Data Quality in Cross-National Surveys

A Longitudinal and Cross-Cultural Analysis of the Quality Indicators Response Rate, Fieldwork Efforts, and Nonresponse Bias

Inauguraldissertation

zur Erlangung des akademischen Grades einer

Doktorin der Sozialwissenschaften

der Universität Mannheim

Vorgelegt von

Verena Roswitha Halbherr

Mannheim, Juli 2016

Hauptamtlicher Dekan der Fakultät für Sozialwissenschaften:

Prof. Dr. Michael Diehl

Erstgutachterin/ Betreuerin: Prof. Dr. Beatrice Rammstedt

Zweitgutacher/ Betreuer: Prof. Dr. Joseph Sakshaug

Drittgutacherin: Prof. Annelies Blom, Ph.D.

Viertgutachter: Prof. Dr. Michael Bošnjak

Tag der Disputation: 26. 10. 2016

Acknowledgments

I would like to express my gratitude to the following people who supported, advised, and encouraged me during my years of work on this dissertation. First and foremost, I would like to thank my advisors Beatrice Rammstedt and Joseph Sakshaug for their advice and support, as well as their encouraging words when I was in doubt. Also, I want to thank my PhD committee especially Michael Bošnjak and Annelies Blom. I especially thank Annelies Blom for her feedback in the colloquium and her expert advice during the whole process of writing. Furthermore, I thank Achim Koch for his continuous support and advice during the process of writing. I thank my colleagues at GESIS-Leibniz Institute for the Social Sciences, my colleagues at the Core Scientific Team of the European Social Survey (ESS), my colleagues at the University of Nuertingen-Geislingen (HfWU), and my former colleagues at TNS Infratest/Kantar Health for their constructive discussions, advice, and support. Especially I want to thank (in first name alphabetical order) Ana Villar, Angelika Scheuer, Bella Struminskaya, Christina Eder, Ineke Stoop, Jessica Walter, Joss Roßmann, Julia Khorshed, Joachim and Anglika Welz, Katrin Hemminger, Klaus Pforr, Regina Jutz, Roswitha Halbherr, Silke Schneider, Tobias Gummer, and Vladimir Bunoza for their constructive critique, discussions, help with Stata and R, proofreading, feedback, and fun disturbances. I also thank Frauke Kreuter and Patricia Moy for supporting me with their network.

Last but not least, I thank the many respondents of the ESS study because without their willingness to share information, this research would not have been possible. In addition, a special thank you goes to all the nonrespondents of the ESS because without them this research also would not have been possible.

This study is based on data from the European Labour Force Survey (EUROSTAT) from 2002, 2004, 2006, 2008 and 2010. The responsibility for all conclusions drawn from the data lies entirely with the author of this dissertation.

Contents

List of Abbreviations 9

I. Introduction 11

II. Literature Review 18

1 Literature on the Development of Response Rates 20

1.1 Response and Nonresponse Rates 20

1.2 Response Rate as Quality Indicator 23

1.2.1 AAPOR Standardization 26

1.2.2 Problems With AAPOR’s Standardization 27

1.2.3 Response Rate Calculation in the ESS 31

1.3 Challenges in Cross-National and Longitudinal Comparison 35

1.4 Hypothesis: Development of Response Rates (Part A) 39

2 Literature on Nonresponse and Nonresponse Bias 42

2.1 Nonresponse and Nonresponse Bias 43

2.2 Reasons for Nonresponse 47

2.2.1 Micro level 48

2.2.2 Meso level 51

2.2.3 Macro level 51

2.3 Hypothesis: Response Rate and Nonresponse Bias (Part B) 53

2.4 Hypothesis: Fieldwork Efforts Response Rate (Part C) 54

2.5 Hypothesis: Fieldwork Efforts and Nonresponse Bias (Part D) 57

2.6 Hypothesis: Variable Specific Effects (Part B and D) 59

3 Discussion 64

III. Data and Methods 68

4 European Social Survey (ESS) 69

5 Labour Force Survey (LFS) 71

6 Harmonization Between the ESS and LFS Datasets 72

7 Nonresponse Bias Calculation 74

IV. Analysis 78

Analysis Part A: Development of Response and Nonresponse Rates 79
8 Development of Response Rates 82
8.1 Trends Within the Country Level 87
8.2 Trends in Countries With Decreasing and Increasing Response Rates 89
8.3 Excurse: Differences in Response Rates due to Sample Design 94
9 Development of Nonresponse Rates 97
9.1 Refusal Rate 100
9.2 Non-Contact Rate 104
9.3 Rate of Not able/Others 108
10 Discussion 113
11 Results for Analysis Part A 114

Analysis Part B: Response Rates and Nonresponse Bias 121
12 Pooled Cross-Sectional Analysis: Overall Bias 125
13 Cross-Sectional Analysis: Variable Specific Bias 131
13.1 Gender (male) 133
13.2 Young Persons 135
13.3 Old Persons 137
13.4 Working Population 139
13.5 Family Status: Married Persons 141
13.6 Low Education 143
13.7 High Education 145
13.8 Nationality 147
13.9 One-Person Household 149
13.10 Five-or-More Person Household 151
14 Discussion 153
15 Results for Analysis Part B 154

Analysis Part C: Fieldwork Efforts and Response Rates 159
16 Fieldwork Efforts 162
16.1 Effect of Separate Fieldwork Efforts on Response Rates 163
16.2 Effects of Fieldwork Effort on Nonresponse Rates 170
17 Fieldwork Effort Index 173
17.1 Construction Fieldwork Effort Index (FEI) 174
17.2 Descriptive Analysis 176

17.3 Quantitative Analysis 177
18 Change of Fieldwork Efforts and Change of Response Rates 181
18.1 Cross-Section Analysis: Overall Level 183
18.2 Separate Analysis for Different Trends in Fieldwork Efforts 187
18.2.1 Increased fieldwork efforts 187
18.2.2 Decreased fieldwork efforts 189
18.2.3 Same level of fieldwork efforts 190
18.2.4 Result 190
19 Longitudinal Qualitative Within-Country Analysis 192
19.1 Germany—Decreasing Response Rate 194
19.2 Spain—Increasing Response Rate 197
19.3 The United Kingdom—Constant Response Rate 199
19.4 Portugal—Constant Response Rate 201
19.5 Result 203
20 Discussion 204
21 Results for Analysis Part C 205

Analysis Part D: Fieldwork Efforts and Nonresponse Bias 212
22 Analysis at the Aggregate Level 216
22.1 Effect of Fieldwork Efforts on Nonresponse Bias: Variable Specific 216
22.2 Effect of Fieldwork Efforts on the Nonresponse Bias Index 218
23 Variable Specific Analysis 220
23.1 Gender (male) 222
23.2 Young Persons 225
23.3 Old Persons 227
23.4 Working Population 229
23.5 Family Status: Married Persons 231
23.6 Low Education 233
23.7 High Education 235
23.8 Nationality 237
23.9 One-Person Household 239
23.10 Five-or-More Person Household 241
23.11 Result 243
24 Discussion 244
25 Results for Analysis Part D 245

V. Conclusion, Discussion and Outlook 247

VI. Literature 255

VII. Appendices 275

Appendix A: Analysis Part A 275

Appendix B: Analysis Part B 282

Appendix C: Analysis Part C 294

Appendix D: Analysis Part D 300

Appendix E: Literature Review 305

List of Abbreviations

AAPOR	American Association for Public Opinion Research
ALLBUS	Allgemeine Bevölkerungsumfrage der Sozialwissenschaften
CAPI	Computer Assisted Personal Interviewing
DIW	Deutsches Institut für Wirtschaftsforschung
DHS	Demographic and Health Surveys
ESS	European Social Survey
EU-SILC	European Union Statistics on Income and Living Conditions
EUROSTAT	Statistical Office of the European Communities
EVS	European Value Survey
FEI	Fieldwork Effort Index
FMI	Fraction of missing information
F2F	Face To Face Interview
ILO	International Labour Organization
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISSP	International Social Survey Programme
LFS	Labour Force Survey
MAR	Missing at random
MCAR	Missing completely at random
NMAR	Not missing at random
NRB	Nonresponse Bias
NSD	Norwegian Centre for Research Data
NTS	National Technical Summary
OECD	Organisation for Economic Co-operation and Development
PACS	Pacte civil de solidarité (English: contract of civil partnership)
PAPI	Paper And Pencil Interviewing
PIAAC	Program for the International Assessment of Adult Competencies

PISA	Program for International Student Assessment
POLS	Pooled Ordinary Least Square Regression
RISQ	Representativity Indicators for Survey Quality
RR	Response Rate
SOEP	German Socio-Economic Panel
TSE	Total Survey Error
TNS	Taylor Nelson Sofres

Country abbreviations (ISO 3166 Alpha-2)

BE	Belgium
CH	Switzerland
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
HU	Hungary
IE	Ireland
NL	The Netherlands
NO	Norway
PL	Poland
PT	Portugal
SE	Sweden
SI	Slovenia
UK	The United Kingdom
US	The United States of America

I. Introduction

Understanding what is happening during data collection is essential to data quality. When commissioning a survey, researchers often have ample knowledge about their topic of interest, the questionnaire, and how to analyze the data. However, they know less about the actual process of data collection and the impact of this process on survey estimates. Large scale interviewer-administered surveys usually are subcontracted to a fieldwork organization that collects the data. After the data are collected, the researcher receives data files with an accompanying technical report describing basic information about the data collection process such as the dates of fieldwork, a few sample characteristics, and the final response rate. Other details about the process of data collection—including information about interviewer training, interviewer recruitment and pay schemes, respondent contact strategies, re-issuing strategies, or response rate enhancing procedures—often go unreported, even though all these factors can impact survey estimates. Without the necessary information about the details of the data collection, a proper evaluation of the data is not possible, especially since fieldwork processes may introduce bias. This dissertation shines a spotlight on the effects of the often unreported aspects of fieldwork efforts made during data collection.

In contrast to those often unreported fieldwork efforts, the response rate often is reported and used as a proxy to describe data quality during fieldwork. Since response rates often are described as decreasing (de Leeuw & de Heer, 2002; Groves, 2011; Peytchev, 2013), a question arises as to whether the data quality is still sufficient, especially since it is financially expensive to achieve or maintain a high response rate in random probability samples when using a large data base (Couper, 2013), and especially since online non-probability access panel surveys are flexible and cheaper (Callegaro et al., 2014).

This dissertation provides an overview of the effects of the often unreported aspects of fieldwork on data quality and the often reported response rate. Specifically, the present study examines the response rate as an indicator of data quality, and analyzes the effects of the response rate and fieldwork process on data quality.

The data quality of a survey has many facets that are conceptualized most comprehensively by the Total Survey Error (TSE) paradigm (Biemer, 2010; Groves & Lyberg, 2010). The focus of the present study is on the nonresponse error, the representation error between the drawing of the sample and the selection of the respondent. Therefore, the main focus is on the process of fieldwork. The analysis examines the component nonresponse error and data quality in cross-national surveys with references to longitudinal surveys and cross-cultural analysis. Importantly, this study provides background for researchers and practitioners interested in the quality of surveys and how survey fieldwork processes impact this quality.

Traditionally, response rates are used to describe several aspects of data quality. Often, they are used as a quick and presumably simple indicator of data quality. Response rates provide an indication of the proportion of the selected sample units that participated in a survey. Low response rate are considered as undesirable, and they serve as indicators of nonresponse bias because the higher the percentage of sample units that participate, the lower is the perceived risk that not measuring these sample units will introduce bias in the survey estimates (Goyder, 1987; Groves et al., 2008; Groves & Peytcheva, 2008; Peytcheva & Groves, 2009). In other words, excluding 5% of a target population may not have, intuitively, as big of an effect on estimates as excluding 50% of the target population. However, research has shown that the relationship between response rates and nonresponse bias is not as strong or as simple as was suspected previously (Curtin, Presser, & Singer, 2000; Groves et al., 2008; Groves & Peytcheva, 2008; Keeter, Miller, Kohut, Groves, & Presser, 2000; Merkle &

Edelman, 2002; Peytcheva & Groves, 2009). Thus, a question arises as to whether a survey with a low response rate has sufficient data quality. If only a limited percentage of a sample participates in a survey, is it still representative of the sampled population?

Response rates also are used as a proxy for the efforts that a survey organization has invested in planning and carrying out fieldwork; and the quality, availability, and ability of their interviewing staff. For example, if a survey achieves low response rates, one may question whether the survey organizations were training their interviewers appropriately or whether each sample unit received enough visits.[1] Within this framework, many survey practitioners equate high efforts during fieldwork to data quality. Thus, when low response rates occur, researchers often request additional investments from the survey organization to obtain the desired response rate.

Over the last few years, an increasing number of concerns have been raised about decreasing response rates. In Germany, the response rates of well-controlled, high-quality surveys are around 30% to 40% (e.g., the European Social Survey [ESS], Allgemeine Bevölkerungsumfrage der Sozialwissenschaften).[2] Other countries suffer from similar low response rates (Brick & Williams, 2013; Dixon & Tucker, 2010; Matsuo & Loosveldt, 2013, Kreuter, 2013).

Survey organizations have claimed that this low response rate does not degrade data quality with respect to nonresponse bias. However, external pressure from researchers means that these organizations must use more fieldwork resources to increase these response rates.

---

[1] The focus of the analyses in this study is on personal interviews conducted by interviewers (face-to-face) only.
[2] For the response rates of the European Social Survey, see the European Social Survey, 2014a, 2014b, 2014c, 2014d, 2014e, 2016a Documentation Reports and Data Assessment report (Matsuo, Billiet, Loosveldt, & Malnar, 2010). For the response rates of the Allgemeine Bevölkerungsumfrage der Sozialwissenschaften, see Wasmer, Blohm, Walter, Scholz, and Jutz (2014).

Some cross-national survey programs (such as the ESS and the Programme for the International Assessment of Adult Competencies) have created rules and guidelines for the target minimum response rate that a country needs to achieve. These survey programs insist that greater efforts should be devoted to increasing or at least maintaining the response rate. However, this additional effort may entail additional survey costs. Also, it may have a negative effect on survey estimates, for example if the additional respondents are very similar to those who already were part of the sample and different from those who remain as nonrespondents. That is, the increased efforts may have an impact not only on how many of the sampled units participate, but also on what type of sampled units participate. If the increased participation is in any way related to the variables of interest, these additional efforts may increase the nonresponse bias.

Thus, the data quality indicators *fieldwork effort, response rate,* and *nonresponse bias* are linked and intertwined in complex ways. Empirical evidence is limited on the relationship of the response rate, fieldwork efforts, and nonresponse bias. A research gap exists as to the possible links between fieldwork efforts and nonresponse bias, and the literature that has examined this relationship is minimal. The present study intends to narrow this gap by providing a basis for previously untested assumptions regarding the response rate, nonresponse bias, and fieldwork efforts. In particular, it focuses on analyzing how fieldwork efforts mediate the relationship between response rates and nonresponse bias. Thus, it focuses specifically on the data on response rates, nonresponse bias, and fieldwork efforts, which often are not available in a comparable manner.

This analysis is comparable cross-culturally and has a longitudinal focus (includes the years 2002 to 2010). The findings of the cross-national study—such as the European Social Survey (ESS)—provides a broad empirical background to help improve the understanding of the data quality of fieldwork in cross-national and national survey contexts. Previous research often has been limited to one survey (Curtin, Presser, & Singer, 2005; Visser, Krosnick, Marquette, & Curtin, 1996) and one point in time (e.g., Brehm, 1994), or to one particular fieldwork effort such as incentives (for example, Pforr et al., 2015). Also, previous research has focused mainly on the United States (Brick & Williams, 2013; Dixon & Tucker, 2010; Groves & Couper, 1998; Groves, Fowler, Couper, Lepkowski, & Singer, 2004) or has used less up-to-date information relating to Europe (Curtin et al., 2000; de Leeuw & de Heer, 2002). The present study analyzed the data from16 European countries from 2002 to 2010 by using metadata and paradata from the European Social Survey (ESS) and the European Labour Force Survey (LFS), the latter of which was used as a reference statistic for the nonresponse bias calculation.

The ESS provides comparable data on fieldwork efforts and response rates over time and across countries. Using the ESS data enables an analysis of the often implicitly assumed link between the response rate and fieldwork efforts. In addition, due to the transparent data structure of the ESS, variables can be harmonized and compared to the Labour Force Survey (LFS), which serves as a benchmark survey because of its large sample sizes and high response rates. By comparing ESS estimates to LFS estimates, a researcher can assess nonresponse bias for those variables that are less sensitive to measurement error, such as socio-demographic variables. Thus, the prevailing assumptions regarding response rate, fieldwork efforts, and nonresponse bias can be tested empirically by evaluating whether certain survey fieldwork features are associated with smaller differences between the ESS and LFS estimates.

On the one hand, the contribution of the present study to the survey literature is a systematic overview of the development of response rates, and on the other hand, provides empirical tests of the assumed links between the three main factors related to survey data quality—response rate, fieldwork efforts, and nonresponse bias. Thus, this study provides useful information regarding the data quality of a cross-national survey by analyzing its fieldwork aspects and cross-national comparability. The present study not only systematically tests the prevailing hypotheses and assumptions regarding the data quality of fieldwork, but also provides a cross-national and longitudinal overview of the issue. It examines data quality in the ESS, which is a flagship survey for other European surveys. Due to the methodological rigor, relative comparability of methods across countries, transparency, and availability of metadata on fieldwork efforts and response rate calculations of the ESS (European Social Survey, 2013; Halbherr & Koch, 2010; Halbherr, Koch, Ederle, & Mayn, 2014), the conclusions of the analyses of the present study, based on the data of the ESS, can be extrapolated to other surveys.

With this goal in mind, the present study has checked empirically the implicit assumptions about the effects of fieldwork efforts on nonresponse bias and data quality by using ESS data, and replacing these assumptions with empirical facts. Therefore, this analysis will help survey practitioners and all researchers interested in the quality of fieldwork to gain deeper insights into fieldwork processes and data quality in cross-national surveys, and into the quality indicators of response rates, nonresponse bias, and fieldwork processes.

The following chapters provide a literature review and a description of the data and methods used. The analysis is divided into four chapters (see Figure 1). Analysis Part A examines the development of response rates. Analysis Part B discusses the relation between response rates and nonresponse bias. Analysis Part C analyses the relation between fieldwork efforts and response rates. Analysis Part D examines the relation between fieldwork efforts and nonresponse bias.

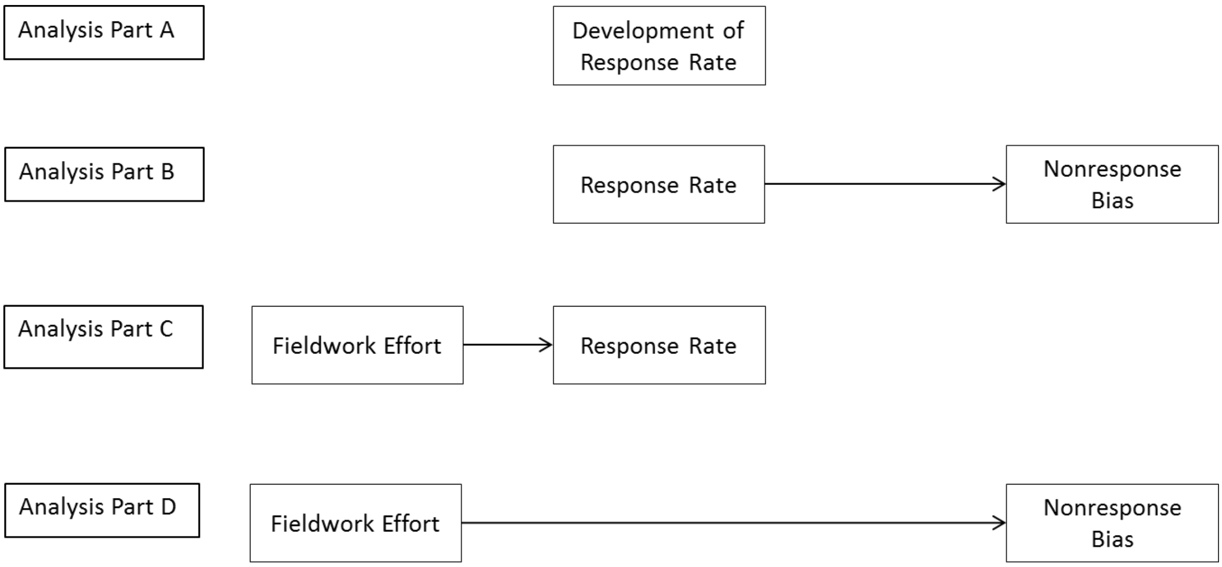

Figure 1. Overview of the analysis parts of the present study.

II. Literature Review

Data quality is a general concern, and, in cross-national surveys, it has become a topic of interest. The data quality of a survey has many facets that are conceptualized most comprehensively by the Total Survey Error (TSE) paradigm (Biemer, 2010; Groves & Lyberg, 2010). Overall, the focus of the present study is on the nonresponse error (Figure 2), the examination of data quality that associated is with the representation error between the drawing of the sample and the selection of the respondent. Therefore, the main focus is on the process of fieldwork.

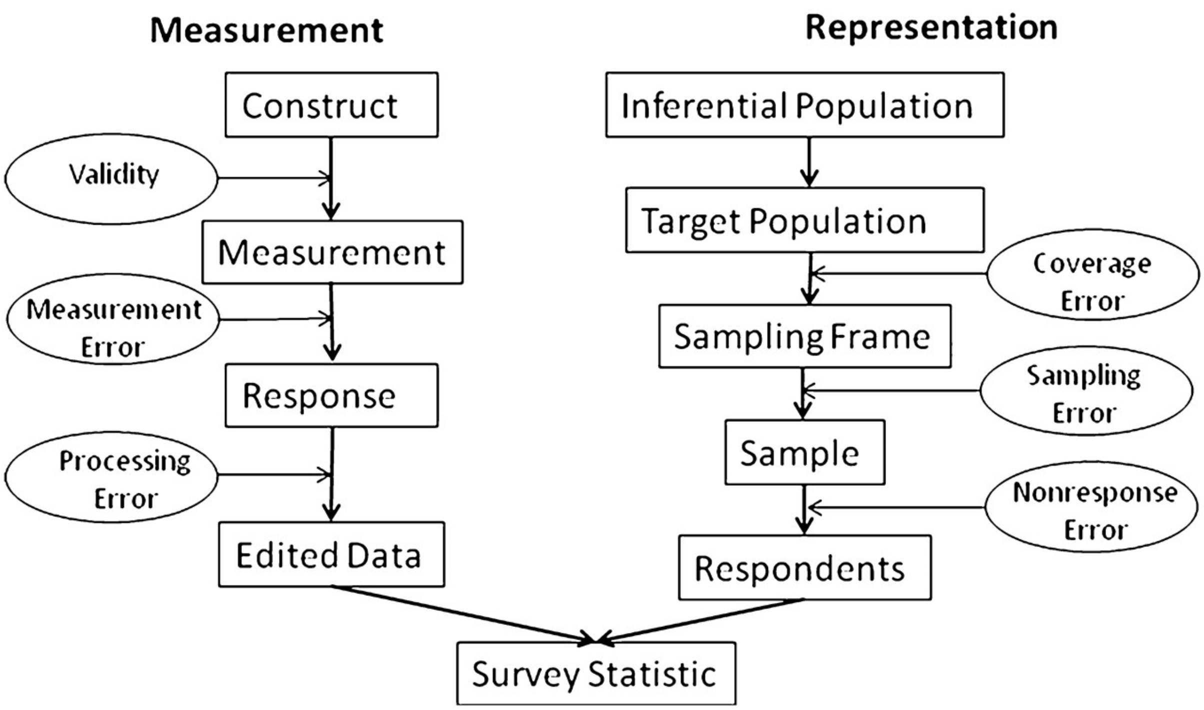

Figure 2. Total Survey Error components linked to steps in the measurement and representational inference process (Groves et al. 2004).

Although the quality of surveys is difficult to measure, one of the most popular and widely used quality indicators is the response rate (Biemer & Lyberg, 2003; Singer, 2006). This chapter describes how other factors besides the response rate—such as nonresponse bias and fieldwork efforts—have become central elements in the discussions about survey quality.

A special focus of this literature review is on the cross-national comparison of fieldwork efforts.

The present chapter provides a literature review that has been used to establish hypotheses on the relationship between three indicators—response rates, nonresponse bias, and fieldwork processes—which are analyzed in the following chapters (Analysis Parts A, B, C, and D; and Figure 1).

Background for the empirical analysis for Part A (Development of Response and Nonresponse Rates) is provided in Section 1, which defines *response rates, nonresponse rates*, and *the response rate as a quality indicator* (Section 1.1). Section 1.2 describes the standardization of the American Association for Public Opinion Research (AAPOR) and the challenges associated with using this standard. Section 1.3 examines the challenges of doing a cross-national comparison of response rates. Section 1.4 elaborates on the hypotheses on the development of response rates.

Section 2 provides background for the Analysis Parts B, C, and D. Section 2.1 discusses the background literature that relates the response rate to the nonresponse bias. Section 2.1 further elaborates on the nonresponse bias and why nonresponse may be problematic. Section 2.2 provides an overview of the different explanations for nonresponse at different levels (micro, meso, and macro).

The explanations at the micro (individual) level provide the background for Part B (Response Rates and Nonresponse Bias). The explanations at the meso level (survey specific level) provide the rationale for the analysis of fieldwork efforts (Part C, Fieldwork Efforts and Response Rates) and Part D (Fieldwork Efforts and Nonresponse Bias).

The present chapter also describes how other factors of quality research—such as the nonresponse bias and fieldwork efforts—have become central elements in the current discussions about survey quality. Section 2.3 discusses the research gap regarding these quality indicators.

In Section 2.3 provides hypotheses on the response rate and nonresponse bias (Analysis Part B); Section 2.4 provides hypotheses on fieldwork efforts and response rates (Part C); and Section 2.5 provides hypotheses on the relationship of fieldwork efforts and nonresponse bias (Part D).

1 Literature on the Development of Response Rates

This section provides background for the analysis of the development of response rates that are conducted in Analysis Part A of this thesis. In the section response rates and nonresponse rates are defined (Section 1.1) and the response rate as quality indicator is discussed (Section 1.2). The standardization process of the American Association for Public Opinion Research (AAPOR) (Section 1.2.1) and the challenges and problems that might arise when applying the AAPOR standards are described in Section 1.2.2. Section 1.1 1.2.3 considers why the calculation of response rates in the European Social Survey (ESS) deviates from the AAPOR definition. In Section 1.3 challenges of the cross-national comparison of response rates are described. And in Section 1.4 hypothesis on the development of response rates are formulated.

1.1 Response and Nonresponse Rates

When conducting a probability-based sample, survey cases can be divided into a set of four main comprehensive groups (American Association for Public Opinion Research, 2016, p. 9)[3]: 1) the interviewees, 2) eligible cases who are not interviewed (nonrespondents), 3) cases of unknown eligibility who are not interviewed, and 4) cases who are not eligible. [4]

In a probability-based survey, the rate of respondents that participate—the response rate—is basically the rate of respondents, divided by the total sample size. [5] (Unit)-Nonresponse is, in general, the other side of the coin of survey respondents. Persons who are selected for a sample (and who also are eligible) and do not provide the required information are defined as *nonrespondents*.[6]

Nonrespondents are persons who cannot be interviewed. These are either persons who cannot be contacted (so called *noncontact*), or persons who refuse to be interviewed (so called *refusals*). The American Association for Public Opinion Research (2016) defines the *refusal rate* as "… the proportion of all cases in which a housing unit or respondent refuses to do an interview, or breaks-off an interview of all potentially eligible cases" (p. 64).[7] The *contact rate* is defined as "… the proportion of all cases in which some responsible member of the housing unit was reached by the survey" (American Association for Public Opinion

[3] The categories are further refined in the full list of the final disposition codes for "In-Person, Household Surveys" (see American Association for Public Opinion Research, 2016, p. 76).
[4] Ineligible cases are not discussed any further, since they are considered "neutral" regarding nonresponse bias. Some literature has suggested that people deliberately misreport their eligibility. For example, a person might lie about kids living in their household to avoid an interview (Demographic and Health Surveys [DHS], 2016, p. 41), or respondents might misreport their employment status (Sinibaldi, 2014, p. 10). However, this evidence is more or less anecdotal, and so an analysis is difficult. Nevertheless, the European Social Survey (ESS) closely monitors the selection process within a household, and so it can be assumed that misreporting of ineligible cases has no effect on the nonresponse bias discussed in the present study.
[5] For more information on the calculation of response and nonresponse rates, see Section 1.1 and 1.2.1.
[6] Survey research recognizes two types of nonresponse. One is the item nonresponse and the other is the unit nonresponse. The item nonresponse focusses on the percentage of nonresponse per item in the questionnaire. The analysis in the present study has focused on the unit-nonresponse—on the respondents who do not respond to the whole questionnaire.
[7] AAPOR provides three definitions of refusal rates that differ in the way they treat cases with dispositions of unknown eligibility. For the definitions of refusal rates, see Appendix Figure E1.

Research 2016, p. 65).[8] The *non-contact rate* is defined in all cases as 100 percent minus the contact rate.

The two subgroups of refusals and noncontacts are the main components of nonresponse (e.g., de Leeuw & de Heer, 2002; Groves et al., 2004; Groves & Couper, 1998, p. 2), and also the main focus of methodological research (Groves, 2006, p. 637; Groves, Cialdini, & Couper, 1992; Groves & Couper, 1998; Singer, 2006). These components have different causes and may, therefore, have different consequences for survey estimates (Singer, 2006, p. 637; Groves, 2006; Groves, Couper, Presser, Singer, Tourangeau, Acosta & Nelson, 2006). A smaller group of nonrespondents who cannot be interviewed include, for example, persons with mental health problems or language problems. Generally, this group is called the “not able/others”. It includes language problems, health and mental problems, and being away during the fieldwork process. This category often is considered to be a residual category, and so it will not be included in the further discussions of the present study.[9] Researchers have evaluated the reasons for, and effects of, nonresponse (Groves & Couper, 1998; Groves, Dillman, Eltinge, & Little, 2002; Keeter et al., 2000), and AAPOR has sponsored an International Workshop on Household Survey Nonresponse.[10]

The three main groups of respondents, nonrespondents (noncontacts and refusals) are discussed in the following sections to show that all these rates (and their changes and entanglements) are important factors for analysis.

[8] The report provides three definitions of contact rates that vary in their definition of the eligibility of undetermined cases (see Appendix Figure E2).
[9] The category “not able/other” is the third type of nonresponse that has been less often researched. Stoop, Koch, and Matsuo (2012) have done some research on this category.
[10] See http://www.nonresponse.org/.

1.2 Response Rate as Quality Indicator

For a long time, researchers considered the response rate as the universal quality measure of survey quality (Biemer & Lyberg, 2003; Singer, 2006, pp. 637-645).[11] This view was based on the assumption that studies with low response rates are not representative of the population. It was assumed that with increasing response rates, the inference is higher because a sample with a high response rate differs less from the gross sample (Alreck & Settle, 1995; Babbie, 1990). Thus, a high response rate was used as a synonym for the validity of results, and thus indicated a higher quality study (Aday, 1996; Babbie, 1990; Backstrom & Hursh, 1963; Rea & Parker, 1997). Based on these assumptions, the textbooks for students and practitioners of survey research defined clear objectives and limits. For example, Alreck and Settle (1995) described the objective of high response rates as follows: “It is obviously important to do as much as possible to reduce nonresponse and encourage an adequate response rate” (p. 184). Babbie (2007) went a step further and defined *acceptable response rates*: “…of at least 50 percent … adequate for analysis and reporting. A response of 60 percent is good; a response rate of 70 percent is very good” (p. 262). Singleton and Straits (2005) were even more demanding and defined a response rate of over 85% as adequate, and any response rate below 70% as producing a serious chance of bias (p. 145).

In textbooks and surveys, the position that the level of response rates is a quality indicator still prevails because survey users and sponsors are familiar with response rates. Intuitively, high response rates seem to be an indicator of good survey processes and practices. Because information on the process of date collection is limited, response rate is

---

[11] Singer (2006, pp. 637-645) provides a comprehensive overview of the development of the response rate between 1995 and 2006.

still widely considered as an indicator for the quality of the fieldwork process (see Biemer & Lyberg,2003, p. 95).

Often, this assumption about the response rate as a quality indicator goes beyond the quality of a survey to the quality of a survey agency: “Indeed, to the users, the nonresponse rate may also be indicative of the competence of a survey organization. High response rates become synonymous with efficient, high-quality data collection operation (Biemer & Lyberg, 2003, p. 95).

Over the last few years, the ideas about response rates as quality indicators have been further developed.[12] The assumption that low response rates are equivalent to low quality has been increasingly questioned. Lelkes, Krosnick, Marx, Judd, and Park (2012) were able to show that studies with low response rates (ca. 20%) can have more exact estimates than studies with higher response rates (ca. 60–70%). Similar results were found by Holbrook, Krosnick, and Pfent (2008) who examined whether low response rates are related to a high nonresponse bias of the unweighted data. They compared the results of 81 national studies with varying response rates between 5% and 54%. Their analyses showed that the studies with low response rates were only slightly less imprecise in the marginal distribution of the demographic variables. In the survey methodological literature, various studies showed that response rates are poor indicators for nonresponse bias (e.g., Curtin et al., 2000; Groves & Heeringa, 2006; Groves, Presser, & Dipko, 2004; Keeter et al., 2000; Merkle & Edelman, 2002; Schouten, 2004). The most cited meta-analysis on the relation of response rates and nonresponse bias was carried out by Groves (2006, p. 659). He could not detect a clear relationship between the level of response rates and the absolute nonresponse bias. Some

[12] In popular literature (Huffington, 1998), the effect of low response rates also has been discussed.

studies have even showed that studies with a high response rate may have a higher nonresponse bias than studies with a low response rate (Koch, Halbherr, Stoop, & Kappelhof, 2014; Kohler, 2007; Stoop, 2005). Thus, an exclusive focus on the response rate should be avoided (Groves & Couper, 1998, pp. 133, 147).

To summarize the current position in the literature with respect to the response rate as a quality indicator, it can be said that the response rate is not a quality indicator by itself, although low response rates may increase the risk of nonresponse bias (Groves, 2006). Or as the American Association for Public Opinion Research (2016) has suggested, response rates "do not necessarily differentiate reliably between accurate and inaccurate data." Moreover, the current literature points to the fact that the response rate is not a quality indicator by itself.[13] Despite all the challenges and obstacles, the response rate is still the most recognized indicator of data quality (Biemer & Lyberg, 2003), although, at the same time, it is one of the most controversial (see Harkness, Braun, Edwards, Johnson, Lyberg, Mohler, Pennell & Smith, 2010, p. 278).

Despite the controversial discussion of response rates as quality indicators, ambitious response rates targets prevail in national and international survey guidelines (Kreuter, 2013, p. 24 f). Examples for these target response rates can be found, for instance, in the US Office of Management and Budget guidelines and in the guidelines provided by the Organisation for Economic Co-operation and Development (OECD) in the US. Similar guidelines and rules for response rate targets also are in use in international surveys such as the Program for

[13] For the discussion of the response rate as only one of many quality indicators see also the discussion of the Total Survey Error (Biemer & Lyberg, 2003 and Section 2).

International Student Assessment (PISA), the Program for the International Assessment of Adult Competencies (PIAAC), and the European Social Survey (ESS).[14]

It can be concluded that it is not only the response rate that can be considered as quality indicator. But many factors need to be taken into account to estimate the quality of data. Wagner (2012, p. 572) sums up the discussion that „No single indicator is likely to replace the response rate. It is necessary to look at the problem from many angles, with different assumptions, in order to provide a plausible case that nonresponse biases for the many statistics produced by any given survey have been diagnosed and remedied." This will be done in the further analysis.

1.2.1 AAPOR Standardization

The definition and standardization of response rates is important for making valid comparisons across time and between countries. Without standardization, these rates are not comparable and an analysis would be biased by other factors in addition to the development of response rates.

This section first describes AAPOR standardization, and in the next Section 1.2.2 the problems with this standardization are examined. Based on these problems, the next section (Section 1.2.3) examines why the ESS deviates from AAPOR standardization and how this deviation enables a better comparability across European countries.

Before the standardization of response rates, researchers used many different methods of calculation. As Groves and Lyberg (1988) have noted, "[t]here are so many ways of

[14] Johnson and Owens (2003) and Kreuter (2013) also point to the fact that in scientific journals there are no strict limits for publishing of surveys with a minimum response rate, but there are unstated limits of response rates under which articles on surveys with low response rates are not published. For example journals such as Public Opinion Quarterly asked authors to report response rates (among other basic information of the survey).

calculating response rates that comparisons across surveys are fraught with misinterpretations" (p. 195). The American Association for Public Opinion Research (2016) has stated that: "Response rates, cooperation rates, and completion rates were often treated as interchangeable in the literature." In the early 1980s, the Council of American Survey Research Organizations (CASRO) made the first attempt to standardize the definition of a *response rate* (CASRO, 1982).

The AAPOR completed this attempt in the late 1990s with their publication of *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys* (revised version American Association for Public Opinion Research, 2016).[15] To overcome this deficiency of differences in response rate calculation, AAPOR provided a definition and a suggestion for calculating response rates.

AAPOR's goal was to establish clear and comprehensive rules, and a standardized way to describe the final disposition of cases and the calculation of the outcome rate, so researchers could compare these rates across surveys of different topics and organizations (American Association for Public Opinion Research, 2016). The AAPOR definitions provide clear guidelines for calculating response rates; cooperation rates; and contact rates across surveys, time, and countries.

1.2.2 Problems With AAPOR's Standardization

Despite the many positive aspects of the AAPPOR standardized definitions, differences still exist in the calculation of outcome codes. The main challenges regarding the AAPOR standardization are described in this section.

[15] For the history of the standard definition of *response rates*, see Blom (2009, p. 41).

The *first potential problem* is AAPOR's focus on the US. In the US for population surveys samples of households are used, in Europe often register data with named individuals or household samples.[16] Therefore, countries using different sample designs than in the US, such as an individual or address sample, have to adapt some categories of the outcome[17] due to the different possibilities of nonresponse in the different sample designs. The AAPOR definitions are not directly applicable to other countries (Lynn, Clarke, Martin & Sturgis, 2002, p. 63). This limitation is due to the fact that "[t]he nature of the sampling methods and sampling frames used for many social surveys in Europe … raises issues that are not dealt with in the AAPOR document" (p. 63). For example, in an individual sample in which the name and address of the potential respondent is known, an empty house is classified as a *noncontact* (in most of the cases). In a household sample, in which the interviewer does the listing of houses, this might not be listed as a household, since it is not occupied. This would count as *ineligible*. So the different types of samples make a difference in the calculation of response rates. Even though response rates are comparable, some slight differences may still exist between the different sample designs for in-person household surveys that could lead to deficiencies in the standardization.

