

Estimation of Nonresponse Bias in the European Social Survey: Using Information from Reluctant Respondents

Jaak Billiet¹, Michel Philippens², Rory Fitzgerald³, and Ineke Stoop⁴

Central to the methodological quality of the first round of the European Social Survey (ESS) was the principle of equivalence in cross-national measurement. The survey was therefore designed with equivalence as its driving force and included features such as the requirement for random probability samples, effective sample sizes, clear specifications for fieldwork institutes, clear rules for interviewers about the mode, number and timing of contact attempts with all sample units and the documentation of all contact attempts using standardised forms. The use of standardised, detailed contact forms has enabled equivalent cross-national comparisons of nonresponse as well as providing some indication of the potential bias in survey estimates. This article seeks to uncover “traces of bias,” that occur as a consequence of nonresponse, by comparing cooperative and more reluctant respondents on different substantive survey estimates. The analytical framework of the article is based upon the assumption that the attitudes of nonrespondents are more like those of reluctant than cooperative respondents. This is tested by analysing the Round 1 contact form data along with substantive data from the main survey questionnaire.

The data were analysed to answer the question whether nonresponse bias was likely to be affecting the parameters in substantive explanatory models for various attitudinal variables. The findings are mixed in this respect. By classifying respondents into “type of respondent,” based upon how easily they agreed to participate in the survey the potential effect of nonresponse on survey estimates was examined. In some of the explanatory models the effect of the “type of respondent” disappeared, but this was not always the case. However, the remaining (significant) effects are small, and do not have serious implications for the parameters of the explanatory variables. This rather optimistic view must be treated with some caution because of the possibility that real “hard” refusals that were not converted may differ from the converted refusals. More research across a range of surveys is needed about this issue.

Key words: Data quality assessment; cross-national surveys; measurement error; European Social Survey.

¹ Katholieke Universiteit Leuven, Department of Sociology, Parkstraat 45, B-3601 Leuven, Belgium. Email: jaak.billiet@soc.kuleuven.ac.be

² SAS Belgium

³ National Centre for Social Research, 35 Northampton Square, London EC1V 0AX, UK. Email: r.fitzgerald@natcen.ac.uk

⁴ I&A – Data Services and IT, PO Box 16164, NL 2500 BD The Hague, The Netherlands. Email: stoop@scp.nl

Acknowledgments: Jaak Billiet is professor in social methodology, Department of Sociology, Katholieke Universiteit Leuven. Michel Philippens was a research assistant at the Department of Sociology, Katholieke Universiteit Leuven, and is currently at SAS Belgium. Rory Fitzgerald is Senior Research Fellow at the Centre for Comparative Social Surveys, City University London. Ineke Stoop is manager of Data Services and IT at the Social and Cultural Planning Office, Den Haag. Contact: Email: Jaak.Billiet@soc.kuleuven.ac.be. The authors wish to thank the anonymous reviewers for their suggestions and comments.

1. Introduction

Nonparticipation in surveys is a major concern of survey researchers all over the world. Cross-national surveys, cross-cultural outcomes and comparisons are all becoming more and more important to understand social trends and to underpin the policies of international organisations. It is therefore of great importance to pay attention to the effects of nonresponse bias in comparative survey research. In a review of the literature pertaining to nonresponse in cross-national surveys, Couper and De Leeuw (2003) comment: “only if we know how data quality is affected by nonresponse in each country or culture can we assess and improve the comparability of international and cross-cultural data” (Couper and De Leeuw 2003, p.157). Despite their obvious importance, nonresponse issues are often ignored in many cross-national surveys. In fact, the strict standards that are applied to the evaluation of national surveys are often suspended when it comes to cross-national studies (Jowell 1998). The ESS aimed to start addressing this problem by focusing not only on the reduction of nonresponse but also on the collection of information to allow analysis of nonresponse for each selected sampling unit.

The concern for nonresponse issues is a key priority for the European Social Survey. From the outset, it was clearly stated that the ESS should become a high-quality research instrument for the social sciences in Europe.⁵ In addition, it was envisaged that it should act as a catalyst for improving social science methodology more generally. In order to achieve these goals, significant attention was paid to the methodological quality of the survey. The aim of a high-quality survey has been realised in a number of steps during the preparation and execution of the survey. Protocols and documents describing all of the methodological quality measures are documented on the ESS home page (www.europeansocialsurvey.org) and on the website of the data archive (<http://ess.nsd.uib.no/>). There were a number of key elements in the design of the survey to maximise methodological excellence. They included:

- Pilot studies in two countries in order to evaluate and improve the questions.
- Methodological experiments based on Multitrait-Multimethod methodology in order to evaluate the reliability and validity of different types of questions (Saris and Münnich 1995; Scherpenzeel and Saris 1997).
- A specification for participating countries providing details of the fieldwork and national coordination requirements of the ESS.
- Detailed protocols for translation procedures and annotated questionnaires.
- The requirement of random probability samples, equal effective sample sizes and a ban on all forms of substitution.
- Clear guidelines for fieldwork institutes.
- Documentation of fieldwork implementation in “National Technical Summaries.”

⁵ The design of the European Social Survey began in 1995, led first by an expert group and then between 1999 and 2001 by a Steering Committee and a Methodological Committee that were set up and supervised by the Standing Committee of the Social Sciences of the European Science Foundation. See: The European Social Survey (ESS) – a research instrument for the social sciences in Europe. Report prepared for the Standing Committee for the Social Sciences (SCSS) of the European Science Foundation (ESF). Strasbourg: ESF. June 1999.

- Uniform contact forms to assess the quality of fieldwork and to facilitate the evaluation of nonresponse.
- Easy access to complete ESS documentation and publication of all errors that occurred.

The first round got under way in 2001 with the title “Monitoring attitude change in Europe.” Fieldwork was conducted in most of the 22 participating countries between September 2002 and mid 2003. By the end of 2003, the data from 21 of the 22 countries was available (Jowell et al. 2007). This article deals with a central component of data quality in Round 1 of the ESS, namely survey participation using information collected in the ESS contact forms, i.e., call record data. Since both the “*process evaluation*” and “*outcome evaluation*” of call record data are equally important, an approach was applied to the ESS covering both aspects. This means that the evaluation of data quality not only needs to deal with the obtained results of a survey (e.g., response rates, comparability of distributions with known distributions in the population, amount of item nonresponse), but also with each step in the process of data collection, and with each attempt to contact a selected sampling unit (see Loosveldt, Carton, and Billiet 2004).

The most important consequence of nonresponse in surveys is the risk that survey estimates will be biased. Bias due to nonresponse is a function of the response rate and the differences (in attitudes or behaviour) between respondents and nonrespondents. This means that conclusions will only be biased if nonresponse is selective and if respondents differ from nonrespondents on survey estimates. To illustrate this, consider the following formula for the respondent mean (mean based on the answers of respondents):

$$\bar{y}_r = \bar{y}_n + \left(\frac{m}{n}\right)[\bar{y}_r - \bar{y}_m] \quad (1)$$

where (m/n) is the proportion of sample members that do not participate in the survey, i.e., the nonresponse rate, and \bar{y}_r , \bar{y}_m , \bar{y}_n are respectively the respondent, nonrespondent and sample means. In the ESS we can expect substantial differences between respondents and nonrespondents. Many of the variables covered in the first round of the ESS (social participation, political interest and involvement, civic duties) have previously been found to be, or are believed to be, highly correlated with survey participation (Voogt and Saris 2003; Groves and Couper 1998). Accordingly, we expect that nonparticipation in the ESS will be highly likely to cause biased estimates and limit the generalisability to each national population.

Comparative cross-national research, such as the ESS, is not mainly focused on simple descriptive statistics such as the country-mean but on the analysis of *differences* between country-means (or proportions). The following formula illustrates the effects of nonresponse on survey estimates of the difference of two country means:

$$\bar{y}_{1r} - \bar{y}_{2r} = (\bar{y}_{1n} - \bar{y}_{2n}) + \left(\frac{m_1}{n_1}[\bar{y}_{1r} - \bar{y}_{1m}] - \frac{m_2}{n_2}[\bar{y}_{2r} - \bar{y}_{2m}]\right) \quad (2)$$

The subscripts 1 and 2 indicate Country 1 and Country 2. This formula shows how country differences in respondent means will not be biased if the respondent means are affected by the same amount of nonresponse bias in each of the countries (Groves and Couper 1998; Couper and De Leeuw 2003). When it comes to two countries with equal nonresponse

rates, there is no bias if and only if nonrespondents differ from respondents in the same way in both countries. If the nonresponse rates differ substantially across countries, estimates will also not be biased if there are no differences between respondents and nonrespondents in both countries. However, if there are differences between respondents and nonrespondents, bias might be even bigger (see Formula 2) when there are differences in the amount of nonresponse. Critically, however, whether the differences between respondents and nonrespondents are equal across countries is hard to assess because most of the time no information is available about the distribution of survey estimates for nonrespondents.

