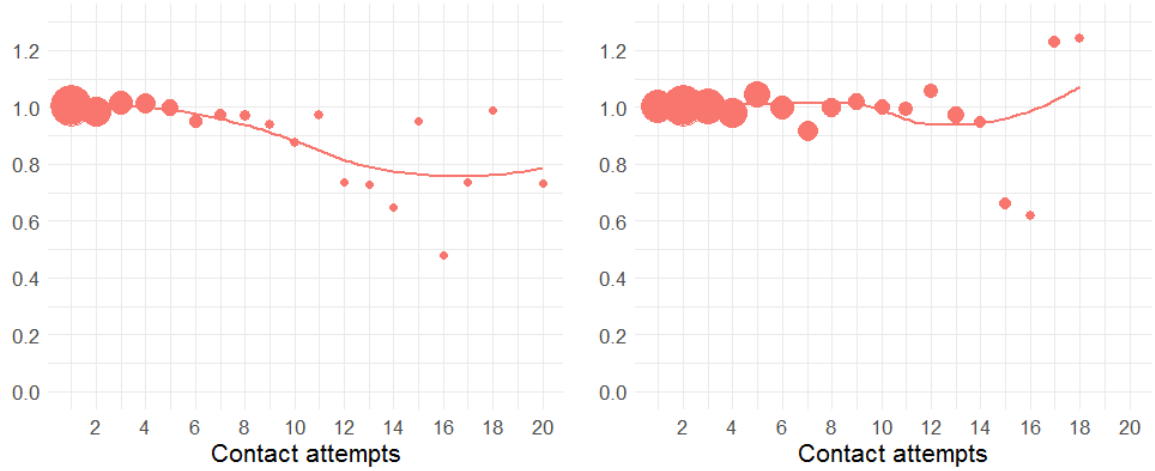
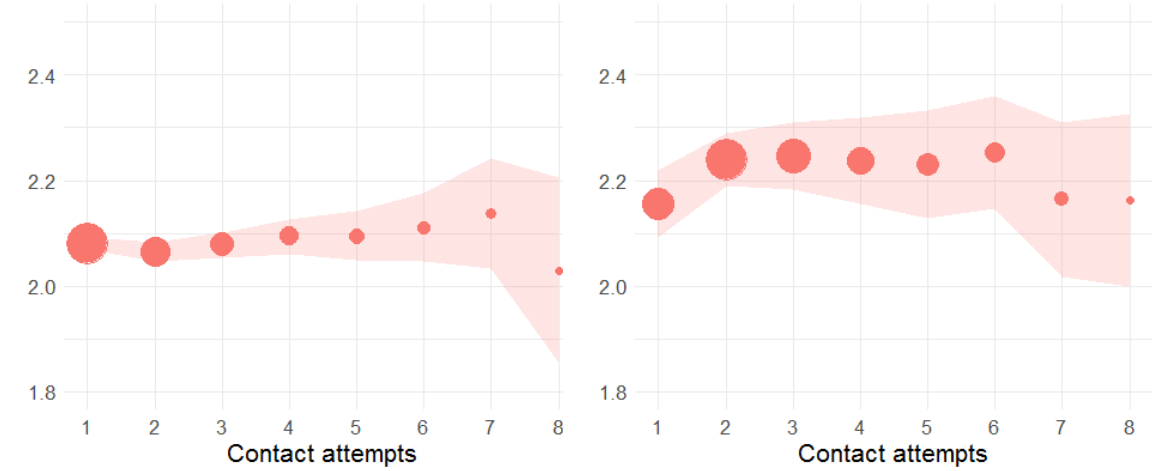


Figure 6: Work-life balance by number of contact attempts required. Scaled from 0-3, showing means and confidence intervals. Left: face-to-face contacting, right: telephone contacting



A key substantive variable in the EWCS is a question on work-life balance. The question is: *In general, how do your working hours fit in with your family or social commitments outside work?* The response categories are 'Very well', 'Well', 'Not very well' and 'Not at all well'. For the purposes of this analysis, these answers have been recoded into a scale from 0-3 where 0 equals 'Not at all well' and 3 equals 'Very well'. For both the face-to-face contacting and telephone contacting countries, the mean of this variable seems to be unrelated to the number of contact attempts required to be able to establish contact with a respondent (Figure 6). Note that the first contact attempt for telephone countries is mostly driven by Finland, so the difference between one required contact attempt and more than one contact attempt required is likely to be the result of country differences in work-life balance.