The *second problem* with the AAPOR standardization is the differences in response rate calculation for different sample designs in cross-national comparison. Response Rate calculation for cross-national surveys with different sampling designs can be problematic in terms of comparability. If response rates are calculated differently in cross-national surveys,

---

[16] The American Association for Public Opinion Research (2016) has described the sampling process of in-person household surveys as follows: "… an in-person household survey is assumed to be one in which housing units are sampled from an address-based sampling frame of some geopolitical area using one of several probability sampling techniques" (p. 23).

[17] AAPOR has created a response rate calculation for samples of named individuals for mail and Internet surveys, which can be adapted for this purpose (American Association for Public Opinion Research, 2016, p. 43 ff.).

problems can arise when analyzing the trend of decreasing response rates. In cross-national surveys without a detailed description of response rates, it is impossible to compare data from different sources or countries (de Heer, 1999, p. 141).

Thus, when comparing response rates cross-nationally, special care must be taken, since the differences in countries may be due to the different sample types used. The ESS as a cross-national comparable survey uses different types of samples. As will be discussed in more detail later, the ESS uses an algorithm for calculating the response rate that takes these differences into account. The ESS provides disposition codes for individual samples and household samples. Therefore, a comparison of response rates between countries is possible.

The *third issue* regarding the AAPOR standardization of cross-national comparable surveys is other possible constraints that may influence available response rate outcomes. These are design features, such as the mode of fieldwork, the definition of what constitutes an eligible population, the location of the population, fieldwork rules and procedures; as well as external factors such as population characteristics, the housing situation, society in general, and the population structure (see Blom, 2009, p. 45). In addition, other constraints that may influence response outcomes can include the mode of case management of potential interviews (paper and pencil or computer assisted); legal requirements such as a respondent's ability to opt out of a register, and permission to re-contact refusals; and the training of interviewers.

The *fourth problem* with the currently available code frames of the AAPOR response outcomes is that they take into account case-level outcomes only. A *case level* is the recorded status and the contact history used. This data typically is collected by an interviewer who decides how and when future calls will be made. In contrast to this *call level* data is usually the data collected in the contact form, which provides additional information for each call.

Since more information is included in the call level data, this data is preferred for the calculation of response rates (see Blom, 2009, p. 9). The code frame of AAPOR does not include this differentiation between case level vs. call level outcome codes (see Blom, 2009; Blom, Lynn, & Jäckle, 2008). As Blom (2009) has pointed out, many response outcomes from the contacting process in the field are initially recorded by the interviewer at the call-level, which means that the interviewer writes down the outcome of each contact attempt (p. 43).[18] Typically, this information provided by the interviewer is transmitted to the survey organization or the interviewer's supervisor to make a decision on the final disposition code of the case. Blom (2009) also has pointed to the need for a comparable definition of response outcomes at the call-level, which is usually provided by the interviewer, supervisor, or the survey agency, and the case-level information that provides a final disposition for each sample unit (i.e., case-level outcome) (p. 43). Blom (2009) also has referred to the standardization challenge with respect to the face-to-face survey (what exactly constitutes a contact attempt) and the problem of standardizing outcome codes for cases in which contact was not achieved (p. 43). The ESS was the first cross-national survey that collected detailed call-level cross-national data on the fieldwork process, and that provided this data to the public (Stoop, Devacht, Billiet, Loosveldt, & Philippens, 2003),[19] which enabled researchers to calculate outcome rates across surveys and countries.[20]

The *fifth problem* with the AAPOR standardization is that it lacks editorial policies. The omission or misreporting of response and nonresponse rates might also be due to the fact that journals do not always have clear guidelines on reporting (see Smith, 2002, p. 36).

---

[18] Of course, the interviewer's focus is on the planning for future calls and not on the definition of standardized outcome codes.
[19] This data can be downloaded free of charge at http://www.europeansocialsurvey.org.
[20] See Blom (2009) for a code frame that takes into account differences in strategies for deriving a final case outcome from a sequence of call-level outcomes, which can be implemented in different countries and surveys.

1.2.3 Response Rate Calculation in the ESS

As discussed previously, the ESS algorithm for computing the response rate takes into account different sample types, and therefore overcomes the unclear definition of AAPOR regarding the cross-national comparability of response rates and nonresponse rates. In addition, the ESS publishes information on the components of nonresponse. By making this information publically available, researchers can make conscious decisions when comparing response rates cross-nationally. The analysis of the ESS data on the development of response rates (see Analysis Part A) enables researchers to draw conclusions about the response rates from a cross-national perspective, without the challenges and problems that a comparison of these rates might include if the data used was from other surveys that do not provide details or response rate calculations, or that differ in many of the aspects previously described.

Due to the challenges of the AAPOR standardization for cross-national surveys and of the surveys conducted outside the United States, the ESS has slightly adapted the calculation of AAPOR response rates[21] (see Equation 1 and 2).[22] The following section describes the calculation of the ESS response rate (European Social Survey, 2013, p. 23f.), and also points out the deviations of the ESS response rate in comparison to the AAPOR response rate.

$$ESS\ Response\ Rate = \frac{\text{interviews}}{\text{number of individuals,households, or addresses selected - ineligibles}} \quad (1)$$

[21] The calculation of the refusal rate and non-contact rate in the ESS does not differ from the AAPOR calculation.

[22] Information about how the response rate is calculated in the ESS can be found in the following documents: European Social Survey (2011, section 7.3 in particular), European Social Survey (2013), Matsuo and Loosveldt (2013a) and in the following documents from the ESS: *ESS6_algorithm_for_computing_outcome_codes, NTech2012xx: Round 6 National Technical Summary* (NTS), *ESS R6 Contact Form Instructions FINAL*, *ESS6-SoF_all_countries*. Those ESS internal documents are available upon request, contact the author Verena Halbherr or the ESS ERIC Headquarter ess@city.ac.uk.

AAPOR provides different definitions of response rates. They vary in their inclusion of partial interviews and of cases of unknown eligibility (American Association for Public Opinion Research, 2015, p. 52f.). The official response rate calculation of the ESS is strongly based on the AAPOR RR 2, since all respondents whose eligibility is not known (those who are never contacted) are assumed to be eligible (see Equation 2).

$$AAPOR\ Response\ Rate = \frac{\text{interviews (complete plus partial)}}{\text{interviews (complete plus partial)} + \text{number of non-interviews} + \text{cases of unknown eligibility}} \quad (2)$$

Minor mapping problems of the AAPOR RR 2 and the ESS response rate calculation are due to the definition of *ineligibles*. In this longitudinal analysis, the relevance of this mapping problem only applies to countries that have changed their sampling design from an address or household type of sample to an individual sample. This is the case for only three countries within the ESS (Spain, Hungary, and Switzerland).[23]

As can be seen in Equation 1, the calculation of response rates in the ESS includes ineligibles. Ineligible cases include (see European Social Survey, 2013, p. 23f.):

Samples of individuals:

- Respondent deceased
- Address not occupied by respondent (not occupied/demolished/not yet built/weekend or second homes)
- Respondent emigrated/left the country long-term (for more than 6 months)
- Respondent resides in an institution

[23] This is the case for Hungary only in the ESS3—a change from an individual to a non-individual sample for just one round. Spain changed from a non-individual sample in ESS1 to an individual sample in following rounds. In ESS5, Switzerland changed to an individual sample from a non-individual sample used in ESS1 through ESS4. Information is based on the ESS sample designs from ESS1 to ESS5.

Samples of households or addresses:

- Address not occupied at all/demolished premises
- Address not yet built/under construction
- Non-residential address (used solely for business/industrial purposes or as an institutional address, for example, a factory, office, or school)
- Address occupied, but no resident household (e.g., weekend or second homes)
- Address occupied by resident household, but no eligible respondents (no one aged 15 years and older)

In the AAPOR definition (American Association for Public Opinion Research, 2016, p. 76), *ineligibles* or *not eligible* include (for in-person household surveys):[24]

- Regular, vacant residences
- Not a housing unit
- Non-residential address (business, government office, other organization, institution, group quarters)
- Vacant housing unit, Seasonal/Vacation/Temporary residence
- No eligible respondents
- Out of sample

The AAPOR response rate definitions for dispositions in "in-person" household surveys are based on area-based probability sampling through which one eligible respondent per housing unit is selected. Problems with mapping the ESS and AAPOR response rates result from the different definitions of eligibility in the samples of individuals in the ESS as compared to those in the AAPOR. In particular, these definitions include deceased

[24] Selection of the applicable disposition codes only.

respondents in samples of individuals. According to the ESS specifications for participating countries in ESS 7 (European Social Survey, 2013, p. 23), deceased respondents in samples of individuals are considered ineligible. The response rate calculation of AAPOR is based on a household sample for which the selection of respondents is done by an interviewer (for example, by the next/last birthday method or a Kish grid). So, a dead person will not be part of the selection process. In an individual sample that is drawn on registers some time before the start of fieldwork, persons who have died between the drawing of the sample and the interview are included. Also, mistakes in registers are possible. Therefore, differences exist between the types of samples, which also means that differences exist in the standardized response rate calculations of AAPOR.

Another difference between the AAPOR response rate and the calculation of the ESS response rate is the classification of the addresses available in the sampling frame that are unoccupied, demolished, or not yet built. In the ESS, they are ineligible (see the following list).[25]

Due to these differences between the samples of named individuals and the household samples, deviations exist between the calculation of the ESS response rate and the AAPOR response rate. Although the response rate calculation of the ESS does not exactly match the AAPOR response rates, the disposition code of samples of named individuals can be adapted to the disposition codes of household surveys. ESS response rates can be mapped onto AAPOR response rates.

[25] This difference seems to stem from a mistake in the specifications, and the ESS has plans to change it for round 8. Countries have raised this issue (e.g., Estonia) when computing response rates during the ESS 6 fieldwork (A. Villar, personal communication, May 6, 2014).

The ESS provides details of disposition codes on a contact-level basis, which enables researchers to calculate their prepared type of response rate. The ESS uses the same setup for all participating countries (mode, specifications for data collection, and so on),[26] as well as a cross-national comparable calculation of response rates based on call level (Matsuo & Loosveldt, 2013). Thus, a comparison of response rates across time and countries is possible with the ESS data.

The described deviation of the ESS response rate calculation from the AAPOR response rate calculation improves the comparability of response rates across countries by taking into account the specific sample design (namely samples of individuals) in European countries. Based on this adaptation, the development of response rates (Analysis Part A) and the relation of response rates and nonresponse bias (Analysis Part B) and the relation of response rates and fieldwork efforts (Analysis Part C) can be analyzed in a comparable way.

1.3 Challenges in Cross-National and Longitudinal Comparison

In comparative surveys, the differences between countries are of special interest (Blom, Jäckle, & Lynn, 2010, p. 336). In addition to the challenges of analyzing response rates and nonresponse rates in single countries or single surveys, comparative research must consider additional issues. When comparing surveys or countries, the following factors must be synchronized: the calculation of the response and nonresponse rates, the definition of the target populations, the nonrespondents, and the ineligibility rates. As described previously the response rate often is calculated differently across surveys, and sufficient information on these calculations is not always provided, which makes a comparison or evaluation difficult. Couper and de Leeuw (2003) have pointed to the importance of comparable response rates

[26] For details, see the ESS 7 specifications for participating counties (European Social Survey, 2013).

definitions in different countries (p.174). With respect to cross-national surveys, de Heer (1999) fromulated his concern about the lack of standards and thus comparability for response rate calculations, especially when analyzing the development of response rates (p. 140f.).

The comparison between countries has two main problems: the causal inference of the country-level data to other countries and problems with the cross-national comparability of response rate calculations (Blom et al., 2010, p. 340f.). The first problem of comparative analysis is that correlations of, for example, response rates and participation at the country-level do not allow a causal inference about the survey participation processes used in different countries. For example, de Leeuw and de Heer (2002) showed that high contact rates in countries with larger average household sizes do not allow an inference that larger household sizes have a higher contact propensity than smaller household sizes. Individual-level inferences require individual-level data (Blom et al., 2010, p. 340f.). The second main problem of comparison between the countries is the comparability of the response rate calculations (see also Section 2.2). Most studies rely on reports or technical summaries that do not provide further details on the disposition codes or the calculation of response rates.

Due to these two problems, previous studies on the cross-national comparability of response rates are problematic. Blom et al. (2010) have suggested that with respect to the current status of research on response rates between countries, “even descriptive inferences about differences between countries in response rates and trends were limited” (p. 340f.).

In addition to the previously described challenges of the comparability of calculation rates and inference, further challenges for cross-national surveys are related to the differences in the characteristics of nonresponse. While response rates may appear similar across countries, these rates may mask differences in the composition of the nonresponse. For example if a group of nonrespondents in one country consists mostly of noncontacts, whereas in another country, it consists mostly of refusal, this could influence the nonresponse bias. The likelihood of the nonresponse bias in the estimates comparing these two countries may be greater when the source of nonresponse (noncontact or refusal) differs in both countries (see Couper and de Leeuw, 2003). Thus, it is not only the calculation but also the composition of nonresponse and response that is important to include in cross-national comparitive analyses.

Another difficulty in comparative research that further complicates the matter is that the reasons for nonresponse may vary across countries, regions, and cultures. From the perspective of cross-national surveys, nonresponse and nonresponse bias need to be analyzed regarding this additional aspect.

To achieve cross-national comparability, not only the survey characteristics need to be examined, but also the processes that might introduce bias. This is the case when response mechanisms differ between countries or when the probability to respond may also be country specific. These additional requirements for cross-national surveys make the analysis of nonresponse and nonresponse bias between countries—with potentially different reasons for nonresponse—especially challenging (Blom et al., 2010, p. 336).[27] Couper and de Leeuw (2003) have described this lack of research as "a paucity of research comparing sources of nonresponse cross-nationally" (p. 170f.).

The following constraints and country specific characteristics may influence cross-national comparison in face-to face surveys. Information based on the discussion of the

[27] To achieve unbiased comparisons between countries, three conditions are required: "(a) nonresponse in each country is MCAR, (b) NMAR processes introduce equivalent bias to each national estimate, or (c) analysts must identify—and use appropriately—a set of additional items that turn the process from NMAR to MAR" (Blom et al., 2010, p. 336).

"ESS 5 Fieldwork Review Meeting" (personal communication, March 28, 2012). The following factors can introduce country specific differences, but are not demonstrably related to differences:

- Different sample frames: individual, address, household.
- Availability of registers and up-to-date information about registers
- Location of population and availability of interviews: e.g., people living on islands, region without any interviewers, dangerous areas (Gaza in Israel)
- Housing situation: skyscraper vs. 1-family houses
- Fieldwork process: e.g., fielding period timing and holiday seasons
- External factors: Christmas time/holiday season, cold winter, hot summer, vacation, and events
- Fieldwork rules: e.g., number of contact attempts, substitution, proxy interviews
- Definition of population and eligibility
- Mode of fieldwork: PAPI, CAPI
- Population characteristics: e.g., survey climate, too many surveys, and so on)
- Convention of interviewers in private houses: e.g., in Japan almost no one is allowed in private apartments, in Portugal it is very unusual to allow interviewers to enter the house
- Kind of back checks: contact, plausibility of data, and so on
- Definition of a full/partial interview—what is considered as a full interview: e.g., if only one question is answered or only the most important questions, or if there is a limit on the maximum percentage of item nonresponses, or if an interview is considered complete when all the questions are answered
- Definition of contact attempts: recording on paper or on computer: Definition of what is exactly considered a contact attempt: call, letter in a mailbox, information from

someone out of the household, information from neighbor? These differences can lead to differences in the types and quantity of contact attempts recorded, and the coded outcomes.

Another challenge is the longitudinal aspect of these surveys. While the majority of cross-national research has assumed that the nonresponse error is constant or stable across countries, Matsuo, Billiet, Loosveldt and Malnar (2010) and Stoop et al. (2010) have suggested that this assumption is not always valid. Matsuo et al. (2010) have shown that nonresponse errors affect descriptive statistics in different countries differently. They have provided examples for cases in which country means differ, as do the estimation of correlations between variables, and the model estimates.

1.4 Hypothesis: Development of Response Rates (Part A)

This section provides a review of the literature on the development of response rates, with a special focus on the cross-national differences in response rate development.

Various authors already have analyzed the development of response rates per se. For example, Groves, Fowler, Couper, Lepkowski, and Singer (2004) analyzed the different trends in US response rates studies since the 1970s (pp. 185–187). They have shown that response rates are decreasing. Groves and Couper (1998) and Brehm (1994) also found a decrease in the response rates of US academic, governmental, business, media and commercial surveys. In Europe the development of response rates was analyzed in de Leeuw and de Heer (2002) study, which found a decline in these rates from the 1980s to the 1990s (p. 52).

At the same time as response rates are decreasing, obviously nonresponse rates are described as increasing. In literature it is described that the nonresponse rate are getting worse and people as less likely to respond or be contacted (Barbier, Loosveldt, & Carton, 2016).

Based on this literature the following hypothesis stated and will be analyzed in Part A:

Hypothesis 1: Response rates are decreasing.

Hypothesis 2: Nonresponse rates are increasing.

Hypothesis 3: Refusal rates are increasing.

Hypothesis 4: Noncontact rates are increasing.

Hypothesis 5: Rates of "not able/others" are increasing.

Previous analyses regarding cross-national comparisons of the development of response rates have focused mainly on the macro level or survey climate. The effect of different survey climates across countries and the changes in survey climates within countries were considered by Lyberg and Dean (1992). Although only based on anecdotal evidence from the Netherlands, Germany, and Sweden, the notion of survey climates has remained an important concept. They showed that a country's survey climate can abruptly change, for example, when people lose trust in surveys and official statistics, as was the case in the 1980s with the Metropolit study in Sweden[28] and the public debates regarding the censuses in Germany and the Netherlands.[29] With respect to the United States, Harris-Kojetin and Tucker (1999, p. 180) found that the indicators of macro-level political and economic conditions, such as decreasing unemployment and increasing consumer expectations for the economy,

[28] Metropolit was a research project that collected register information for a sample of 15,000 persons born in 1953 without their consent and knowledge. When this data collection became public, it caused response rates in all Swedish surveys to decrease.

[29] The census debates in Germany and the Netherlands emanated from World War II scars and the fear of Big Brother, which resulted in enormous public and media attention.

were related to refusal rates in a current Population Survey between 1960–1988. Also Groves and Couper (1998) supplemented the data with metadata on information that might relate to the survey climate, such as the social environmental correlates of survey participation (p. 173).

Models and studies that inlcude the factors of the meso (survey specific level) and macro (indiviudal level) levels in their analysis of the development of response rates are found less often. Research on the effects of the macro level (individual level) on response rates is rare. Blom (2009) also has described differences on the macro level, such as in the population distribution of the characteristics associated with a propensity to respond (e.g., household size, urbanity) that need to be taken into account when analyzing the development of response rates (p. 9).

Blom (2009) has identified, among other factors, the meso level characteristics of surveys, such as the particular survey agency, different interviewer skills, and so on (p. 9). De Leeuw and de Heer (2002, p. 44), who also performed research on the cross-national comparabilty of response rates, have added another difference based on survey characteristics, such as the type of survey (panel or cross-sectional). They found differences between countries with respect to the extent of the decline of response rates.[30] Based on this study, de Leeuw and de Heer (2002) concluded that “Nonresponse is indeed an increasing problem in the developed world” (p. 52). Couper and de Leeuw (2003) also examined the response rates of three cross-national surveys (the International Social Survey Programme, International Adult Literacy Survey, and the Trends in International Mathematics and Science Study).

[30] The time series of the analyzed survey differed. In some countries data was available from 1972 to 1980 in other from 1990 to 1994.

They also found different levels of response rates, which they considered to be caused by the survey designs and differences in societies.

The above described important studies reflect a *research gap*. The overview of the previous studies in this research field as described above lack a European context, since most of them are based on US American studies. De Leeuw and de Heer (2002, p. 52) found the same trend in Europe. However, de Leeuw and de Heer (2002) analyzed surveys that were conducted at different points in time and on different topics, as well as different variables and studies. Thus, the effect of the response rate on nonresponse bias may be overlapped by this noise in the data. Moreover, the comparability of response rates is not provided since response rates are not calculated in a cross-national standardized way. In addition, the time series data ends in the mid-1990s. Therefore, these studies might be less suited for comparability across surveys, time, and countries. In addition, most of the previous research has focused on the macro level to explain the development of response rates. A focus on the meso level, including survey characteristics and fieldwork efforts, has been very limited. Also, the analysis of effects of the micro level (individual level) on response rates has been limited. The analysis conducted in the present narrows the research gap by systematically analyzing the effect of factors on the micro and meso level on response rate.

2 Literature on Nonresponse and Nonresponse Bias

This section provides background information for the analysis part related to nonresponse bias. These are Part B (Response Rates and Nonresponse Bias), Part C (Fieldwork Efforts and Response Rates) and Part D (Fieldwork Efforts and Nonresponse Bias).

First in Section 2.1 problems related to low response rates or high nonresponse rates are described. In Section 2.2 the literature on reasons for nonresponse is summarized. This is followed by the three sections (Section 2.3, 2.4 and 2.5) where the hypotheses for the analysis chapters are formulated and the literature background for them is described.

2.1 Nonresponse and Nonresponse Bias

In short, *respondents* are those individuals with whom an interview has been achieved. *Nonrespondents* "are those missing from a probability sample" (Singer, 2006, p. 637).[31] In this section, problems related to low response rates and high nonresponse rates are introduced. It examines why nonresponse might be a problem for data quality, and provides a special focus on nonresponse bias.

Low response rates or nonresponse is a problem when respondents and nonrespondents systematically differ in the variable(s) of interest. In this case, nonresponse is selective, and certain groups may be underrepresented (Couper & de Leeuw, 2003, p. 165). This difference between respondents and nonrespondents is referred to as *nonresponse bias* (Groves, 2006; Schupp & Wolf, 2015, p. 13).

The differences between respondents and nonrespondents depend on the variables of interest, which differ between surveys. Depending on the interests of a survey researcher, certain variables will be central, for example, income in a survey about poverty, and attitudes towards Europe in a political survey. Nonresponse only affects the quality of survey data when the composition of a sample with a low response rate differs from the composition of a sample with high or 100 percent response rates. For example, this may be the case in a survey on income when only persons with low or medium income take part, and rich individuals do

[31] For the definition, calculations, and discussion of response and nonresponse rates, see Section 2.2 in this chapter.

not participate.[32] Thus, nonresponse bias is not a survey characteristic (as is the overall response rate); rather, it is a characteristic of the variable of interest. Nonresponse bias is always variable specific (Groves, 2006).[33]

Not only the variable or interest is important, but also the correlation of the variable of interest with the probability to respond (Merkle and Edelman, 2002). In cases in which the probability to respond may be related to the probability to participate, nonresponse cannot be ignored. In cases in which the topic of a survey is not related to the probability to participate, nonresponse bias might be ignored. Little and Rubin (2002) have suggested that when nonrespondents are missing completely at random (MCAR) from a sample, an (unadjusted) estimate will be unbiased, since it does not differ systematically with respect to the survey items that contribute to the estimate. So in this case, nonresponse is the result of pure chance, and MCAR nonresponse is not a problem because nonresponse bias does not exist (Couper and de Leeuw, 2003, p. 166).

In cases in which nonrespondents differ systemically from respondents, nonrespondents are not missing at random (NMAR), and so estimates will be biased. This is the case unless the differences between respondents and nonrespondents can be fully explained (in a statistical sense) by other available data that correlates with the survey estimate and the propensity to respond. If this is the case, estimates can be adjusted to produce an unbiased estimate where the nonrespondents are missing a random (MAR). It can be said that: “Bias flows from nonresponse when the causes of the nonresponse are linked to

[32] Another example of nonresponse that cannot be ignored is a US study on HIV positive individuals that showed that the central characteristic (of being HIV positive) was related to the probability to respond. HIV positive persons tend to respond less often to surveys on HIV due to the stigma of the virus. Thus, the proportion of people who are HIV positive is underestimated (Groves, Fowler, Couper, Lepkowski, & Singer, 2004, p. 180).

[33] For the methods used in the present study to detect nonresponse bias, see Groves, 2006, pp. 654-656; Schnell, 1997, p. 135; Stoop, Billiet, Koch, & Fitzgerald, 2010, p. 280.

the survey statistics measured" (Groves, Fowler, Couper, Lepkowski, & Singer, 2004). Thus, it cannot be said that higher nonresponse rates lead to greater nonresponse bias: "response rates are in fact a poor indicator of nonresponse bias" (Peytchev, 2013, p. 107).

Unfortunately, in most cases, the researcher does not have enough information to know whether nonresponse is caused by some of the attributes important to a survey.[34] So, researchers have to decide, in absence of this information, whether processes that are linked to the probability to respond might be related to survey variables.

In addition to the issue of nonresponse bias just described, the inference and precision of estimates are also important issues in the discussion of nonresponse. Even if the differences between respondents and nonrespondents are small, nonresponse reduces the number of completed cases available for analyses, and thus may threaten inference by increasing confidence intervals and standard errors relative to a survey of the same initial sample size with no nonresponse (Peytchev, 2013). The smaller the variance is, the higher the precision of the outcome (Bethlehem, 2002; Stoop et al., 2010, p. 29 f.).

Another aspect related to nonresponse and how this might affect data quality is related to the sample design. The probability of obtaining an interview depends on the selection probabilities specified in the sampling design, and the unknown probabilities to respond. When adjusted or responsive designs are used, these assumptions for nonresponse, selection probabilities are used in the estimation procedures. These may affect the sample design and lead to biased estimates.[35]

[34] This might be the case especially in multi-purpose surveys.
[35] See http://www.risq-project.eu/indicators.html.

Another aspect that might be introduced by nonresponse is a post-survey adjustment to compensate for nonresponse, which may affect the variance of estimates in another way. The application of differential weighting factors tends to inflate the variance of the estimates (Little & Vartivarian, 2005). A post-survey adjustment to compensate for nonresponse usually is based on only a few key demographic variables. Weighting with these particular characteristics produces a sample that resembles the population that embodies these characteristics. Thus, this procedure does not guarantee representation of the full population on the substantive variables of interest. Full representation is only possible when nonrespondents are “missing at random” (see Couper & de Leeuw, 2003, p. 168).

The extent of nonresponse bias on an unadjusted design-based estimate depends, on the one hand, on the difference between respondents and nonrespondents, and on the other, on the response rate. So to reduce nonresponse bias, two options are available: a) have a nonresponse rate of zero, which means no nonrespondents or b) no difference, with respect to the variables of interest, exists between respondents and nonrespondents. Both options are more or less hypothetical, since a 100% response rate and a lack of difference between response rates and nonresponse rates do not exist in empirical practice. High response rates limit nonresponse bias, but “… a high response rate does not guarantee there will be no error” (Couper & de Leeuw, 2003, p. 166).

Besides all methodological aspects nonresponse may increase survey costs. To compensate for nonresponse, the initial sample needs to be increased, which increases the cost of the survey (Couper & de Leeuw, 2003, p. 166; Groves, Fowler, Couper, Lepkowski & Singer, 2004). As Groves, Fowler, Couper, Lepkowski, and Singer (2004) have argued, it is “common practice to maximize response rates within the cost constraints of the survey” (p. 181).

After many years of ongoing research, the question of whether nonresponse harms the quality of a survey has not been determined. Groves, Fowler, Couper, Lepkowski, and Singer (2004) have summarized this long discussion with their finding that “Sometimes, nonresponse harms the quality of the survey statistics; sometimes, it does not” (p. 178).

2.2 Reasons for Nonresponse

Over the last decades, several explanations have been proposed nonresponse (for example Dillman, 1978; Groves et al., 1992; Groves & Couper, 1998; Hox, de Leeuw, & Vorst, 1996). To systematically provide an overview of the explanations for nonresponse, this study has grouped them in different levels (Figure 3).[36] Basically, in the literature, three different levels have been proposed as influences on response and nonresponse. These factors may exist on the macro level (e.g., society, culture, economic situation, social and cultural context), on the meso level of the survey itself (e.g., survey design, type of survey organization [university, commercial, statistical office], the number of available interviewers), and the micro level (e.g., respondent, interviewer) (cf. de Leeuw & de Heer, 2002; Groves & Couper, 1998).

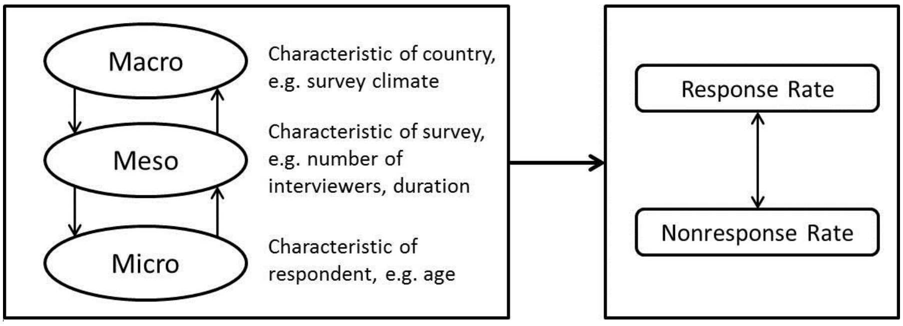

Figure 3. Reasons for response and nonresponse at the macro, meso, and micro levels.

[36] This is the basic for the hypothesis of the effect of response rates and on different socio-demographic variables and the effect of fieldwork efforts as meso variable on response rate (Analysis Part B).

As de Leeuw and de Heer (2002) have suggested, “from a theoretical point of view, influences on all levels are important” (p. 42). They also point out that most of the models for nonresponse use a “multilevel” conceptual framework that includes different levels of influences.

The following sections provide information for the hypotheses analyzed in the analysis in Part B (Response Rates and Nonresponse Bias), Part C (Fieldwork Efforts and Response Rates) and Part D (Fieldwork Efforts and Nonresponse Bias). Analyses in Part B on nonresponse bias are conducted for the socio-demographic variables for the respondents, which are at the micro level. Analysis in Part C is conducted at the meso level (survey level). Analysis in Part D is a combination of effects on the meso level (fieldwork efforts) and the micro level (socio-demographic variables at the respondent level). This section provides the rational why the micro and meso level is further analyzed. For literature that allows drawing hypothesis see Section 2.3, 2.4 and 2.5.

2.2.1 Micro level

On the micro-level, the personal characteristics of a respondent are used to explain the response or nonresponse to a survey. Groves and Couper (1998) have suggested that the personal characteristics and behavioral basis of the nonrespondents are essential to understanding the consequences of nonresponse (p. 25). They have found that the socio-demographic characteristics of the household or sample person—such as age, gender, family status, education, income, or urbanity—are not directly related to participation in a survey. In a sense, these researchers have acknowledged that the decision of nonrespondents not to take part in a survey is well founded: “In contrast to the view that nonrespondents actions are in some sense based on ignorance and lack of sophistication, this approach attempts to identity costs and benefits of responding from the sample persons perspective" (Groves & Couper, 1998, p. 45).[37] In contrast to the assumption that nonresponse is a well-founded decision of a respondent, Schnell (1991) has suggested that participation is not linked to any characteristics of a respondent; rather, the decision to respond is more or less an ad-hoc decision.

In addition to the respondents' characteristics and decision to take part or not in a survey, participation also depends on the at-home patterns of the sampled persons and on the interviewers' contact attempt strategy (Groves & Couper, 1998; Kortmann & Halbherr, 2009).[38] The explanation for the low participation rates of some socio-demographic groups may lie in the fact that they are at home less often, and thus interviewers' opportunities to contact these potential respondents is low.

Research has shown that the following groups of people are more difficult to contact: younger respondents (Lynn, 2003), individuals living in one-person households, people living in big cities, people with high education or high income (Durrant & Steele, 2009; Goyder, 1987). Immigrates also are considered to be difficult to contact (Koch, 1997) because they spend prolonged time periods in their home countries (Blohm & Diehl, 2001) and are also more likely to live in urban areas where contactability is more difficult (Feskens, Hox, Lensvelt-Mulders, & Schmeets, 2007). The respondents who are assumed to be at home more often are women because they often are taking care of children or the elderly (Groves & Couper, 1998; Stoop, 2005). Also, persons with low education and non-nationals are underrepresented in surveys (Hartmann & Schimpl-Neimanns, 1992; Helmschrott & Martin, 2014; Koch, Halbherr; Stoop & Kappelhof, 2014; Stoop, 2005). The low participation rates

[37] Many theories that attempt to explain survey participation are based on rational choice theory, for instance, the "opportunity cost hypothesis" (Groves & Couper, 1998), social exchange theory (e.g. Dillman, 1978; Goyder, 1987), or social isolation theories (Groves & Couper, 1998).

[38] The focus on this study is on face-to-face surveys only.

of persons with low education has been explained by suggesting that that they are less interested in surveys, especially when they think they may not know enough about the survey topic. The underrepresentation of non-nationals is explained by suggesting that they move more often, and so interviewers have difficulties tracking their new addresses, and also language problems when contacting non-nationals. In case of a non-contact, the eligibly cannot be determined. In addition to tracking the new addresses, interviewers have also reported language problems regarding the contacting of non-nationals (Stoop, 2005; Koch, 1997).[39]

Blom et al. (2010) have described cross-national findings regarding survey-specific effects on nonresponse bias and the response rate (p. 340f.). They found that factors that were positively associated with contact rates included demographic factors like larger average household size and a higher proportion of young children; and also factors at the survey level (meso level) such as more lenient rules for sampling and respondent selection.

In any case, the effects of socio-demographic characteristics are central to discussions about the response rate and nonresponse bias. Although Peytcheva and Groves (2009) have carried out a study related to demographics and substantive variables, a more systematic analysis of these factors in relation to the response rate would provide additional insights to the existing literature.

The effect of response rates on nonresponse bias with respect to socio-demographic variables (such as age, gender, nationality, education, household size and family status) is discussed in Analysis Part B of the present study. The effect of fieldwork efforts on nonresponse bias for socio-demographic variables is analyzed in Part D.

---

[39] For further literature review on the inclusion of certain subgroups see Section 2.6.

2.2.2 Meso level

On the meso level, explanations are related to the characteristics of a survey, the field organization, and the interviewer—for example, the topic of a survey, the duration and timing of fieldwork, the saliency of the topic, the mode, the survey design, and the incentives. Also, the process quality of the survey organization (sufficient number of interviewers, interviewer workload, interviewer training and briefing) plays a role. The survey management and the implementation of the survey itself are associated with country differences in outcomes (see de Heer, 1999; de Leeuw & de Heer, 2002). Among others de Heer (1999); Groves and Couper (1998), Couper and de Leeuw (2003, p. 165), Sakshaug, Yan, and Tourangeau (2010) have described (at the meso level) the effects of the characteristics of surveys, field organization, and interviewers on nonresponse.

For example, research on the effects of incentives (e.g., Pforr et al., 2015; Singer & Ye, 2013) or the leverage-salience of survey topic (Groves, Singer, & Corning, 2000) or interviewed behaviour (Blohm, Hox, & Koch, 2007) has found that factors at the meso level might have influence on the propensity of respondents to participate in the survey. . This finding points to an assumed relation of the effects of survey-releated characteristics to nonresponse bias.

The effects of these meso level characteristics or fieldwork efforts on the response rate are systematically analysed in Analysis Part C of the present study. In addition, the implicit association between fieldwork efforts and nonresposne bias with respect to socio-demographic variabels are examined in Analysis Part D.

2.2.3 Macro level

The macro level can be described as the level of society, culture, and other country specific factors like the economic situation, the unemployment rate, the discussions that are going on, and the media context in a country. This level is also often, as mentioned above, described as the survey-taking climate of a country.

Many studies have examined the macro level effects on response rates. For example, the effects of certain societal, cultural, and economic factors on response rates have been analyzed for a variety of variables. The following paragraphs describe some of these macro level factors, but since the researcher cannot directly affect this level of influence, these factors are not further examined in this study.

Previous research on the effects of the survey climate has been conducted by various researchers. For example, studies have shown that response rates vary systematically with the economic and political conditions within a country (Harris-Kojetin & Tucker, 1999).[40] Also, the general differences between response rates in rural and urban regions have been a well-documented phenomenon in several countries (e.g., Foster & Bushnell, 1994; Goyder, 1987).

Further examples of the survey climate regarding the potential for interview fatigue can be found in various studies. For example, the SOEP revealed a rapid decline in the response rate in Eastern Germany after reunification. After an eagerness to participate in surveys in 1990, which resulted in very high response rates, these rates decreased in following years to a level similar to the former Western Germany (TNS Infratest Sozialforschung, 2011, p. 23).[41] Bronner (1988) drew a similar conclusion when investigating the response trends of immigrant groups in the Netherlands. [42] Gengler (2016) also showed a

[40] For example, they found evidence of a short-term relationship between the refusal rate and political and economic conditions.
[41] In the first wave of the SOEP in Eastern Germany, a response rate of 70% was achieved for first time respondents (TNS Infratest Sozialforschung, 2011, p. 23ff.).
[42] Recent immigrants to the Netherlands (mainly from North Africa) had a higher response rate than immigrants who lived in the country for years (mainly from Southern Europe).

similar effect prevailing in Qatar. Couper and de Leeuw (2003, p. 170) have observed the same trend in declining response rates for countries and regions in transition.

Also, softer factors like public debates may potentially influence response and nonresponse rates. Couper and de Leeuw (2003, p. 171) have described this effect in Germany and Sweden with the public debate about privacy issues and research intrusion, and in the USA where telemarketers have rapidly spread.

2.3 Hypothesis: Response Rate and Nonresponse Bias (Part B)

In general, researchers have assumed that the higher the response rate is, the better the data quality. Surveys with low response rates are considered as having problematic data quality because of a higher risk of high nonresponse bias (Goyder, 1987; Groves et al., 2008; Groves & Peytcheva, 2008; Peytcheva & Groves, 2009). Thus, many survey guidelines and specifications require high response rates (European Social Survey, 2009; Johnson & Owens, 2003; Kreuter, 2013; Organisation for Economic Co-operation and Development, 2010).

However, some researchers have questioned the assumption that low response rates are equivalent to low quality. For example, Lelkes, Krosnick, Marx, Judd, and Park (2012) and Holbrook, Krosnick, and Pfent (2008) were able to show that studies with low response rates can have more exact estimates than studies with higher response rates (e.g., Curtin et al., 2000; Groves & Heeringa, 2006; Groves, Presser, & Dipko, 2004; Keeter et al., 2000; Merkle & Edelman, 2002; Schouten, 2004). Other studies have shown that studies with higher response rates might even have a higher nonresponse bias than studies with lower response rates (Koch et al., 2014; Kohler, 2007; Stoop, 2005). The most cited meta-analysis of the relationship of response rates and nonresponse bias was carried out by Groves (2006, p. 659), which showed that low response rates may increase the risk of nonresponse bias (Groves, 2006, p. 659).[43]

Based on these previous studies, the hypotheses on the relation between response rates and nonresponse bias, which is analyzed in Part B, can be formulated as follows:

Hypothesis 6: The higher the response rate is, the lower the nonresponse bias.