According to Couper and De Leeuw (2003), the composition of nonresponse may give an important indication of country differences in the “nonresponse mechanism” – i.e., the differences between respondents and nonrespondents. The different kinds of nonresponse, *refusals* and *noncontacts*, have a different effect on the distribution of survey estimates. Hence, if the nonresponse group in one society is made up entirely of noncontacts, while in another it is mostly refusals, the likelihood of nonresponse bias in the estimates of differences between countries’ mean scores may be greater than if the nonresponse mechanism is similar across countries. Despite all efforts in ESS Round 1, we have observed large differences in both refusal and noncontact rates, and we may therefore expect marked bias when making comparisons concerning the kind of variables mentioned.

In this article we will focus on the obtained response rates (outcome) and evaluate the possibility of estimating the direction of nonresponse bias in some substantive variables using information about response conversion efforts. This study on nonresponse is thus restricted to refusals and does not include respondents who were not contacted despite having been subject, at least in theory, to at least four contact attempts (for this aspect see Billiet, Koch, and Philippens 2007).

Voogt argues that the distinction between cooperative and reluctant respondents is a promising tool when seeking to estimate nonresponse bias (Voogt 2004, pp. 100–102). On a different survey than the ESS he reapproached a sample of respondents who twice refused to participate (final refusals) by phone and asked them some key questions. This research found that their answers were in the direction that was expected (Voogt 2004, pp. 92–94). In other words, the responses from these respondents were similar to the ones you would expect from the type of respondents who typically refuse to take part in surveys. Only to a limited extent is a similar pattern identified in this study on data from ESS Round 1. The analysis of the reluctant respondents is limited to five countries because in the other 17 country samples the number of converted respondents was too small to enable reliably taking advantage of the distinction between cooperative and reluctant respondents. Before focusing on the analysis of the reluctant respondents, we will assess the nonresponse reduction strategies, and the outcomes of these strategies in all first round countries.

2. Nonresponse Reduction Strategies in ESS Round 1

The methodological standards for the ESS were devised by a team consisting of survey methodologists from a wide range of European countries. A key component of the ESS is that each participating nation should follow these high methodological standards, adapting

them where necessary but always taking care not to compromise equivalence. A key characteristic of the ESS is that the methodological standards set are those of the countries that normally achieve the highest quality rather than being set at the lowest common denominator (Lynn 2003). The protocols and guidelines made available to “National Coordinators” (NCs) for the purpose of improving response rates provide, among other things, clear rules regarding the mode of data collection, the number and time of the visits to selected sample units, the requirements of the sample, the minimum and maximum duration of the fieldwork period and the maximum workload of the interviewers. Alongside these standards, challenging targets were set (see the fieldwork instructions on www.europeansocialsurvey.org).

Countries were required to achieve a *minimum sample size* of 2,000 realised interviews and an *effective sample size* of 1,500. In countries with a population of less than two million a sample size of 1,000 completed interviews was deemed to be sufficient. In Round 1, this applied only to Luxembourg and Slovenia. “Effective sample sizes” were used to make random samples equivalent even when they were characterised by differences in design effects. Simple random sampling served as a useful benchmark against which to compare other sample designs (Biemer and Lyberg 2003, pp. 347–349). The larger the design effect of a particularly complex sample, the larger the size required in order to obtain an effective sample size of 1,500 (the benchmark) given an estimation of the design effect for that sample.

With respect to nonresponse, countries were asked to aim for a response rate of at least 70%. Although it was clear from the beginning that achieving a 70% response rate was extremely unlikely in some countries (e.g., Switzerland) and would be a challenge for others (e.g., the Netherlands), the aim was to motivate countries to obtain better results than they would normally achieve on similar national surveys. To help countries reach this target response rate they were encouraged to implement a series of measures associated with current “best practice.” These included: the selection of experienced interviewers where possible; boosting interviewers’ confidence in their abilities; personal briefing sessions for all interviewers (including a session on doorstep introductions and persuasion skills); the reissuing of all “easy-to-convert” refusals and as many “hard” refusals as possible.

It is important to remember, however, that setting standards and challenging targets will not always guarantee that these will be met (Park and Jowell 1997). It is therefore necessary to introduce a further stage of monitoring, evaluation and feedback. It is important that compliance with the specified standards is monitored in an objective way and that the results are reported publicly. Ideally, the results of monitoring should feed back into the survey process, producing actions designed to result in the improvement of standards (Lynn 2003). To accommodate monitoring and evaluation of fieldwork processes, the “Central Coordinating Team” (CCT) of the ESS decided to carefully document nonresponse. All NCs were asked to fill in a “small-scale” National Technical Summary concerning some important aspects of fieldwork. As a result, the length of the fieldwork period, number of required visits, the use of special refusal conversion strategies, response rates and other aspects of fieldwork are documented at an aggregate level per country.

As outlined earlier, the ESS goes even further by also collecting information about each individual sample unit. Developing uniform “contact forms” with the aim of documenting

each attempt to make contact with all sample units and attempt an interview was a complex task. And this task was made considerably more challenging in the cross-national context. The first step was to make an inventory of contact forms used by several European survey organisations. In turn several versions of contact forms to account for different sample frames and selection procedures had to be developed and a balance between the fieldwork burden associated with registering and keying in contact data and methodological data needed to be found (Stoop et al. 2003; Devacht et al. 2003). In the end, these efforts resulted in a standardised contact form specification and the construction of a standardised data file comprising information on:

- Day, date, month and hour of visit.
- Nonresponse outcome of each visit.
- Neighbourhood characteristics of each sample unit.
- Reason for refusal, estimated age and gender for each refusal.
- Information on selection procedure.
- Mode of visit (telephone vs face-to-face).
- Interviewer identification.

Unfortunately not all countries delivered a complete call record dataset. The reasons for this varied: survey agencies not familiar with the collection of call record data found the burden too heavy whilst others could not deliver some of the information because of confidentiality laws in their countries (Stoop et al. 2003; Devacht et al. 2003). In the end, comparable and complete call record datasets were available for seventeen countries: Austria (AT), Belgium (BE), Germany (DE), Finland (FI), Great Britain (GB), Greece (GR), Hungary (HU), Ireland (IE), Israel (IL), Italy (IT), Luxemburg (LU), Poland (PL), Portugal (PT), Spain (ES), Switzerland (CH), the Netherlands (NL), and Slovenia (SI). The call record datasets for Norway (NO) and Denmark (DK) do not contain all variables but it is possible to compute the final outcomes. Three samples are more problematic for our aim: a large number of sampling units are missing in the call record dataset of the Czech Republic (CZ), the Swedish (SE) dataset holds only information about the final outcome of the contacts and no call record data set is available for France (FR). Despite the practical problems, and despite occasional inaccuracies and errors, we have no reason to doubt the veracity of the call record data.

We will focus now on the analysis of these rather unique call record data, and in particular on the information about the response and contact rates. Section 3 deals with the achieved response and nonresponse rates. These results are discussed in the context of the fieldwork efforts to see if they explain the big differences in response rates between countries. The refusal conversion efforts and their usefulness for obtaining information about nonresponse bias will be examined in Sections 4 and 5.

3. Achieved Response and Nonresponse Rates in ESS Round 1

The call record data offer the advantage that the same nonresponse outcome definitions and nonresponse rate formulae can be used across countries, thereby enabling equivalent nonresponse comparisons. The call record data were checked for internal consistency, and then compared with data in the main datasets in order to find out whether each record

coded as “interview completed” had its match with a record in the interview datasets. If necessary, NCs were consulted when inconsistencies occurred. As mentioned above, not all countries delivered a dataset containing the necessary information. For most countries for which we have comparable call record data, the response rates computed by means of these data are very close to the figures that were provided by the national survey organisations (see Table 1). We may therefore assume that we obtained reliable estimates of the response and nonresponse rates. For countries with no suitable call record data, we report response rates that were calculated by the national survey organisations and which were reported in the “National Technical Summary.” Comparisons with these countries should therefore be treated with caution.

Before turning to the actual response rates we need to discuss the definitions and rules that were used to calculate nonresponse rates based on the call record data. First we need to

Table 1. Achieved response rates according to National Technical Summary (NTS), and response rates, refusal and noncontact rates according to call record data (percentages)

Country	Response rate		Noncontact rate	Refusal rate	Eligible sample size	Total sample size
	NTS	Call record data				
GR	80.0	79.6	1.7	16.9	3,222	3,227
FI	73.2	73.3	1.4	20.9	2,728	2,766
HU	69.9	70.3	3.2	15.1	2,398	2,484
IL	71.0	70.9	3.0	21.3	3,523	3,600
PL	73.2	72.2	0.8	19.6	2,921	2,978
SI	70.5	71.8	2.4	15.3	2,114	2,222
SE ¹	69.0	—	4.0	21.0	2,878	3,000
PT	68.8	68.8	3.2	26.9	2,196	2,366
DK	67.6	68.4	4.6	23.0	2,143	2,243
NL	67.9	67.8	2.5	26.2	3,486	3,570
NO	65.0	65.0	3.0	25.0	3,109	3,215
IE	64.5	64.4	8.1	22.9	3,179	3,185
AU	60.4	60.6	10.1	27.0	3,725	3,828
BE	59.3	59.3	4.5	25.6	3,204	3,340
GB	55.5	55.0	3.5	30.6	3,730	4,013
DE	55.7	53.7	5.9	29.3	5,436	5,796
ES	53.2	53.6	7.9	35.3	3,227	3,657
IT	43.7	43.4	2.8	45.8	2,778	3,000
LU	43.9	43.2	6.9	37.0	3,589	3,773
FR ²	43.1	—	16.2	36.5	3,159	3,748
CZ ³	43.0	—	11.9	20.6	3,051	3,330
CH ⁴	33.5	33.0	2.0	55.1	4,652	5,086

¹ No call record data available for Sweden.