Breaking down by country (not shown) reveals that the patterns differ by country. For some countries, work-life balance is unaffected by the number of contact attempts and in some countries the pattern slightly goes up, while in others it slightly goes down.

Several substantive variables other than work-life balance show the same lack of association with the number of contact attempts required.

Predicting contact

At each contact attempt, paradata is recorded by the interviewer. These include the time of contact, the type of dwelling and the degree of urbanisation. In addition, country, region and interviewer ID are known. Using a multilevel cox proportional hazards model, we are analysing if these variables are related to the probability of being contacted by estimating the effect on the

odds of being contacted at a given contact attempt. Because we also include the variables about the time in between contact attempts, the first contact attempt is effectively excluded from this analysis.

The results in table 1 show that contact is less likely in the weekend and more likely at the edges of the day. The time in between contact attempts positively affects the odds of being contacted. Contacts are more likely in rural areas and also for farms in comparison to other dwelling types. Region level variables on educational levels, income, labour force participation and working hours are not related to the odds of being contacted.

Table 1: Results multilevel cox proportional hazards model with random intercepts for country and region.

Weekend	-0.1365	**
Another day	0.0344	
Time: 8 - 10 (ref)		
Time: 10 - 12	-0.0724	*
Time: 13 - 15	-0.1158	**
Time: 16 - 18	-0.1303	**
Time: 19 - 20	-0.0851	*
Time: 20 - 22	-0.0338	
Time since last attempt	0.0003	**
Time since last attempt squared	0	
Interviewer workload	-0.0004	**
Rural (ref)		
Intermediate	-0.054	**
Urban	-0.1005	**
Other dwelling (ref)		
Multi unit dwelling	0.0817	*
Terraced dwelling	0.2487	**
Semi-detached dwelling	0.2954	**
Detached dwelling	0.2771	**
Farm	0.5803	**
Share of low educated in region	0.01	
Share of middle educated in region	0.007	
Share of high educated in region (ref)		
Regional income	0.003	
Regional labour force participation	-0.004	
Regional average working hours	0.014	
** p<0.001		
* p<0.05		

Predicting cooperation

Contacted eligible individuals were asked to participate in the survey. For those, cooperation or refusal was recorded in the contact data. Using a multilevel logit model, we estimate the probability of cooperating for the individuals that were contacted. The results (table 2) show patterns similar to the results of the model that estimates contact probability.

Table 2: Results multilevel logit with random intercepts for country, region and interviewer ID.

Weekend	-0.2102	**
Time: 8 - 10 (ref)		
Time: 10 - 12	-0.3385	**
Time: 13 - 15	-0.1655	**
Time: 16 - 18	-0.2591	**
Time: 19 - 20	-0.1363	**
Time: 20 - 22	0.4383	**
Rural (ref)		
Intermediate	-0.2699	**
Urban	-0.3892	**
Other dwelling (ref)		
Multi unit dwelling	-0.0263	
Terraced dwelling	0.1361	
Semi-detached dwelling	-0.0196	
Detached dwelling	0.019	
Farm	0.3957	**
Share of low educated in region	0.0022	
Share of middle educated in region	0.0108	
Share of high educated in region (ref)		
Regional income level	-0.0072	*
Regional labour force participation	0.0101	
Regional average working hours	0.0666	**
** p<0.001		
* p<0.05		

The predictions from both models can be related to substantive survey variables. However, the predicted contact and cooperation probabilities do not seem to be significantly related to key substantive variables in the survey (e.g. work-life balance).

Discussion

This paper shows little evidence for any association of the number of contact attempts, contact probability and cooperation probability with substantive variables in the survey. A naïve conclusion would be that non-response does not bias substantive outcome variables in the EWCS.

However, two problems may inhibit this analysis:

1. The share of final outcomes on the first contact attempt is very high. This may indicate that the recording of the contact data is not accurate and that there may have been

unrecorded contact attempts. This seems to be supported by the contact data from the countries that used telephone contacting. For those countries, there are more final outcomes after the first attempt (Figure 3).

2. The number of variables that can be used to predict contact and cooperation and is available for both respondents and non-respondents is limited. Although most of the variables currently available do show significant associations with contact and cooperation, the overall fit of the models is poor.