This hypothesis can be applied to the overall level of bias of a survey (for the pooled cross-sectional analysis of the overall level of bias, see Section 12). Additionally, this hypothesis can be applied to variable-specific nonresponse bias (for analyses, see Section 13, for a further discussion of variable-specific effects, see Section 2.6). The *hypothesis for nonresponse bias of separate socio-demographic variables* is the same as the general hypothesis (Hypothesis 6):

The higher the response rate, the lower the nonresponse bias for the variables gender, age, working population, family status, education, nationality and household size.

2.4 Hypothesis: Fieldwork Efforts Response Rate (Part C)

Most high quality surveys, such as the ESS, have guidelines and specifications for researchers and fieldwork organizations relating to how to implement fieldwork (e.g., European Social Survey, 2013; Organisation for Economic Co-operation and Development, 2010; Stoop, Koch, Halbherr, Fitzgerald, & Widdop, 2014). In the survey literature, the use of incentives, brochures, and interviewer briefings in general are considered to be additional efforts of fieldwork. In particular, these measures are used to enhance the response rate (Koch, Fitzgerald, Stoop, Widdop, & Halbherr, 2012; Stoop et al., 2014).

[43] For further details, see Section 1.2 and the discussion on response rates as a quality indicator.

The characteristics of good fieldwork processes are based on the findings of Dillman (1978), Stoop et al. (2010) and Luiten (2013), which include the following: the method used to contact respondents (advance letters, brochures, and incentives), the length of the fieldwork, and interviewer-related factors (length of the interviewer briefing, personal briefing of interviewers, payment of interviewers, interviewer training, and interviewer experience).[44]

The general hypothesis prevailing in the survey literature is that the higher the fieldwork efforts of a survey is, the higher the response rate (Couper & de Leeuw, 2003; de Heer, 1999; Groves & Couper, 1998; Pforr et al., 2015; Sakshaug et al., 2010). A relationship between well-implemented fieldwork and high data quality is assumed. Thus, a hypothesis, which is analyzed in Part C, can be formulated as follows:

Hypothesis 7: The more fieldwork efforts—such as the use of incentives, interviewer briefings, and so on—are implemented in a survey, the higher the response rate.

This following hypothesis can be applied to an effort-specific relationship between separate fieldwork efforts and the response rate:

Hypothesis 8: Countries that use additional fieldwork efforts (e.g., send an advance letter, or use a brochure or incentives) achieve higher response rates than countries that do not use these fieldwork efforts.

Studies often have distinguished between groups in society who are hard to contact and those who are reluctant to cooperate. For example, the studies of Goyder (1987), Groves and Couper (1998), and Stoop (2005) have provided overviews of these kinds of studies. The

[44] For details on the fieldwork effort and the fieldwork effort index, see Section 16 and 17.

distinctions between these groups is based on the likelihood that a member of a household is at home (Smith, 1983), and on her/his willingness and ability to cooperate. Factors that are mentioned in relation to survey cooperation are labor force participation and socio-economic status, as, for example, indicated by age, income, education, and urbanicity (Goyder, 1987, p. 84). Based on this literature, fieldwork efforts are considered to have different effects on different respondent groups.[45] For further details on the effects of fieldwork efforts on the response rates of certain socio-demograhic groups, see Section 2.6.

Therefore, hypotheses on the different effects of refusals and non-contacts can be stated as follows:

Hypothesis 9a: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentives) achieve lower <u>*non-contact rates*</u> *than countries that do not use these fieldwork efforts.*

Hypothesis 9b: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentives) achieve lower <u>*refusal rates*</u> *than countries that do not use these fieldwork efforts.*

Also, a hypothesis can be formulated on the overall effects of fieldwork efforts (measured by a compound index)[46] and response rates:

Hypothesis 10: The higher the fieldwork effort (the score on the compound fieldwork effort index) is, the higher the response rate.

[45] The differentiation of respondent groups is based on different socio-demographic variables.
[46] For details on the construction of the compound fieldwork effort index, see Section 17.1.

Applying this hypothesis to a longitudinal analysis—the expected change over time regarding the implemented fieldwork effort and the response rate—the following hypotheses can be formulated:

> *Hypothesis 11: A correlation exists between the change in fieldwork efforts (in the preceding compared to the subsequent round) and the change in response rates (in the preceding compared to the subsequent round).*

> *Hypothesis 11a: When countries put <u>more effort</u> into fieldwork than in the previous round, the response rate increases compared to the previous round.*

> *Hypothesis 11b: When countries put <u>less effort</u> into fieldwork than in the previous round, the response rate decreases compared to the previous round.*

> *Hypothesis 11c: When countries put <u>the same effort</u> into fieldwork as in the previous round, the response rate remains the same as in the previous round.*

2.5 Hypothesis: Fieldwork Efforts and Nonresponse Bias (Part D)

The methodological literature analyzes the link between fieldwork efforts and response rates (see Section 2 and 2.3) or the link between response rates and nonresponse bias (see Section 2 and 2.4). Little research is available on the direct link between fieldwork efforts [47] and nonresponse bias. The survey literature mainly includes this direct link assumption but provides very little empirical research. Research into the relation of the fieldwork effort and nonresponse bias has been conducted only for very specific fieldwork efforts, such as the relationship of the number of contact attempts and nonresponse bias (Wood, White, & Hotop, 2006) or the number of contact attempts and the coefficients of

---

[47] For details on the fieldwork effort and the fieldwork effort index, see Section 16 and 17.

variations (Durrant, 2016). A systematic analysis of the effects of fieldwork efforts on nonresponse bias is lacking. To close this research gap, Part D provides a systematic analysis.

In the survey literature, two assumptions prevail with respect to the relation of fieldwork efforts and nonresponse bias. First, an increase in the fieldwork effort is said to decrease the underrepresentation of the “difficult to get” respondents, groups who are less likely to take part in surveys and thus are underrepresented (Goyder, 1987; Stoop, 2005).[48] An assumption is made that with more fieldwork efforts, a data sample becomes more balanced and thus less biased (Couper & de Leeuw, 2003, p. 165). The second assumption in the literature is that an increase in fieldwork efforts leads only to the inclusion of more “easy to get” respondents in the achieved sample (Stoop, 2005). Fieldwork efforts include the practice that interviewers go back and contact initially uncooperative households, with an aim to convince them to take part in the survey (“refusal conversion”). Including more of the “easy to get” respondents in a survey would mean “more of the same”—more respondents with characteristics similar to those already included in the sample. Respondent groups who already are well included in an achieved sample could be overrepresented to an even greater extent by increasing the effort. Therefore, an increase in the fieldwork effort could lead to an increase in the response rate, but only a certain group of respondents—“more of the same”—would be included, and so the nonresponse bias would increase.[49]

[48] In many surveys, post-stratification weights are constructed using socio-demographic variables to adjust for the nonresponse bias. The present analysis does not include a post-stratification weight, only design weights.

[49] Besides this aspect or (unit) nonresponse bias, the quality of the answers of the unengaged respondents who are recruited during the refusal conversion might be of poorer quality, compared to the answers of respondents of earlier stages (Kreuter, Müller & Trappmann, 2010). Also, the persuasion of respondents during refusal conversion raises ethical questions about the voluntary nature of participation in the survey and the right to privacy (Dutwin et al., 2014).

Based on the first assumption found in the literature (previously described), the following hypothesis can be formulated on the relation of the fieldwork effort and nonresponse bias:

Hypothesis 12: The higher the fieldwork effort is, the lower the nonresponse bias.

Hypothesis 12 is tested for a variety of socio-demographic variables.[50] The purpose of doing this variable-specific analysis is that previous research, such as the meta-analysis on the response rate and nonresponse bias by Groves and Peytcheva (2008), has found that most of the variance in the nonresponse bias is within-survey rather than across-surveys. Based on this finding, a variable-specific analysis is carried out to determine the relationship of fieldwork efforts and different nonresponse bias indices of the socio-demographics. The variables analyzed are *gender*, *age*, *working population*, *family status*, *education*, *nationality* and *household size* (see Section 23).

For the *variable level analysis*, the same hypothesis as described above is utilized:

The higher the fieldwork effort is, the lower the nonresponse bias for the variables gender (male), age, working population, family status, education, nationality and household size.

2.6 Hypothesis: Variable Specific Effects (Part B and D)

Based on previous studies, it can be concluded that the effects of response rates on nonresponse bias are variable specific (see Analysis Part B). Also, the effects of fieldwork efforts on nonresponse bias are variable specific (see Analysis Part D). Therefore, a variable-

[50] As background for the selection of the variables for this specific analysis, I refer to Section 2.6 and also to 2.4 Part D Section 23 where these literature references are further described.

specific analysis seems a reasonable approach to providing deeper insights into the overall process. A comprehensive overview across countries and over time—using data from the ESS ranging from 2002 to 2010—with respect to the effects of fieldwork efforts enabled an analysis of the effects of these efforts on variable-specific nonresponse bias.

Based on the literature described below, hypotheses for certain socio-demographic subgroups were formulated. The findings of these studies concerning nonresponse bias and contactability have been mixed regarding whether a significant correlation exists between the tested variables. Also, the nature of the relationship (positive or negative) differs across studies (for an overview of this correlation, see Luiten, 2013, p. 83). The common thread in all these studies is that the effects of the response rate or fieldwork efforts affect the various socio-demographic groups differently. The following paragraphs provide an overview of the literature on the effects of fieldwork efforts on variable-specific nonresponse bias.

Literature on the relationship of the socio-demographic variable *gender (male)* and nonresponse bias can be found in previous studies by Koch, Halbherr, Stoop, and Kappelhof (2014), Kohler (2007), and Sodeur (1997). The at-home patterns of males differ from the average at-home patterns of females. Males are at home less often (Stoop, 2005, p. 69 f.), and so they are more difficult to include in surveys. Based on these findings, the variable *gender* is included in the analysis. Hypotheses on the positive effects of fieldwork efforts on males in the net sample, and the effects of the response rate on the nonresponse bias with respect to the variable *male* are formulated as follows:

Part B: The higher the response rate is, the lower the nonresponse bias of the variable gender (male).

Part D: The higher the fieldwork effort is, the lower the nonresponse bias for the variable gender (male).

The same rationale of the at-home pattern applies to younger persons who, in general, are less often at home, and therefore, are more difficult to contact. Analyses of the effects of the variable *age* on nonresponse bias have been provided by the studies of Kortmann and Halbherr (2009), Hoolbrook, Green, and Krosnick (2003, pp. 94, 110), and Lynn (2003). In addition, Groves and Couper (1998, p. 133) did an inventory of the literature on the correlation of *age* and overall response rate and found mixed results—different authors provided different explanations for the underrepresentation of young or older respondents. Elderly persons may be socially isolated, and thus have lower cooperation rates (Groves & Couper, 1998, p. 133). A higher incidence of health problems make the elderly physically or mentally less able to participate (Cohan & Duffy, 2002, p. 21f.). Hoolbrook et. al. (2003, pp. 94, 110) have found that the elderly are reluctant to admit strangers into their homes, and thus have a higher risk of refusal.

In summary, a significant relationship between the variable *age* and a nonresponse bias may be potentially possible. Therefore, an analysis of this relationship seems appropriate. Hypotheses of this potential relationship are as follows:

Part B: The higher the response rate is, the lower the nonresponse bias of the variable young persons/old persons.

Part D: The higher the fieldwork effort is, the lower the nonresponse bias for the variable young persons/old persons.

With respect to the variable *working population*, Lynn et al. (2002, p. 142) have pointed to the difficulty of contacting employed persons due to their at-home pattern. Therefore, persons with an occupation may be more likely to be underrepresented in surveys. Based on the literature, the following hypotheses for variable *working population* can be formulated:

Part B: The higher the response rate is, the lower the nonresponse bias of the variable working population.

Part D: The higher the fieldwork effort is, the lower the nonresponse bias for the variable working population.

Different explanations have been provided for the underrepresentation of low educated persons and the overrepresentation of highly educated persons; for example, see the discussions on the relationship of the variable *education* and nonresponse bias in the research of Helmschrott and Martin (2014), Holbrook et al. (2003, pp. 94, 110), and Koch, Halbherr, Stoop, and Kappelhof (2014). One explanation for the underrepresentation is that interviewers are more attracted by persons with higher education (Brehm, 1993, p. 31). Groves and Couper (1998, p. 128) have argued that persons with low education are less prone to cooperate. The cognitive burden may be less for highly educated persons, and so they are more likely to cooperate (Tourangeau & Smith, 1996; Holbrook et al., 2003, p. 82). Thus, an examination of the effects of fieldwork efforts to contact or convince respondents to participate in a survey is essential. Based on the previous findings in the literature, the present study included the variables *high education* and *low education* in its analysis. The hypotheses are formulated as follows:

Part B: The higher the response rate is, the lower the nonresponse bias of the variable low education/high education.

Part D: The higher the fieldwork effort is, the lower the nonresponse bias for the variable low education/high education.

Based on the findings of the previous literature, the present study also examined the relationship between the variable *nationality*—which includes nationals and non-nationals—

and nonresponse bias (e.g., see Blohm & Diehl, 2001; Feskens et al., 2007; Holbrook et al., 2003, pp. 94, 110; Helmschrott & Martin, 2014; Koch, 1997; Koch et al., 2014). These studies have shown that immigration groups are perceived as less likely to respond than an indigenous population. Voogth (2004, p. 106) has argued that in the Netherlands, this lower response is due to the fact that non-nationals are generally concentrated in specific areas, and probably are less integrated and thus less accessible to survey invitations. More current studies also found an underrepresentation of migrants (e.g., with respect to Germany, see Kortmann & Halbherr, 2009).

The hypotheses are for variable *nationality* are formulated as follows:

Part B: The higher the response rate is, the lower the nonresponse bias of the variable nationality.

Part D: The higher the fieldwork effort is, the lower the nonresponse bias for the variable nationality.

Previous studies that examined the relationship between the socio-demographic variable *household size* and the nonresponse bias have found two major effects. First, *household size* affects the likelihood of refusal, and second, the likelihood of being contacted. Studies by Durrant and Steele (2009), Goyder (1987), and Campanelli, Sturgis, and Purdon (1997, p. 3ff.) have shown that people living alone have a higher risk of refusal, and thus, are more likely to be underrepresented. Second, Goyder (1987), Koch (1993), Lynn et al. (2002), Lynn and Clarke (2002), and Stoop (2005, p. 70) have conducted studies on household size and its effects on contactability, and found that the accessibility to the households of their sampled populations was related to the probability that a household member would be at home. This probability is related to the number of household members and also the occupational status of these individuals. It can be concluded that household size plays a role in contactability, and that higher non-contact rates may be associated with 1-person households. For these two reasons, the variable *household size* is included in the analysis of the present study on the relationship of fieldwork efforts and nonresponse bias.

Two different household sizes were analyzed—first, persons living in *one-person households*, and second, larger households with *five-or-more persons households*. The hypotheses are formulated as follows:

Part B: The higher the response rate is, the lower the nonresponse bias of the variable one-person household/five-or-more persons household.

Part D: The higher the fieldwork effort is, the lower the nonresponse bias for the variable one-person household/five-or-more person household.

3 Discussion

The European Social Survey (ESS) is the first survey dealing with many of the previously described challenges that have undermined the comparability of response rates between countries. In particular, the following analyses help to close the research gap because this survey provides cross-national comparable and longitudinal data.

The data of the ESS enables to close the research gap for various aspects. First, the ESS allows analyzing response rates in European countries, whereas previous research was focused mainly on the US. Second, more up- to- date information from 2002 to 2010 is included in this analysis. Third, because of the cross-national comparable set-up, the ESS data enabled a comparison of the calculation of response and nonresponse rates and the composition of these rates.

A further research gap that can be identified is the relation between response rate and nonresponse bias. As discussed previously, response rates have an effect on the maximal nonresponse bias. The current literature provides some analysis of this relationship, but a systematic test has been missing in the research. Analysis Part B of the present study has addressed this issue. ESS data provides a unique opportunity for researchers to access comparable response rates across time and countries, and to consider variables included in the analysis of nonresponse bias that are comparable across time and countries as well. Thus, Analysis Part B helps to close the research gap, and in addition, extends the literature by adding a cross-national analysis.

Especially relevant is the fact, that the ESS was the first cross-national survey to collect detailed call level disposition codes[51] in cross-national data and made available to the public for analysis (Stoop et al. 2003). This call-level disposition codes allows calculating response rates comparable between countries. This aspect of comparability only recently has gained more attention: "… in cross-national or cross-cultural survey research, nonresponse issues have been largely ignored" (Couper & de Leeuw, 2003, p. 157). Blom et al. (2010) have suggested that: "Research on cross-national comparisons of nonresponse is still in its infancy" (p. 341).

The analysis of one survey with the same topic is also a positive aspect of this comparison. Previous research has included different surveys, at different points in time on various topics. In contrast, the ESS disposition codes and calculation of response rates are harmonized across time and countries. This comparable data across time and countries allows

[51] Call-level disposition codes are the data collected in the contact form, which provides additional information for each call (Blom, 2009). For further details, see Section 1.2.2.

analyzing the development of response rates (Analysis Part A) over the last 10 years in Europe with a comparable data set.

The ESS has overcome issues in comparative survey research that have long been ignored in methodological research. In addition, it also has enabled the analysis of nonresponse issues. Thus, the analyses performed in Analysis Parts A, B, C and D of the present study are based on comparable data from the ESS that enables, for the first time, a valid comparison of response and nonresponse rates, and nonresponse bias cross-nationally and cross-culturally.

The research gap regarding the effect of fieldwork efforts and response rates also can be analyzed. Literature that examines the factors that affect response rates-especially at the meso-level (fieldwork and survey-specific level)- have not yet been systematically analyzed. These meso-level factors are especially important, since researchers can influence them. Although information on the details and efforts of fieldwork usually are not publically available, Analysis Part C has addressed this lack by analyzing the effects of fieldwork efforts on response rates.

Another research gap can be closed with the analysis in Analysis Part D regarding the relation of fieldwork efforts and nonresponse bias. In the discussions in the literature—on the interplay of response rates, nonresponse bias, and fieldwork efforts—only the effects of response rates on nonresponse bias, and the effects of fieldwork efforts on nonresponse bias have been analyzed. The missing link, only implicitly assumed, is the effect of fieldwork efforts on nonresponse bias. The literature assumes that the level of the response rate—whether it is high or low—moderates the effect of the nonresponse bias. The relation of fieldwork efforts on nonresponse bias as well as the effects of different levels of response rates is systematically analyzed in Analysis Part D of the present study.

Thus, the present study has provided a systematic overview of the central quality indicators for survey research (response rate, nonresponse bias, and fieldwork efforts). Thus, the longitudinal and cross-sectional nature of the analysis of the present study has provided further information on these basic concepts of data quality in cross-national comparisons.

III. Data and Methods

For the present analyses, data from the European Social Survey (ESS) and European Labour Force Survey (LFS) (for comparison purposes) were used. The datasets of the ESS were used for information on response rates, and fieldwork efforts and processes. LFS data were employed to calculate the nonresponse bias. In addition to describing these two datasets, this section describes the justifications for choosing these datasets, the harmonization of the variables between the two datasets, and the method used for the nonresponse bias calculations.

The European Social Survey (ESS) was selected because the data set included high quality data collected in a comparable standard across time and countries. The standardization of response rates (see Section 2.2) was especially important, as was the harmonization of the data collection of the socio-demographic variables (see Section 5). The Labour Force Survey (LFS) was chosen because this dataset is also cross-national and comparable across time and countries. In addition, the LFS and the ESS were selected because they collect socio-demographic variables in a comparable way across the two data sets. Also, the target populations of the LFS and ESS can be harmonized.

Negative aspects regarding the comparability of these two datasets are that not all countries that are included in the ESS are in the LFS. The European LFS does not include some participating countries of the ESS, such as Israel, Ukraine, Turkey, and Russia. Other problems are related to the different sampling frames, different legal status (opt-out option in the legal census), and the compulsory participation in the LFS versus and the voluntary participation in the ESS. In addition, in some countries, more comprehensive studies are available (e.g., the German Micro-census, Russian Census data). The problem with using these more comprehensive data files is that they are collected with a national focus, and therefore, are not comparable across countries. Using census information or information from registers would perhaps provide a more comprehensive sample, although other challenges would arise. First, the datasets from different countries and the variables would need to be harmonized. The harmonization of different national datasets (such as the German Micro-census and other national datasets) would have made the harmonization process vulnerable to national differences. This might be especially challenging when mapping information is not available, e.g., for the harmonization of education variables. Also, quality issues could arise as well; registers are not free of mistakes.[52]

Since the analyses performed by the present study focused mainly on the comparability of countries, the cross-national comparable dataset of the ESS was chosen. In addition, the following factors are essential for comparison: the verification of variables between national and international statistical offices, the implementation of international standards for the measurement of education (ISCED 97), the availability of data at different points in time and in the English language, a comparable data collection mode (face-to-face), the availability of the population, and the possibility for harmonizing the target population (age 15 and up without an upper age limit). Even though the LFS may have some deficits "it is still the far best source for our needs" (Vehovar & Slavec, 2012).[53]

4 European Social Survey (ESS)

The European Social Survey is a bi-annual survey conducted since 2001 in more than 30 nations. It is an academically driven cross-national survey carried out using face-to-face

[52] LFS data quality information is available in the document *Quality Report* (Eurostat, 2014).
[53] Due to this reason, the LFS data was selected for the post-stratification weights of the ESS data. Post-stratification weights have been available since 2015 from http://www.europeansocialsurvey.org) as well.

interviews. The samples are conducted with newly selected, cross-sectional samples in each round. The number of countries that participated in each round differs.

The ESS measures attitudes, beliefs and behavior patterns, and one of its main aims is “to achieve and spread higher standards of rigour in cross-national research in the social sciences, including, for example, questionnaire design and pre-testing, sampling, data collection, reduction of bias and the reliability of questions”[54] The central aspects of the ESS are its methodological rigor, cross-national comparability of the substantive and methodological variables, and the comparability of the processes of the survey itself.

The ESS is unique because it collects and publishes details on the process of data collection, and among other things, on fieldwork. Also, the socio-demographic variables can be made comparable with the LFS. Thus, the ESS dataset provides lots of opportunities for methodological research in general, and for research on fieldwork and response rates in particular.

ESS data can be downloaded after registration at the Norwegian Centre for Research Data (NSD) in Bergen, Norway free of charge (http://www.europeansocialsurvey.org/data/). The analysis carried out by the present study is based on the cumulative data files *ESS1e06_3. Ed 6.3* (download at 08.05.2013), which includes round 1 to 5 (data from 2002 to 2010). The data is weighted by the design weight (dweight). This weight corrects for the differences in the selection probabilities between sampling units in a country, which are computed as a normed inverse of the inclusion probabilities. Due to the longitudinal focus of the present study, only countries were included that participated in all ESS rounds between 2002 and 2010. These countries are Belgium, Denmark, Finland, France, Germany, Hungary, Ireland,

[54] See http://www.europeansocialsurvey.org/about/.

the Netherlands, Norway, Poland, Portugal, Slovenia, Spain, Sweden, Switzerland, and the United Kingdom.

5 Labour Force Survey (LFS)

The European Labour Force Survey (LFS) is a household survey designed to obtain information on the labor market and related issues. It is a rotating random sample survey covering the population living in private households. The European LFS is conducted in the 28 Member States of the European Union, two candidate countries, and three countries of the European Free Trade Association (EFTA). The data collection started in 1968. Between 1968 and 1972, the LFS was conducted as an annual survey, from 1973–1981 as a biennial survey, and since 1983 as an annual survey. From 1998, the EU LFS has step-by-step become a continuous quarterly survey.[55]

The LFS sample units are individuals, households, and dwellings, depending on the available sample frame of the countries. The data collection is conducted by national statistical institutes. To achieve a comparable data set harmonized at the European level, the data is centrally processed by the Statistical Office of the European Communities (EUROSTAT).[56] The definitions used are common to all EU Member States, and are based on international recommendations by the International Labour Organization (ILO).[57]

As recommended by the Statistical Office of the European Communities (2012), LFS data is weighted with the standard weight variable COEFF. This weight is used to correct for

[55] See http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey.
[56] For details on national implementation see Statistical Office of the European Communities (2012).
[57] http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Labour_force_survey_%28LFS%29.

differences in selection probabilities. In addition, it includes a post-stratification adjustment to adapt the LFS data to population characteristics (such as gender, age, region, and so on).[58]

Access to the LFS microdata was granted by EUROSTAT and downloaded from the GESIS-Leibniz Institute for the Social Science data archive.[59] The 2012 release version of the yearly dataset for 2002, 2004, 2006, 2008 and 2010 were used.

6 Harmonization Between the ESS and LFS Datasets

The target population of the LFS was adjusted to the ESS target population. Therefore, only individuals aged 15 years and older, who are living in private households, were included. The datasets of the ESS and LFS were aggregated per country and rounds and matched to achieve a harmonized dataset between the different points in time and the different datasets.

The different socio-demographic variables were harmonized between the different points in time and the datasets. The variables included in the harmonization process were gender, age, family status, employment status, household size, education, and nationality (see Table 1). For the harmonization process of the variables, different factors need to be considered. First, the variables need to be comparable between times within the same country. For example, if the questions in the ESS have changed, which was the case for family status, the variables need to be made comparable across time. Also, if the definition of variables, such as the laws regarding what is defined as married has changed, this needs to be considered as well. For example, in France the new status of PACS (Pacte civil de solidarité) has been introduced, which can be described as a living arrangement comparable to marriage,

[58] In addition to the correction to population statistics, several countries use additional information for weighting, such as information on unemployment or nationality (Statistical Office of the European Communities, 2012).
[59] The raw data files were first downloaded in computer software SAS. The SAS-system files were then imported to the computer software STATA, SPSS and R (depending on the analysis performed).

but also open to homosexual couples. This PACS was included in the ESS, defined as "married". Also, with respect to the LFS, the inclusion of certain groups, like the old age group (65 years and older), has changed in certain countries over time.

The second part of the harmonization process was comparability across time. Especially challenging was the harmonization of education variables and the definition of what can be considered as low and high education. The ISCED coding [60] provides comparability in general, but at the country level, especially in the United Kingdom, this definition varies. The same education, e.g., for a nurse has different types of school systems in different European countries. In Switzerland and the United Kingdom, a university degree is needed, whereas in Germany, a nurse needs a vocational training. [61]

The final dataset includes ESS metadata—such as the response rate, non-contact rate, refusal rate, rate of not able/others—and information on the socio-demographics variables from the ESS and the LFS.

[60] For further details on ISCED coding, see Section 13.6 and Organisation for Economic Co-operation and Development, 1999; and Schneider & Kogan, 2008.
[61] For further information about the harmonization between the ESS and LFS dataset, see Halbherr (2017, manuscript in preparation).

Table 1

Description of socio-demographic variables in the ESS and LFS.

Variable	Categories	ESS source variable	LFS source variable
Gender	Male Female	gndr	sex
Age	15-24 years 25-34 years 35-44 years 45-54 years 55-64 years 65-74 years 75 years and older	agea	age
Family status	Not married Married (incl. registered partnership)	marital[a] maritala[b] maritalb[c]	marstat
Work population	Not in paid work In paid work (for at least one hour in the last seven days)	pdwrk	wstator
Nationality	National of country No national of country	ctzcntr	national[d]
Education	Low education (ISCED 0-2) Medium education (ISCED 3-4) High education (ISCED 5-6)	edulvla[e] edulvlb[c]	hatlevel
Household size	Respondent lives in household comprising 1-person 2-persons 3-persons 4-persons 5 and more persons	hhmmb	hhnbpers

[a]ESS 1 and ESS 2. [b]ESS 3 and ESS 4. [c]ESS 5. [d]non-nationals recoded in one category. [e]ESS 1 to ESS 4.

7 Nonresponse Bias Calculation

This section provides an overview of the different methods that can be used for nonresponse bias calculation. It also provides the rationale why the method of the comparison with external statistics was selected.

Different methods can be used to detect nonresponse bias (see Groves, 2006, pp. 654-656; Schnell, 1997, p. 135; Stoop et al., 2010, p. 280). The most popular methods are the R-indicator (Schouten & Cobben, 2007), the Dissimilarity index (Blohm, 2006; Duncan & Duncan, 1955; Holbrook & Krosnick, 2010, p. 42; Koch et al., 2014), the use of paradata (Andridge & Little, 2011; Kreuter, 201a3; Kreuter & Casas-Cordero, 2010; Kreuter et al., 2010; Sakshaug & Kreuter, 2011; Sinibaldi, 2014), the comparison of “hard to reach” and “easy to get respondents” (Billiet, Philippens, Fitzgerald, & Stoop, 2007; Curtin et al., 2000; Fricker & Tourangeau, 2010), the Fraction of Missing Information (FMI) (Kreuter et al., 2010; Wagner, 2010; Wagner, 2012; Wagner & Couper, 2011), internal quality criteria such as the gender distribution (Koch et al., 2014; Kohler, 2007; Sodeur, 1997), the combination of data from different surveys or the integration of administrative data with surveys (Kreuter, 2013a, p. 32), and the comparison with external statistics (“gold standard”) (Groves, 2006; Hartmann & Schimpl-Neimanns, 1992; Holbrook & Krosnick, 2010; Koch, 1998; Krosnick, 1991; Krosnick et al., 2002; Levy & Lemeshow, 2013; Massey & Roger, 2013). Each of these methods has it pros and cons and can be applied to different situtations. Stoop et al. (2010) have pointed out that: “Each individual approach to measuring and correction for nonresponse bias has its flaws, sometimes serious flaws, but they can… alert the analyst to the presence of bias” (p. 280).

The present study has used the “gold standard” of comparing ESS data with external data (of the LFS). The socio-demographic sample composition used in the ESS was compared with the distribution of more accurate benchmark data to provide an indication of the degree of over- and underrepresentation of certain demographic subgroups in the ESS samples. This deviation of the ESS to the LFS data can be used as an indicator of nonresponse bias.

This comparison with external statistics ("gold standard") about population parameters is widely used and is a well-known method for analyzing the degree of nonresponse bias (Groves, 2006; Hartmann & Schimpl-Neimanns, 1992; Holbrook & Krosnick, 2010; Koch, 1998; Krosnick, 1991; Krosnick et al., 2002; Levy & Lemeshow, 2013; Massey & Roger, 2013). Other ways of calculating nonresponse bias, which are described above, are only applicable if the population statistics from the gross sample are available (for example this is a prerequisite of the R-indicator). Other forms of calculation would limit the number of nonresponse bias calculations for the variables (e.g., the internal criteria of the gender distribution). Other methods of analyzing the nonresponse bias, would limit the number of countries that could be included in the analysis due to the availability of data (e.g., if datasets were combined with administrative data, which is not available in many countries). Paradata would not be available for all countries and all rounds of the ESS, for example, since many countries still use CAPI. Therefore, the comparisons with external statistics was chosen, which enables a calculation of the nonresponse bias for a variety of variables, in many countries, and for many points in time.

Besides the comparison for nonresponse detection, another reason for selection of the comparison with the LFS is that this method of comparison of the population parameters is also used for nonresponse adjustment and the weighting of the survey data according to population statistics. This is the case when post-stratification weights are produced. For example, the ESS used the LFS to adjust ESS data to LFS data.[62]

Of course the choice of the method of comparison with external statistics has certain limitations. Obviously, to calculate nonresponse bias, it is necessary to have information

---

[62] For a discussion of the quality of the LFS and for issues concerning comparability, see Section 5 and Section 6.

about the “true” value of the variable (e.g., from the Census or from other reliable sources). Therefore, comparison is restricted to available variables, namely the socio-demographic variables. Socio-demographic characteristics are important, since they are potentially related to substantive variables, such as attitudes and behaviors (Koch et al., 2014). Thus, the findings of the present study go beyond these variables.

IV. Analysis

The analysis conducted by this study provides a comprehensive overview of the development of response rates and the relation between response rates, fieldwork efforts, and nonresponse bias in a cross-national perspective. The ESS provides data that enables this correlation to be systematically and empirically tested. Specifically, the analysis is divided into four parts. Analysis Part A analyzes the development of response rates. Analysis Parts B, C, and D further examine the indicators of data quality which are generally associated with the response rate. Analysis Part B analyzes the relation between response rates and nonresponse bias. Analysis Part C analyzes the relation between fieldwork effort and response rate and Analysis Part D investigates the relation between fieldwork effort and nonresponse bias (see Figure 1). Thus, this analysis provides the background for an empirical analysis of the three aspects of data quality in fieldwork, the response rate, nonresponse bias, and fieldwork efforts.

Different types of analysis are used in the cross-sectional analysis, such as the Pooled Ordinary Least Square Regression (POLS) analysis, descriptive analysis, OLS regression analysis, and the interpretation of the Pearson Correlation Coefficient. For the longitudinal analysis, the change between rounds is the focus. In addition, a qualitative analysis was conducted at the country level.

Analysis Part A: Development of Response and Nonresponse Rates

The response rate is considered to be a central indicator or almost synonym for survey quality.[63] Thus, the development of response rates is central in survey research. As described previously, the concern about a decline in response rates is related to the fear of decreasing quality of fieldwork (Biemer & Lyberg, 2003; Blom, 2009), the increasing risk of nonresponse bias (Goyder, 1987; Groves et al., 2008; Groves & Peytcheva, 2008; Peytcheva & Groves, 2009), and the increasing costs of doing surveys (Groves, Fowler, Couper, Lepkowski, & Singer, 2004, p. 181; Groves & Heeringa, 2006).

The hypothesis for the development of response rates described in literature is that response rates are decreasing (Brick & Williams, 2013; de Leeuw & de Heer, 2002; Goyder, 1987). Previous studies that found a decreasing trend did not necessarily have comparable setups: they had diverse fieldwork procedures and specifications, different survey arrangements (e.g., mode, timing, interviewer instructions, and so on) or different ways to calculate response rates, and often they focused on different topics. Any of these factors could bias results. In addition, most of these studies were conducted in the United States, and it remains to be seen whether their findings could be applied to European countries as well. Most of the analyses on the development of response rates were based on data gathered before the year 2000 (de Leeuw & de Heer, 2002), and so questions have been raised as to whether this development is still found in more recent data.

---

[63] This is the case despite the various problems with the response rate as a quality indicator (Groves et al., 2008; Groves & Peytcheva, 2008; Peytcheva & Groves, 2009).

The research question in Analysis Part A examines whether a trend is occurring in the development of response rates. Hypotheses are tested (see Table 2) regarding the different trends of response rates and at the different response rate levels. In addition, the (unit) nonresponse rate—the complement to the response rate—is analyzed.

The ESS does provide comparable setup and response rates calculations.[64] The "input" harmonized setup of fieldwork and the questionnaire,[65] the clear specification[66] of the survey setup, and the general cross-cultural standardized and comparable setup enable a comparison of response rates in Europe. Also the ESS does not include a variety of different surveys like previous studies (e.g., Groves & Peytcheva, 2008). An important aspect regarding the comparability of response rates is that even though the American Association for Public Opinion Research (2015) has standardized the calculation of response and nonresponse rates, these rates are not fully comparable.[67] The data from the ESS—with comparable response rates across time and country—provide an opportunity to examine the trend of response rates in Europe.[68] Using data from one comparative survey, all these factors can be kept constant.

Due to the high methodological quality, and the transparency and comparative setup of the ESS (European Social Survey, 2009, 2011, 2013a; Halbherr & Koch, 2010; Halbherr, Koch, Ederle, & Mayn, 2014; Koch, 2006, 2007, 2010), the trend of response rates in the

[64] Many European studies like the European Value Survey (EVS), European Union Statistics on Income and Living Conditions (EU-SILC), European Labour Force Survey (LFS), International Social Survey Programme (ISSP), and EUROBAROMETER do not provide the public with information on the details of response rates calculations, nor on the comparability between countries, time, and across different modes. For example, the EVS does not publish response rates. In the LFS and ISSP, the calculation of response rates differs between countries and over time.
[65] See Koch, Blom, Stoop, and Kappelhof (2009).
[66] See European Social Survey (2009).
[67] For details on the problems with the standardization and comparability of response rates, see Sections 1.2.1 and 1.2.2.
[68] For details on the response rate calculation in the ESS see Section 1.2.3

ESS can be used as a proxy for other studies that might not provide these details. Thus, the results of these analyses of the ESS are transferrable to other surveys.

Not only the survey response rate, but also the complement namely survey nonresponse rate is an important consideration for analysis (Section 8.3). Often, in literature it is described that the nonresponse rate are getting worse and people as less likely to respond or be contacted (Barbier, Loosveldt, & Carton, 2016). To understand the development of response rates, the development of the elements of nonresponse rates also needs to be analyzed. The same challenges that apply to the calculation of response rates, apply to nonresponse rates for the subgroups “refusal rate,” ”non-contact rate,” and the rate of “not able/others.” Complementary to Hypothesis 1: “Response rates are decreasing,” Hypothesis 2 regarding the development of nonresponse rates can be formulated as follows: “Nonresponse rates are increasing” (see Table 2).

Table 2

Overview Hypothesis Analysis Part A

Response Rates—in general
Hypothesis 1: Response rates are decreasing.
Nonresponse Rates—in general
Hypothesis 2: Nonresponse rates are increasing.
Nonresponse rate—detail
Hypothesis 3: Refusal rates are increasing.
Hypothesis 4: Noncontact rates are increasing.
Hypothesis 5: Rates of “not able/others” are increasing.

The following sections analyze the development of response and nonresponse rates with a focus on development at a country level. Hypothesis 1 on the development of response rates was analyzed with three perspectives in mind drawn from the literature on survey response rates. The first perspective (Section 8.1) is the general trend in response rates. What trend in response rates can be seen at the country level? Do response rates decrease as the literature suggests (de Leeuw & de Heer, 2002; Groves & Couper, 1998; Groves, Fowler, Couper, Lepkowski, & Singer, 2004: 185-187) (see Hypothesis 1)? Since the research question examines whether response rates decrease, Section 8.2 further investigates the trend of response rates—whether a continuous trend (decreasing or increasing) in the development of response rates over the ESS rounds is found at the country level (Hypotheses 1 and 2).

By looking at these three different perspectives on response rate development, a comprehensive overview of the development of response rates in the ESS at the country level is examined and then compared to most prevailing assumption regarding the development of response rates in the survey literature.

8 Development of Response Rates

This section analyzes the development of response rates in the ESS from 2002 to 2010. It includes a descriptive analysis of the mean response rates over time and per country. Also, a *t*-test is used to describe the differences in mean response rates between 2002 and 2010. A pooled ordinary least square (POLS) regression using robust standard errors is calculated, and the adjusted predictions for the mean response rates pooled in the rounds over time are provided. The estimated mean changes in response rates for each country between rounds are calculated and displayed in graphs to differentiate the various trends in response rates. Bar charts are used to visualize the trends within these groups. Regression analyses and line charts are used to further investigate the trend of response rates at different levels.