² Figures based on reported final outcome variable in call record dataset for France.

³ A large number of sampling units (481) are missing in the call record dataset of the Czech Republic; outcome based on information about 2,849 units.

⁴ For Switzerland, two approaches were followed. The first included face-to-face recruitment and the second telephone recruitment. In this article we only report on the telephone part of the survey, since the contact form data for the face-to-face part were not suitable for analysis.

examine the construction of an *overall nonresponse disposition* for each sample unit. Since the call record dataset did not contain a variable that expressed the “final nonresponse disposition code” of each sample unit, we had to merge or combine the separate outcomes for the different visits into one final code. Essentially there are two methods to do this: (1) the outcome of the last contact (with any member of the household) can be considered as the final nonresponse code (see AAPOR 2004) or (2) a priority system of visit outcomes can be constructed to select the outcome with the highest priority (see Lynn et al. 2001). For example, a refusal code that comes earlier in a sequence of visits has priority over a noncontact code on the occasion of a subsequent and final visit. We chose to use a combination of both methods, and the outcome of the last contact was used as the final nonresponse code. The exception was when a refusal occurred at an earlier visit and subsequent contacts with the household resulted in other eligible nonresponse outcomes. In this case, the final nonresponse code was “refusal to participate” (Philippens and Billiet 2004). When a nonresponse code has been followed by an interview because of successful nonresponse conversion, the final outcome is of course a response code because it has higher priority in the coding procedure.

With respect to the *definition of outcome codes*, the “refusal” code includes proxy, household/respondent refusals, broken appointments, respondent at home but did not answer the door and interviews that were broken off before being completed. Noncontacts are defined as those addresses or households for which no contact with anyone was made at any visit. Respondents that moved within the country and were not reapproached were excluded from the noncontact category to enhance comparability between household and individual-named samples on the one hand and address samples on the other. The response, refusal, and noncontact rates are reported in Table 1 and are all expressed as percentages of the total eligible sample. Ineligibles comprise addresses or households that are not residential, not occupied, other ineligibility, respondent deceased and respondent moved abroad.

The figures in Table 1 illustrate that about half of the participating countries obtained response rates close to or higher than the specified target rate of 70%. On the other hand, we do observe rather large differences with respect to nonresponse rates. Some countries (Slovenia, Poland, Hungary, Israel, Greece, and Finland) achieved response rates higher than 70% while others achieved response rates lower than 50% (Italy, Luxemburg, France, Czech Republic, and Switzerland). These large differences obviously raise questions with respect to the validity of cross-national comparisons. The decomposition of nonresponse seems rather similar across countries. For all countries except one (Czech Republic), refusal to participate is the most important reason for nonresponse. The aim to keep noncontact rates to a strict minimum of 3% or less was achieved in most countries. In general, noncontact rates are lower than 5%, with the exception of Germany (6%), Ireland (8%), Spain (8%), Austria (10%), the Czech Republic (12%), and France (16%). The data suggest that in future rounds of the ESS these countries might lower their nonresponse rate by further increasing contact rates. The sampling units that are not classified in one of the three mentioned categories (response, noncontact, and refusal) belong to an additional category, “other.” This category contains between 1.1% and 9.9% of the sampling units in 15 countries, but it is between 10% and 12.9% in six other countries. One country even counts 24.5% sampling units in the category “other.” The size

of this category depends largely on characteristics of the sample, and on decisions taken at the data checking stage. It contains sampling units for which no call record data are found (e.g., Czech Republic), or completed interviews that were not accepted as valid interviews by the NCs (e.g., Belgium).

There is one particularly remarkable observation. The well-documented problem of nonresponse in the Netherlands is not replicated here (see e.g., De Heer 1999; Hox and De Leeuw 2002; Stoop and Philippens 2004). In the ESS, the response rate achieved in the Netherlands is close to the specified target rate of 70%. In this article we will demonstrate how this result was achieved through the implementation of a resource-intensive refusal conversion strategy.

There are many factors that may be responsible for the observed differences in response rates. We can distinguish between factors that can be influenced by the researcher and those that are fixed and cannot be manipulated. Amongst the latter we can distinguish between the survey-climate and the at home patterns. The survey-climate refers to the survey practice and the general extent to which people consider survey research and survey interviews to be useful and legitimate (Groves and Couper 1998). Survey-climate might strongly influence survey cooperation and refusal rates. At home patterns refer to the patterns of time use that may influence the number of hours that people spend at home. These at home patterns influence the ability to contact households and the amount of effort needed to bring down noncontact rates. Given the large demographic differences between countries (e.g., with respect to birth rates, women working, outdoor life-style), these at home patterns are likely to differ greatly across countries (see De Heer and De Leeuw 2002). Survey-climate and at home patterns are interesting and important because they may inform both researchers and fieldwork institutes how best to organise the fieldwork. Most interesting, however, are those factors that are, at least in principle, under the control of the researcher. According to De Heer (1999 pp. 136–137) these factors can be divided into three groups:

- (a) *General design factors*: e.g., mode of data collection, survey method (panel vs cross-sectional) and observational unit (household vs. individual).
- (b) *Fieldwork efforts*: the number of contact attempts, refusal conversion efforts, interviewer and respondent incentives, and interviewer training.
- (c) *Survey organisation*: e.g., the employment conditions of the interviewers, and the arrangements concerning supervision.

The standards and outlines provided to the NCs are related to each of these factors. However, in the remainder of this article we will focus on one aspect of fieldwork differences between countries, namely the differences in the efforts made towards refusal conversion, and the outcomes of these efforts.

4. Refusal Conversion Efforts in ESS Round 1

In order to estimate nonresponse bias we will use information about reluctant respondents to see how they differ from those who took part more easily. To ensure that the conclusions drawn are accurate a minimum number of reluctant respondents are needed in order to be able to search for traces of nonresponse bias in the data. Five ESS country samples

(AU, CH, DE, GB, and NL) contain sufficient numbers of reluctant respondents and can be used in the analysis of bias in Section 5, but we first offer an overview of the results of the efforts made in all countries that contain information about converted refusals and the obtained outcomes.

4.1. Country Differences in Refusal Conversion Efforts

Survey researchers use many techniques to increase survey participation. One of these is the implementation of a *refusal conversion programme*. Refusal conversion means reapproaching initially reluctant respondents to persuade them to reconsider participating in the survey. Much of the likely success of refusal conversion procedures is attributable to the “easiness” of the initial refusal. Refusals often occur because of temporal circumstances such as bad timing or mood swings, suggesting that the group of consistent die-hard refusers may constitute only a small part of the total group of refusals. The refusal conversion procedures in the ESS allow the examination of differences in the practice and implementation of refusal conversion procedures and the assessment of the effects of refusal conversion on response rates.

The ESS Central Coordinating Team recommended that countries reissue all presumed “easy-to-convert” refusals and as many presumed “hard” refusals as possible to a senior interviewer in order to increase participation. However, different funding levels between countries meant that this issue was governed by guidelines and suggestions rather than strict rules with set targets. As a result, the effort put into refusal conversion varied considerably. Furthermore, given that the specifications of the refusal conversion procedure were rather general we can expect that the refusal conversion practice will diverge to a considerable extent in different countries. In the case of a refusal, the interviewers were asked to estimate how likely future cooperation was in view of later decisions concerning reapproaching the refusals. However, it was not clearly defined centrally on what basis a case should be reissued. In addition, the differences between countries in initial response rates (before refusal conversion) can also be responsible for differences in refusal conversion efforts. In fact, there may be less (or even no) need to implement refusal conversion in countries in which the minimum target response rate (70%) was realised without refusal conversion practices.

We will however see later in this article that information about the converted respondents may be of major importance in order to assess response bias. It should be remembered that reissuing a refusal to another interviewer on the basis of process information collected by a previous interviewer requires significant effort. Whatever refusal conversion practice is implemented, each procedure creates some practical organisational problems. Not all survey organisations will have the means and capacity to organise an effective refusal conversion practice (Loosveldt et al. 2003) and ESS Round 1 involved 23 survey institutes.