Descriptive analysis (Table 3) of response rates shows that the rates vary between countries and rounds of the ESS.[69] Response rates range between 30% (DE in ESS 5) and 76% (in PT ESS 4). The mean response rate in the ESS for countries that participated in all first five rounds of the ESS is 60%.[70]

Table 3

Response rate ESS 1 to ESS 5 (percentages)

Country	ESS 1 (2002)	ESS 2 (2004)	ESS 3 (2006)	ESS 4 (2008)	ESS 5 (2010)	Mean
Belgium	58.4	61.2	61.0	59.0	53.4	58.6
Denmark	68.4	64.2	50.8	53.6	54.9	58.4
Finland	73.3	70.7	64.4	68.4	59.4	67.2
France	43.1	43.6	46.0	49.9	47.1	45.9
Germany	51.7	51.0	52.9	42.7	53.2	50.3
Hungary	69.3	66.6	66.0	61.3	60.7	64.8
Ireland	64.4	62.0	50.4	49.2	59.8	57.2
Netherlands	67.8	64.3	59.8	49.8	60.0	60.3
Norway	65.0	66.2	64.4	60.4	58.5	62.9
Poland	72.1	73.7	70.0	70.9	70.0	71.3
Portugal	68.8	71.3	72.7	75.8	67.1	71.1
Slovenia	70.5	70.2	64.9	58.8	64.4	65.8
Spain	51.5	54.9	66.2	66.8	68.6	61.6
Sweden	69.0	65.4	65.5	62.2	51.8	62.8
Switzerland	32.5	48.5	50.0	48.8	53.2	46.6
United Kingdom	55.0	50.6	52.1	54.5	56.3	53.7
Mean	61.3	61.5	59.8	58.3	58.7	59.9

In a longitudinal perspective (Figure 4), a slight but steady decline of the mean response rate occurs from the first two rounds in 2002 and 2004 from 61% (ESS 1) and 65% (ESS 2) to the last round 59% (2010-ESS 5). This shows a decline in the response rate of over 1% every two years (ESS 1 to ESS 2: -0.2%; ESS 2 to ESS 3: -1.7%; ESS 3 to ESS 4:

[69] The response rates are derived from different gross sample sizes across time and countries. Since the unit of analysis in the present study is the realized response rates at the county level across time, the differences in the gross sample sizes are ignored, and instead, the unweighted response rates are used.
[70] The specification for participating countries (European Social Survey, 2011) in the ESS is a target response rate of 70% that every participating country should achieve. However, not all countries have achieved this goal.

-1.5%; ESS 4 to ESS 5: -1.1%). But no significant differences were detected between the first and the last round (p = .284)[71] and between the consecutive rounds.

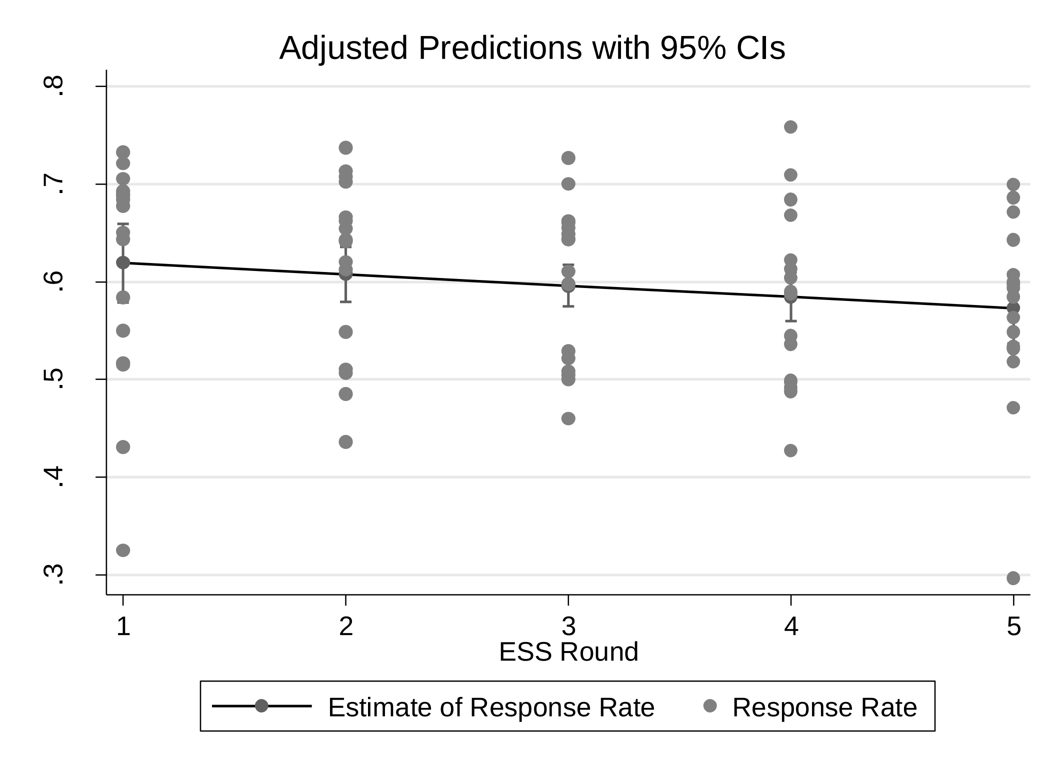

Figure 4. POLS of the development of response rates controlled for ESS rounds.

A further analysis that included the variable *time* was carried out using a pooled ordinary least square (POLS) regression. This approach was suggested by Brüderl and Ludwig (2015, p. 535) who, for descriptive purposes, advised taking into account the between variation by using a POLS regression.

The statistical assumptions of the POLS regression are based on the assumptions of the ordinary least square (OLS) regression (Wolf & Best, 2015, p. 67f., p. 75 f.; Meulemann, Loosveldt, & Emonds, 2015, p. 83; Heeringa, West, & Berglund, 2015, p. 230). These assumptions include that the relationship between the independent and dependent variables

[71] Levene test (p = .247); equal variance is assumed. A t-test of independent samples was conducted.

needs to be linear, and that the variables are of metric (ratio or interval) scale. A scatterplot can be used to check whether the relationship between the independent and dependent variables are linear, for example, by using a lowess curve, residual and partial residual plots, or a statistical "lack-of-fit-test." A scatterplot is also used to check for outliers, since linear regression is sensitive to outlier effects.

The second assumption, which is central to linear regression models, is the homoscedasticity or constant variance assumption. Homoscedasticity describes a situation in which the error term (the "noise" or random disturbance in the relationship between the independent variables and the dependent variable) is the same across all values of the independent variables. This assumption can be diagnosed by using a scatterplot, a studentized residual plot, or a Goldfeld-Quandt test.

The third assumption is that there is little or no auto-correlation. Autocorrelation occurs when the residuals are not independent from each other. This assumption can be tested by a scatterplot or a Durbin-Watson test. For a time-series cross-sectional analysis, a Wooldridge test or Modified Wald test can be used (Fortin-Rittberger, 2015).

The fourth assumption is that of normality, which can be checked with a histogram and a fitted normal curve, a Quantile-Quantile-Plot, or a goodness of fit test, e.g., the Kolmogorov-Smirnov test or Shapiro-Wilk. These four assumptions of the Gauss-Markow theorem guarantee that the OLS coefficients are the best linear unbiased estimators (BLUE) (Elff, 2015, p.21).

In addition to these four assumptions of the OLS, the fifth associated assumption is that there is no or little multicollinearity. Multicollinearity occurs when the independent variables are not independent from each other. The independence assumption is that the error of the mean has to be independent from the independent variables. Tests for multicollinearity are the correlation matrix, tolerance, variance inflation factor (VIF) and the condition index (Meulemann et al., 2015).

Also the absence of influential observations is an important aspect in a linear regression analysis (Meulemann et al., 2015, pp. 101ff.). Datasets may contain observations that are separated from the rest of the observations or data, which means that these data points have values that deviate strongly from the other observations. These cases could affect the results of a regression analysis (Allen, 1997, p. 177). Meulemann et al. (2015, p. 101) have differentiated between extreme cases: first extreme observations with a deviating score on the dependent variable, and second, cases with an extreme score on the independent variable, which are called *leverage points*. The outliers and leverage points do not necessarily influence the results in general, but can have an impact under certain conditions. Meulemann et al. (2015, p. 102) have pointed out that observations influence results if they possess leverage, and if they also are inconstant with the regression relation of other observations. Diagnostics for influential observations can be done by using DFFITS, Cook's D distance, DFBETA (Meulemann, 2015, p. 102), or a scatterplot with outliers and leverage points. Due to the correlation of response rates across time (autocorrelation), the present study estimated robust standard errors and the respective confidence intervals.[72]

Different opinions and solutions exit with respect to what constitutes outliers or how influential observations should be identified (see Freund & Wilson, 1998, p. 143). Due to the different explanations for outliers and leverage points, a further investigation of these observations are necessary, for example, to determine whether an outlier is based on a mistake, an error during data collection, or on an actual data point. Also, some influential

---

[72] Stata command: vce (robust) uses the robust or sandwich estimator of variance. This estimator is robust for some types of misspecification, given that the observations are independent.

observations may be caused by a factor not included in the model. Thus, these observations may be used to provide additional input for further specifications of a regression model. If an explanation for an influential observation is not clear, Meulemann et al. (2015, p. 106) have suggested that it is not good practice to simply remove this observation from the data set. Instead, they suggest using robust regression estimators, which is a technique that is less sensitive to outliers. Thus, the present study has used robust regression estimators.[73]

Due to assumptions just mentioned about the regression analysis, the following analysis uses scatterplots to provide a visualization of the data points so that the empirical results of the regression analysis are supported by a visual display.

Figure 4 illustrates the adjusted predictions for the mean response rates pooled in rounds. Although the analysis of the mean response rate shows a decreasing tendency (b = -.01), the trend is not significant (t = -1.43; p =.157).[74] Thus, Hypothesis 1 (*Response rates are decreasing.)* cannot be supported.

8.1 Trends Within the Country Level

Looking at the mean response rate (Figure 4) provides a good overview of the general development of response rates. However, at the same time, the mean may mask the development at the level of individual countries. Therefore, the next step is to examine the development of response rates at the country level by analyzing the general trend within countries (Section 8.1); analyzing the trends in countries with respect to decreasing, increasing, and constant response rates (Section 8.2).

---

[73] Robust estimation is implemented in the analysis by using the Stata command vce (robust).
[74] Since the inclusion of countries with missing data points would have biased the longitudinal development in further analyses, the present analysis, as mentioned previously, is limited to the countries that have information for all ESS rounds.

An analysis of the general trend of response rate development at the country level shows (Figure 5) that no clear tendency exists across countries. In some countries, like in Switzerland, from 2002 to 2010, response rates have increased; in some countries, response rates have remained constant; and in most countries, response rates have decreased. In addition, in some countries, response rates have gone up and down.[75] [76]

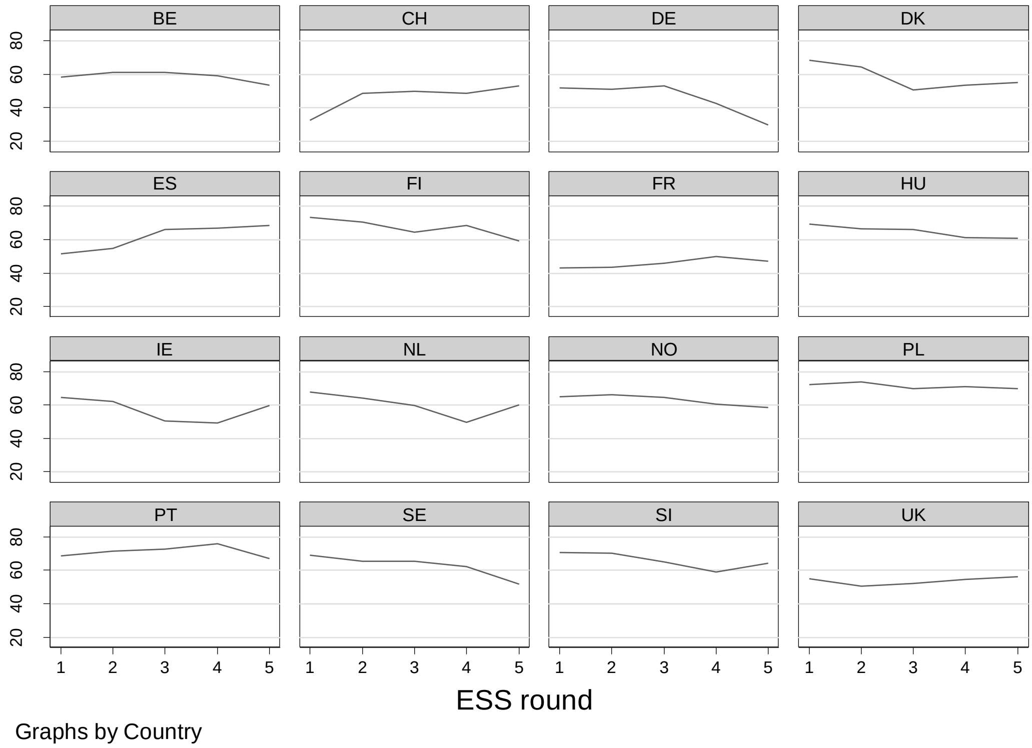

Figure 5. *Response rates (in percentages) per country from ESS 1 to ESS 5 (2002 to 2010).*

To further investigate the development of response rates at the country level, for each country, a regression model was estimated with the response rate as the dependent variable

---

[75] A maximum increase of 16.0 percentage points occurred in Switzerland from 2002 to 2004. The maximum decrease in response rates occurred in Denmark from 2004 (ESS 2) to 2006 (ESS 3) with 13.4 percentage points (see Table 3). For an overview of the changes in response rates between the rounds see Appendix Table A1.
[76] For the development of response rates according to response rates levels (low, medium, high) in the first round of the ESS see Appendix Figure A1 to A4.

and time as the independent variable (ESS round).[77] The results of each regression model is displayed in Figure 6. The coefficient of the regression analysis is marked with a dot (e.g., Belgium (BE): *b* = -.012; Spain (ES): *b* = .046). The coefficient shows the mean increase of response rates for each country between the rounds. In countries with a positive coefficient, the mean response rates have increased, whereas in countries with a negative coefficient, the response rates have decreased.

In most countries (11 out of 16), the average change in response rates across rounds was negative.[78] The line shows the confidence interval. The significance of this trend can be detected by looking at the 0-line at the *x*-axis. In five cases, the O line is included, which shows that no significant differences were detected between the rounds. In seven countries—Germany, Denmark, Finland, Hungary, Norway, Poland, and Sweden—a significant decrease occurred in response rates.[79] However, against the described negative trend in the survey literature and Hypothesis 1, in three countries—Switzerland, Spain and France—a significant increase in response rates occurred. Thus, it can be concluded that Hypothesis 1 is not supported.

8.2 Trends in Countries With Decreasing and Increasing Response Rates

In this section, countries are examined separately according to their different trends in response rates between 2002 and 2010. The focus is to determine whether the trends within the countries over the rounds are constant or if changes between rounds can be detected.

[77] The fact that the confidence intervals are large must be taken into account, meaning that the change in response rates across rounds follows a linear trend only imperfectly.
[78] However, the fact that the confidence intervals are large must be taken into account, meaning that the change in response rates across rounds follows a linear trend only imperfectly.
[79] In contrast to the other reported significance level (5%), here the level of significance is 10%.

Based on the results of the coefficients from the POLS regression (see Figure 6), countries were grouped as those with increasing, decreasing, or without a tendency response rates. Countries with an increase in response rates are defined as those countries for which the effect of survey rounds on response rate are positive and significant at the 10% level. Countries with decreasing response rates are defined as those countries with a significantly negative effect of survey rounds on response rates.[80]

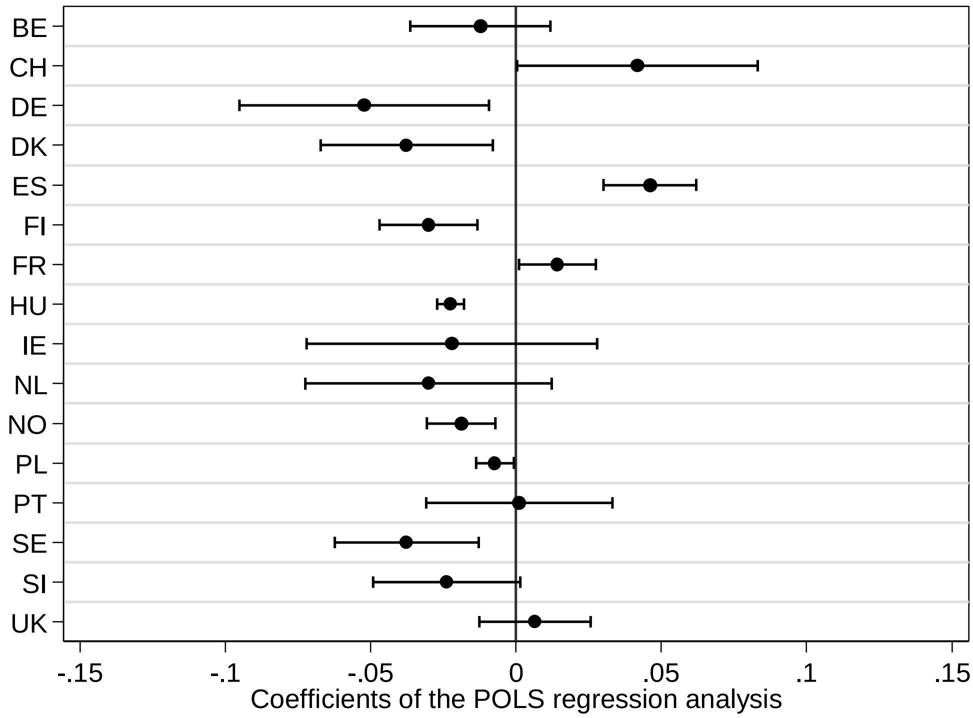

Figure 6. Estimated mean changes in response rates for each country between rounds.

Trends in countries with decreasing response rates: As shown in Figure 6 and Figure 7, out of 16 countries had response rates that decreased significantly. These countries were Germany, Denmark, Finland, Hungary, Norway, Poland, and Sweden.

---

[80] For the regressions of response rates on survey rounds, I estimated robust standard errors.

When comparing the response rates in the first round of the ESS in 2002 to the response rates in the fifth round in 2010, large differences can be detected. On average for all countries with decreasing response rates, the decline was 12 percentage points from the first to the last round of the ESS.[81] In Denmark, Finland, Sweden, and Germany, response rates declined by over 10 percentage points. In Germany and Sweden, the decrease was over 17 percentage points (Sweden 17% and Germany 22%). Hungary had a decline of 9 percentage points, and Norway had a decline of 7 percentage points. In Poland the decline in response rates was smaller at 2 percentage points.

A further analysis of these seven countries (Figure 7) showed a clear downward trend in response rates in Germany, Hungary, Sweden; whereas in Denmark, Finland, Norway, and Poland, response rates went up and down over the rounds. So about half of these countries (four out of seven) had decreasing response rates, which is not a clear downward trend.

[81] Countries with decreasing response rates usually had a rather high response rate in ESS 1, with the exception of Germany.

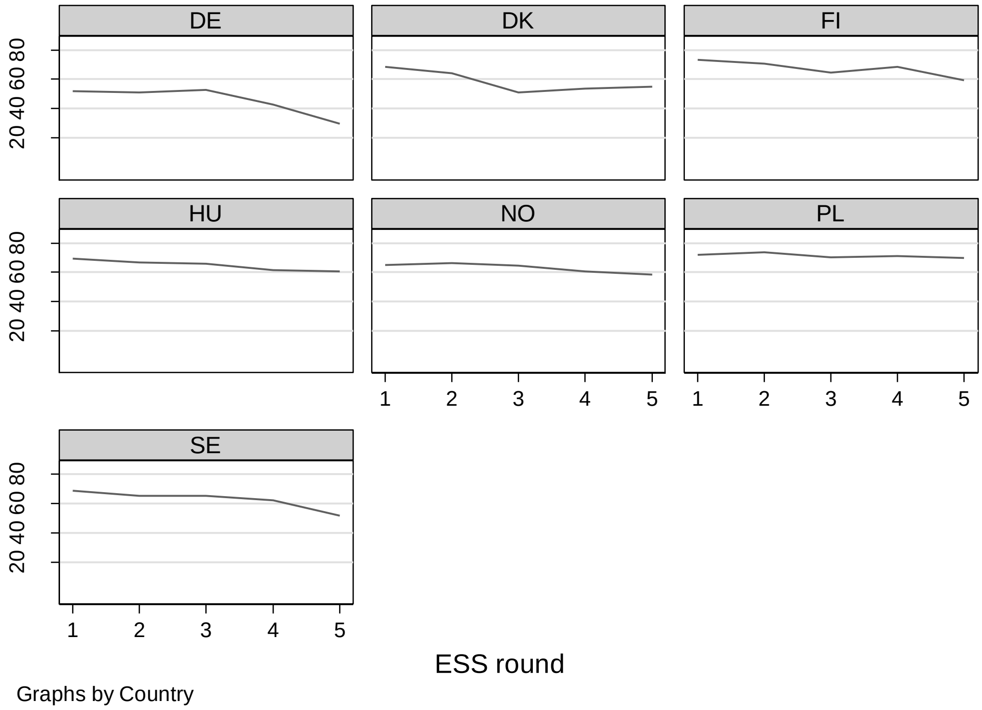

Figure 7. Countries with decreasing response rates from ESS 1 to ESS 5 (2002 to 2010).

Countries with no tendency in response rates: In 6 out of 16 countries, no significant changes were detected in response rates (see Figure 6 and Figure 8). Countries with no tendency in response rates were Belgium, Ireland, the Netherlands, Portugal, Slovenia, and the United Kingdom.

The level of response rates in ESS 1 was rather high for some countries (Slovenia, the Netherlands, Portugal, and Ireland), and for other countries, the mean response rate in ESS 1 was at a medium level (Belgium and the United Kingdom). The response rates in these countries increased and decreased with no clear tendency.

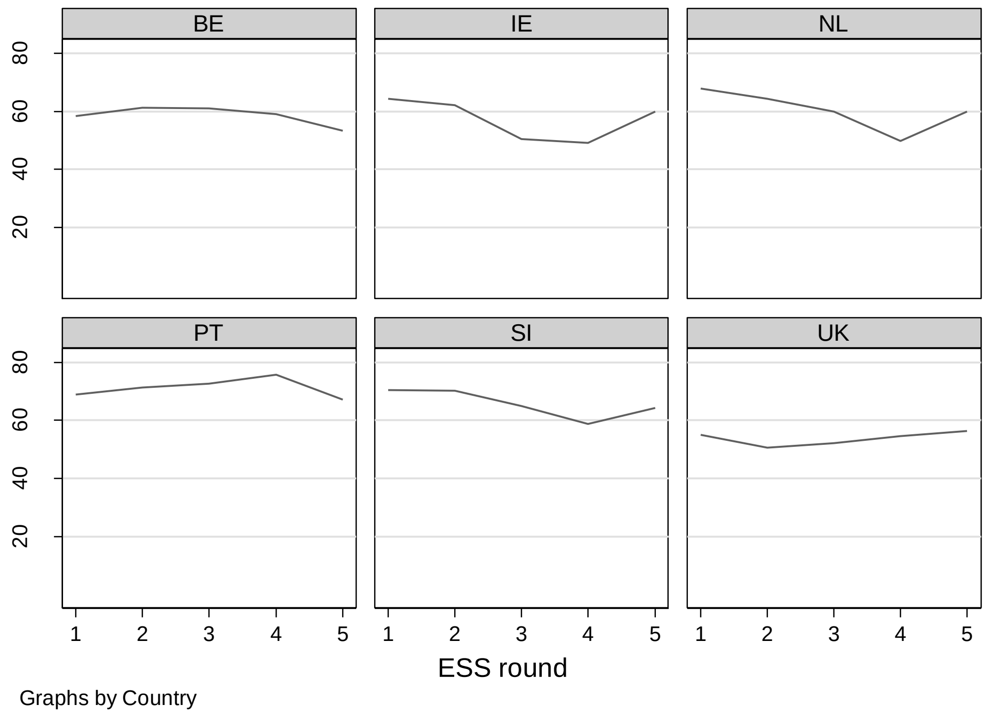

Figure 8. Countries with no tendency in response rates.

Trends for countries with increasing response rates: In 3 out of 16 countries, the response rate increased (Figure 6 and Figure 9). For the countries with increasing response rates, on average their response rates increased from ESS 1 to ESS 5 by 14 percentage points. The countries with increasing response rates had different starting levels: Switzerland started ESS 1 with a response rate at a little above 30%; France started at a little above 40%, and Spain started at a little above 50%. In France, Switzerland, and Spain, response rates increased steadily over the rounds, meaning that if response rates increased, they continued to increase over the rounds.

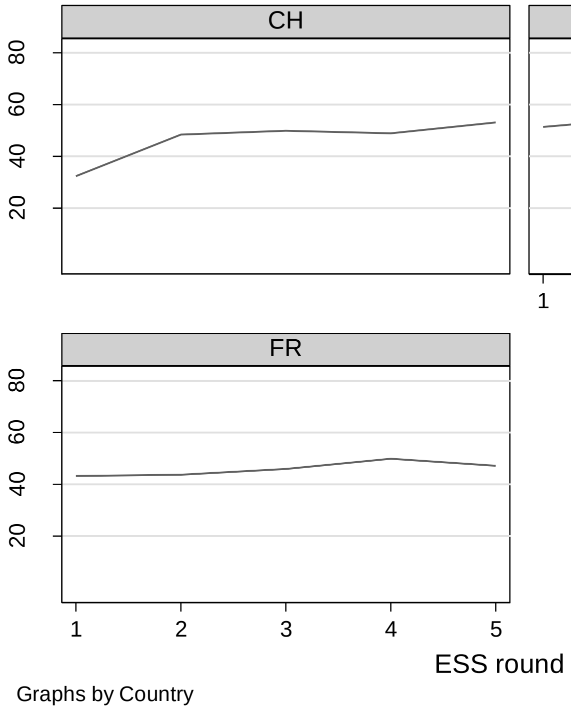

Figure 9. Countries with increasing response rates.

8.3 Excurse: Differences in Response Rates due to Sample Design

In addition to the explanations of the influence of micro, meso and macro levels, in cross-national comparisons, another factor may affect the extent of the nonresponse rate. In part, this difference has arisen due to the design of the sample. Different sample types may have affected the general level of the non-contact rate.

In the ESS, probability-based sampling is a requirement, but the type of this probability-based sampling differs. In some countries, individual samples are used, meaning individuals are drawn from a register. In other countries, samples of addresses or households are drawn.

In these address and household samples, an interviewer selects a respondent within a sampled address or household. The selection process is done in two steps. First the address or households is selected, and then the respondent is selected. So for countries with individual samples (in which the name of the respondent is already selected), not being able to contact a respondent involves only the step of contacting the selected individual. In countries that use address or household samples, the household or address is selected first to define the eligibility of a potential respondent. This eligibility is a prerequisite of the outcome code non-contact. The interviewer "must determine that the sample unit is an occupied unit with an eligible respondent and no contact with members of the housing unit is achievable" (American Association for Public Opinion Research, 2015, p. 23). Under certain circumstances, the selection process for the address or household sample may lead to different non-contact rates for samples of households or addresses.

In different samples, the *final disposition* codes may vary. For example, the definition of a dead person and thus the eligibly status of the dead person differs between different types of samples. In register samples, a person is selected as a potential respondent at a certain point in time. In case a respondent dies between the drawing of the sample and the interview (for example, in the ESS in Germany, the drawing of the sample takes place in the spring of the survey year, while the interviewer contacts the respondent in the fall of the survey year), the dead respondent is defined as ineligible (see American Association for Public Opinion Research (2015, p. 23). In cases where an address or more than one person household is selected in the drawing of the sample, the interviewer makes the selection of the potential respondent at the doorstep before the interview. The interviewer does a listing of all persons in a household. Dead persons are not included. In a case where no contact can be made within the household, this will be defined as non-contact.

Also, *timing issues* may have different effects on different sample types. For example, a housing unit that was occupied on the status date and then became vacant because the household moved before any other contact was attempted would be a non-contact case (AAPOR code 2.20) if no interview was obtained, and not a vacant housing unit, and therefore not a not-eligible case (AAPOR code 4.60). These small effects might potentially lead to differences in the response rate.

The use of an address sample may help to partly explain a *high non-contact rate*. With respect to an address sample, the interviewers receive addresses, and the potential respondent is selected in two steps. First, the interviewer selects the household by a procedure that includes mailboxes and address. Then the potential respondent is selected using the "last birthday method." Selecting an address within a house can be challenging because in large cities (such as in Paris). According to A. Cournilleau, the National Coordinator of France (personal communication, June 12, 2014), access to larger houses is not possible, since many buildings, especially in Paris, are not accessible due to gate keepers, doormen, or other devices. Thus, the selection of a household is not possible. In Ireland, the high number of non-contacts, especially in round 5 (2010), was explained by Michael Breen (the National Coordinator for Ireland) as being caused by the financial crisis, which had led to many houses being empty (M. Breen, personal communication, July 26, 2011). Therefore, contact with these households could not be established, and so the eligibility status of potential respondents could not be known for these addresses. So, the coding for eligible cases as potentially non-contact or ineligible is not always clear.[82] For example, a housing unit that was occupied on a status date and then became vacant because the household moved before any other contact was attempted would be listed as a non-contact case (AAPOR definition 2.20) if no interview was obtained (and not a vacant housing unit, and therefore not a not-eligible case (AAPOR definition 4.60). The large amount of empty houses during the financial crisis in 2010 explain according to personal communication with the National

[82] AAPOR advises that for non-contacts, researchers "must determine that the sample unit is an occupied unit with an eligible respondent and no contact with members of the housing unit is achievable" (American Association for Public Opinion Research, 2015, p. 23).

Coordinator for Ireland M. Breen (personal communication, July 26, 2011) is nationally around 20% and this leads to an increase of the non-contact rate, as well as the increase of the rate of "not able/others".

9 Development of Nonresponse Rates

To understand the trends in the response rate, the development of the nonresponse needs to be further investigated to determine whether one subgroup of non-respondents (refusal, non-contact, not able/others) might be the driver for the change in the response rate. In the previous section, the research question set out to determine whether response rates decrease over time. In contrast to this question, this section sets out to determine whether the nonresponse rate has increased over time, and whether one nonresponse subgroup in particular has increased.

Non-respondents are defined as "those missing from a probability sample" (Singer, 2006, p. 637). The three elements of nonresponse are refusal, non-contact, not able/others. One type of non-respondent is a person who cannot be contacted and therefore cannot be interviewed; another is a person who refuses to be interviewed (see Groves & Couper, 1998; Groves, Fowler, Couper, Lepkowski, & Singer, 2004). The American Association for Public Opinion Research (2015) defines *refusals* as: "The proportion of all cases in which a housing unit or the respondent refuses to be interviewed, or breaks-off an interview of all potentially eligible cases" (p. 6). [83] The second main part of nonresponse, in addition to refusals, is non-contacts. The American Association for Public Opinion Research (2015, p. 23) categorizes

[83] The American Association for Public Opinion Research (2015, p. 55) provides three definitions of refusal rates, which differ in the way they treat cases with dispositions of unknown eligibility (p. 55). The ESS refusal rate is comparable to the AAPOR REF1: the number of refusals divided by the interviews (complete and partial) plus the non-respondents (refusals, non-contacts, and others) plus the cases of unknown eligibility.

the non-contacts in in-person household surveys as belonging to three types: a) unable to gain access to the building (which includes homes behind locked gates and guarded apartment buildings); b) no one reached at a housing unit; and c) respondent away or unavailable.[84] These two components, the refusals and non-contacts, are described as having different reasons, and thus are assumed to have different consequences for a survey estimate (see Groves & Couper, 1998; Singer, 2006, p. 637). The third nonresponse category, in addition to refusal and non-contacts, is the “not able/others” cases. This element of nonresponse includes respondents who have been contacted but are not able to be interviewed. These include the contacted cases in which “there is a respondent who did not refuse the interview, but no interview is obtainable” (American Association for Public Opinion Research, 2015, p. 23). According to AAPOR, this includes: a) death,[85] b) respondents physical and/or mental inability to do an interview, c) language problems, and d) miscellaneous other reasons.[86]

For a deeper understanding of the development of response rates, the nonresponse categories need to be analyzed. The largest nonresponse category is refusal with 27%, the

[84] The eligibility of a potential respondent is a prerequisite of the outcome code non-contact. AAPOR advises that for non-contacts, researchers “must determine that the sample unit is an occupied unit with an eligible respondent and no contact with members of the housing unit is achievable” (American Association for Public Opinion Research, 2015, p. 23). Also, for the second category b) where no one is reached at the housing unit and so no contact is made with a responsible household member, the presence of an eligible household member needs to be ascertained. In addition “the unavailability of the designated respondent means that enough information is obtained to determine who the respondent is, but the respondent is never available when the interviewer attempts an interview“ (American Association for Public Opinion Research, 2015, p. 23).

[85] Dead people can be coded differently. According to the American Association for Public Opinion Research (2015, p. 23) “Whether death makes a case a non-respondent or an ineligible respondent depends on fieldwork timing. Surveys should define a date on which eligibility status is determined. This would usually be either the first day of the field period or the first day that a particular case was fielded. Thus, for example, if a person were alive and selected as the respondent from a sampled housing unit in an in-person household survey on this status date but died before an interview was completed, the case would be classified as a non-response due to death (2.31). Similar time rules would apply to other statuses. For example, a housing unit that was occupied on status date and then became vacant because the household moved before any other contact was attempted would be a non-contact case (2.20) if no interview was obtained (and not a vacant housing unit, and therefore not a not-eligible case, 4.60).“

[86] Nonresponse is not only due to the unwillingness of respondents, but also to, for example, organizational issues (such as no interviewers available, fieldwork period ended, too many surveys to be conducted by one interviewer, and so on). Literature on these organizational issues is rarely available (Stoop, 2015). Also, although research in general on this group of “not able/others” is rare, Stoop, Koch, and Matsuo (2012) have carried out some.

second is “not able/others” with 9%, and the smallest is non-contact with 5%. Figure 10 shows the percentage for the three subgroups of nonresponse (refusal, noncontacts, and not able/others) compared to the overall response rates across the different ESS rounds. The development of the subgroups of nonresponse rates is used to explain the development of response rates, to determine whether the development of response rates may be driven by the increase of a particular subgroup of non-contact rates.[87]

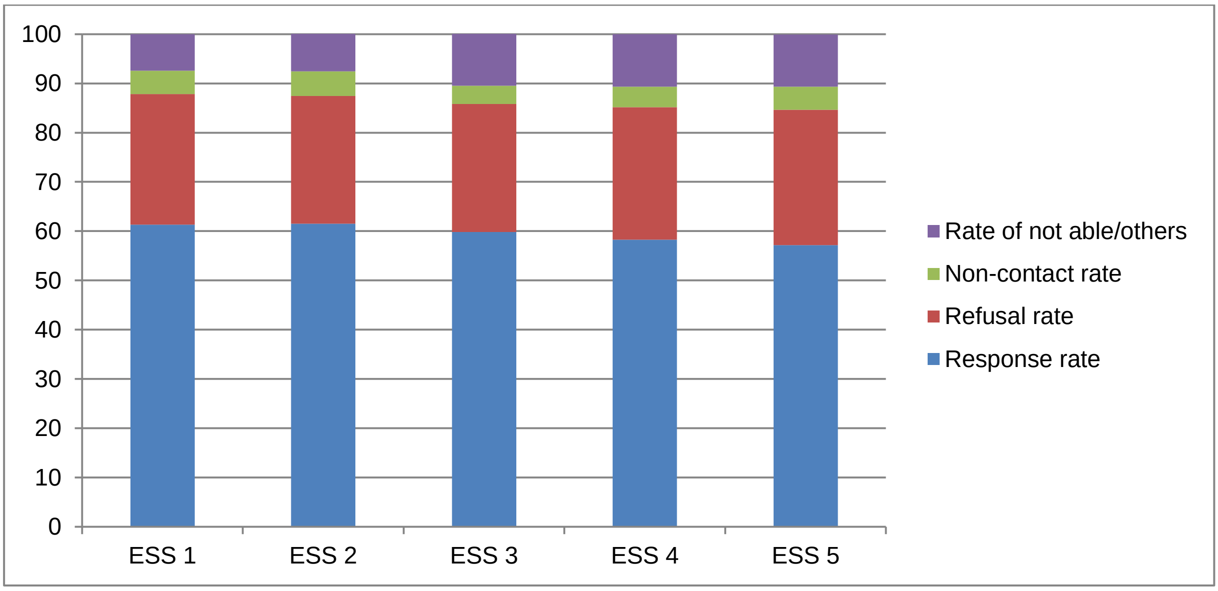

Figure 10. Outcome codes: response rate, refusal rate, non-contact rate, rate of “not able/others” (mean across rounds).

Complementary to Hypothesis 1 concerning the development of response rates—“Response rates are decreasing.”—the hypothesis regarding the development of nonresponse rates has been formulated as:

Hypothesis 2: Nonresponse rates are increasing.

In more detail for different types of nonresponse (refusal, noncontacts and not able/others), the following hypothesis are tested (see Table 2).

[87] For more details see Appendix Figure A5 to A8.

Hypothesis 3: Refusal rates are increasing.

Hypothesis 4: Noncontact rates are increasing.

Hypothesis 5: Rates of "not able/others" are constant.

In the survey literature, the decreasing tendency in response rates is explained by the survey climate getting worse and people being less likely to respond (Barbier et al., 2016; Harris-Kojetin & Tucker, 1999; Lyberg & Dean, 1992). This leads to Hypothesis 3 stating that refusal rates are increasing. Another hypothesis discussed (Barbier et al., 2016; de Heer, 1999; de Leeuw & de Heer, 2002) is that survey organizations put less effort in their surveys, and thus interviewers do not contact all respondents, which increases the non-contact rates (Hypothesis 4). The third category of nonresponse, the "not able/others" is less frequently discussed in the literature (Stoop, 2015; Stoop et al., 2012). Since the category is described as a residual category comprised of different groups, it is not expected to change over the rounds (Hypothesis 5).

9.1 Refusal Rate

In this section, the development of refusal rates is described. As a complement to the development of response rates, which is described in the literature (Brehm, 1994; de Heer, 1999; de Leeuw & de Heer, 2002; Groves, Fowler, Couper, Lepkowski, & Singer, 2004) as decreasing, nonresponse rates—and thus refusal rates—are described as increasing (Alreck & Settle, 1995; Barbier et al., 2016; Brick & Williams, 2013). Based on literature the tested hypothesis (see also Section 1.4) is that refusal rates are increasing (Hypothesis 3).

The average refusal rate in the ESS across all rounds has ranged between 13% and 51% with an average of 27%. Over the rounds, the mean refusal rate was rather constant, ranging between 26% (ESS 2) and 28% (ESS 5) (see Table 4 and Figure 10). A *t*-test for the equality of the mean for independent samples comparing the first and last round showed no significant decrease of the mean refusal rate (p = .680).[88]

Table 4

Refusal rates (in percentages)

Country	ESS 1 (2002)	ESS 2 (2004)	ESS 3 (2006)	ESS 4 (2008)	ESS 5 (2010)	Mean
Belgium	25.2	26.4	24.1	25.9	32.5	26.8
Switzerland	51.2	44.0	40.7	33.4	29.4	39.7
Germany	28.2	32.8	25.4	32.6	39.6	31.7
Denmark	23.0	24.6	37.9	32.9	29.5	29.6
Spain	33.9	25.1	21.7	16.1	15.4	22.4
Finland	20.8	22.7	23.2	20.9	27.2	23.0
France	38.5	39.3	40.6	36.3	35.3	38.0
Hungary	14.9	15.0	26.4	27.5	26.5	22.1
Ireland	22.9	21.6	13.8	24.2	13.2	19.1
Netherlands	26.2	29.1	33.3	40.6	32.7	32.4
Norway	24.2	25.9	25.9	30.9	31.7	27.7
Poland	19.6	19.1	16.3	18.0	19.1	18.4
Portugal	26.9	18.2	21.0	20.1	21.0	21.4
Sweden	21.0	22.0	22.9	25.3	35.7	25.4
Slovenia	17.3	15.3	15.9	21.8	22.5	18.6
United Kingdom	30.6	33.2	26.7	24.4	29.4	28.7
Mean	26.5	25.9	26.0	26.9	27.5	26.6

The POLS of the development of the refusal rate pooled by round showed that the estimate of refusal rates is increasing over time (Figure 11). A regression analysis of the mean response rate (dependent variable) and the round (independent variable) showed that the refusal rate increased, although it is rather low and not significant (b = .00; t = 0.54; p = .591).[89] However, the variance of the refusal rate is rather high at all the different points in time.

[88] Levene test (p = .477), thus equal variance is assumed. A t-test of independent samples was conducted.
[89] A regression analysis was performed with the dependent variable “refusal rate” and the independent variable “ESS rounds” with robust standard errors (Stata command vce).

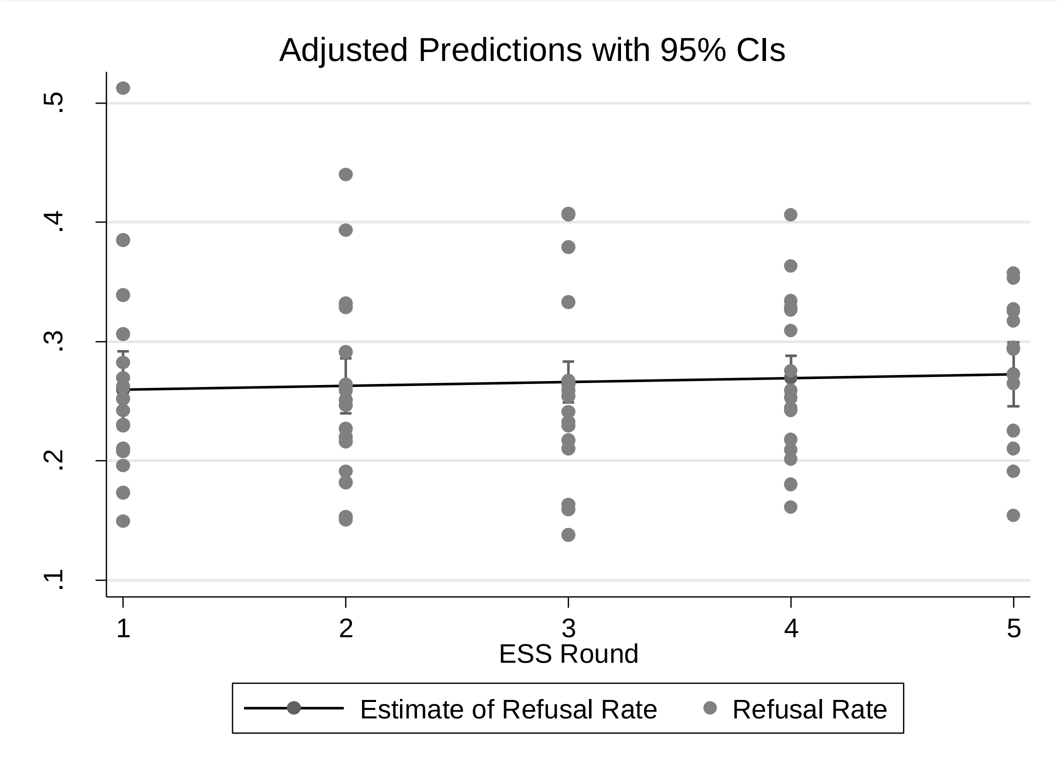

Figure 11. POLS of the development of refusal rates controlled for rounds.

Analysis at the country level (Figure 12) showed that the mean refusal rate per round varied widely within countries over the rounds. In some countries (Switzerland, Spain, and France), the refusal rate decreased steadily; in other countries (Belgium, Germany, Denmark, Ireland, Netherlands, Poland, Portugal, Slovenia, and the United Kingdom), the refusal rate went up and down over time; and in some countries, the mean refusal rate increased steadily (Finland, Sweden, Norway, and Hungary).

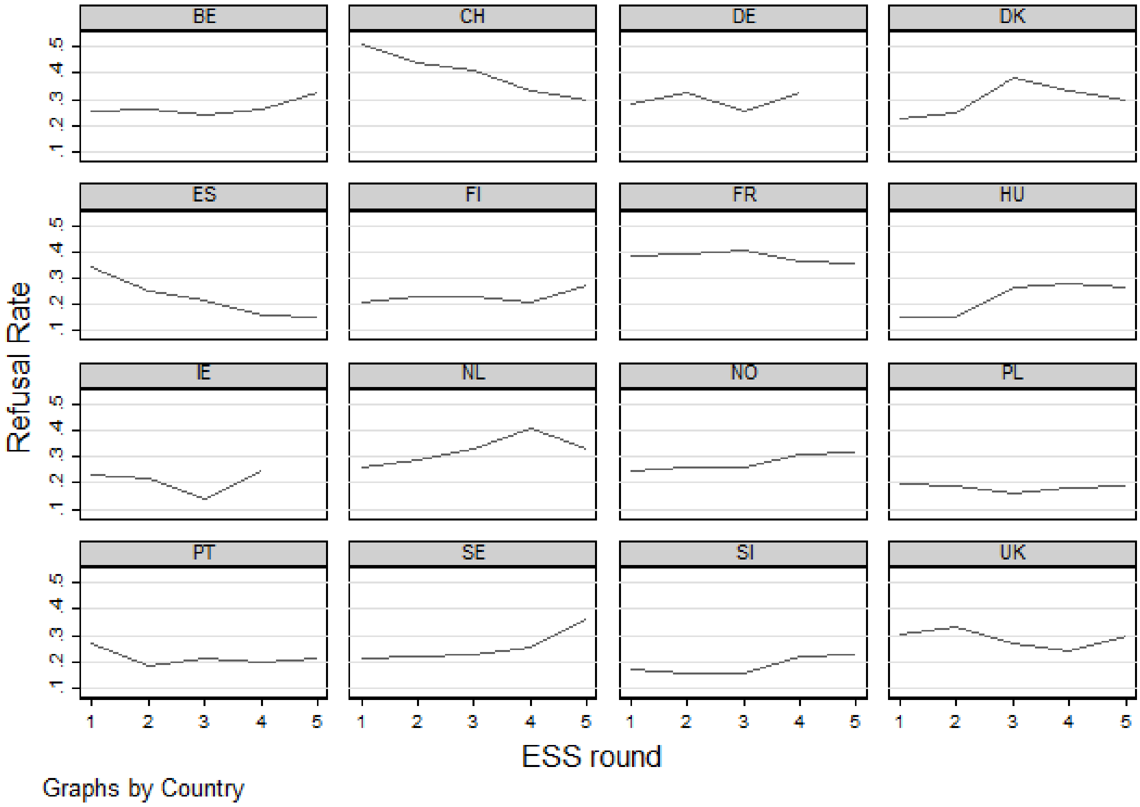

Figure 12. Refusal rates (in percentages) per country from ESS 1 to ESS 5 (2002–2010).

To detect country specific tendencies, a further analysis was conducted. A regression model for the refusal rate was estimated. Figure 13 shows the mean refusal rate between the rounds for each country. The dot marks the coefficient. For the majority of countries, the refusal rate increased over the rounds, but the trend was not significant for all countries. The line shows the confidence interval. The significance of this trend can be detected by looking at the 0-line at the *x*-axis. If the O line is included, this shows that no significant differences were detected between the rounds. A significant increase in refusal rates was detected in Hungary, Norway, Sweden, and Slovenia. In a few countries, the refusal rates decreased (Ireland, Poland, Portugal, and the United Kingdom), but they only significantly decreased in

Switzerland and Spain.[90] The analysis of the development of the refusal rate as part of the nonresponse rate showed that Hypothesis 3 (*Refusal rates are increasing.*) cannot be supported for all of the countries.

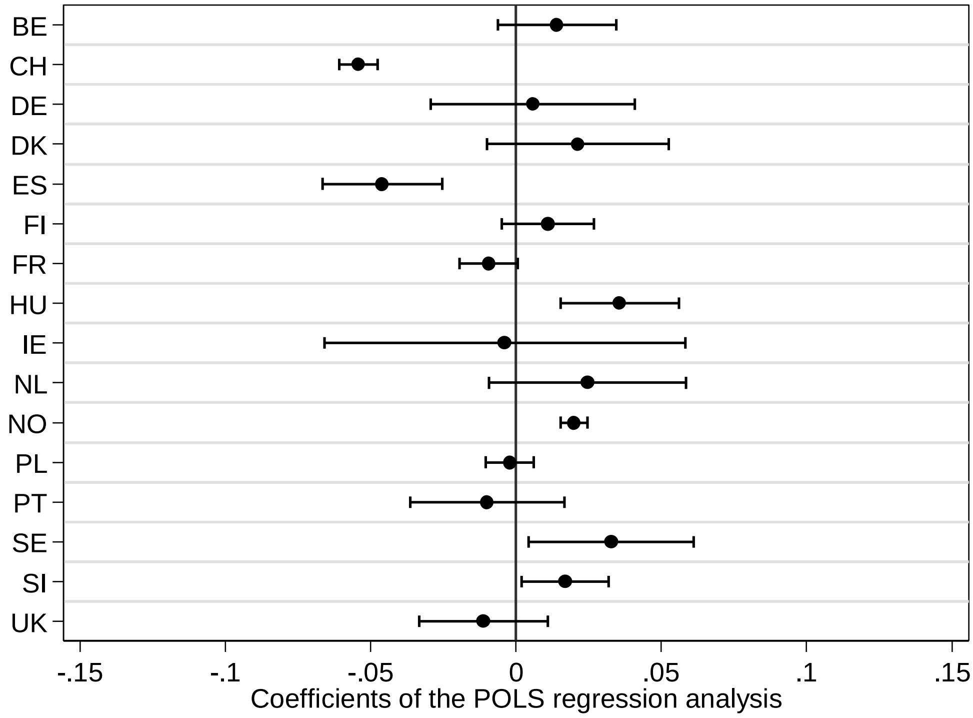

Figure 13. Estimated mean changes in the refusal rates for each country between rounds (10% confidence interval).

9.2 Non-Contact Rate

In this section, the development of the non-contact rate is described. As described previously, non-contacts are defined in in-persons households as a) unable to gain access to the building (which includes homes behind locked gates and guarded apartment buildings); b) no one reached at a housing unit; and c) respondent away or unavailable (American Association for Public Opinion Research, 2015, p. 23). For the above described reasons, the nonresponse rates—and thus non-contact rate—are described as increasing (Alreck & Settle,

[90] Significant changes can be seen in case the confidence intervals do not cover the 0 line. At the 5% significance level, Sweden and Slovenia do not show significant changes between the years 2002–2010.

1995; Barbier et al., 2016; Brick & Williams, 2013). Thus, the tested hypothesis is that non-contact rates are increasing (Hypothesis 4).

At the across country level, the mean non-contact rate in the ESS across all rounds has ranged between 4% (ESS 3) and 5% (ESS 2), with an average non-contact rate of 5% (see). So, for all rounds of the ESS, less than 5% of the potential respondents could not be contacted.[91] From the country perspective, the mean non-contact rate ranged between 0% (Denmark in ESS 4) and a maximum of 22% (Ireland in ESS 5) (see Table 5).[92] A *t*-test for the equality of means for independent samples that compared the first and last rounds showed no significant decrease of the mean non-contact rate (p = .929).[93]

Table 5

Non-contact rate ESS 1 to ESS 5 (in percentages)

Country	ESS 1 (2002)	ESS 2 (2004)	ESS 3 (2006)	ESS 4 (2008)	ESS 5 (2010)	Mean
Belgium	4.5	3.5	2.9	2.2	1.3	2.9
Switzerland	3.8	2.1	2.2	8.7	7.3	4.8
Germany	5.7	7.0	5.0	6.4	7.4	6.3
Denmark	4.6	4.9	3.3	0.3	3.7	3.4
Spain	7.6	7.1	3.3	2.6	1.6	4.4
Finland	2.9	2.1	2.7	2.6	1.7	2.4
France	14.7	8.6	6.6	7.7	9.3	9.4
Hungary	3.1	5.7	2.9	2.5	2.6	3.4
Ireland	8.1	10.6	9.1	9.1	21.6	11.7
Netherlands	2.5	2.7	2.6	3.0	2.7	2.7
Norway	3.0	1.7	0.8	0.9	1.4	1.6
Poland	0.8	0.9	1.3	1.6	1.1	1.1
Portugal	3.2	2.7	3.8	3.0	3.2	3.2
Sweden	4.0	2.4	2.0	3.9	0.7	2.6
Slovenia	5.1	10.2	2.9	4.7	3.4	5.3
United Kingdom	3.5	7.9	7.2	7.9	5.9	6.5
Mean	4.8	5.0	3.7	4.2	4.7	4.5

[91] The maximum non-contact rate in the ESS is 5% as indicated in the specifications for participating countries (European Social Survey, 2013).
[92] Three extreme cases had a high percentage of non-contact rate: in ESS 3, Ireland had 9% and in ESS 5 22%[92]; and in ESS 1, France had 15%
[93] Levene test (p = .221), thus equal variance is assumed. A *t*-test of independent samples was conducted.

As can be seen in Table 5 and Figure 14, the variance of the non-contact rate varied widely across countries. To investigate this variance, I conducted further analysis at the country level. The POLS regression (Figure 14)[94] showed a slight decreasing tendency over time, which was not significant (b = -.00; t = -0.34; p = .734).[95]

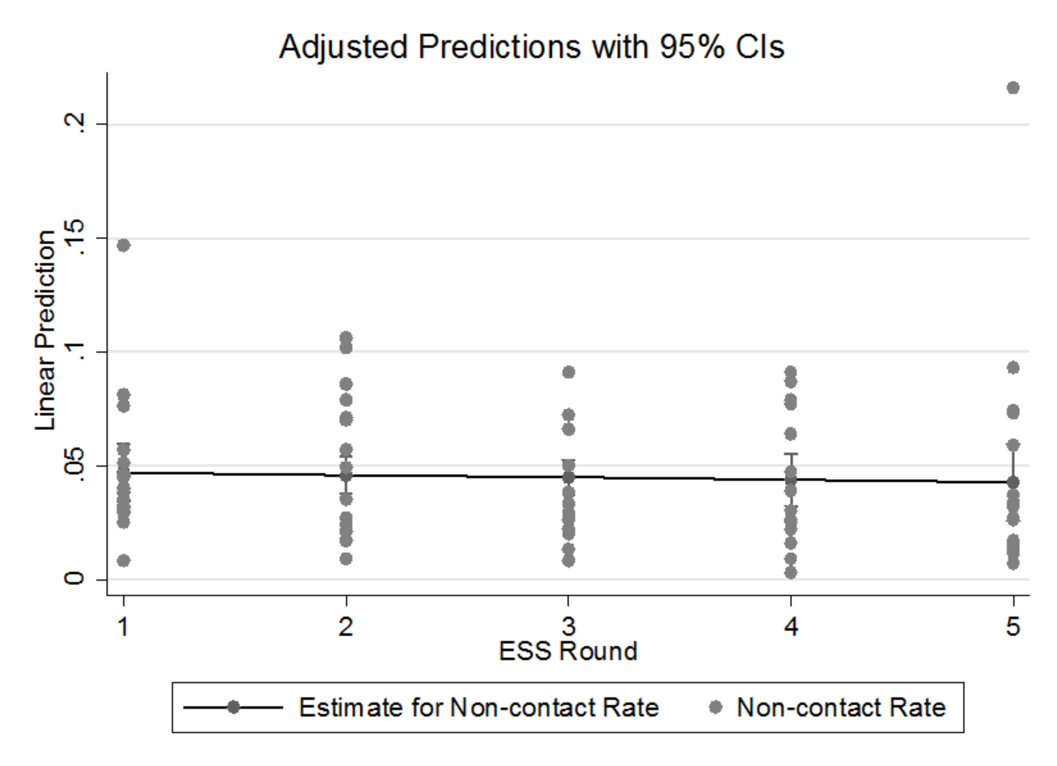

Figure 14. POLS of the development of non-contact rates controlled for rounds.

At the country level, different trends occurred in the development of the response rate (Figure 15). In some countries, non-contact rates increased steadily (Belgium, Spain, Finland, and Hungary) while in other countries, the non-contact rates varied over the rounds. A slight increase in non-contact rates was detected only in two countries (the Netherlands and Poland). So although for the majority of countries (11 out of 16) the non-contact rates increased, this was not a general trend for all countries.

---

[94] For the rationale for using a multilevel model using POLS, see section 8 and Brüderl and Ludwig (2015).
[95] A regression analysis was performed with the dependent variable “non-contact rate” and the independent variable “ESS rounds” with robust standard errors (Stata command vce).

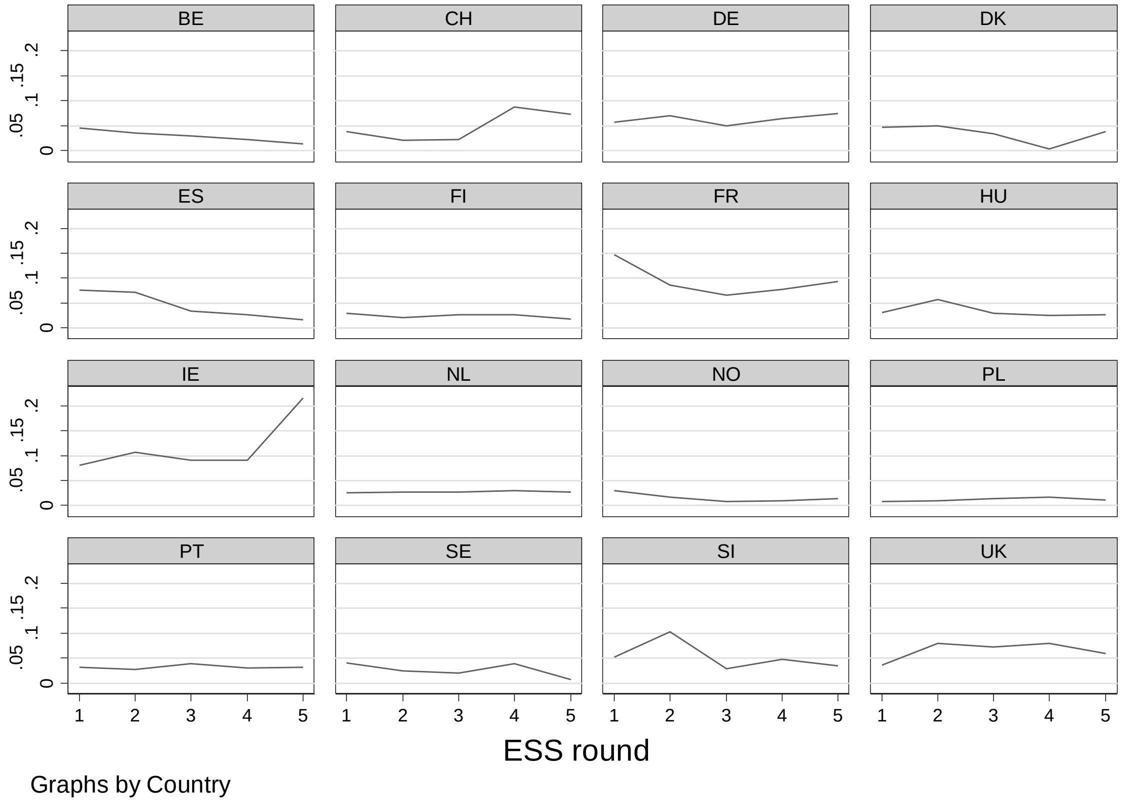

Figure 15. Non-contact rate (in percentages) per country from ESS 1 to ESS 5 (2002 to 2010).

The regression model (Figure 16) supports this finding. It highlights that at the country level, both positive and negative developments were found. At the 10% significance level, only Belgium (BE) and Spain (ES) had a significant negative trend. For the majority of countries, no significant effect was found. Our analysis of the development of non-contact rates as part of the nonresponse rates showed that Hypothesis 4 (*Non-contact rates are increasing.*) cannot be supported.

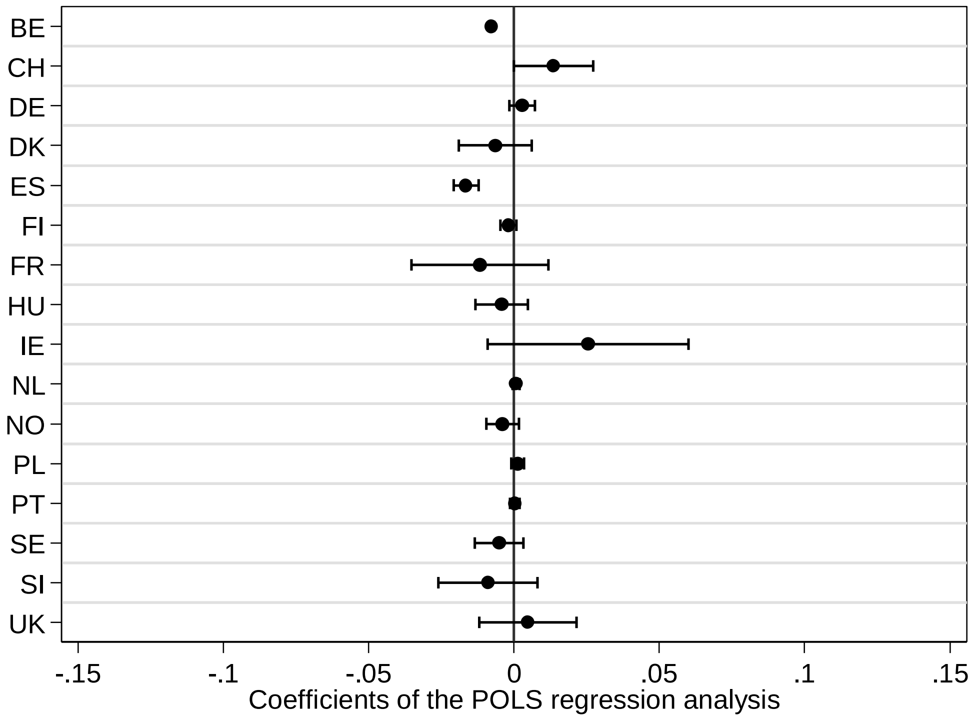

Figure 16. Estimated mean changes in non-contact rates for each country between rounds (10% confidence interval).

9.3 Rate of Not able/Others

The third category of nonresponse is not able/others. This category comprises the contacted cases in which "there is a respondent who did not refuse the interview, but no interview is obtainable" (AAPOR, 2015, p. 23). According to AAPOR, this includes: a) death,[96] b) respondent's physical and/or mental inability to do an interview, c) language problems, and d) miscellaneous other reasons.[97] This category, similar to the category of non-contact, is not due to the unwillingness of respondents only, but also can include country

[96] Dead people can be coded differently. According to the American Association for Public Opinion Research (2015, p. 23), "Whether death makes a case a non-respondent or an ineligible respondent depends on fieldwork timing. Surveys should define a date on which eligibility status is determined. This would usually be either the first day of the field period or the first day that a particular case was fielded. Thus, for example, if a person were alive and selected as the respondent from a sampled housing unit in an in-person household survey on this status dates, but died before an interview was completed, the case would be classified as a non-response due to death (2.31)."

[97] Nonresponse is not only due to the unwillingness of respondents, but also to, for example, organizational issues such as the lack of available interviewers, fieldwork period ended, the need for interviews to be conducted by one interviewer, and so on.

specific factors such as the survey design and the selection of the sample, [98] and organizational issues such as the lack of interviewers, fieldwork period ended, and so on (Stoop, 2015; Stoop et al., 2012).

The tested hypothesis is (see Section 8.3) that the rates of not able/others is increasing (Hypothesis 5). The average rate of not able/others at the country level ranges between 7% and 11%. As can be seen in Table 6, the minimum rate of others is 1% (Portugal in ESS 1 and ESS 4), and the maximum rate of others of all rounds in the ESS is 27% (Ireland in ESS 3).[99] A *t*-test for equality of mean for independent samples comparing the first and last round showed a significant lower the mean the rate of not able/others. At the 10% significance level the mean rate of "not able/others" in 2002 (ESS 1) is significantly lower than in 2010 (ESS 5) (p = .071).[100]

---

[98] As described previously (American Association for Public Opinion Research, 2015, p. 23), the time issue regarding the definition of a dead person and an occupied house depends on the sample type (individual sample vs. household or address sample). For example, a housing unit that was occupied on the status date and then became vacant because the household moved before any other contact was attempted would be a non-contact case (AAPOR code 2.20) if no interview was obtained, and not a vacant housing unit, and therefore not a not-eligible case (AAPOR code 4.60).
[99] Regarding ESS 5, Germany is an outlier with a rate of not able/others of 23%.
[100] Levene test (p = .240), thus equal variance is assumed. A *t*-test of independent samples was conducted.

Table 6

Rates of not able/others (in percentages)

Country	ESS 1 (2002)	ESS 2 (2004)	ESS 3 (2006)	ESS 4 (2008)	ESS 5 (2010)	Mean
Belgium	11.9	8.9	12.0	12.9	12.8	11.7
Switzerland	12.5	5.4	7.1	9.1	10.1	8.8
Germany	14.4	9.2	16.7	18.3	23.3	16.4
Denmark	4.0	6.3	8.0	13.2	11.9	8.7
Spain	7.0	12.9	8.8	14.5	14.4	11.5
Finland	3.0	4.5	9.7	8.1	11.7	7.4
France	3.7	8.5	6.8	6.1	8.3	6.7
Hungary	12.7	12.7	4.7	8.7	10.2	9.8
Ireland	4.6	5.8	26.7	17.5	5.4	12.0
Netherlands	3.5	3.9	4.3	6.6	4.6	4.6
Norway	7.8	6.2	8.9	7.8	8.4	7.8
Poland	7.5	6.3	12.4	9.5	9.8	9.1
Portugal	1.1	7.8	2.5	1.1	8.7	4.2
Sweden	6.0	10.2	9.6	8.6	11.8	9.2
Slovenia	7.1	4.3	16.3	14.7	9.7	10.4
United Kingdom	10.9	8.3	14.0	13.2	8.4	11.0
Mean	7.4	7.6	10.5	10.6	10.6	9.3

A POLS regression pooled for the different points of time was calculated.[101] Increasing rates of not able/others over time was detected (Figure 17). A regression analysis also showed a positive trend and a significant increase in the rate of not able/others (b = 1.61; t = 2.30; p = .024).[102]

[101] For the rationale for using a multilevel model using POLS, see Section 8 and Brüderl and Ludwig (2015).

[102] A regression analysis was performed with the dependent variable “rate of others” and the independent variable “ESS rounds” with robust standard errors (Stata command vce).

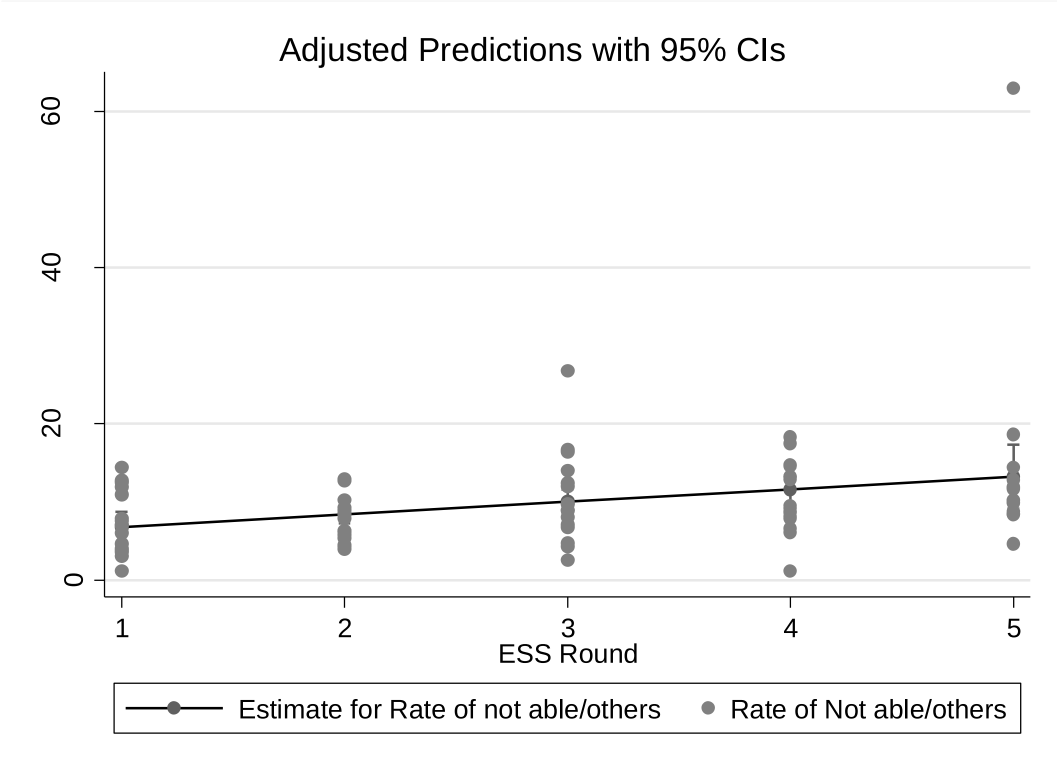

Figure 17. POLS of the development of rates of not able/others controlled for rounds.

An analysis at the country level (Figure 18) suggested that in contrast to the refusal and non-contact rates, the mean level of the outcome category not able/others increased significantly over the rounds. However, within countries, the variance is rather high (e.g., Ireland ESS 1: 4.6%; ESS 2: 5.8%; ESS 3: 26.7%; ESS 3: 17.5%).[103]

[103] Other countries with a high variance of rate of others were Portugal and Germany.

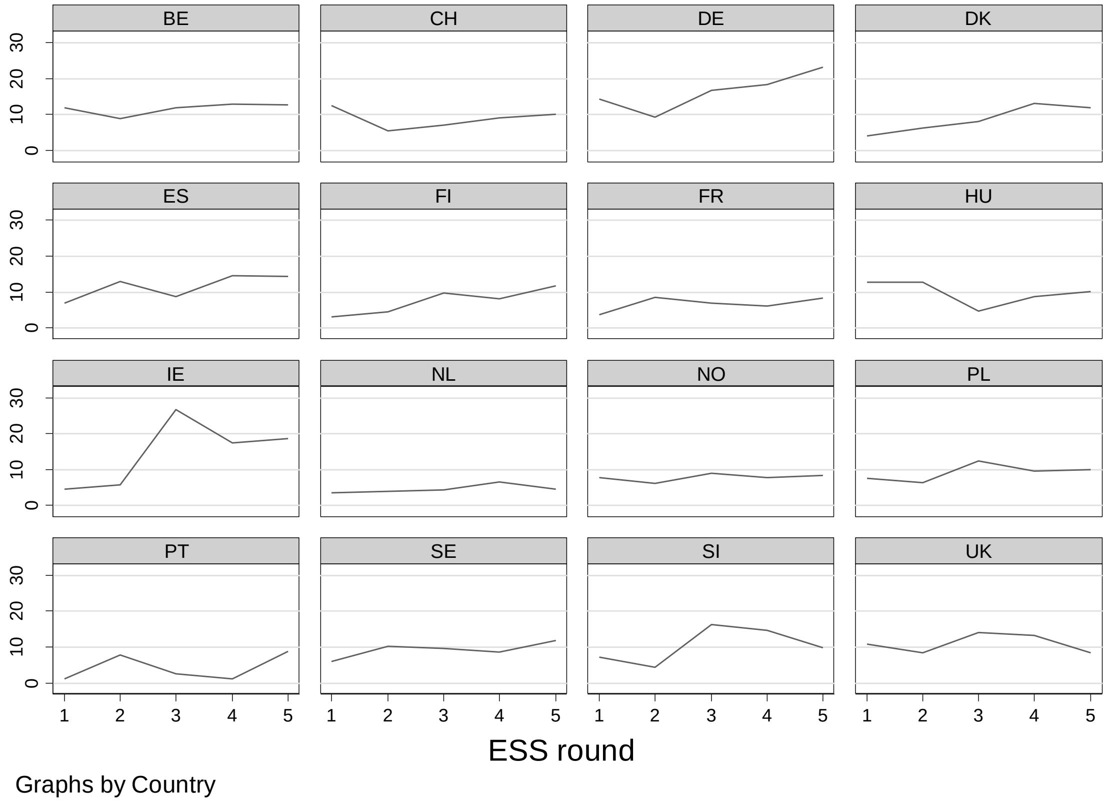

Figure 18. Rate of not able/others (in percentages) per country from ESS 1 to ESS 5 (2002 to 2010).

The regression model (Figure 19)[104] for the rate of not able/others for each country shows that in general, the estimate of the rates of not able/others are positive. As can be seen in Figure 18 and Figure 19, the variance and the confidence intervals are rather large, for example, for Germany, Ireland and Portugal. Even though a significant increase in the mean response rates was seen, at the country level, the increase of the rates of not able/others was not significant for most countries. The rates increased significantly only in Denmark, Spain, Finland, and Ireland.[105]

---

[104] Due to the large variance in Germany, the scale of the *x*-axis of Figure 19 for the rate of "not able/others" differs to the refusal rate (Figure 13) and the non-contact rate (Figure 16).
[105] At the 5% significance level, the rates of "not able/others" in Spain and Ireland were not significant.

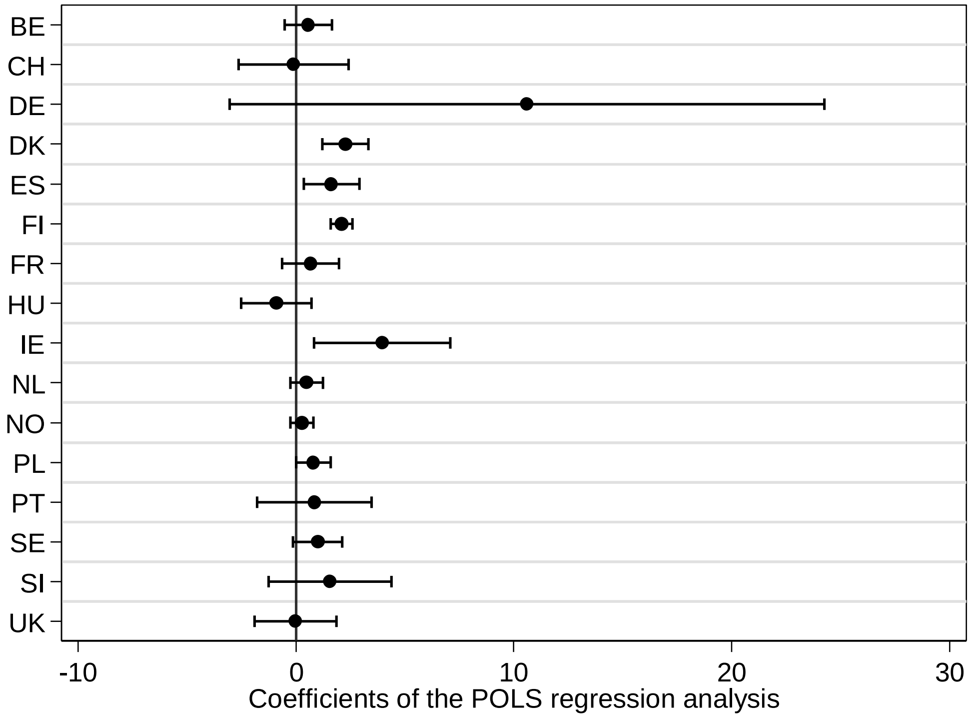

Figure 19. Estimated mean changes in the rates of not able/others for each country between rounds (10% confidence interval).

10 Discussion

The results of the above analyses are applicable and conclusions might be drawn to other high quality survey. Since the ESS is very well documented regarding the outcome codes (response rates, refusal rate, non-contact rate and rate of not able/other), the reliability of other face-to-face studies on the development of those rates can be increased by comparing their results to this study based on ESS data. The results of analysis might be less well suited for other surveys following less rigorous models. The response rates might differ in other studies with different samples.[106] Comparability of the above analysis to surveys with random

[106] Another difference that may affect the response rate comparability is that the EUROBAROMETER collects data within two weeks with no re-issuing period.

route processes might be difficult, because the differences of the selection procedure (see discussion above on the effect on samples on response, refusal and non-contact rate).

In general the comparability of the response rates across time and countries is something unique in the ESS. The documentation of the outcome codes and the information on the calculation of response and nonresponse rates enable researchers to perform analyses on the development of response rates. Due to the public documentation of the data, the information can be made comparable between time and countries.

11 Results for Analysis Part A

This section analyzes the development of the outcome codes of the European Social Survey (ESS). The development of response rates and nonresponse rates (non-contact, refusal, and not able/others) was examined. This analysis was carried out for the mean and the changes at the country level. The hypotheses for the development of response rates and nonresponse rates (refusal, non-contact, and not able/others) were tested. The expected trend found in the literature was that response rates are decreasing over time (Hypothesis 1) and, at the same time, nonresponse rates (refusal and non-contact rate) are increasing (Hypotheses 2, 6, and 7) and the rate of not able/others is constant (Hypothesis 5).

The first analysis is at the country level. A stacked bar chart (Figure 20) of the mean response rates and nonresponse rates per countries shows that at the country level, the mean response, refusal, non-contact and not able/others rates differ between countries.[107] Therefore, a more detailed analysis of the trend within countries was conducted.