Figure 1 shows the decomposition of the percentage of eligible sample units that explicitly refused to participate at least once into a) refusals that were not reapproached, b) refusals that were reapproached but not converted and c) successfully converted refusals. In line with our expectations we observe that refusal conversion efforts vary greatly across countries. In the Netherlands, Great Britain, and Switzerland 88%, 84% and 77% of all

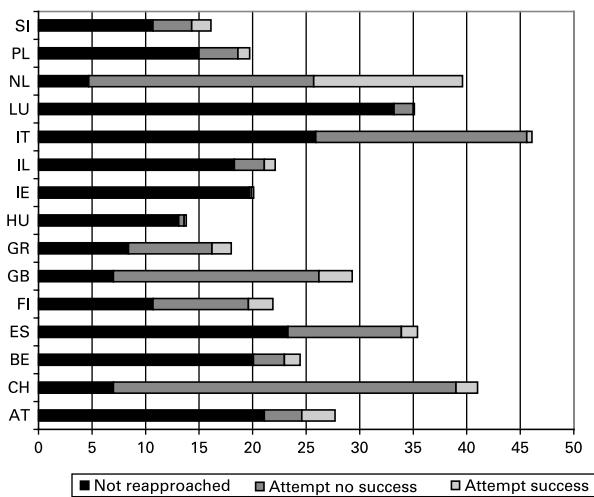


Fig. 1. Decomposition of eligible sample units that refused at least once

refusals were reapproached, respectively. Extensive conversion efforts were also made in Greece, Finland, and Italy, with 54%, 50% and 44% of all refusals being reapproached, respectively. The majority of countries, including Spain, Slovenia, Poland, Belgium, Austria, and Israel, made moderate efforts at conversion, with between 20% and 34% of all refusals being reapproached. Ireland, Hungary, and Luxembourg made hardly any efforts to convert refusals. For Germany and Portugal missing information about some cases in the contact form dataset means we cannot ascertain the percentage of cases that were reissued.

4.2. Effectiveness of Refusal Conversion Strategies

The refusal conversion success rate was highest in Austria followed by the Netherlands, Belgium, and Slovenia. In these countries, respectively, 47%, 39%, 33%, and 32% of all reapproached refusals were successfully “converted.” For Greece, Finland, Poland, and Israel, conversion success rates lie between 20% and 30%. In Great Britain and Spain conversion rates were closer to 15%, while the lowest conversion success rates were achieved in Switzerland and Italy – at around just 5%.

These conversion success rates cannot simply be compared in a straightforward manner across all countries. Some countries focused their efforts on a relatively small and perhaps “easy-to-convert” group of refusals (e.g., Belgium) while others have reissued almost all refusals (e.g., the Netherlands and presumably Germany). Without doubt, high conversion rates will be harder to obtain in the latter case. In this respect the result in the Netherlands is quite remarkable. When we compare conversion success rates among countries that reissued the majority of refusals, i.e., Switzerland, the Netherlands, and Great Britain, we observe that the success rate of the Netherlands (39.7%) is much higher than that of Great Britain (13.8%) and Switzerland (6.1%). Inspection of the “National Technical Summary” reveals that the Dutch survey organisation implemented a range of special refusal conversion strategies. This included a second letter (sent halfway through fieldwork) to

previously-refusing respondents in which they were again asked to participate. It also included financial incentives to reluctant sample members of up to five euros and other monetary prizes (e.g., a quiz). There were also efforts to convert refusals by telephone. It is not unreasonable to assume that this combination of methods contributed greatly to the success of refusal conversion in the Netherlands.

4.3. Effects on Response Rates and Differentiation of the Respondents

In general, the effects of refusal conversion strategies on response rates are rather marginal. For most countries the effects vary between one and three percentage points. For countries such as Switzerland and Great Britain that invested significant efforts, the returns were somewhat disappointing. In the Netherlands, however, an intensive and effective refusal conversion strategy managed to increase response rates by 16 percentage points (from 51.8% to 67.8%). The results in the Netherlands demonstrate that, with significant and targeted efforts, acceptable response rates can be obtained in countries with less favourable survey climates.

On the basis of the information about converted respondents in the contact forms, it is possible to distinguish three kinds of respondents: *cooperative* respondents, reluctant respondents who were *easy-to-convert*, and reluctant respondents who were *hard-to-convert*. Cooperative respondents are those who decided to participate at first contact with an interviewer. Easy-to-convert refusals are originally reluctant respondents who were easy to persuade and who decided to cooperate after one new attempt. Hard-to-convert refusals are reluctant respondents who decided to participate after several attempts or after special incentives had been used in order to persuade them. We have information about 19 countries. Sweden, Norway, and France did not provide the necessary information on the contact forms.

In the Netherlands, three strategies were used in order to persuade the reluctant respondents. About 7,000 Euro was spent to raise the interviewer fee for an interview with a former refusal. Additionally, the same total amount was spent on incentives for the respondents. There was also a telephone conversion exercise by the telephone unit that cost about 6,600 Euros. In sum, the total budget for the response conversion attempts in the Netherlands was at least 20,500 Euros. However, isolating the specific costs for refusal conversion efforts that are not part of standard efforts by a survey agency is very difficult. And in the case of the Netherlands there were additional costs that could not be itemised which were probably associated with the refusal conversion efforts. In Germany, a budget of 25,000 Euros was spent on incentives, and another 11,750 Euros on telephone refusal conversion.

In the next section we study nonresponse bias by analysing the characteristics of respondents in the dataset. Table 2 shows that this kind of analysis is limited by the small number of converted respondents. Differentiation between easy-to-convert and hard-to-convert refusals for which we have complete data is only possible in the Netherlands and Germany where the proportion of converted respondents is over 15%. In three other countries, Great Britain, Switzerland, and Austria, the proportion of converted respondents is over 5% (more than 100 observations), allowing some analysis. Analysis of converted respondents is problematic in all other countries.

Table 2. Distribution of cooperative and reluctant respondents in 19 country samples of ESS Round 1 (horizontal percentages)

Country*	Kind of respondents			Total (100%)
	Cooperative	Easy-to-convert refusals	Hard-to-convert refusals	
AT	94.9	4.7	0.4	2,244
BE	97.6	2.2	0.2	1,895
CH	92.6	3.5	3.9	2,004
CZ	100.0	0.0	0.0	1,352
DE	83.0	10.5	6.6	2,918
DK	95.2	4.9	0.0	1,468
ES	97.2	2.2	0.6	1,727
FI	96.9	3.0	0.2	1,999
GB	94.3	3.6	2.1	2,052
GR	97.7	2.0	0.2	2,564
HU	99.8	0.2	0.0	1,684
IE	99.7	0.0	0.3	2,020
IL	98.5	1.2	0.3	2,493
IT	98.8	1.2	0.0	1,205
LU	99.7	0.3	0.0	1,547
NL	79.6	10.7	9.7	2,361
PL	98.4	1.1	0.5	2,109
PT	96.6	2.4	1.0	1,511
SI	96.4	3.2	0.4	1,518
Total	95.1	3.4	1.6	34,667

*No information about the kind of refusals was available from Sweden, France, and Norway.

5. Usefulness of Refusal Conversion for Estimating Nonresponse Bias

5.1. Expectations, Data, and Methodology

The actual aim of refusal conversion techniques is not only to decrease nonresponse bias in survey estimates, but also to obtain information about nonrespondents. In terms of the effectiveness of refusal conversion in reducing bias and obtaining information about final refusals, there are two main directions of thought (Lin and Schaeffer 1995).

The first direction, the so-called “continuum of resistance model,” assumes that the difficult-to-approach respondents and the reluctant respondents have much in common with the sample units who finally refuse to cooperate. Because of this, persuading reluctant sample units yields a more representative or balanced sample as response rates increase (Stoop 2004; Teitler et al. 2003; Voogt 2004). This viewpoint is supported by evidence in postal surveys in which respondents are distinguished according to the moment of participation resulting in three groups: respondents that participated before or after a first recall (cooperative), those who participated after a second or third recall, and respondents who did not participate but answered some crucial questions by telephone afterwards. Analysis of a crucial question shows that the last group is closer to the second call respondents than to the first group of cooperative respondents (Billiet and Waege 2003, pp. 309–314).

The second direction, the “*classes of nonparticipants model*,” assumes that people do not participate for various reasons. There are thus several classes of nonrespondents. Some classes (e.g., difficult to contact, reluctant respondents) have much in common with the final nonrespondents (e.g., the noncontacted sample units, the final refusals). Other classes of nonrespondents are related to characteristics of the survey (e.g., topic, sponsorship, burden), or to situational characteristics (e.g., no time, illness). The nonrespondents belonging to these classes have little in common with the converted refusals. Under this viewpoint, distributions of survey estimates will remain relatively stable as the pool of respondents increases (Keeter et al. 2000). Research results presented by Curtin et al. (2000), Teitler et al. (2003) and Stoop (2005) confirm the latter viewpoint and suggest that refusal conversion efforts or extended interviewer efforts often have only small effects on cross-sectional survey estimates. This may be due to the fact that the reluctant respondents who are persuaded to participate differ both from the cooperative respondents and from the refusals that are never converted (Stoop 2004, p. 50).

It is impossible to test the hypotheses about differences between cooperative respondents, reluctant respondents and final refusals with information based on refusal conversion. Other approaches are necessary, such as approaching a small sample of firm refusals by telephone with a very small set of crucial questions (Voogt 2004), and even then the problem of nonresponse to this exercise remains. Such an exercise was not possible for ESS Round 1. However, it is possible to indirectly test this hypothesis by using the information about substantive variables that are, according to existing theories, related to nonresponse and by differentiating between easy-to-convert and hard-to-convert refusals in countries that have large enough amounts of such cases.