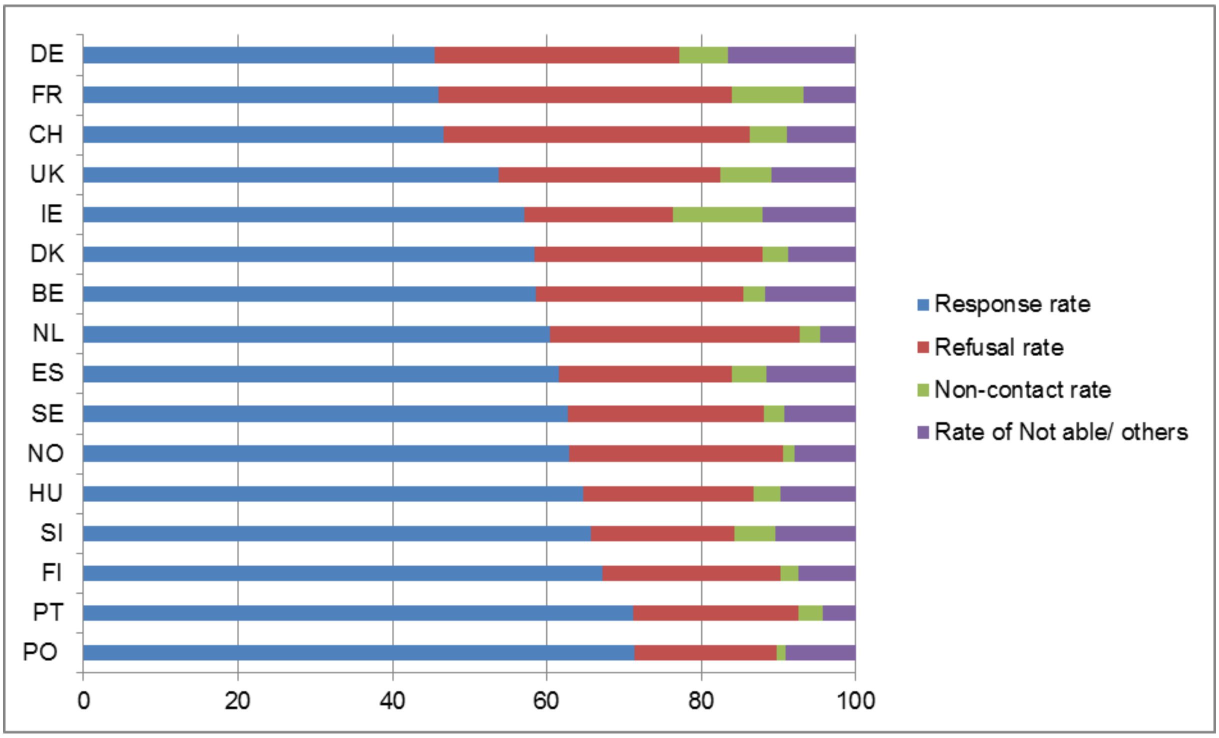

Figure 20. Stacked bar chart of mean response and nonresponse rates (in percentages) per country.

For the mean response rates including all countries, a slight negative tendency from the first round of the ESS in 2002 to the last round 2010 was detected (Figure 21). Even though the coefficient suggested a negative trend (1 percentage point in every two years or half a percentage point every year), the decline in response rates between 2002 and 2010 was not significant.

---

[107] For an overview of the trend of the outcome of the survey per round, see Appendix Figure A5, for round and country, see Appendix Figure A6, Figure A7, and Figure A8.

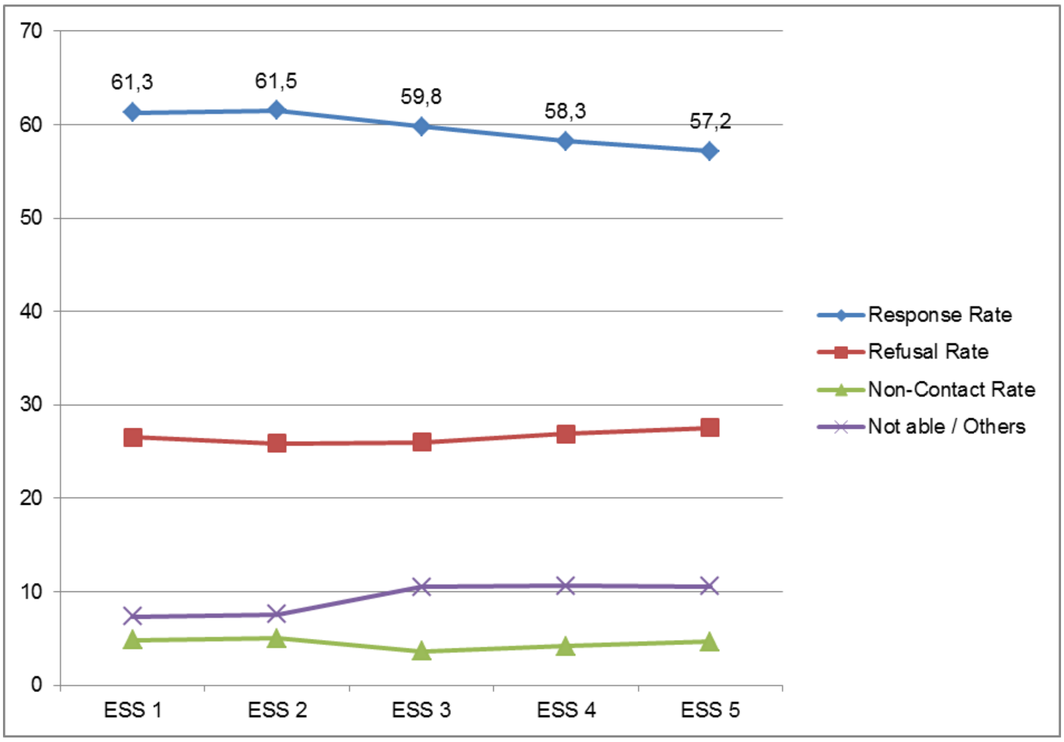

Figure 21. Response rate, refusal rate, non-contact rate, rate of not able/others (in percentages) for ESS 1 to ESS 5.

The development of the response rates varies between countries. In seven countries (Germany, Denmark, Finland, Hungary, Norway, Poland, and Sweden), a decrease in the response rate was significant. However, in Switzerland, Spain, and France, the opposite was the case and the response rate significantly increased.[108] It can be said that at the country level, response rates are changing, but not in the same direction for all countries. So a significant decrease of response rates, that is described in the survey literature can be found in almost half of the countries (7 out of 16), but not for all of those countries analyzed.

The second analysis of the development of response rates was done by examining whether the development of response rates at the country level was stable over time. The analysis of the countries with decreasing response rates showed that the decline was rather

[108] Based on the results of the coefficients from the POLS regression (Figure 6), countries were defined as increasing, decreasing, or without a tendency. Countries with increasing response rates are defined as those for which the effect of survey rounds on response rates are positive and significant at the 10% level; countries with a decreasing tendency are defined as those with a negative and significant rate at the 10% level; countries with no significant change are defined as those with no tendency.

large—11.7 percentage points from the first to the last round of the ESS. For the seven countries with decreasing response rates, a constant downward tendency could only be detected for Denmark, Finland, Norway and Poland. In the other countries, the response rate went up and down over time (see Figure 7).

To more fully analyze the trend in response rates development, the nonresponse also was examined. The nonresponse rates were examined to determine whether any drivers (refusal, non-contact, not able/others) might be influencing the development of response rates and whether a decrease in the response rate might trigger an increase in the nonresponse rate. Hypothesis 2 was tested to determine whether nonresponse rates are increasing.

With respect to the refusal rate at the country level, no general significant difference was detected for the mean in 2002 and 2010. Only Hungary, Norway, Sweden, and Slovenia showed a significant increase in refusal rates. In contrast, a significant decrease was found in the refusal rates in Switzerland and Spain. Thus, Hypothesis 3 that the refusal rate increases over time could not be supported for the majority of countries.

With respect to the non-contact rate (Hypothesis 4), as for the refusal rate, no significant difference was found for the mean non-contact rate in 2002 and 2010. This finding was seen at the overall mean level and the country level. The two expectations are Belgium and Spain where a significant decrease in the non-contact rate occurred. Thus, Hypothesis 4 that over time non-contact rates increase cannot be supported.

Table 7

Results of hypotheses tests

Hypotheses	Result
Response Rates: In general	
Hypothesis 1: Response rates are decreasing.	Not supported for all countries, but for some countries.
Nonresponse Rates: In general	
Hypothesis 2: Nonresponse rates are increasing.	No support of hypothesis.
Nonresponse rate: Different types of nonresponse	
Hypothesis 3: Refusal rates are increasing.	No support of hypothesis.
Hypothesis 4: Noncontact rates are increasing.	No support of hypothesis.
Hypothesis 5: Rates of “not able/others” are increasing.	Support of hypothesis.

In contrast to the findings for the refusal and non-contact rates, the regression model for the rate of not able/others showed a significant increase of the mean between 2002 and 2010. This finding does not support Hypothesis 5 (*Rates of not able/others are constant*). At the country level, a significant change was detected in Denmark, Finland, Spain, and Ireland where the rates of not able/others increased significantly.

To sum up, with respect to the development of the outcome codes of refusal, non-contact, and not able/others in the ESS between 2002 and 2010 at the country level, significant changes were detected for only a limited number of cases. For the mean response regarding the refusal and non-contact rates, no significant changes were found between 2002 and 2010. For the rate not able/others, a significant increase of the mean over time was found. So, the question about whether the response rate decreased in the ESS between 2002 and 2010 can be answered: no significant decrease was found in the mean response rate, and no significant increase was found in the mean refusal rate and the mean non-contact rate. In some countries, the response rate, the refusal and non-contact rates decreased or increased significantly.

The findings of the present study expand on those in the literature on the development of response rates. First and most important, a cross-national and across-time comparable data set was used. Thus, the response rate calculation is comparable across time and countries. In previous research, different surveys were used, and details about how the response rate was calculated were not transparent. Thus, the surveys included in the previous literature did not necessarily have comparable setups, ways to calculate response rates, or topics, which all had the potential to bias the results. In the ESS, the calculation of response rates is comparable, and also the information on response rates is accessible, available, and comparable across time and countries. Thus, these analyses are based on comparable high quality data basis. In addition, most of the research on the development of response rates has been carried out in the USA, and the ESS provides information for the European context. For Europe literature is only available on response rate before 2002 (de Leeuw & de Heer, 2002) and this study provides an update on this topic.

The ESS research is relevant because it shows that a general tendency of decreasing response rates cannot be detected in the data. However, country-specific trends in response rates do occur, and analyses have shown that over time response rates are becoming similar across Europe ("regression to the mean"). Not all response rates decrease, but in the majority of cases, the response rates from countries with high rates to begin with decrease, while countries with low response rates to begin with increase their response rates over time. The development of response rates is not an overall trend, but rather country-specific with respect to the development of response rates and also the development of nonresponse rates, including refusal, non-contact, and not able/others.

The findings of the ESS can be applied to other studies that may not provide as transparent, comparable, accessible, and documented information on the calculation of the response rate or that might not have standardized response rates across time and countries. Also, by using a European flagship survey with a high emphasis on response rates, other national surveys may be able to compare their rates with the ESS.

Analysis Part B: Response Rates and Nonresponse Bias

This chapter analyzes the relationship between response rates and nonresponse bias. In discussions about the development of response rates, the response rate itself often is used as a presumably easy indicator of data quality (see chapter II especially Section 1.2). In general, surveys with low response rates are considered as having problematic data quality because of a higher risk of high nonresponse bias (Goyder, 1987; Groves et al., 2008; Groves & Peytcheva, 2008; Peytcheva & Groves, 2009). Thus, in general, it is assumed that the higher the response rate, the better the data quality. For this reason, many survey guidelines and specifications require high response rates (European Social Survey, 2009; Johnson & Owens, 2003; Kreuter, 2013; Organisation for Economic Co-operation and Development, 2010). However, the relation between the response rate and nonresponse bias often is implicit.

Information from non-respondents is, as the term suggests, difficult to obtain because non-respondents do not take part in surveys, and thus their information is not available.

It is difficult to find comparable information or benchmark data that can be harmonized between the two datasets regarding variables, time, and survey populations to calculate nonresponse bias. Also the testing of the hypotheses on the relation of nonresponse bias and response rate is challenging because of the problems with the comparability of the calculations of the response rates (see chapter II especially Section 1.2 and 1.3). Data from the European Social Survey (ESS) provides comparable calculations of response rates. In addition its data can be harmonized and made comparable with external benchmark data and therefore provides information about nonresponse bias. Thus, the ESS enables an analysis of the hypothesis as to whether data quality regarding a nonresponse bias is getting larger when response rates are decreasing.

The following research question is systematically investigated in Analysis Part B: Does a correlation exist between response rates and nonresponse bias (as a measure of data quality)? By looking at these relations, we can answer the question as to whether an association exists between low response rates and high nonresponse bias.

In Analysis Part B, the following hypothesis is tested:

Hypothesis 6: The higher the response rate is, the lower the nonresponse bias.

The analyses of the nonresponse bias in Analysis Part B are based on data from the ESS and the Labour Force Survey (LFS). The *nonresponse bias* is defined as the deviation between the ESS and the LFS—as an external benchmark or "gold standard" (see Literature, Section 1.1 and Data and Methods, Section 7 and Equation 3 and 4).[109] The nonresponse bias is calculated for the socio-demographic variables gender, age, occupation, family status, education, nationality, and household size.[110]

Two forms of nonresponse bias measures were calculated. [111] The first is the *relative nonresponse bias*, which is the deviation of the ESS to the LFS, standardized to the "gold standard" LFS (see Equation 3). In the following text the relative nonresponse bias is also referred to as relative bias. For example, if the proportion of males in the ESS is 40% and the proportion in the LFS is 50%, the relative nonresponse bias for the variable "gender (male)" would be ([0.4 - 0.5] / 0.5 = -0.2). This relative nonresponse bias of -0.2 can be interpreted to

[109] As described in (Chapter III Data and Methods, Section 7) the method of comparison of the ESS data with external statistics ("gold standard") is used. This method of comparison is also used by (Groves, 2006; Hartmann & Schimpl-Neimanns, 1992; Holbrook & Krosnick, 2010; Koch, 1998; Krosnick, 1991; Krosnick et al., 2002; Levy & Lemeshow, 2013; Massey & Roger, 2013).

[110] Before I compared these data, I recoded the variables to make them comparable across the two data sets. For more information on this harmonization of the socio-demographic data from the ESS and LFS, see chapter Halbherr (2017, manuscript in preparation) and Chapter III Data and Methods, Section 6.

[111] In this analysis, other measurement errors or systematic bias, e.g., context or sampling effects, were not considered. As in the previous research of Groves et al. (2008); Peytcheva and Groves (2009, p. 196), the analysis is based on the assumption that no other error or bias (i.e., measurement error or coverage error) exist.

mean that if the ESS and LFS are compared, a relative bias of the ESS compared to the LFS of minus 20% can be detected. The algebraic sign shows that the proportion of the variable “gender (male)” in the ESS is, compared to the LFS, underrepresented. In some of the following analyses, the direction of the bias (over- or underrepresentation) is not primarily important.

The second measurement of nonresponse bias is the absolute value of the relative bias (Equation 4), which is in short also referred to as absolute relative bias. The absolute value of the relative nonresponse bias does not show the direction of the nonresponse bias, only its magnitude. In the previous “gender (male)” example, this would be a bias of 20%.

$$Relative\ Nonresponse\ Bias = \frac{\mathrm{ESS-LFS}}{LFS} \qquad (3)$$

$$Absolute\ value\ of\ relative\ Nonresponse\ Bias = \left|\frac{\mathrm{ESS-LFS}}{LFS}\right| \qquad (4)$$

The relative nonresponse bias and the absolute value of the relative nonresponse bias are used in the present study to analyze the different aspects of nonresponse bias. The advantage of the relative nonresponse bias is that it can provide information on both the overrepresentation and underrepresentation of the respective variables in the ESS, compared to the “gold standard” LFS (Equation 3), for example, whether *males* are overrepresented in a sample (see, for example, Figure 24). The relative nonresponse bias can be used to provide information about underrepresented variables whose inclusion in a sample might be difficult to accomplish (a “hard to get population”). An overrepresentation suggests that these repondent groups are easier to contact (an “easy to get population”).

Increasing response rates do not mean that all cases are equally represented (see Analysis Part D). Increasing response rates may lead to an overrepresentation of an “easy to get population” (see Chapter II, Section 2). Respondent groups who already are well included in an achieved sample could become even more overrepresented, and “more of the same” respondents could be included in the sample. This increase in response rate would increase the nonresponse bias, since only a certain group of respondents, namely “more of the same”, would be included in the net sample. while some groups who are considered to be a “hard to get population” would be less likely to take part in the survey. Thus, groups who are more likely to take part in surveys often are overrepresented, and the “hard to get” population” often is underrepresented (Goyder, 1987; Stoop, 2005).

In contrast to the information provided by the relative nonresponse bias that describes overrepresentation and underrepresentation, the absolute value of the relative bias (Equation 4) only includes the absolute value without its direction (see, for example, Figure 25). This reduction enables an interpretation of the hypothesis on the relation of the response rate and nonresponse bias, for example, “The higher the response rate is, the lower the nonresponse bias.” In this hypothesis, the focus is on the level of bias without considering its direction (overrepresentation or underrepresentation).

In the following chapters in Analysis Part B[112], the relationship between the response rate and nonresposne bias is examined in two steps. First, the anlaysis includes all variables. A pooled analysis of socio-demographic variables was calculated to test the hypothesis as to whether an overall pattern exists (see Section 12). Second, a variable-specific analysis was conducted to examine the effects of the separate socio-demographic variables on the nonresponse bias (see Section 13). These analyses provide a comprehensive overview as to whether a relationship exists between response rates and the socio-demographic variables of *gender, age, occupation, family status, education, nationality,* and *household size*. These

[112] In many surveys, post-stratification weights are constructed using socio-demographic variables to adjust for the nonresponse bias. The present analysis does not include a post-stratification weight, only design weights.

analyses were conducted using the relative nonresponse bias to provide an overview of whether the respective variable was overrepresented or underrepresented in the ESS, compared to the LFS (Equation 3). In a second step, the analysis used the absolute value of the nonresponse bias (Equation 4), which enabled the testing of the hypotheses on the relationship of response rates and nonresponse bias. These analyses helped to determine whether certain variables had a larger impact than other variables on the relationship of the response rate and the nonresponse bias.

12 Pooled Cross-Sectional Analysis: Overall Bias

This section tests Hypothesis 6 (*The higher the response rate is, the lower the nonresponse bias.*). As described previously, in a first step, the general relation of response rates and nonresponse bias is analyzed by pooling all the variables.

In this section, pooled cross-sectional analyses are conducted. In these multiple regressions, the dependent variables are the nonresponse bias of all socio-demographic variables. As described previously, the nonresponse bias was calculated by comparing the ESS data to the benchmark or “gold standard” of the LFS (see Equation 3 and Equation 4). The independent variable is the response rate in the ESS. The pooled cross-sectional analyses of all bias variables followed the pioneering article written by Groves and Peytcheva (2008) who analyzed the relation of response rates to nonresponse bias for US surveys.

The analyses (Figure 22) performed by the present study—based on the European context—are grounded in and extend the findings of Groves and Peytcheva (2008) and Groves (2006). Also, the present study extends the previous research by using intercomparable surveys with the same set-up, the same topic, and the same calculation of response rate. Also, the bias variables are kept constant over the different points in time and surveys. By keeping constant these factors and others, the noise in their models of the relation of response rate to nonresponse bias was reduced.

Two multiple regressions were run to predict the response rate from the nonresponse bias variables. The first one included the relative bias, the second the absolute value of the relative bias. Due to the correlation of response rates across time (autocorrelation), the regressions were conducted using robust standard errors.[113] The respective confidence intervals were estimated as well. The variables included as dependent variables were the relative nonresponse bias (see Equation 3) from the data from the ESS compared to the external benchmark of the LFS.[114]

The nonresponse bias variables included in the multiple regression were gender (male), young respondents, old respondents, working population, married persons, low education, high education, persons having the nationality of the country in which the study was conducted, persons living in one-person household, and persons living in a five-or-more person household. These socio-demographic variables were selected because in literature, they often are used for weighting and nonresponse bias analysis. In addition, these selected variables are the general socio-demographic variables used for general substantive analyses. Therefore, these variables play a central role in surveys.

The first multiple regression included the relative nonresponse bias (see Equation 3) of all of the nonresponse bias variables (see Table 8). The analyzed nonresponse bias

---

[113] Stata command vce (robust) uses the robust or sandwich estimator of variance. This estimator is robust to some types of misspecification given the observations are independent.

[114] Extreme outliers were detected based on the calculation of the box-and-whisker plot. An *outlier* is defined as any case with a value that lays more than one and a half times the length of the box from either end of the box. That is, if a data point is below Q1 – 1.5×IQR or above Q3 + 1.5×IQR, it is defined as being too far from the central values to be reasonable. Thus, those cases were excluded from the further analysis. These cases excluded as outliers were the education variables in the United Kingdom and the variable nationality in Poland.

variables statistically significantly predicted the response rate (F = 7.93, p < .000).[115] The variables gender (male), married, working population, and low education added statistically significantly to the prediction, p < .05. The coefficient of determination (R^2) showed that 64.2% of the variance in the dependent variable (nonresponse bias variables) can be explained by the independent variable, which is the response rate. This finding leads to the conclusion that the regression mode of the relative nonresponse bias and the response rate is a good fit for the data.

Figure 22 illustrates the first regression analysis in a scatterplot. The x-axis shows the response rate (independent variable) and the y-axis shows the nonresponse bias variables (dependent variables). The deviation from the 0% y-axis describes the difference between the percentage in the ESS data and the LFS data. A 0% deviation means that the percentage in the ESS and LFS are equal. A value below the 0 line describes an underrepresentation of a proportion in the ESS compared to LFS. If the value is above the 0 line value on the y-axis, the variable is overrepresented in the ESS compared to the LFS. The further away from the y-axis, the bigger is the difference between the ESS and LFS. The graphs include the same variables for the relative nonresponse bias as in the previous regression analysis, namely gender (male), young respondents, old respondents, working population, married persons, low education, high education, person having the nationality of the country in which the study was conducted, one-person household, and five-or-more persons household; and includes all 16 countries that took part in all rounds of the ESS.

[115] The analysis that included the outliers (United Kingdom education variables and Poland nationality) provided similar results (F = 10.13, p < .000; R^2 = .58). The variables male, occupation, and nationality added statistically significantly to the prediction, p < .01.

Table 8

Regression analysis nonresponse bias and response rate

					95% CI	
	Coef.	*t*	*p*		Lower bound	Upper bound
Relative Nonresponse Bias						
gender (male)	-.007	-2.76	.009	***	-.013	-.002
young respondents	.001	1.21	.233		-.001	.003
old respondent	.001	0.97	.340		-.001	.003
working population	.002	1.80	.080	+	-.000	.005
married persons	-.004	-2.53	.016	**	-.007	-.001
low education	.003	3.55	.001	***	.001	.004
high education	.001	0.87	.391		-.001	.002
nationality	-.009	-1.53	.134		-.021	.003
one-person household	.000	0.12	.908		-.002	.002
five-or-more person household	-.001	-0.46	.648		-.003	.002
_cons	.654	28.45	.000		.608	.701
Relative Nonresponse Bias (absolute value)						
gender (male)	.004	1.37	.179		-.002	.010
young respondents	-.000	-0.10	.920		-.005	.004
old respondent	-.000	-0.42	.678		-.004	.002
working population	-.001	-0.52	.610		-.006	.004
married persons	-.003	-1.44	.159		-.008	.001
low education	-.004	-4.58	.000	***	-.005	-.002
high education	-.001	-0.72	.474		-.003	.001
nationality	-.010	-1.29	.206		-.026	.006
one-person household	-.001	-0.53	.596		-.003	.002
five-or-more person household	-.001	-0.44	.664		-.004	.002
_cons	.736	16.08	.000		.643	.828

Note. CI = confidence interval. Source: Response Rate from the ESS; Nonresponse bias: comparison of the harmonized socio-demographic variable in the ESS to the LFS for the 16 countries that took part in all rounds of the ESS. Number of observations n = 46.
+ p < .1. *** *p* < .01. ** *p* < .05. * *p* < .10.

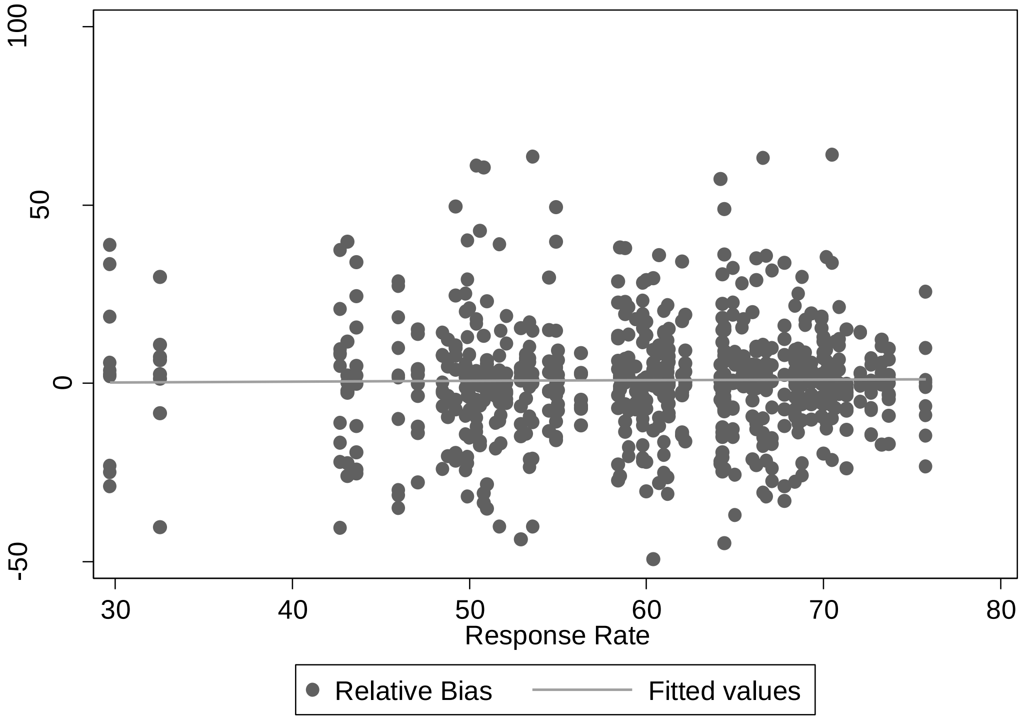

Figure 22. Relative nonresponse bias and response rate.

To calculate the correlation coefficient (*r*) and do a regression analysis, the structure of the data needed to be rearranged. The data was transferred to a dataset with a long data structure. This data structure for all countries and rounds included all the different nonresponse bias variables. The Pearson Correlation Coefficient (r = .01; p = .757) showed a low strength of association between the two variables—response rate and the nonresponse bias. In addition, a linear regression analysis was calculated to display the slope (Figure 22 and Appendix Figure B1).[116] The regression analysis showed a positive but not significant correlation (coef = .02; t = .28; p = .776).

The second multiple regression (Figure 23 and Appendix Figure B2) included the absolute value of the relative nonresponse bias (Equation 4). Using the absolute value of the

[116] See Appendix Figure B1 for a graph in which the nonresponse bias variables are displayed separately for each variable in different colors.

relative bias enabled an interpretation of the increase or decrease of response rates, without taking into account over- or underrepresentation. The absolute value of the relative bias provided similar results as the relative bias (see Table 8). The analyzed nonresponse bias variables statistically significantly predicted the response rate (F = 9.33, p < .000).[117] The coefficient of determination (R^2) showed that 65.5% of the variance of the absolute value of the relative nonresponse bias variables was explained by the response rate. The findings of the multiple regression analysis showed that the regression model was a good fit for the data.

The Pearson Correlation Coefficient (r = -.16; p = .000) showed a small and significant strength of association between the two variables—response rates and the nonresponse bias (absolute value of the relative bias). As can be seen in Table 8, for the absolute value of the relative nonresponse bias, the variable “low education“ adds statistically significantly to the prediction (p < .05).

The regression analysis for the absolute value of the nonresponse bias is shown in Figure 23. This analysis suggests a general negative tendency: The higher the response rates are, the lower the nonresponse bias. The linear regression analysis[118] supports these findings of a negative and significant correlation (coef = -.17; t = -3.85; p = .000), which also supports Hypothesis 6 (*The higher the response rate is, the lower the nonresponse bias.*).

The variance of the nonresponse bias for some variables differs. As can be seen in Figure 23, some differences exist between the variables regarding the amount of nonresponse bias. For some variables the nonresponse bias is rather low, whereas for others, the bias is

[117] The analysis that included the outlier (the variable education in the United Kingdom and the variable nationality in Poland) provided similar results (F = 10.18, p < .000; R^2 = .62). The variables male and occupation added statistically significantly to the prediction, p < .01.

[118] See Appendix Figure B2 for a graph in which the different nonresponse bias variables are portrayed in different colors.

high. The general pattern supports the conclusion that no clear relationship exists between the response rates and nonresponse bias for all variables. The results of the regression analysis and the visual interpretation (Figure 22, Appendix Figure B2, and Table 8) support Hypothesis 6, which states that higher response rates are correlated with a lower nonresponse bias.

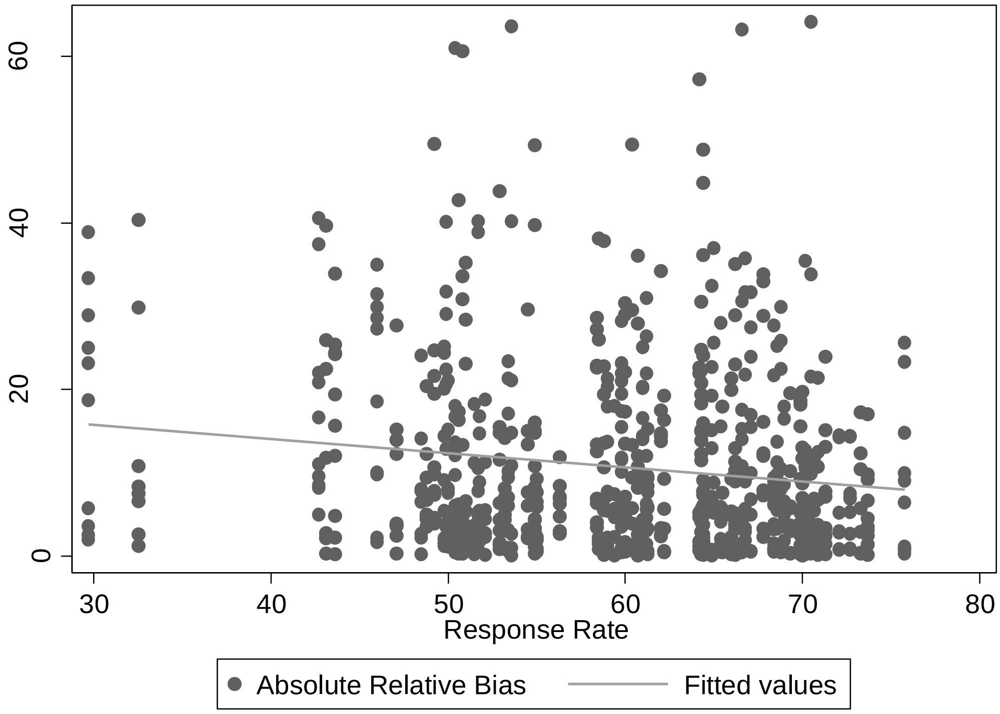

Figure 23. Absolute value of the relative nonresponse bias and response rate.

13 Cross-Sectional Analysis: Variable Specific Bias

The general question asked in this section focuses on whether a relation exists between response rates and nonresponse bias. Do countries with low response rates have a higher nonresponse bias? Following the analysis of the overall bias in the previous section, this section further examines the nonresponse bias at the variable level. This analysis further extends the research of Groves and Peytcheva (2008) and Groves (2006) who analyzed the relationship of a mix of different variables from different studies. The present study took a closer and separate look at each of the socio-demographic variables—gender, age, occupation, family status, education, nationality, and household size. Thus it provides information about whether response rates affect the different variables differently.

We tested Hypothesis 6 (*The higher the response rate is, the lower the risk of nonresponse bias.*). We analyzed the correlation between response rates and nonresponse bias, the relative nonresponse bias (see Equation 3), and the absolute value of the relative nonresponse bias (see Equation 4). As described previously, the nonresponse bias was calculated by comparing the ESS data to the benchmark or "gold standard" of the LFS (see Equation 3 and Equation 4). Regressions were run for each variable for the relative and absolute values of the nonresponse bias. Due to correlation of response rates across time (autocorrelation), regressions were conducted using robust standard errors. In addition, a visual analysis was conducted and the strength of the binary association between the response rate and the respective socio-demographic variable was calculated. The relation of the response rate (independent variable, displayed on the *x*-axis) to the nonresponse bias (dependent variable, displayed on the *y*-axis) was analyzed and depicted in scatterplots.

The analysis was carried out in two steps. First, for the relative nonresponse bias, a linear regression analysis was calculated with the response rate as the independent variable and the relative nonresponse bias for the respective variables as the dependent variable.[119] Analyzing the relative bias provided a general overview of the distribution of the nonresponse bias across different response rates, which showed the over- and under-representation of the ESS data in comparison to the LFS.

[119] Outliers were identified by calculating a box plot and identifying extreme outliers. This was the case for the education variable in the United Kingdom, in Sweden for the old respondent and in Poland for the nationality.

In a second step, the same analysis of the absolute values of the relative nonresponse bias was conducted. The absolute value did not consider the over- or underrepresentation; rather, only the absolute nonresponse bias without the direction was analyzed. For this absolute value, a binary correlation using the Pearson Correlation Coefficient was calculated. Also a linear regression analysis was calculated to display the slope. Due to the low number of cases, not only the level of significance was used to describe the correlation, but also the general pattern. By using both the pattern and the test of significance, a more comprehensive picture of the absolute values of the relative bias was obtained.

Due to comparability reasons, the same y-axis is used. The of course can lead to the problem that labels are not readable anymore. Therefore for each of the analysis additional graphs are included in the Appendix Figure B3 to Figure B22.

13.1 Gender (male)

The general Hypothesis 6 was tested at the variable level. When applied to the variable "gender (male)," the hypothesis is that the higher the response rates are, the lower the bias for the variable.

First, the relative nonresponse bias was analyzed. The relation between the relative bias of the variable "gender (male)" and the response rate is displayed in Figure 24. Over the 0 line of the *y*-axis, an increase in the response rate could be described as an increase in bias and overrepresentation which means an overrepresentation in the ESS data in comparison to the LFS. Under the 0 on the *y*-axis nonresponse bias, an increase in the response rate could be described as a decrease in the nonresponse bias, meaning a lower underrepresentation. For the variable "gender (male)," cases with a relative nonresponse bias over as well under the 0 line at the *y*-axis were found. A regression analysis suggested that as response rates increase, the nonresponse bias gets larger (see Figure 24 with comparable *y*-axis across all graphs and

Appendix Figure B3). The slope of the regression is negative. The negative correlation suggests that with increasing response rates, men are becoming increasingly underrepresented. The correlation of response rate and nonresponse bias for this variable is significant (coef = -.01; t = -2.54; p = .013; R^2 = .07). In comparison to other variables (see Figure 26 to Figure 38), the relative nonresponse bias for the variable “gender (male)” is rather low.

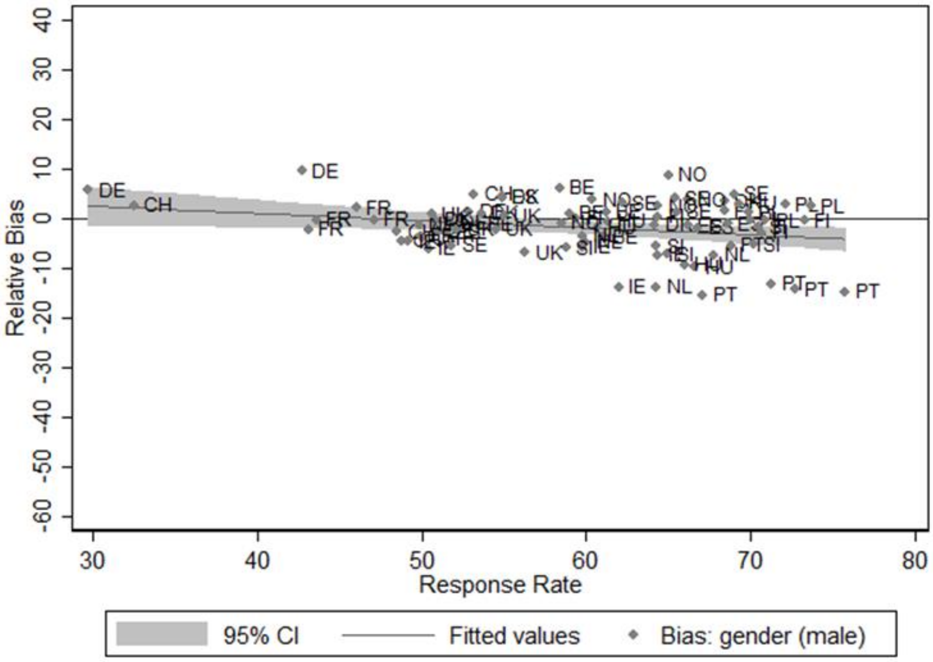

Figure 24. Relative bias for the variable “gender (male)” and the response rate.

In a second step, an analysis of the absolute value of the relative bias was conducted (see Figure 25 with comparable y-axis across all graphs and Appendix Figure B4). The regression showed a slight but significant correlation between the absolute value of the nonresponse bias for the variable “gender (male)” and the response rate. The Pearson Correlation Coefficient (r = .22; p = .049) showed a low strength of association between the two variables response rate and nonresponse bias. A positive correlation was detected (coef = .01; t = 2.48; p = .015). Higher values of relative nonresponse bias for the variable “gender (male)” were associated with higher response rates, which does not support Hypothesis 6

(*The higher the response rate is, the lower the nonresponse bias.*). In fact, the data suggested the contrary: the higher the response rate is, the higher the nonresponse bias.

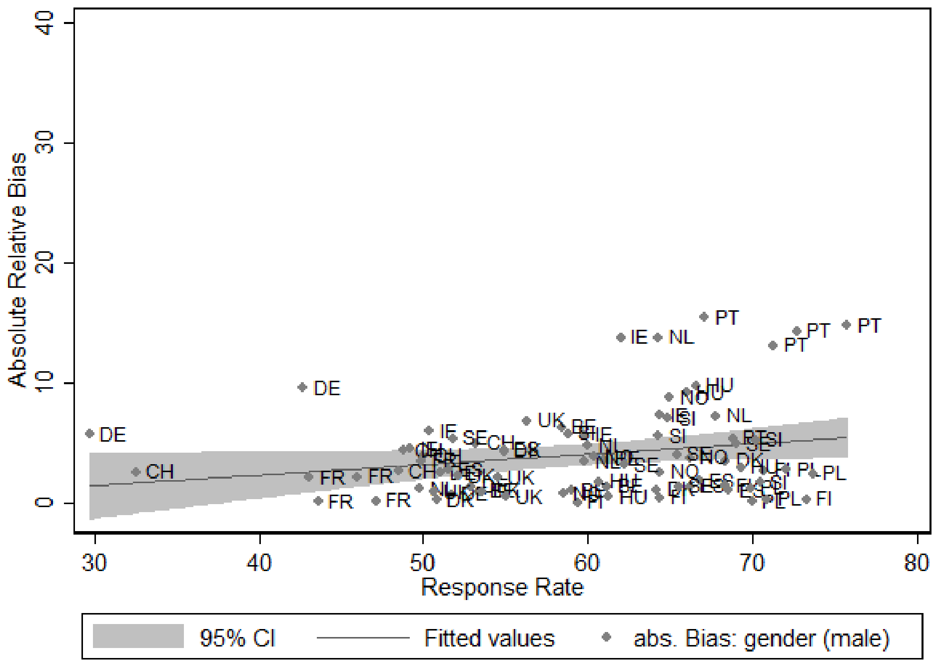

Figure 25. Absolute relative bias for the variable “gender (male)” and the response rate.