In a preliminary study Loosveldt, Philippens, Stoop, and Billiet (2003) studied the differences between initially cooperative respondents and converted refusals over a range of 26 survey estimates in seven countries of the ESS. Four types of survey estimates were studied: socio-demographic variables, social integration, political involvement and attitudes towards immigration. From previous studies the researchers expected to find a relationship between these variables and nonresponse. It was expected that respondents who are less integrated in society, participate less in politics, have more individualistic attitudes and have negative views toward immigrants are less likely to participate in surveys (Brehm 1993; Groves and Couper 1998; Loosveldt and Carton 2002a; Voogt and Saris 2003). A number of differences in the expected direction between cooperative and reluctant respondents were found, indicating the presence of nonresponse bias. However, because of low conversion rates, reapproaching refusals did not result in a substantial decrease in bias in most of the countries they observed. The analysis in this article differs in several respects from that conducted by Loosveldt et al. (2003).

Firstly, the analysis is done with five countries that converted at least 100 refusals: Austria ($N = 2,257$; 115 converted), Great Britain ($N = 2,052$; 117 converted), Switzerland ($N = 2,041$; 151 converted), Germany ($N = 2,995$; 497 converted) and the Netherlands ($N = 2,364$; 481 converted).

Secondly, we focus on latent variables (or constructs) that are constructed by means of multiple observed attitudinal indicators, and not on individual items. Latent variables are not measured directly. They are hypothetical constructs invented by the researcher for the purpose of understanding a research area. These constructs are related to each other in

certain ways as specified by the investigator's theory (Bentler 1980). The measurements of latent variables are much more reliable and valid than the particular items on which their measurements are based, and they reflect the intended theoretical concepts much better (Bollen 1989). At the current time most substantial analysis in the social sciences in which attitudes or values play a role, is multivariate analysis in which latent variables or constructs are used rather than individual items. The individual items (opinion statements or questions) are specifically designed as sets of indicators for underlying latent variables and not as isolated entities. Researchers are therefore much more interested in the bias in these constructs than in the bias in particular items and should not presume that bias in specific indicators affected by nonresponse, disappears in constructs based on sets of indicators. The constructs we use are *perceived threat from immigration* (Round 1 ESS questionnaire items D18-D24), *trust in politics* (items B7-B10), *interest in politics* (items B1-B4), *political participation* (items E23, E26 and E27) and *social trust* (A8, A9 and A10). We added one single behavioural variable to this list, *participation in the last national legislative election* (B13), because we have information at the aggregate level on voter turnout to compare it with. This is useful for estimating the overreporting of participation in elections. The literature suggests that we can expect those who do not participate in elections to be more likely to refuse cooperation with surveys (Voogt and Saris 2004).

Thirdly, the data were analysed to answer the question as to whether nonresponse bias was likely to be affecting the parameters in substantive explanatory models for various attitudinal variables. To answer this question, the attitudes were the dependent variables in the regression models rather than the kind of respondent. It would certainly be possible to turn this completely around and to focus on the variables that predict the ratio "reluctant/cooperative," using logistic regression models. However, our way of analysing these data is a better expression of the major concern of researchers: whether the parameters in their explanatory models are biased. We thus wish to ascertain the net effects of the *kind of respondent* (cooperative, easy-to-convert refusal, hard-to-convert refusal) on the dependent latent variables. The effects on these attitudinal variables are studied within multiple regression models in which a number of theoretically relevant predictors are included. The effect on the reported voter turnout is studied in a multiple logistic regression model because the dependent variable in this case is the ratio "participation/nonparticipation" with nonparticipation as the reference category.

The control variables are selected because of their significant effect on the dependent variables identified in several substantive studies on ethnocentrism and on indicators of social integration (see Billiet 1995 and 1998). The social background variables (control variables) in all the models are: *gender*, *level of education* (primary, lower secondary, upper secondary, post-secondary but nontertiary, second stage of tertiary), *age* (centred around the mean), *level of urbanisation* (five points: countryside – urban), *job experience* (ever had a job vs none), *experience of more than three months of unemployment during past five years* (yes/no), *active participation in social organisations*, *total time of TV-watching on average weekday* (A1), *religious involvement* (C13, C14, C15, and E18), and *religious denomination* (Catholic, Reformed, Orthodox, Other Christian, Other religion and None as baseline).

The analysis was carried out in four steps. Firstly we examined five background variables that, according to previous studies (Groves and Couper 1998, p. 32; Loosveldt and Carton 2001; 2002a; 2002b), are related to nonresponse to see whether the percentages

or mean estimates differ for cooperative and reluctant respondents. The variables are: gender, age, education, degree of urbanisation, and participation in voluntary associations. According to these studies, higher nonresponse rates are expected for the lower educated, those who live in metropolitan areas, and those who do not participate in voluntary associations. The expectations are not so obvious for gender and age: older women for example may be much easier to contact but at the same time more reluctant to cooperate.

Secondly, we examined the means of five attitudinal variables and reported participation in the last national election for reluctant and cooperative respondents. It was hypothesised that there would be significantly lower mean scores ($p < 0.05$) among reluctant respondents, except in relation to threat from immigrants where we expected a higher mean score.

Thirdly, regression analysis was performed with a dichotomous type of respondent (reluctant versus cooperative) as a predictor, with the attitude scales as dependent variables. Participation in elections (yes = 1; no = 0) was analysed using logistic regression. The results of simple regressions with the kind of respondent as a factor (Model 1) were then compared with those of multiple regressions (Model 2) in which all relevant background variables are included so that we are controlling for them in the model. We want to find out whether the effect of respondent type (cooperative, reluctant) disappears after inclusion of and controlling for the background variables in the models.

Fourthly, the samples in the Netherlands and Germany provided even richer data. Here easy-to-convert and hard-to-convert refusals were analysed separately. In this fourth step, the regressions were carried out with the three categories of the type of respondent as factor: easy-to-convert versus cooperative and hard-to-convert refusals versus cooperative respondents. We only report the standardised regression parameters and the t -values for this variable because we are only interested in the effects of respondent type on the dependent variables (attitudes and voter turnout). These effects are reported in simple regressions (Model 1) and in multiple regressions (Model 2) with the relevant background variables included.

5.2. Findings: Traces of Bias

5.2.1. Step 1: Differences in Social Background Variables

Table 3 does not confirm our hypothesis since there is no clear, uniform pattern. According to existing theories about nonresponse we expect a larger proportion of respondents with lower levels of *education* among the converted respondents. The differences for Great Britain, Germany and the Netherlands are in the expected direction but are statistically significant in the Netherlands only. In Austria and in Switzerland the differences are actually in the opposite direction. We expected proportionally more respondents living in big *cities* and in the suburbs among the converted refusals, but this was only the case in Germany. The difference (which is significant) is once again in the opposite direction in Austria, with people living in towns and small cities less likely to be among the reluctant respondents. We found no substantive difference in the other countries. The kind of response is independent of *social participation*, except in Austria where the difference between cooperative and reluctant respondents is in the expected direction. The differences in mean *age* are all in the expected direction (higher among the reluctant

Table 3. Differences in education level, urbanisation, participation in social activities, and mean age according to kind of respondent in five countries (column percentages and means)

Background variables	Austria		Great Britain		Switzerland		Germany		The Netherlands	
	Coop.	Reluct.								
Gender										
% male	48.2 $\chi^2_1 = 6.465, p = 0.011$	36.1	47.8 $\chi^2_1 = 6.833, p = 0.009$	59.7	49.4 $\chi^2_1 = 0.415, p = 0.519$	52.2	48.0 $\chi^2_1 = 0.027, p = 0.869$	48.5	46.2 $\chi^2_1 = 0.683, p = 0.409$	44.1
Education										
% Lower and lower secondary	31.0	21.8	54.8	60.7	21.8	13.8	15.8	18.4	42.1	48.6
% Upper and post-secondary	57.2	61.1	22.3	21.5	64.8	74.0	62.7	59.5	35.4	33.7
% Second stage of tertiary	11.9 $\chi^2_2 = 5.85, p = 0.054$	12.1	23.0 $\chi^2_2 = 2.25, p = 0.324$	17.8	14.47 $\chi^2_2 = 6.71, p = 0.035$	12.2	21.5 $\chi^2_2 = 2.42, p = 0.298$	22.1	22.5 $\chi^2_2 = 8.17, p = 0.017$	17.7
Urbanisation										
% Countryside and village	44.6	36.0	26.2	20.6	61.3	60.2	27.6	20.0	46.8	46.2
% Town and small city	20.1	44.5	44.8	46.8	20.2	19.8	39.7	40.5	26.2	25.6
% Big cities and suburbs	35.3 $\chi^2_2 = 40.45, p < 0.000$	19.5	29.1	32.7	18.5	20.0	32.8	39.5	27.0	28.2
Participation in associations										
% No	39.2 $\chi^2_1 = 6.82, p = 0.009$	51.4	42.5 $\chi^2_1 = 1.30, p = 0.524$	47.6	36.7 $\chi^2_1 = 0.62, p = 0.432$	33.3	47.2 $\chi^2_1 = 2.80, p = 0.094$	51.4	40.0 $\chi^2_1 = 1.79, p = 0.181$	43.4
Mean age	43.82 $t = -0.48, p = 0.631$	44.57	46.28 $t = -1.42, p = 0.157$	48.63	45.78 $t = -1.38, p = 0.166$	47.88	45.90 $t = -3.49, p < 0.001$	49.03	45.13 $t = -1.65, p = 0.661$	46.54

($p < 0.05$) = significant difference.

respondents), but are only significant in Germany. Finally, we found significant differences in *gender* distribution only in Austria and Great Britain, and these are in different directions (Table 4).