13.2 Young Persons

As for the other variables, Hypothesis 6 (*The higher the response rate is, the lower the risk of nonresponse bias.*) was applied to the variable “young persons” to determine whether higher the response rates are related to a lower nonresponse bias. Again, first an analysis of relative bias was carried out, which provided information on the over- or under-representation of the variable “young persons” in the ESS in comparison to the LFS. Young persons were underrepresented in the ESS in most countries. With higher response rates, the relative bias did not change significantly (coef= -.00; t = 1.14; p = .259; R^2 = .02). An analysis of the correlation of the response rate and nonresponse bias for young people found that the relative bias in general was small (Figure 26 and Appendix Figure B5), but larger than for the variable “gender (male)” (Figure 24).

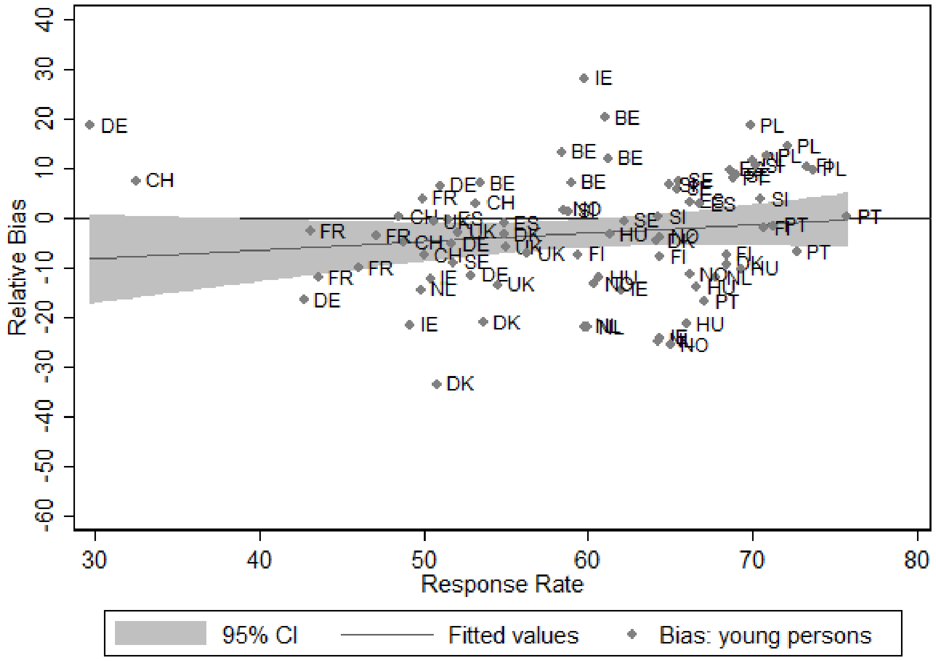

Figure 26. Relative bias for the variable “young persons” and the response rate.

In a second step, the relation of the absolute value of the relative nonresponse bias was analyzed. A regression analysis of the absolute value of the relative bias was conducted by calculating the Pearson Correlation coefficient and a regression. The Pearson Correlation Coefficient (r = -.13; p = .906) showed a small negative, but not significant, association between the two variables. A regression analysis (Figure 27 and Appendix Figure B6) showed no significant change of the bias when response rates changed (coef = -.01; t = -1.36; p = .117; R^2 = .02). This finding was contrary to the expectation formulated in Hypothesis 6. The data did not show a correlation between the higher the response rates and lower nonresponse bias.

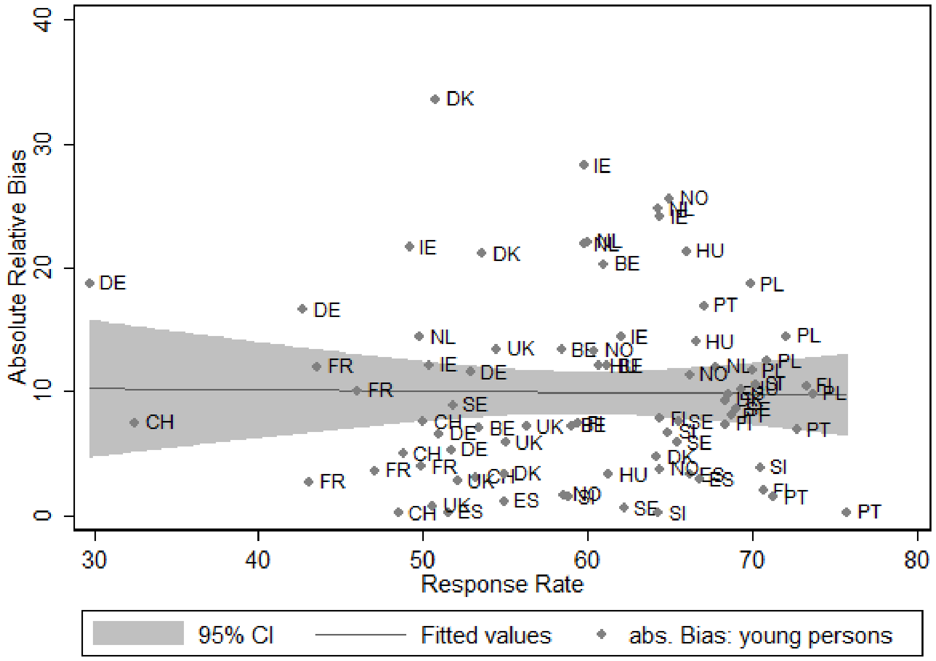

Figure 27. Absolute relative bias for the variable “young persons” and the response rate.

13.3 Old Persons

Again, Hypothesis 6 was tested stating that the higher the response rates the lower the nonresponse bias. The hypothesis was applied to the variable “old persons.” Older respondents were defined, as described previously, as age 75 and older.

As can be seen in Figure 28 and Appendix Figure B7, the analysis of the relative bias, in most countries people aged 75 and older are underrepresented in the ESS in comparison to the LFS.[120] A regression analysis showed a highly significant correlation between the level of response rates and the nonresponse bias. The higher the response rate was, the lower the bias (coef = .00; t = 4.03; p = .000; R^2 = .22). In general, in comparison to the age group of the young respondents, the underrepresentation of the older respondents is larger.

---

[120] Information from Norway for this variable was not available. Based on the analysis of the boxplots, extreme outliers in Sweden were deleted in the above analyses.

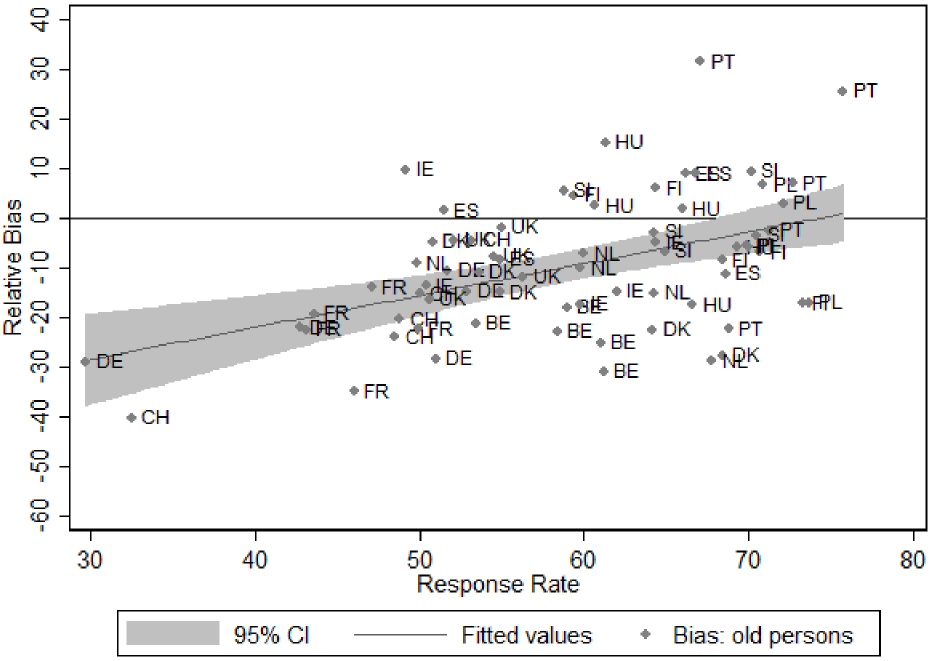

Figure 28. Relative bias for the variable “older persons” and the response rate.

As in the previous analyses, we looked at the absolute value of the relative bias, and only analyzed the absolute value of the bias without considering the direction (over- or underrepresentation). The Pearson Correlation Coefficient showed a small negative and significant correlation between the bias and the response rate (r = -.35; p = .003). In addition, a regression analysis was calculated (see Figure 29 and Appendix Figure B8). The analysis of the absolute value of the relative bias showed the same results (coef = -.04; t = -4.46; p = .000; R^2 = .26) as the Pearson coefficient and the analysis of the relative bias. For the variable “old persons,” the higher the response rate is, the lower the nonresponse bias. This result supported Hypothesis 6 (*The higher the response rate is, the lower the nonresponse bias.*).

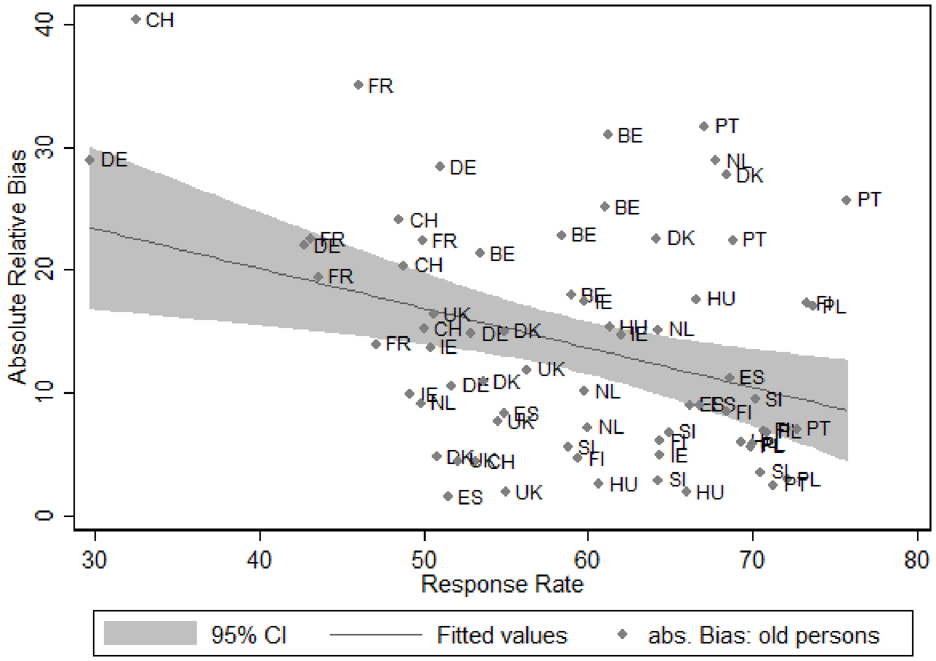

Figure 29. Absolute relative bias for the variable “old persons” and the response rate.

13.4 Working Population

Similar to the variables above, Hypothesis 6 was tested for the variable “working population”: the hypothesis that the higher the response rate is, the lower the bias for the variable.

In most countries studied by the ESS, the working population is underrepresented (see Figure 30 and Appendix Figure B9). Yet, the level of underrepresentation differs between countries. In some countries like in Norway and Switzerland, the bias is rather low, whereas in Portugal and Ireland, the bias is high. In comparison to the other analyzed variables in this section, the level of nonresponse bias is intermediate for this variable.

A regression analysis showed that for the relative nonresponse bias for the variable “working population,” no significant correlation existed between the response rate and the nonresponse bias (coef = -.00; t = -1.46; p = .149; R^2 = .02).

Figure 30. Relative bias for the variable "working population" and the response rate.

Furthermore, the analysis of the absolute value of the relative bias showed the same result (Figure 31 and Appendix Figure B10). The Pearson Correlation Coefficient showed a weak negative and significant correlation (r = -.35; p = .003) between the variables response rate and nonresponse bias. A regression analysis also showed a negative tendency, which was not significant (coef = - .00; t = 1.38; p = .172; R^2 = .02). So, the hypothesis that "the higher the response rate is, the lower the nonresponse bias" was not supported by our results, since no significant correlation was found between the two variables.

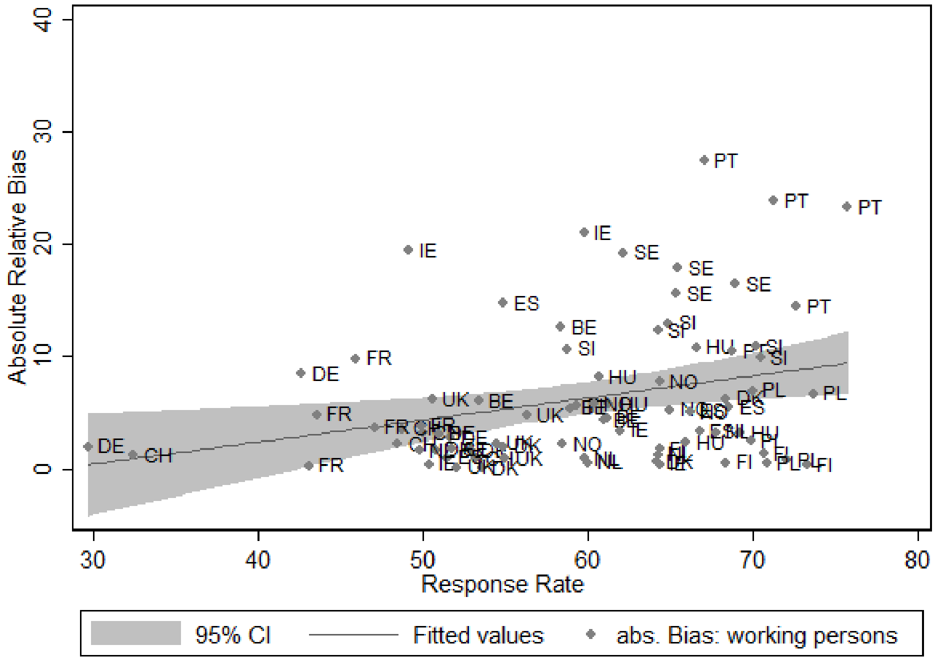

Figure 31. Absolute relative bias for the variable “working population” and the response rate.

13.5 Family Status: Married Persons

As for the previous variables, Hypothesis 6 was tested for the “married persons” variable: the higher the response rate is, the lower the nonresponse bias. Figure 32 and Appendix Figure B11 shows that in most of the countries, married persons are overrepresented in the ESS compared to the LFS, which means that in the ESS, the percentage of “married persons” is higher compared to the LFS. The level of the nonresponse bias for the “married persons” variable was at an intermediate level compared to the other variables analyzed in this section. A regression analysis of the relative nonresponse bias for “married persons” and response rate showed a highly significant correlation between the response rate and the nonresponse bias for the variable “married persons” (coef = -.00; t = -3.13; p = .003; R^2 = .09). With increasing response rates, the overrepresentation decreased, which was a decrease in the nonresponse bias. Figure 32 shows that up to a level of response rate of 70%, the nonresponse bias decreased.

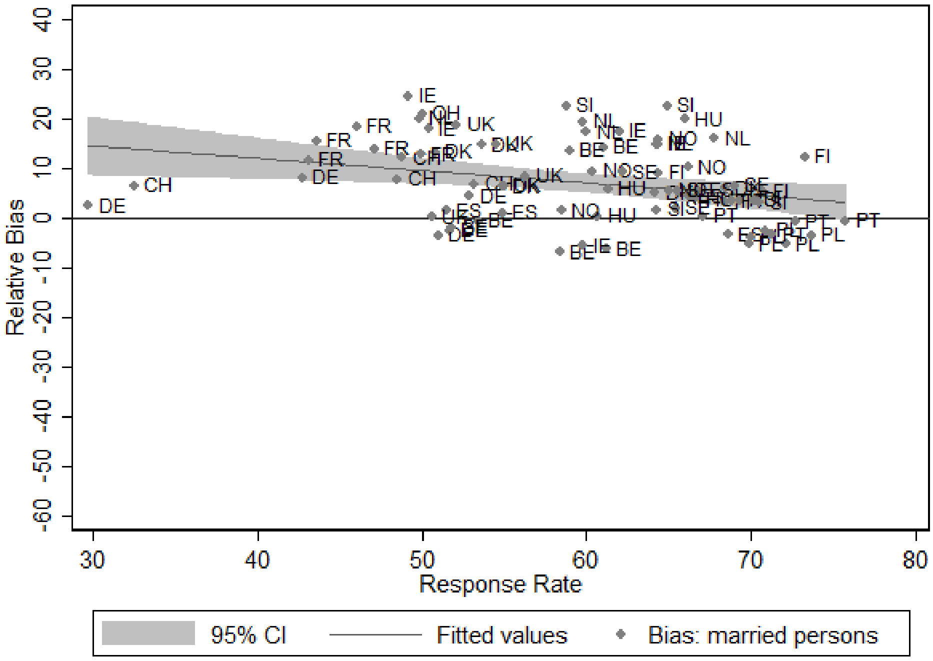

Figure 32. Relative bias for the variable “married persons” and the response rate.

An analysis of the absolute value of the relative bias, which as described previously only considered the level of bias without taking into account the nonresponse bias, supported these findings. The Pearson Correlation Coefficient (r = -.26; p = .0220) showed a small and significant correlation of association between the two variables absolute value of the relative bias and response rate. Also, the absolute value of the relative bias (coef = -.007; t = 3.31; p = .001; R^2 = .10) supported Hypothesis 6. As can be seen in Figure 33 and Appendix Figure B12, the higher the response rate is, the lower the bias for the variable “married persons.”

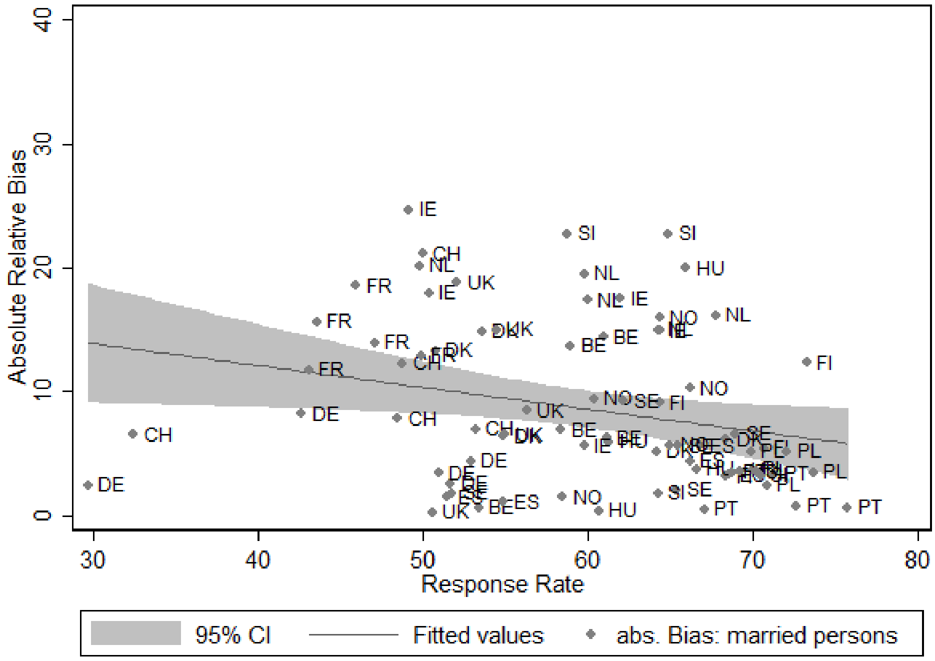

Figure 33. Absolute relative bias for the variable “married persons” and the response rate.

13.6 Low Education

As for the other socio-demographic variables, Hypothesis 6 (*The higher the response rate is, the lower the nonresponse bias.*) was tested for the variable “low education.” The level of education is classified by the ISCED level.[121] *Low educated persons* are defined as those with an ISCED level of 0-2.[122]

As shown in Figure 34, the proportion of low educated persons in the ESS in comparison to the LFS differs. In comparison to the other variables analyzed in this section, the level of nonresponse bias is rather high for the variable “low education.”[123] An over- as

[121] The ISCED is the acronym for the International Standard Classification of Education. It is a statistical framework for organizing information on education maintained by the United Nations Educational, Scientific and Cultural Organization (UNESCO). Low levels of education are described as O–2, medium educational level as 3–4, and high educational levels are defined as 4–6. For further information on the ISCED, see (Organisation for Economic Co-operation and Development, 1999; Schneider & Kogan, 2008).
[122] Extreme outliers were detected by boxplot analysis. The extreme outliers in the United Kingdom data were not included in this analysis.
[123] In addition to the explanation of an increasing nonresponse from less educated persons (Helmschrott & Martin, 2014), harmonization issues between the LFS and the ESS also might help to explain the high level of

well as underrepresentation was found, which suggests that in some countries, the proportion of low educated persons in the ESS data was higher, and in other cases, lower. For the majority of countries, an underrepresentation of persons with low education was detected.

A regression analysis (Figure 34 and Appendix Figure B13) showed a highly significant correlation between the response rate and nonresponse bias. With an increasing response rate, the underrepresentation of respondents with low education decreased (coef = .00; t = 4.49; p = .000; R^2 = .19).[124] This finding suggests that with higher response rates, the nonresponse bias for the variable "low education" decreases, and the data is more accurate.

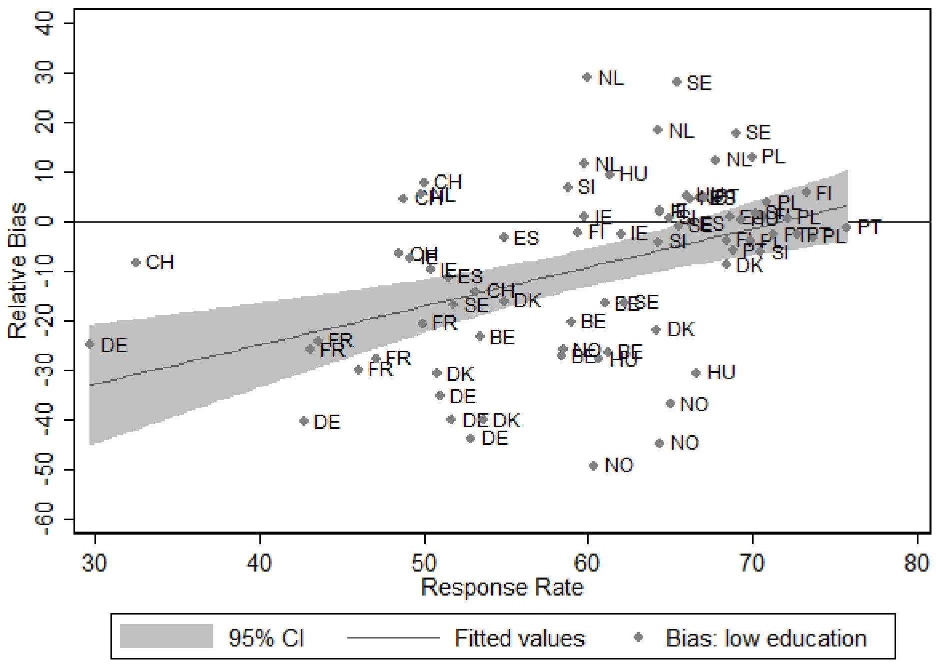

Figure 34. Relative bias for the variable "low education" and the response rate.

bias. For details, on the harmonization between the LFS and ESS and the problem of the mapping of the United Kingdom in the LFS to the ISCED groups is unclear in some details (see Schneider, 2008). Due to the United Kingdom being an extreme outlier, the data of this country has been excluded.

[124] When the outlier United Kingdom was included, no significant correlation was detected (coef= -.00; t = 1.27; p= .207; R^2= .31).

A second step analyzed the absolute value of the relative bias. The Pearson Correlation Coefficient showed a medium negative and significant correlation between the nonresponse bias and response rate (r = -.42; p = .000). A regression analysis supported this finding (Figure 35 and Appendix Figure B14). With an increasing response rate, the nonresponse bias decreased significantly (coef = -.01; t = -4.92; p = .000; R^2 = .19).[125] Thus, for the variable "low education," Hypothesis 6 can be supported—"the higher the response rate is, the lower the nonresponse bias."

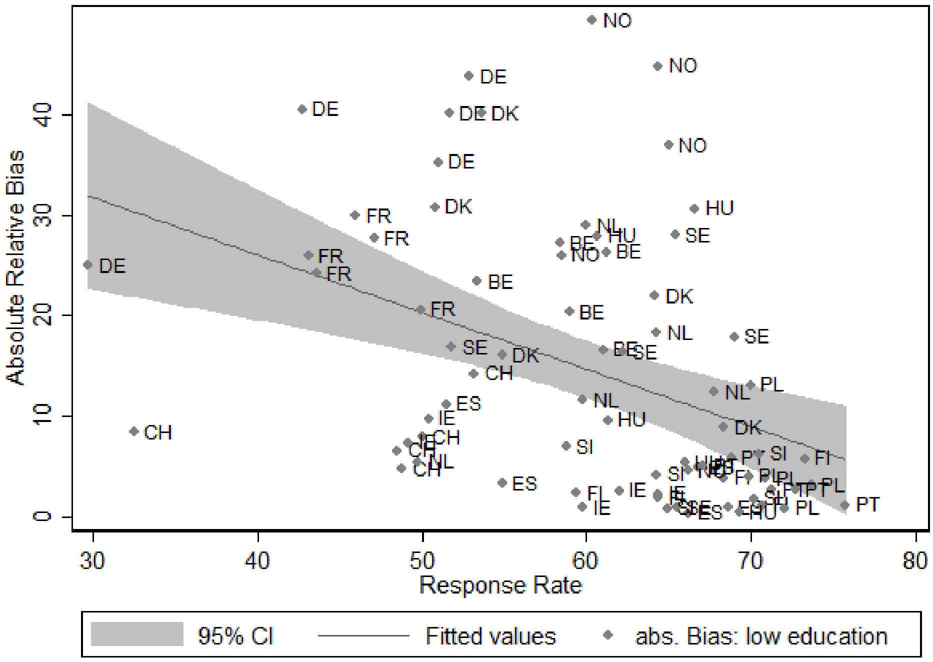

Figure 35. Absolute relative bias for the variable "low education" and the response rate.

13.7 High Education

The relation of response rate and nonresponse bias was analyzed with respect to the variable "high education." Highly educated persons are defined as those having an ISCED level of 5 or 6.

---

[125] When the outlier United Kingdom was included, a significant correlation was detected, too (coef= -.00; t = -4.21; p= .000; R^2= .16).

First, the relation of the relative bias for the variable “high education” and the response rate was analyzed (Figure 36 and Appendix Figure B15).[126] The graph shows a pattern that highly educated people are in most countries overrepresented in the ESS compared to the LFS. Compared to the other variables, the level of nonresponse bias for “high education” is rather high. A regression analysis showed a highly significant correlation between response rates and nonresponse bias for the variable “high education” (coef = -.00; t = -3.56; p = .001; R^2 = .13).[127] The overrepresentation decreases with an increasing response rate. This finding showed that higher response rates are correlated to a lower nonresponse bias.

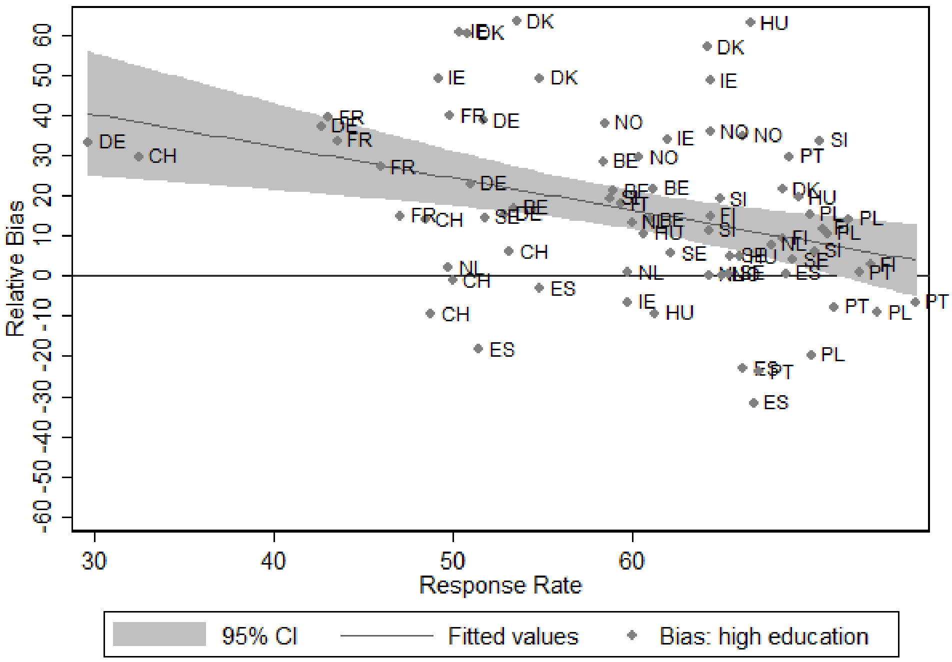

Figure 36. Relative bias for the variable “high education” and the response rate.

A regression analysis of the absolute value of the relative bias (Figure 37 and Appendix Figure B16) showed the same pattern as for the relative bias. Also, a negative and

[126] As for the analysis of “low education,” the extreme outlier for the “high education” variable in the United Kingdom was excluded.
[127] If the outlier United Kingdom is included, the result is the same (coef = -.00; t = -3.46; p = .001; R^2 = .12).

significant correlation was detected (coef = -.01; t = 4.20; p = .000; R^2 = .14). The strength of association was measured by the Pearson Correlation Coefficient, which showed a negative medium and significant strength of association (r = -.33; p = .004) between the two variables response rate and nonresponse bias for the variable “high education”. So, the results for the variable “high education” support Hypothesis 6: the higher the response rate is, the lower the nonresponse bias for the variable “high education.”

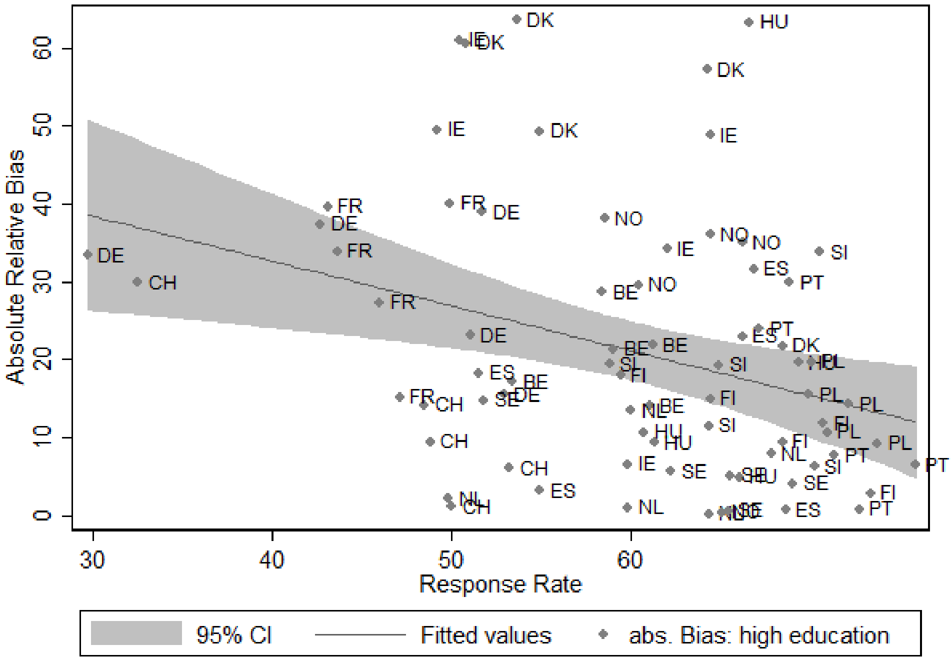

Figure 37. Absolute relative bias for the variable “high education” and the response rate.

13.8 Nationality

Hypothesis 6 also was tested for the socio-demographic variable “nationality,” which includes the percentage of respondents with the nationality of the analyzed country.

Immigration groups are perceived as less likely to respond than the indigenous population. Voogth (2004, p. 106) states that in the Netherlands this is due to the fact, that they are generally concentrated in specific, areas, and probably less integrated and less accessible to survey invitations. More up to date studies also found underrepresentation of migrants (e.g. in Germany see Kortmann, & Halbherr (2009). Compared to the variables "low and high education," the nonresponse bias for the variable "nationality" is low (see Figure 36). Figure 38 and Appendix Figure B17 show that persons with the nationality of the country analyzed are slightly overrepresented in the ESS compared to the LFS. With decreasing response rates, the nonresponse bias also decreased. A regression analysis supported the findings of the visual analysis.[128] A significant correlation between response rates and relative nonresponse bias was found (coef = -.02-; t = -5.52; p = .000; R^2 = .30).

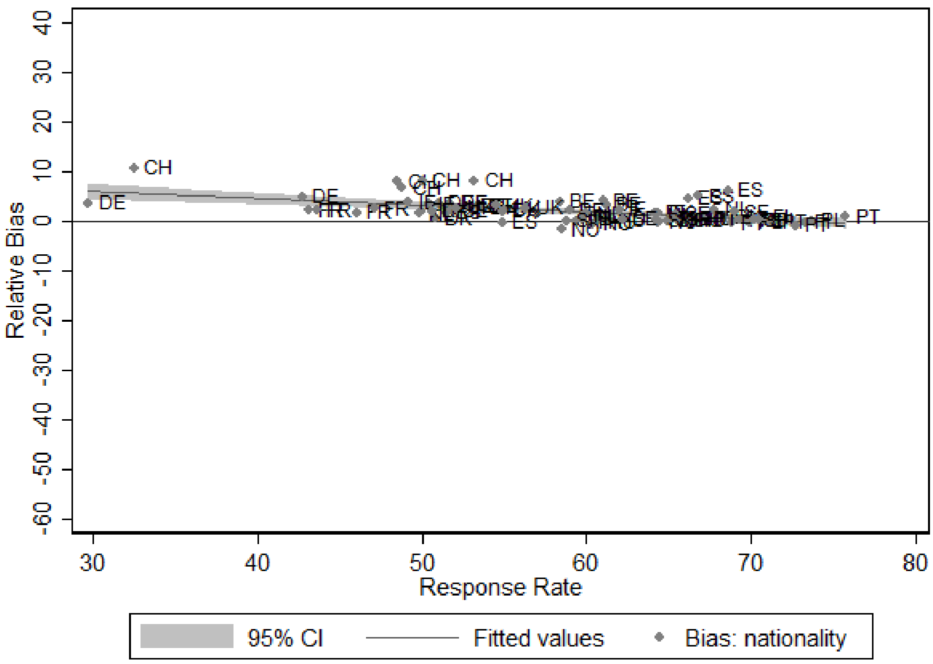

Figure 38. Relative bias for the variable "nationality" and the response rate

The analysis of the absolute value of the nonresponse bias provided similar results. The Pearson Correlation Coefficient on the strength and the direction of the relation between the response rate and the absolute value of the relative bias showed that higher response rates are related to a lower nonresponse bias (r = -.57; p = .000). The same conclusion can be

[128] The extreme outlier Poland was excluded from this analysis. Including Poland in the regression analysis provided similar results (coef = -.23; t =-5.97; p = .000; R^2 = .32).

drawn for the regression analysis on the absolute value of the relative bias and response rate (Figure 39 and Appendix Figure B18). The negative coefficient describes a significant negative correlation (coef = -.03; t = 5.39; p = .000; R^2 = .31). This supports Hypothesis 6. For the variable "nationality", higher response rates are associated with a lower nonresponse bias.

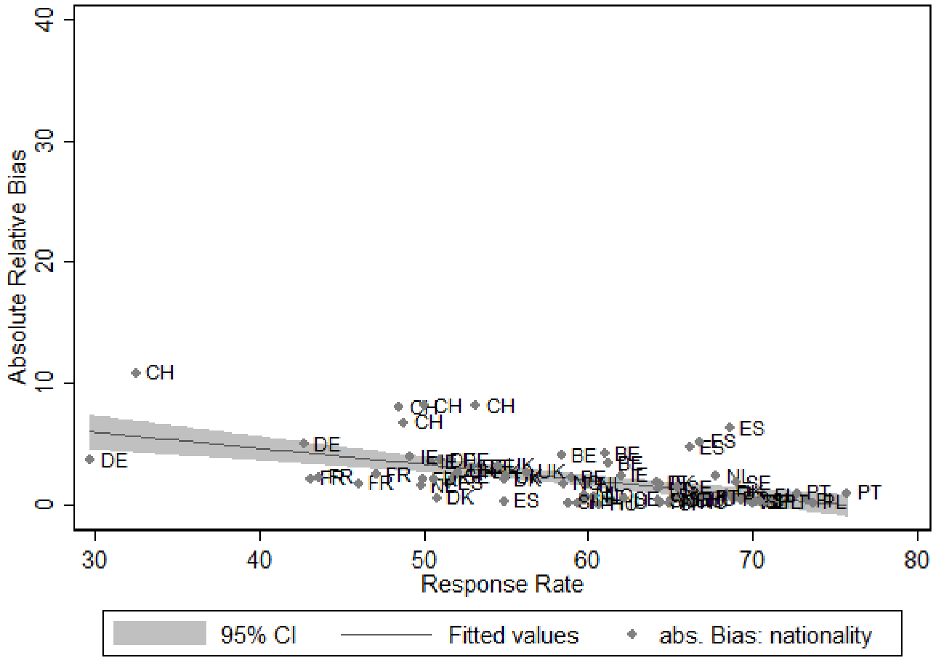

Figure 39. Absolute relative bias for the variable "nationality" and the response rate.

13.9 One-Person Household

Similar to the analysis of the other socio-demographic variables, the "one-person household" variable was analyzed with respect to the potential correlation between the response rate and the nonresponse bias.

As shown in Figure 40, in most of the countries, persons living in one-person households are underrepresented in the ESS compared to the LFS. Compared to the other variables analyzed in this section, the level of nonresponse bias for the variable "one-person household" is high. With increasing response rates, the nonresponse bias decreased, but at the

60% response rate level, the underrepresentation of people from one-person households began to turn into an overrepresentation. A regression analysis showed that a significant positive correlation between the response rate and relative nonresponse bias (coef = .00; t = 2.34; p = .023; R^2 = .10). As shown in Figure 40 and Appendix Figure B19, as response rates increased, the underrepresentation of the variable “one-person household” decreased.