5.2.2. Step 2: Differences in Attitudinal and Behavioural Variables

So what are the traces of bias in the attitudinal variables and reported participation in the last election? Our hypothesis is that we shall find significantly lower mean scores ($p < 0.05$) among reluctant than among cooperative ones, except in relation to the threat from immigrants, where we expect to find the reverse pattern.

All of the attitudinal variables are continuous, ranging from 0 to 10, whilst participation in the last elections is a dichotomous variable (coded 0, 1). Apart from the *perceived threat from immigrants* there were hardly any differences between cooperative and reluctant respondents on these estimates in Austria, Great Britain, and Switzerland. The only significant difference was found in political participation in Austria but this was in the opposite direction than expected. The proportion of voters that participated in the last elections in three countries was somewhat closer to the figures of actual voter turnout in these countries after inclusion of reluctant respondents, but cooperative and reluctant respondents (converted refusals) did not differ significantly in this respect. Feeling threatened by immigrants (*perceived threat from immigrants*) was the only variable that showed substantial differences in the expected direction in four of the five countries. The converted refusals have higher scores on that variable and we assume that this is also the case among the nonobserved firm refusals. This probably indicates that the perceived threat from immigrants was somewhat underestimated in the ESS because of nonresponse and that the bias is probably larger in countries with lower response levels.

5.2.3. Step 3: Reluctant Versus Cooperative Respondents

The effect parameters (β -coefficients) and *t*-values of the simple regressions do not offer any extra information. Significant differences in mean scores are reflected in a significant decrease or increase in units of standard deviation as a result of belonging to the category of cooperative or reluctant respondents in their attitudes. In addition, we use simple regression parameters for the type of respondent in order to compare these with the effects, after controlling for the relevant social background variables in the regression models.

The analyses are only reported for Germany and the Netherlands because the numbers of reluctant respondents are too small for this kind of analysis in the other countries. All effects are in the expected direction in the sample of Great Britain but they are mostly not significant (at the 0.05 level) because of the small number of reluctant respondents. In Austria and Switzerland, some of the parameters are in the opposite direction but they are not statistically significant. Focusing on the two countries with large numbers of converted refusals made it possible to repeat the analysis for easy and hard-to-convert refusals. In this case, the category of reluctant respondents is split into these two subcategories. The term "traces" was used in the title of this section on findings because we have no direct estimation of the effect of real nonresponse (final refusals and noncontacts) on the attitudinal variables. Rather we base it on the assumption that nonrespondents are similar to reluctant respondents. The observed effects are named "bias" because there is no theoretical explanation of why respondents' answers with regard to these attitudinal

Table 4. Means and differences between means-test of attitudinal variables among cooperative and reluctant respondents in five countries

Variable	Austria		Great Britain		Switzerland		Germany		The Netherlands	
	Coop.	Reluct.	Coop.	Reluct.	Coop.	Reluct.	Coop.	Reluct.	Coop.	Reluct.
Political trust	5.29 <i>t</i> = 0.72, <i>p</i> = 0.47	5.16	4.90 <i>t</i> = 1.36, <i>p</i> = 0.17	4.67	5.93 <i>t</i> = 0.94, <i>p</i> = 0.35	5.80	5.16 <i>t</i> = 3.24, <i>p</i> < 0.01	4.88	5.36 <i>t</i> = 2.08, <i>p</i> = 0.04	5.18
Political participation	5.81 <i>t</i> = -2.63, <i>p</i> < 0.01	6.29	5.28 <i>t</i> = 0.38, <i>p</i> = 0.70	5.21	5.82 <i>t</i> = -0.81, <i>p</i> = 0.42	5.95	5.60 <i>t</i> = 3.20, <i>p</i> < 0.01	5.29	5.89 <i>t</i> = 2.69, <i>p</i> < 0.01	5.66
Political interest	5.14 <i>t</i> = 1.00, <i>p</i> = 0.32	5.25	5.48 <i>t</i> = -0.25, <i>p</i> = 0.80	5.51	5.25 <i>t</i> = 0.02, <i>p</i> = 0.98	5.25	5.39 <i>t</i> = 1.18, <i>p</i> = 0.23	5.32	4.96 <i>t</i> = 0.99, <i>p</i> = 0.32	4.90
Social trust	5.32 <i>t</i> = 0.23, <i>p</i> = 0.82	5.28	5.35 <i>t</i> = 0.65, <i>p</i> = 0.52	5.25	5.73 <i>t</i> = 0.61, <i>p</i> = 0.54	5.64	5.13 <i>t</i> = 2.46, <i>p</i> < 0.05	4.92	5.72 <i>t</i> = 0.03, <i>p</i> = 0.98	5.71
Perceived threat from immigrants	5.04 <i>t</i> = 1.33, <i>p</i> = 0.18	4.81	5.30 <i>t</i> = -3.00, <i>p</i> < 0.01	5.74	5.29 <i>t</i> = -2.04, <i>p</i> = 0.04	5.54	5.57 <i>t</i> = -3.02, <i>p</i> < 0.01	5.81	5.00 <i>t</i> = -3.51, <i>p</i> < 0.001	5.28
Participated last election	0.887 <i>t</i> = 0.32, <i>p</i> = 0.75	0.878	0.720 <i>t</i> = -1.29, <i>p</i> = 0.20	0.775	0.690 <i>t</i> = 0.22, <i>p</i> = 0.82	0.681	0.856 <i>t</i> = 1.15, <i>p</i> = 0.25	0.835	0.861 <i>t</i> = -0.35, <i>p</i> = 0.73	0.868
Population*	0.843		0.594		0.434		0.791		0.791	

* Mean: proportion voter turnout (participation = 1; nonparticipation = 0).

Source: <http://www.idea.int/vt/western.cfm>

variables or their participation in elections should be different depending upon the nature of their cooperation in a survey (cooperative, reluctant, easy-to-convert/hard-to-convert).

In Table 5, the *t*-values of the corresponding effect parameters (β coefficients or odds ratios) for reluctant respondents (easy and hard-to-convert together) are reported. The cooperative respondents are the reference category (baseline). The upper row (Model 1) contains details of the parameters in a simple regression with only the kind of respondent as a predictor, the second row (Model 2) contains the same information in the complete model with all relevant background variables included. All important regression assumptions were tested in the samples of the two countries and no serious violations were observed. The explained variance (R^2) of the complete models is rather low. It varies between 3% for political interest to about 15% for ethnic threat.

We can reasonably assume that the bias among the real nonrespondents is in the same direction as among the reluctant respondents but we have no direct evidence and we do not know how pronounced the real bias is. Several observations can be made. First, for the attitudinal variables, all observed significant effects are in the expected direction and the likelihood of bias in the latent variables varies little between the two countries. Second, the presence and extent of the traces of bias varies between attitudes. As expected, the traces of bias were strongest for perceived threat from immigrants, political participation and trust in politics. There were hardly any traces of bias for political interest and social trust. The same conclusion holds for participation in elections. Contrary to expectations, there was no significant trace of bias observed in that variable. Third, in nearly all situations where significant effects were observed, the bias became smaller after controlling for the relevant background variables. The reason for this is that the nonresponse bias is not completely independent of background variables such as education, urbanisation, and age (see Table 3). However, for attitudinal variables such as political trust and political participation, the bias does not disappear after controlling for the background variables in Germany although it does disappear for perceived threat from immigrants.

In the Netherlands, the effect of kind of respondent on perceived threat from immigrants does not disappear in the complete model once we have controlled for the background variables. This means that we cannot exclude the possibility that the parameters in the explanatory models for ethnic prejudice are somewhat biased because of nonresponse, even after the use of population weights on background variables such as gender, age, and education. However the bias appears to be negligible.

5.3. Step 4: Cooperative Versus Easy- and Hard-to-convert Refusals

In Table 6 a distinction is made between two categories of reluctant respondents, the easy- and hard-to-convert refusals. As was already explained in our comment on Table 2, the distinction between easy- and hard-to-convert refusals is not based on the evaluation of the interviewer in the contact forms but on the actual investment needed in order to convince a respondent who initially refused. The reference category remains "cooperative respondents." As expected, where significant effects were found, they were mostly stronger among hard-to-convert refusals than among easy-to-convert refusals (Table 6), but interestingly enough this was not always the case. The exceptions were political trust and threat from immigration in Germany and political interest in the Netherlands.

Table 5. Standardised estimates (β -coefficients) and t -values of reluctant respondents (versus cooperative respondents) in simple (Model 1) and multiple (Model 2) regressions in Germany and the Netherlands

Explained variable	Kind of respondent	Germany		The Netherlands	
		β	t -value	β	t -value
Political trust	Reluctant (Model 1)	-0.060**	-3.24	-0.043*	-2.08
	(Model 2)	-0.053**	-2.91	-0.004	-0.20
Political participation	ref: cooperative				
	Reluctant (Model 1)	-0.060**	-3.20**	-0.056**	-2.69
Political interest	(Model 2)	-0.047**	-2.56	-0.039	-1.91
	ref: cooperative				
Social trust	Reluctant (Model 1)	-0.002	-1.18	-0.056	-0.99
	(Model 2)	-0.011	-0.59	-0.005	-0.27
Perceived threat from immigrants	ref: cooperative				
	Reluctant (Model 1)	-0.046*	-2.49	-0.002	-0.03
	(Model 2)	-0.042*	-2.22	0.008	0.38
	ref: cooperative				
Participation in elections	Reluctant (Model 1)	0.056**	3.02	0.072***	3.51
	(Model 2)	0.030	1.70	0.047*	2.34
	ref: cooperative				
		Odds ratio	Wald Chi-square	Odds ratio	Wald Chi-square
	Reluctant (Model 1)	1.189	1.509	0.951	0.109
	(Model 2)	1.088	0.150	0.887	0.551
	ref: cooperative				

$p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 6. Standardised estimates (β -coefficients) and *t*-values for easy-to-convert and hard-to-convert refusals (versus cooperative) in simple (Model 1) and multiple (Model 2) regressions in Germany and the Netherlands

Explained variable	Model	Kind of respondent	Germany		The Netherlands	
			β	<i>t</i> -value	β	<i>t</i> -value
Political trust	1	Easy-to-convert	−0.047*	−2.53	0.004	0.21
		Hard-to-convert	−0.042*	−2.26	−0.072***	−3.48
	2	Easy-to-convert	−0.036	−1.94	0.029	1.42
		Hard-to-convert ref: cooperative	−0.046*	−2.47	−0.038	−1.87*
Political participation	1	Easy-to-convert	−0.032	−1.69	−0.009	−0.42
		Hard-to-convert	−0.062***	−3.32	−0.078***	−3.76
	2	Easy-to-convert	−0.019	−1.01	0.001	0.05
		Hard-to-convert ref: cooperative	−0.057**	−3.09	−0.063**	−3.06
Political interest	1	Easy-to-convert	0.009	0.49	−0.046*	−2.21
		Hard-to-convert	−0.050**	−2.68	0.017	0.84
	2	Easy-to-convert	0.012	0.65	−0.036	−1.73
		Hard-to-convert ref: cooperative	−0.035	−1.88	0.031	1.47
Social trust	1	Easy-to-convert	−0.041*	−2.23	0.018	0.88
		Hard-to-convert	−0.025	−1.35	−0.021	−1.03
	2	Easy-to-convert	−0.033	−1.77	0.028	1.33
		Hard-to-convert ref: cooperative	−0.028	−1.50	−0.018	−0.86

Perceived threat from immigrants	1	Easy-to-convert	0.052***	2.80	0.019	0.89
		Hard-to-convert	0.028	1.49	0.093***	4.50
	2	Easy-to-convert	0.032	1.79	-0.003	-0.14
		Hard-to-convert	0.008	0.56	0.071***	3.52
		ref: cooperative				
Participation in elections	1	Easy-to-convert	Odds ratio	Wald Chi-square	Odds ratio	Wald Chi-square
		Hard-to-convert	0.989	0.007	0.948	0.017
	2	Easy-to-convert	1.149	0.873	0.953	0.010
		Hard-to-convert	0.920	0.343	0.885	0.077
		ref: cooperative	1.154	0.769	0.888	0.058

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Many researchers are most focused upon the relationship between two variables in the context of the other variables in multiple regression or path models. It is important to stress that the effects of the kind of respondent (cooperative/reluctant) remain small for all attitudinal variables. Even for the two attitudes that are most affected by the “kind of respondent” (political participation and perceived threat from immigrants), excluding this variable from the models has no effect either on the total explained variance or on the values of the substantive predictors that were included in all models. Remember that these predictors are: gender, level of education, age, level of urbanisation, job experience, past unemployment, active participation in social organisations, total time of TV-watching on an average weekday, religious involvement and religious denomination. The parameters of the complete models are available on demand. Moreover, conclusions about the effects of the substantive predictors on the attitudes remain unchanged in the samples where the reluctant respondents are not included. This means that the effect of response on the substantive relationships between the predictors and dependent variables is very small.

6. Discussion and Conclusion

Our ability to generalise about “traces of bias” is largely narrowed by the lack of data on reluctant respondents in many countries. Apart from the five countries that had between 115 and 497 respondents who were initially reluctant to cooperate, the rest had 71 or fewer respondents converted and some had no more than 15 hard-to-convert refusals who were converted. That is not problematic for countries that obtained response rates of over 70% but it is problematic for the countries that obtained lower response rates. Attempts to increase the response rates substantially by reapproaching and convincing reluctant sample members are vitally important for two reasons. Firstly they increase the sample size and secondly they provide information about the presumed magnitude and direction of the bias in a large number of variables. The quality of the data would have increased considerably if there had been this kind of data for all countries.

Such a strategy, however, would require a substantial increase in survey resources that is not likely to be forthcoming in all countries participating in the ESS. The experience of the Netherlands suggests that reapproaching reluctant respondents is worthwhile if one can use the best interviewers (possibly with bonus arrangements) and encourage respondents by offering larger (and perhaps imaginative) incentives. The findings also suggest that on some variables – and especially on perceived threat from immigrants – the quality of the estimates in the survey was improved by response conversion. However, there are other obstacles to increasing response rates. In some countries, the National Coordinators have to apply very strict rules in regard to information about reluctant respondents. In fact in some cases the addresses of respondents that formally refuse have to be instantly deleted and no information about them can be retained. Such arrangements are a significant threat to serious survey research.

In the last part of the article, the term “traces of bias” was used because we only have an indirect way of estimating the likely bias in survey estimates. The “traces” hypothesis is valid on the assumption that the converted respondents have attitudes and behave in ways that are closer to those of persons who remain nonrespondents than to those of actual respondents. However, it is possible that the final nonrespondents differ much more from

cooperative respondents than the reluctant respondents for whom we have full information. A real test for this would be a very short interview among a sample of nonrespondents in which some crucial questions are used (Kersten and Bethlehem 1984). Voogt and Saris (2003) propose a question about interest in politics in surveys on political elections. This is not easy to implement in a Europe-wide survey and unless the response rate to the question is very high it is likely to lead to a new nonresponse category about which there is no information. In the ESS, the questions on political participation, trust in politics and threat from immigrants are good candidates for estimating likely bias. But these questions are even more sensitive and delicate to ask nonrespondents than the question on political interest. The analysis of bias will certainly be clearer once small samples of follow-up interviews among nonrespondents are performed but this will increase the cost considerably and is simply not possible in some countries. Despite these practical and financial challenges, a follow-up study among nonrespondents is foreseen in a small number of countries in Round 3 of the ESS, made possible by additional infrastructure funding from the EU.

Perhaps most important in the present study is the finding that the relationship between the type of respondent (cooperative, reluctant) and the attitudinal and background variables was not all in the same direction in all countries. This needs further research and discussion because it creates a serious challenge to any scholar who believes there is a theory of nonresponse that applies cross-nationally. It would be very unfortunate if an interaction exists between kind of response, explanatory variables and country, because it then becomes impossible to make inferences from the findings in a small number of well-documented countries to other countries where there is less information about converted respondents. When the factors that are related to nonresponse are not stable over countries, efforts need to be made to obtain a substantial number of reluctant respondents in all countries. Without such efforts data corrections for nonresponse will remain less than optimal. Furthermore, whilst we have a view on the direction of bias by using information from reluctant respondents that are converted, a number of factors cloud our analysis. This is because not all refusals are reapproached, the selection of cases to reapproach is rarely (if ever) random and a fraction of respondents still refuse to cooperate (see Fig. 1). In addition, we have no information about the distribution of easy- and hard-to-convert refusals among selected units that did not cooperate at all and we have only very limited information about the characteristics of these units (i.e., information about the local area and the property).

The data were analysed to answer the question as to whether nonresponse bias was likely to be affecting the parameters in substantive explanatory models for various attitudinal variables. Because of this question, the attitudes were the dependent variables in the regression models and not the kind of respondent. Our question is crucial from a cross-national point of view where one wants to compare explanatory models using country samples characterised by large differences in nonresponse. Bias is likely when there is still a relation between the dependent variable (attitude) and the kind of respondent (cooperative/reluctant) in a model in which the relevant predictors are included. The findings are mixed in this respect. In some of the explanatory models the effect of the kind of respondent disappeared, but this was not always the case. In the two countries in which we could do multivariate analysis, easy-to-convert refusals did not have a significant effect and hard-to-convert refusals only mattered on two of the six measures.

Furthermore the remaining (significant) effects are small and do not have serious implications for the parameters of the substantial explanatory variables. This rather optimistic view must be treated with some caution because of the possibility that real refusals may differ from the converted refusals.

Even when response rates have been enhanced substantially, the question remains as to whether higher response rates actually decrease nonresponse bias. Groves (2006) reports recent studies in which the linkage between nonresponse rates and nonresponse bias is absent. Stoop (2005) has also shown that higher response rates do not necessarily result in better data quality and also that respondents who require the greatest fieldwork efforts are not necessarily similar to final nonrespondents on the variables of interest. We could not provide evidence on the refusals that were not converted. Expending a great deal of effort in converting easy-to-convert refusals may not be the best use of scarce resources. More research across a range of surveys is needed to answer the question as to whether higher response rates decrease nonresponse bias. Indeed, in the light of our mixed results, we are not able to decide which of the two models, the "*continuum of resistance mode*" or the "*classes of nonparticipants model*" finds most support in our data. Further research on the differences and similarities in reasons for refusing cooperation between the two kinds of reluctant respondents (easy- and hard-to-convert refusals) and the refusals who were reapproached and who still refused to participate in a survey is needed. Planned follow-up studies to be conducted by the ESS team in 2007 may help us to answer this. The call records data contain the necessary information for further research in this direction.

Finally, we should remember that the Netherlands and Germany do not have more severe problems with traces of nonresponse bias than other countries. These countries presumably followed the ESS standards more closely partly to make it possible to detect traces of bias from nonresponse. In the case of these countries we have information about the direction of nonresponse bias. It is very likely that this bias also exists in the other countries but we have no way of detecting it for variables that are not documented in population statistics, which is always the case for attitudes.

7. References

- AAPOR (2004). Standard Definitions. Final Dispositions of case Codes and Outcome Rates for Surveys. AAPOR.
- Bentler, P.M. (1980). Multivariate Analysis with Latent Variables: Causal Modeling. *Annual Review of Psychology*, 31, 419–456.
- Biemer, P.P. and Lyberg, L.E. (2003). *Introduction to Survey Quality*. NY: Hoboken, Wiley and Sons.
- Billiet, J. (1995). Church Involvement, Individualism, and Ethnic Prejudice among Flemish Roman Catholics: New Evidence of a Moderating Effect. *Journal of the Scientific Study of Religion*, 34, 224–233.
- Billiet, J. (1998). Social Capital, Religious-philosophical Involvement and Social Integration in Belgium: An Empirical Investigation. In Secularization and Social Integration. Papers in honor of Karel Dobbelaere, R. Laermans, B. Wilson, and J. Billiet (eds). Leuven: Leuven University Press.

- Billiet, J. and Waege, H. (2003). Een samenleving onderzocht. Methoden van sociaal wetenschappelijk onderzoek. Brussel: de Boeck. [In Dutch]
- Billiet, J., Koch, A., and Philippens, M. (2007). Understanding and Improving Response Rates. In Measuring Attitudes Cross-nationally: Lessons from the European Social Survey. R., Jowell, C. Roberts, R. Fitzgerald, and G. Eva (eds). London: Sage.
- Bollen, K.A. (1989). Structural Equations with Latent Variables. New York: John Wiley.
- Brehm, J. (1993). The Phantom Respondents: Opinion Surveys and Political Representation. Ann Arbor: University of Michigan Press.
- Couper, M.P. and De Leeuw, E. (2003). Nonresponse in Cross-cultural and Cross-national Surveys. In Cross-cultural Survey Methods. J. Harkness, F. Van de Vijver, and P. Mohler (eds). New Jersey: John Wiley.
- Curtin, R., Presser, S., and Singer, E. (2000). The Effects of Response Rate Changes on the Index of Consumer Sentiment. *Public Opinion Quarterly*, 64, 413–428.
- De Heer, W. (1999). International Response Trends: Results of an International Survey. *Journal of Official Statistics*, 15, 129–142.
- De Heer, W. and De Leeuw, E. (2002). Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In Survey Nonresponse, R. Groves, D.A. Dillman, J.L. Eltige, and R. Little (eds). New York: John Wiley.
- Devacht, S., Loosveldt, G., Philippens, M., and Billiet, J. (2003). Procesevaluatie van de dataverzameling in de eerste ronde van het European Social Survey, Onderzoeksverslag van het Departement Sociologie, Afdeling Dataverzameling en analyse, DA/2003–33. [In Dutch]
- Groves, R.M. (2006). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly*, 70, 646–675.
- Groves, R.M. and Couper, M.P. (1998). Nonresponse in Household Interview Surveys. New York: John Wiley.
- Hox, J. and De Leeuw, E. (2002). The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison. In Survey Nonresponse, R. Groves, D.A. Dillman, J.L. Eltige, and R. Little (eds). New York: John Wiley.
- Keeter, S., Miller, C., Kohut, A., Groves, R.M., and Presser, S. (2000). Consequences of Reducing Nonresponse in a National Telephone Survey. *Public Opinion Quarterly*, 64, 125–149.
- Kersten, H.M.P. and Bethlehem, J.G. (1984). Exploring and Reducing the Nonresponse Bias Basic Question. *Statistical Journal of the United Nations*, ECE 2, 369–380.
- Jowell, R. (1998). How Comparative Is Comparative Research? *American Behavioral Scientist*, 42, 168–177.
- Jowell, R., Roberts, C., Fitzgerald, R. and Eva, G. (eds) (2007). Measuring Attitudes Cross-nationally: Lessons from the European Social Survey. London: Sage.
- Lin, F. and Schaeffer, N.C. (1995). Using Survey Participation to Estimate the Impact of Nonparticipation. *Public Opinion Quarterly*, 59, 236–258.
- Loosveldt, G. and Carton, A. (2001). An Empirical Test of a Limited Model for Panel Refusals. *International Journal of Public Opinion Research*, 13, 173–185.
- Loosveldt, G. and Carton, A. (2002a). Utilitarian Individualism and Panel Nonresponse. *International Journal of Public Opinion Research*, 14, 428–438.

- Loosveldt, G. and Carton, A. (2002b). Kwaliteitsevaluatie van surveys. Een toepassing op de surveys naar culturele verschuivingen in Vlaanderen. In Vlaanderen gepeild. Brussel: Ministerie van de Vlaamse Gemeenschap, 2001 (Evaluation of Survey Quality. An Application on the Surveys on Cultural Change in Flanders). [In Dutch]
- Loosveldt, G., Philippens, M., Stoop, I., and Billiet, J. (2003). Reluctance and Refusal Conversion Procedures in the ESS. Paper presented at the International Workshop on Household Survey Nonresponse, Leuven, 22–24 September, Belgium.
- Loosveldt, G., Carton, A., and Billiet, J. (2004). Assessment of Survey Data Quality: A Pragmatic Approach Focused on Interviewer Tasks. *International Journal of Market Research*, 46, 65–82.
- Lynn, P. (2003). Developing Quality Standards for Cross-national Survey Research: Five Approaches. *International Journal of Social Research Methodology*, 6, 323–336.
- Lynn, P.J., Beerten, R., Laiho, J., and Martin, J. (2001). Recommended Standard Final Outcome Categories and Standard Definitions of Response Rate for Social Surveys. Working Paper 2001, 23 October.
- Park, A. and Jowell, R. (1997). Consistencies and Differences in a Cross-national Survey. The International Social Survey Programme, ISSP Working Paper. Cologne: Zentral Archive.
- Philippens, M. and Billiet, J. (2004). Monitoring and Evaluating Nonresponse Issues and Fieldwork Efforts in the European Social Survey. Paper presented at the European Conference on Quality and Methodology in Official Statistics. Mainz.
- Saris, W. and Münnich, A. (1995). The Multitrait-Multimethod Approach to Evaluate Measurement Instruments. Budapest: Eötvös University Press.
- Scherpenzeel, A. and Saris, W. (1997). The Validity and Reliability of Survey Questions. A Meta-Analysis of MTMM Studies. *Sociological Methods and Research*, 25, 341–383.
- Stoop, I. (2004). Surveying Nonrespondents. *Field Methods*, 16, 23–54.
- Stoop, I. (2005). The Hunt for the Last Respondent: Nonresponse in Sample Surveys. Social and Cultural Planning Office, The Hague.
- Stoop, I., Devacht, S., Loosveldt, G., Billiet, J., and Philippens, M. (2003). Developing a Uniform Contact Description Form. Paper presented at the 14th International Workshop on Household Survey Nonresponse. Leuven.
- Stoop, I. and Philippens, M. (2004). Nonresponse in Nederland: van zwart schaap naar witte raaf. In Hollandse taferelen. Den Haag: SCP (Nonresponse in the Netherlands: From Black Sheep to White Raven). [In Dutch]
- Teitler, J.O., Reichman, R.E., and Sprachman, S. (2003). Costs and Benefits of Improving Response Rates for a Hard to Reach Population. *Public Opinion Quarterly*, 26, 126–138.
- Voogt, R. and Saris, W. (2003). Political Interest, The Key for Correcting to Correction Methods for Nonresponse Bias in Election Studies. Paper presented at the 14th International Workshop on Household Survey Nonresponse. Leuven.
- Voogt, R. (2004). I Am Not Interested: Nonresponse Bias, Response Bias and Stimulus Effects in Election Research. Amsterdam: University of Amsterdam (PhD Dissertation).

Received January 2006

Revised January 2007