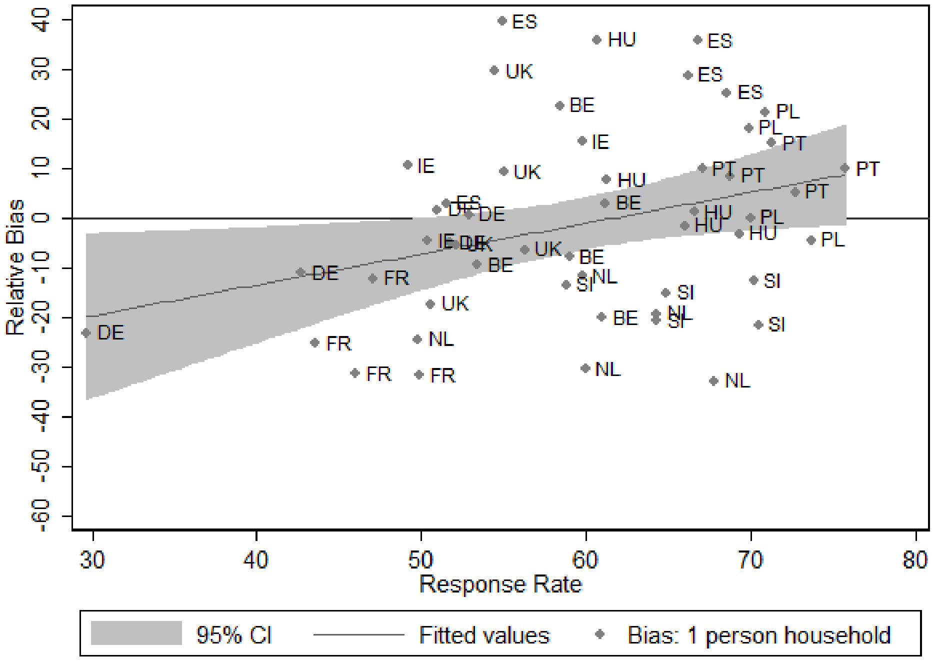

Figure 40. Relative bias for the variable “one-person household” and the response rate.

An analysis of the absolute value of the relative bias supported these findings. A further regression analysis of the absolute nonresponse bias provided the same results (coef = -.01; t = -2.49; p = .016; R^2 =.14). As shown in Figure 41 and Appendix Figure B20, the distribution of the bias is rather wide. The Pearson Correlation Coefficient described a weak non-significant correlation between the response rate and nonresponse bias (r = -.10; p = .476). This finding supports Hypothesis 6 for the variable “one-person household,” an increase in the response rate is associated with a decrease in the relative nonresponse bias.

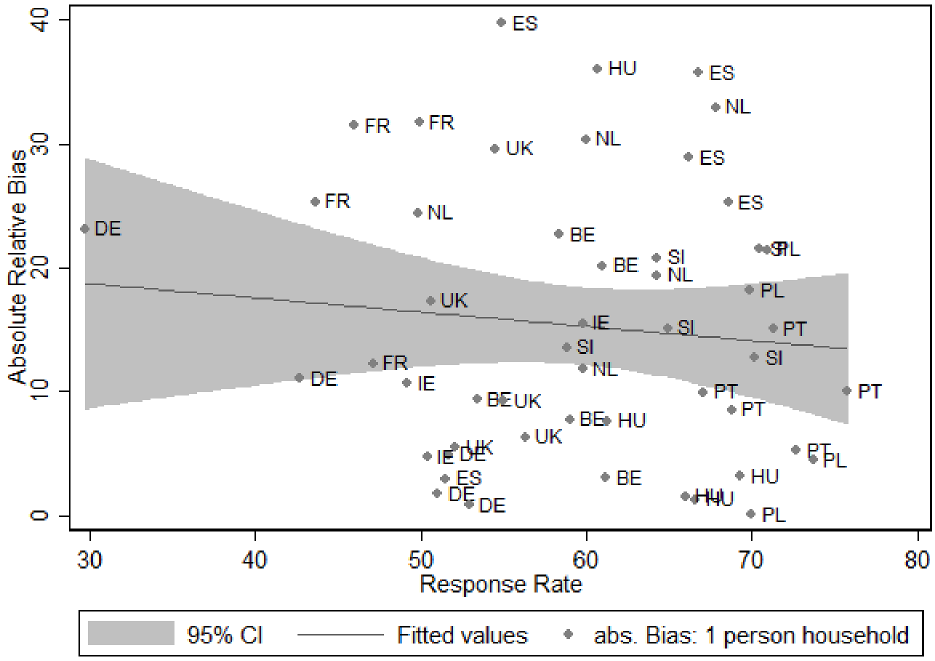

Figure 41. Absolute relative bias for the variable "one-person household" and the response rate.

13.10 Five-or-More Person Household

Hypothesis 6 also was tested for the variable describing a large household. The correlation of the relative bias for this variable and the response rate is displayed in Figure 42 and Appendix Figure B21, which shows that in general, people living in a five-or-more person household are overrepresented in the ESS compared to the LFS. A regression analysis showed that even though a pattern of decreasing nonresponse bias occurred when response rates increased, no significant correlation exists between the response rate and nonresponse bias for the variable "five-or-more person household" (coef = -.00; t = -1.77; p = .084; R^2 = .11). Compared to other variables analyzed, the nonresponse bias for the variable "five-or-more person household" was at a high level.

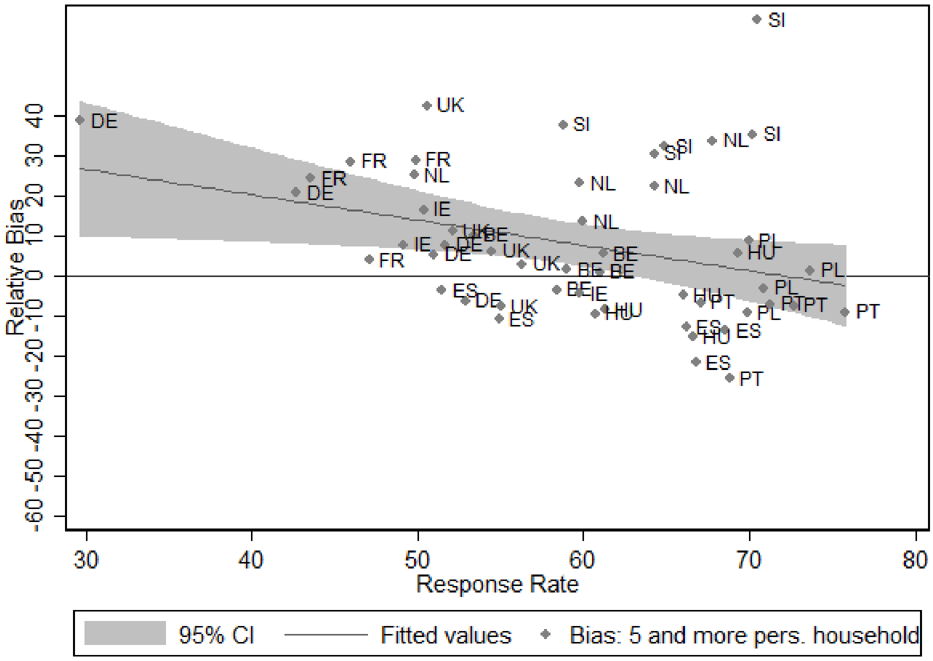

Figure 42. Relative bias for the variable "five-or-more person household" and the response rate.

The binary correlation of the absolute value (Figure 43 and Appendix Figure B22) of the nonresponse bias and the response rate was measured with the Pearson Correlation Coefficient. It showed a weak and non-significant correlation exists between the nonresponse bias and response rate (r = -.11; p = .434). The regression analysis showed a negative but non-significant correlation between the two variables (coef = -.01; t = -1.55; p = .128; R^2 = .06). Thus, Hypothesis 6 cannot be supported for this variable.

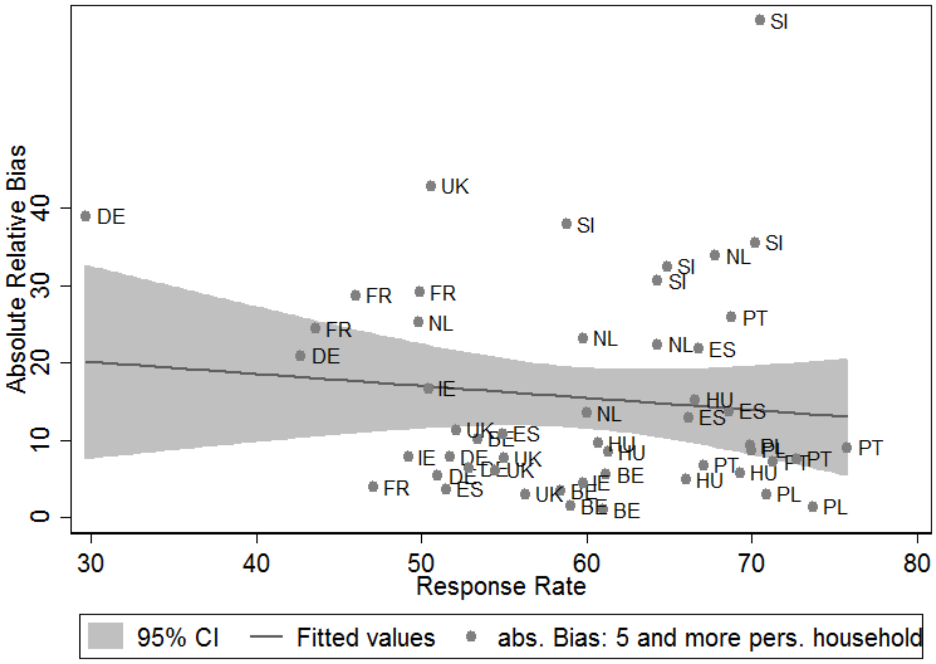

Figure 43. Absolute relative bias for the variable “five-or-more person household” and the response rate.

14 Discussion

The limitation of this kind of study is that only the available variables in the external statistics can be compared. Therefore, the analysis can be only conducted for socio-demographic variables, where comparable external statistics are available and where these external statistics can be harmonized. For other variables, such as substantive variables this is not possible since reference statistics are very rare or not available. To analyses nonresponse bias for substantive variables other types of nonresponse bias analysis is appropriate.[129] Other means of analyzing the nonresponse bias are the “R-indicator” (Schouten et al., 2012; Schouten, Schouten, Shlomo, Da Silva, & Skinner, 2016; Shlomo, & Skinner, 2011), the

[129] One method is the comparison of early and late respondents within a survey. Another method is the comparison of respondents who answered at the first contact attempt in comparison to respondents who answered later, where the respondents who answered late are considered to be more similar to nonrespondents (see Chapter III, Section 7).

"Fraction of Missing Information" (Little, 1986; Wagner, 2012), and the use of paradata (Krueger & West, 2014). These types of analysis could not be performed for the present study because they require auxiliary information beyond the gross sample. For example, the R-indicator needs information from the register, but this information is not available for all countries.[130] An application of the R-indicator would limit the number of analyzed variables and countries to only a very few for which this information is available. Also the use of, but in these cases selection of, available paradata would be limited.

Another point of critique for the method nonresponse bias calculation using the method of comparison with external statistics is that the external statistic is assumed to be error free, which is not always the case. The selected type of analysis using external data like the Labour Force Survey data has benefits, for example, because that data is available for all countries, has comparable fielding periods, and can be harmonized between the ESS and the LFS. Thus, using this external data for comparison enables the inclusion of more countries in the analysis, and helps to create a comprehensive picture.

15 Results for Analysis Part B

Analysis Part B of the present study analyzed the effect of response rate rates on the nonresponse bias by testing Hypothesis 6 (*The higher the response rate is, the lower the risk of nonresponse bias.*) on a variety of variables. The nonresponse bias was calculated by comparing the data from the ESS to the "gold standard" of the LFS. The deviations of the two datasets were defined as the nonresponse bias. This nonresponse bias analysis was carried out

[130] This information is from a personal communication with B. Schouten (personal communication, April 18, 2016) in which he describes how he and his colleagues (Annamaria Bianchi, Natalie Shlomo, Damiao Da Silva, and Chris Skinner) worked on an estimation of response propensities and Indicators of representative responses using population-level information. This paper was submitted to a journal, but is of an experimental nature. A calculation of representativeness and nonresponse bias are not possible at this stage for all countries in their study.

for the socio-demographic variables gender, education, occupation, nationality, household size, and family status. The testing of the hypothesis was done in two steps: Section 12 conducted a pooled cross-sectional analysis, and Section 13 provided an analysis at the variable level.

The pooled cross-sectional analysis for all variables (Figure 22) shows that in general, a lower response rate level was associated with a higher nonresponse bias level. The regression model was a good fit for the data. The results of the regression analysis supported Hypothesis 6. The Pearson Correlation Coefficient of the variable response rate and the (absolute value of the) nonresponse bias showed a small and significant strength of association (r = -.16; p = .000). Thus, Hypothesis 6—stating a correlation between the response rate and nonresponse bias—was supported. The higher the response rate is, the lower the (absolute value of the relative) nonresponse bias.

In a second step, a further analysis was conducted at the variable level (Section 13). Analyses at the variable level showed that the relationship of the response rate and nonresponse bias did not apply to all variables. Some of the socio-demographic variables supported Hypothesis 6, but for other variables, no correlation, or a relation opposite the expectation, was found (Table 9). For the majority of the variables, the relation of the response rates and nonresponse bias is as stated by the Hypothesis 6. For these variables—"old persons," "married persons," persons with "low education," persons with "high education," "nationals of the country," and persons live in "one-person households"—a higher response rate is indeed correlated with a lower nonresponse bias. However, for the variables "gender (male)" and persons living in a "five-or-more person household," a significant, reversed correlation was detected. For these two variables, a higher response rates was correlated with a higher nonresponse bias, which contradicts Hypothesis 6. For the variables "young persons (age 15–24)" and the "working population", no significant correlation was detected between the response rate and nonresponse bias. Thus, the picture is mixed. For the majority of the analyzed variables (6 out of 10), the hypothesis was supported; for two variables, the effect was contrary to what was the expected; and for another two variables, no significant correlation was detected.

Table 9

Variable-specific overview of the correlation of the response rate and nonresponse bias

Analyzed nonresponse bias for the variable	**Over- or under-representation**	**Does a relationship exist between the nonresponse bias and response rate? If so, in what direction?**	**Hypothesis 6 supported by results?**
	Relative nonresponse bias	**Absolute (value of the relative) nonresponse bias**	
gender (male)	With a low response rate, a low nonresponse bias is detected. With an increasing response rate, the nonresponse bias increases and is more under-represented.	The higher the response rate is, the larger the bias.	No support for hypothesis. Relation is contrary to assumed direction.
young persons	Slight under-representation	No change in bias with increasing response rates.	No support for hypothesis. No correlation.
old persons	High under-representation at a low response rate level.	The higher the response rate is, the smaller the bias.	Support of hypothesis.
working population	Underrepresentation of the working population.	No significant correlation.	No support for hypothesis. No correlation.
family status: married persons	Overrepresentation of married persons at a low response rate level. Nonresponse bias decreases with increasing response rate.	The higher the response rate is, the smaller the bias.	Support of hypothesis.

low education	Persons with low education are under-represented at low response rate levels. Decreases of nonresponse bias with higher response rate.	The higher the response rate is, the smaller the bias.	Support of hypothesis.
high education	Persons with high education are over-represented at low response rate levels. Decrease of nonresponse bias at higher response rate levels.	The higher the response rate is, the smaller the bias.	Support of hypothesis.
nationality of country	Persons with the nationality of the country analyzed are overrepresented. The over-representation decreases with higher response rates.	The higher the response rate is, the smaller the bias.	Support of hypothesis.
one-person households	Persons living in one-person households are underrepresented in countries with low response rates. With an increasing response rate, the bias decreases.	The higher the response rate is, the smaller the bias.	Support of hypothesis.
Five-or-more person households	Persons living in 5-or-more person households are over-represented in countries with low response rates. With increasing response rates, the bias decreases.	The higher the response rate is, the larger the bias.	No support for hypothesis. Relation is contrary to assumed direction.

The level of nonresponse bias differs between the variables. For the variables “gender (male)” and “nationality,” the level of nonresponse bias is rather low. For the variables “young respondents," “working population” and “married respondents,” the nonresponse level is medium. For the variables “older respondents,” “education (low and high),” and “household size” (one-person household and 5-or-more person household), the nonresponse bias overall is very large. Thus, we can conclude that the level as well as the pattern of nonresponse bias is variable-specific.

This analysis extends the previous findings of Groves and Peytcheva (2008) by keeping constant the survey, the survey-set up, the topic, and the bias variables. The present study has shown that even when many factors are constant, a relation can be detected between the response rate and nonresponse bias. Therefore, we can say that response rates affect the nonresponse bias. In general, a higher response rate leads to a lower nonresponse bias, but not for all countries and all variables.

Analysis Part C: Fieldwork Efforts and Response Rates

Response rate is one of the most used indicators of high quality (Biemer & Lyberg, 2003; Singer, 2006, pp. 637-645). Due to the assumed relationship of well-implemented fieldwork and high data quality, specifications and guidelines for surveys are set to advise fieldwork organizations on how to implement a survey (e.g., European Social Survey, 2013a; Organisation for Economic Co-operation and Development, 2010; Stoop, Koch, Halbherr, Fitzgerald, & Widdop, 2014). An important aspect of these survey specifications—in addition to minimizing the Total Survey Error (TSE) and increasing cross-cultural comparability—is setting standards for a minimum response rate or strategies to enhance the response rate. The strategies to enhance the response rate include fieldwork efforts, such as incentives, interviewer briefings, and advance letters. The general hypothesis prevailing in the survey literature is that the higher the fieldwork efforts of a survey, the higher the response rate (Couper & de Leeuw, 2003; de Heer, 1999; Groves & Couper, 1998; Pforr et al., 2015; Sakshaug et al., 2010).

However, the hypothesis about the relationship between fieldwork effort and response rate is challenging to analyze because often information about fieldwork is not publically available or is not comparable across time and across countries. Thus, it is difficult to systematically test this hypothesis, especially in a cross-national and longitudinal context. As described previously, the availably and the calculation of a harmonized response rate (see Analysis Part A) is very rare in surveys. The European Social Survey (ESS) is the only survey that provides this information on response rates in a public and cross-country comparable format. Also, the availability, transparency, and comparability of fieldwork efforts are something unique to the ESS.[131] Since the ESS provides an opportunity for researchers to analyze the data on fieldwork efforts and response rates, a fundamental research question in the field of survey methodology can be empirically answered.

In this chapter (Analysis Part C), the relationship between fieldwork efforts (independent variable) and response rates (dependent variable) is analyzed. This empirical testing of this hypothesis can confirm or deny the implicit assumption of the relationship between fieldwork efforts and response rates. Systematic analysis of the effects of fieldwork efforts on response rates helps to answer the basic question in survey methodology about whether increased fieldwork efforts are associated with higher response rates.

This analysis adds to the examination of the general indicators for fieldwork: response rates, fieldwork efforts, and nonresponse bias. In the previous part of the present study (Analysis Part B), the effect of the response rate on data quality, measured by nonresponse bias, was analyzed. The data showed no clear relationship between the response rate and nonresponse bias—the nonresponse is variable and country specific. Depending on the variable and the country, the effect of the response rate on nonresponse bias differs.

Considering these findings, one might ask whether certain measures carried out in the fieldwork affect the response rate. This research question is examined in a cross-national and longitudinal context in Analysis Part C.

The general hypothesis is that the more fieldwork effort is put into a survey, the higher the response rate. The hypothesis for individual fieldwork efforts is the same as the general hypothesis: the more advance letters are sent to a respondent, the more incentives are

[131] The ESS provides information on fieldwork efforts for all participating countries on its website www.europeansocialsurvey.org in a publically available report (European Social Survey, 2014a, 2014b, 2014c, 2014d, 2016a).

used, the more interviewers attend briefings, the higher the percentage of experienced interviewers, and the better the payment of interviewers, the higher the response rate. These hypotheses are tested in the following paragraphs. The analyzed hypothesis on the relation of fieldwork efforts and response rates are listed in Table 10.

Table 10

Hypothesis on fieldwork efforts and response rates

General hypothesis on fieldwork effort and response rates

Hypothesis 7: The more fieldwork efforts—such as the use of incentives, interviewer briefings, and so on—are implemented in a survey, the higher the response rate.

Fieldwork efforts and response rates (variable specific effects)

Hypothesis 8: Countries that use additional fieldwork efforts (e.g., send an advance letter, or use a brochure or incentives) achieve higher response rates than countries that do not use these fieldwork efforts.

Fieldwork efforts and nonresponse rates

Hypothesis 9a: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentives) achieve lower non-contact rates than countries that do not use these fieldwork efforts.

Hypothesis 9b: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentives) achieve lower refusal rates than countries that do not use these fieldwork efforts.

Fieldwork effort (as a compound “fieldwork effort index”) and response rates

Hypothesis 10: The higher the fieldwork effort (the score on the compound fieldwork effort index) is, the higher the response rate.

Longitudinal analysis: Change between the rounds of the fieldwork effort index and response rate

Hypothesis 11: A correlation exists between the change in fieldwork efforts (in the preceding compared to the subsequent round) and the change in response rates (in the preceding compared to the subsequent round).

Different trends

Hypothesis 11a: When countries put more effort into fieldwork than in the previous round, the response rate increases compared to the previous round.

Hypothesis 11b: When countries put less effort into fieldwork than in the previous round, the response rate decreases compared to the previous round.

Hypothesis 11c: When countries put the same effort into fieldwork as in the previous round, the response rate remains the same as in the previous round.

In the first and second section (Section 16 and 17), cross-sectional analyses were conducted to test whether a general relationship exists between fieldwork efforts and response rates or nonresponse rates. Section 16 provides a general analysis of whether a relation between fieldwork efforts and response rates or nonresponse rates can be seen (Hypothesis 7, 8, 9a, and 9b). Section 17 constructs a compound index of fieldwork efforts (fieldwork effort index [FEI]) to determine whether countries with a high fieldwork effort have high response rates and low nonresponse bias (Hypothesis 10). Section 18 provides a longitudinal analysis by looking at change between rounds with a focus on change from one round to the subsequent round. By looking at the change between rounds and keeping country specific issues constant between rounds, the effects of the fieldwork efforts can be isolated.[132] Thus, questions can be answered as to whether additional fieldwork effort increases the response rate at the country level (Hypotheses 11, 11a, 11b, 11c). Section 19 provides a more qualitative analysis of the fieldwork efforts at the country level of selected countries. This comprehensive analysis of the various aspects helps to determine further whether fieldwork efforts affect the response rate.

16 Fieldwork Efforts

This section analyzes the hypotheses about the effects of fieldwork efforts (such as the use of incentives, interviewer briefings, advance letters, and so on) on the response rate. This enables a testing of a general assumption (see Hypothesis 7) of survey research—whether high fieldwork efforts in a survey lead to higher response rates. The data is pooled to analyze the general relation at a cross-sectional level. The effects of separate fieldwork efforts (such as the use of incentives, interviewer briefings, and so on) on the response rate

[132] These hypotheses assume that fieldwork efforts are the only influence on the response rate. This *ceteris paribus* assumption is plausible, since the same countries are analyzed at different points in time.

are tested, for example, whether the use of incentives leads to higher response rates (Section 16.1). The research question is answered whether the use of a certain fieldwork effort correlate with the higher response rates. In addition, the effects of fieldwork efforts on the nonresponse rate (refusal, non-contact, and the rate of others) are tested (Section 16.2).

16.1 Effect of Separate Fieldwork Efforts on Response Rates

The first step analyzes the relationship between the separate aspects of fieldwork efforts on response rates at a cross-sectional level. General Hypothesis 7 (*The more fieldwork efforts—such as the use of incentives, interviewer briefings, and so on—are implemented in a survey, the higher the response rate.*) was tested to determine whether differences exist in the response and nonresponse rates between the countries that use a particular survey procedure (such as providing incentives, using brochures, or using long interviewer briefings) compared to countries that do not implement this procedure.

By applying general Hypothesis 7 to separate fieldwork efforts, the following hypothesis can be formulated:

> *Hypothesis 8: Countries that use additional fieldwork efforts (e.g., send an advance letter, or use a brochure or incentives) achieve higher response rates than countries that do not use these fieldwork efforts.*

In the survey literature, the use of incentives, brochures, and interviewer briefings in general are considered as additional efforts of fieldwork. In particular, these measures are used with the goal to enhance the response rate (De Leeuw, Callegaro, Hox, Korendijk, & Lensvelt-Mulders, 2007; Koch, Fitzgerald, Stoop, Widdop, & Halbherr, 2012; Stoop et al., 2014). The factors constituting a fieldwork effort include: the contact of respondents (the use of advance letters, brochures, and incentives), the length of the fieldwork and interviewer-related factors (length of the interviewer briefing, personal briefing of interviewers, payment of interviewers, interviewer training, and interviewer experience).[133]

The selection of fieldwork effort factors was based on the data of the ESS, which provides a differentiation between high and low fieldwork efforts (Stoop et al., 2010, p. 104ff.). Additionally, these efforts have been described in the survey literature as response enhancing methods (see Dillman, 1978; Stoop et al., 2010). In Table 11, the high fieldwork efforts are mentioned first and low efforts second. For example, one indicator of fieldwork effort is the number of experienced interviewers working on the survey. A high fieldwork effort within a country is defined as that 90% or more of all survey interviewers have previous experience with interviewing (measured by length of time they work for the survey company). A low fieldwork effort is defined as that less than 90% of the survey interviewers have work experience with interviewing.

The following fieldwork efforts were selected: the use of incentives (yes/no), use of brochures (yes/no), use of advance letters (yes/no), length of the interviewer briefing (more than half a day/less than half a day), personal briefing of interviewers (90% or more briefed/less than 90% briefed), payment of interviewers (per hour or per interview + bonus per interview), interviewer experience (90% or more experienced/ less than 90% experienced), interviewer trained in refusal conversion (yes/no), and length of fieldwork (90 days and more/89 days and less) (see Table 11).[134]

---

[133] The differentiation between high efforts and low efforts often is dichotomous (e.g., the use of advance letter yes/no). For the definition of high and low efforts of above described continuous variables (e.g., the percentage of briefed interviewers), see Stoop, Billiet, Koch, & Fitzgerald (2010).
[134] The cut-off points were selected based on the analyses of Stoop et al. (2010) study, which based the cut points on the distribution of this variable in the ESS round 1 to 3.

For a selection of fieldwork efforts, the mean response rate for countries using this effort (e.g., use of an incentive) in contrast to countries not using it was analyzed. This case-control design (Lacy, 1997) enabled an analysis of whether, for example, in countries that used an incentive, the average response rate is higher than in the countries that did not use an incentive. To determine whether a positive relationship exists between the use of a certain fieldwork effort (such as an incentive) and the response rate, the average response rates for the cases using this response-enhancing measure and for those not using this measure were calculated.

Table 11

Average response and nonresponse rate by use of fieldwork efforts

	n	Response rate (%)	Noncontact rate (%)	Refusal rate (%)	Not able/ other rate (%)
Incentive					
Yes	48	57.9	4.4	27.9	11.0
No	32	62.2 p = .050*[a]	4.5	24.8	8.5
Use of brochures					
Yes	63	59.3	4.2	26.8	10.1
No	17	60.7	5.5	25.7	9.5
Advance letters					
Yes	74	60.2	4.0	26.5	10.1
No	6	52.8	9.9 p = .008**[b]	28.1	9.2
Length of interviewer briefing					
More than half a day	51	58.7	4.8	28.1	8.9
Less than half a day	29	61.2	4.0	23.8 p = .008**[c]	11.9
Personal briefing of interviewer					
90% or more of briefed	73	59.7	4.5	26.9	9.2
Less than 90% briefed	6	58.6	4.0	20.2	20.5
Payment of interviewer					
per hour or per interview + bonus	56	59.5	3.7	26.8	10.6
per interview	24	60.0	6.3 p = .013**[d]	26.1	8.7
Interviewer experience					
90% or more experienced	53	59.7	4.5	27.0	9.9
Less than 90% experienced	23	61.2	3.9	25.1	9.7

Interviewer trained in refusal conversion					
Yes	71	59.4	4.3	26.7	9.5
No	8	57.5	5.7	25.6	14.5
Length of fieldwork					
90 days and more	70	58.5	4.6	27.6	10.1
89 days and less	10	67.5	3.7	19.8	9.1
		p = .000**[e]		p = .002**[f]	

Note. RR = Response Rate.
[a]t = - 1.99; p = .05. Levene test (p = .178), thus equal variance is assumed. A t-test of independent samples was conducted.
[b]t = - 2.75; p = .008. Levene test (p = .025), thus no equal variance is assumed. A t-test of independent samples was conducted. Please note that the number of cases with no advance letter is low (n = 6) in comparison to the cases that use an advance letter (n = 74).
[c]t = - 4.51; p = .008. Levene test (p = .716), thus equal variance is assumed. A t-test of independent samples was conducted.
[d]t= - 3.33; p = .013. Levene test (p = .003), thus no equal variance is assumed. A t-test of independent samples was conducted.
[e]t = - 5.03; p = .000. Levene test (p = .016), thus no equal variance is assumed. A t-test of independent samples was conducted.
[f]t = 3.21; p = .002. Levene test (p = .215), thus equal variance is assumed. A t-test of independent samples was conducted.
*** p < .01. ** p < .05. * p < .10.

Contrary to the expectations of Hypothesis 7 (*The more fieldwork efforts—such as the use of incentives, interviewer briefings, and so on—are implemented in a survey, the higher the response rate.*) and Hypothesis 8 (*Countries that use additional fieldwork efforts [e.g., send an advance letter or use a brochure or use incentives] achieve higher response rates than countries that do not use this fieldwork effort.*), in most cases that use higher fieldwork efforts, the response rate is lower than in cases that do not use higher fieldwork efforts. The results described in Table 11 seem to suggest a pattern that the use of incentives, brochures, and longer interviewer briefings are correlated with a lower response rate and a higher refusal rate. For example, in countries that did not use an incentive, the aggregate response rate was higher (62.2%) than in the countries that used an incentive (57.9%). This correlation was found for all the mentioned survey procedures. The only exception was the training of interviewers to convert refusals, which produced a level of response rate higher in the countries that used this strategy than the countries that did not. Hypothesis 8 was supported only by the variable “use of interviewer training”.[135] Additionally, an analysis was conducted to determine whether in cases where the fieldwork effort was implemented, the mean response rate was higher than in the cases where it was not used.

A *t*-test of independent samples was conducted.[136] By doing this test, a significant difference was found in the response rates between cases with and without the fieldwork effort “incentive” (t= -1.99; p = .050) and for the variable “length of fieldwork” (t = -5.03; p = .000). The difference between the cases using certain fieldwork efforts and those that do not use them is in the same direction as described previously, meaning that higher fieldwork effort is correlated with lower response rate. Contrary to the assumptions made with respect to those cases where these fieldwork efforts are used, the mean response rate is lower.

A more detailed look into a particular fieldwork effort was done for the variable “length of fieldwork.” The relationship between the length of fieldwork and response rates are visualized in Figure 44. Each data point shows the results for each county in each round. For example, Hungary in round 1 (HU1) had a response rate of 69.3% and 29 days of

[135] It might be that the level of fieldwork efforts is based on the response rate in previous rounds or in comparable surveys. The decision to use certain fieldwork procedures that cause additional fieldwork effort, such as an incentive, is not made at random within countries. Countries that expect difficulties in gaining cooperation will use an incentive to improve the response rate. So the decision to use an incentive is not independent of the experience to obtain high response rates (Brehm, 1993, pp. 128-130). Countries differ in their level of response rates in previous rounds of the ESS. Additionally, countries with an expected low level of response rate might choose to increase efforts in fieldwork (e.g., prolong the fieldwork period to increase the response rate). So, the reversed causality might be that countries that expect a low response rate increase their fieldwork efforts. I make the assumption that countries with a lower response rate will put more effort into their fieldwork to increase their response rate in the next round. For countries that achieve an above the mean response rate, the need to put more effort into their fieldwork is low, since the response rate is “sufficiently” high. So, the causal mechanism would be that low response rates in previous rounds lead researchers to do more to increase the response rate in the following round. This would be the reversed causality of the described hypothesis that examines whether that high level of fieldwork efforts leads to a high response rate. So the line of argument is endogenous.

[136] The analysis of Stoop et al. (2010) considers the selection of the countries not as a random sample, whereas the analysis including a *t*-test assumes that the selection of countries is not based on a selection process.

fieldwork. As can be seen in Figure 44, the length of the fieldwork period shows the same pattern of negative correlation between the fieldwork effort and response rate as described previously. Normally, one would expect that the longer the fieldwork period is, the higher the chance that all target respondents could be reached, since the longer time span would allow for repeated contact attempts with people who are difficult to contact (Groves & Couper, 1998, pp. 272-274). The longer a survey is in the field, the more time is available to re-contact reluctant respondents or refusals. If surveys have problems in achieving high response rates or if they have other problems in the field, often the fieldwork period is extended. In many surveys, including the ESS, the end of the fieldwork period is planned, but in case of difficulties, the length of the fieldwork period may be extended. Thus, if countries have difficulties in achieving enough interviews, they may decide to extend the fieldwork period.[137] As the comparison of the mean response rate suggests, in the cases in which the fieldwork effort was implemented, the response rate in general was lower. This result is contrary to Hypothesis 8, which suggests that countries that use additional fieldwork efforts achieve higher response rates than countries that do not use these additional efforts.

[137] This result does not mean that a positive correlation exists at the "within country level," which suggests that at the country level, a positive effect on the response rate may occur if a country decides to extend the fieldwork effort.

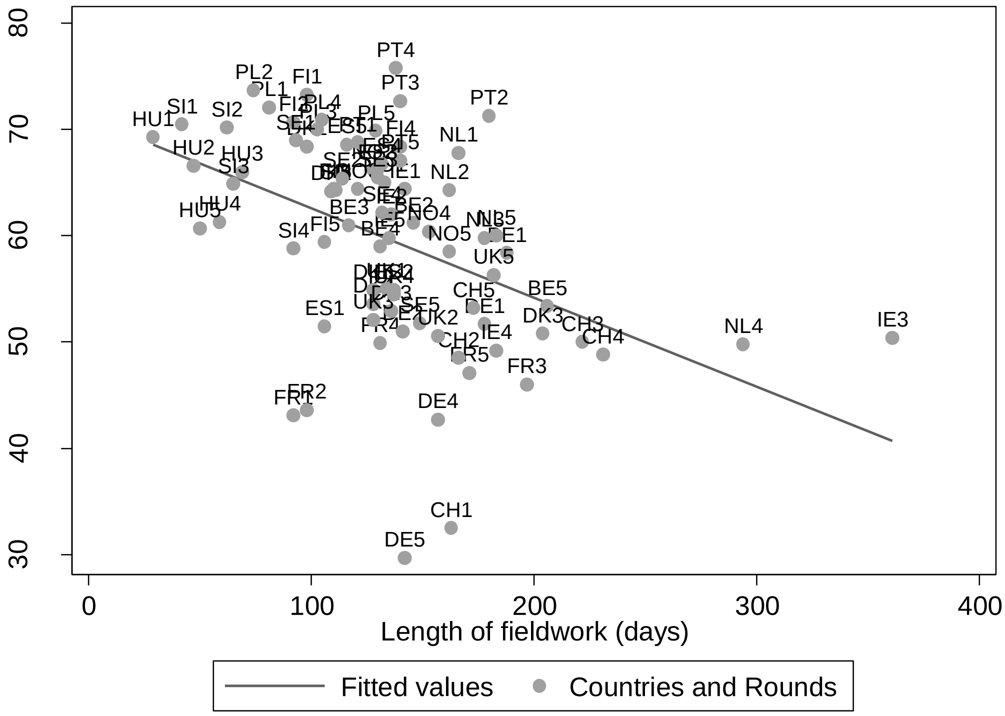

Figure 44. Length of fieldwork period (days) and response rate.

The traditional expectation (Hypothesis 7) would be that the more fieldwork efforts (such as the use of incentives, interviewer briefings, and so on) are implemented in a survey, the higher the response rate. However, a reversed causality seems to be the case. This contra-intuitive finding for the variable “length of fieldwork period” does not support the general expectation for fieldwork (Hypothesis 7), which assumes that countries that put lots of effort into their survey fieldwork obtain higher response rates than countries make low fieldwork efforts.

Two conclusions can be drawn from this contra-intuitive finding. First, a plausible interpretation is that due to the non-experimental setup of the data collection (the fieldwork efforts were not randomly distributed to the countries but were decided by fieldwork managers[138]), country specific aspects at the micro and macro level—such as the survey climate[139] within a country—might play an important role. The findings reveal a paradox of the correlation of fieldwork efforts and response rates. Countries that expect a high response rate do not need a lot of additional effort to achieve a high response rate. Thus, an extra effort and additional budget are not needed to extend the current fieldwork effort. Therefore, low effort might be related to a high response rate.

In the next section, this analysis is extended from examining separate factors to an index of fieldwork efforts. The following sections examine the changes in the response rate from a previous round to a next round. The longitudinal perspective is provided in Section 18. The paradox (described above) that low fieldwork efforts are correlated with high response rates can be avoided when the change of the response rate and fieldwork efforts within a country is analyzed.

16.2 Effects of Fieldwork Effort on Nonresponse Rates

In this section, the other side of the response rate coin, namely the nonresponse rate, was analyzed. Following the above described argument supporting the positive effect of fieldwork efforts on response rates, one would expect that higher fieldwork efforts lead to higher response rates and thus to lower nonresponse rates. Thus, with respect to the non-contact rate and the refusal rate, one also would expect lower rates in countries where additional fieldwork efforts were used.[140]

---

[138] In the ESS, the specifications for participating countries (European Social Survey, 2011, 2013) and the guidelines for enhancing response rates (Koch et al., 2012; Stoop et al., 2014) advise countries to implement certain fieldwork efforts. The actual implementation of these fieldwork efforts is decided by the National Coordinators of the countries.

[139] The phrase *survey climate* describes the general willingness of respondents to take part in surveys (Lyberg & Dean, 1992).

[140] A hypothesis for the rate of “not able/others” is absent in the literature. As described earlier (Chapter II Literature Review, Section 8.3), this rate often is used as a residual category.

The reasons for the nonresponse category refusal and non-contact are different. In the first case, persons choose not to respond to the survey request. In the second case the persons were not contacted (see Section 9 and 16.2). Since those are different reasons for nonresponse, the two reasons are further analyzed separately.

The hypotheses on the relation of the fieldwork effort to the refusal and non-contact rates can be formulated as follows:

> *Hypothesis 9a: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentive) achieve lower* <u>*non-contact rates*</u> *than countries that do not use these efforts.*

> *Hypothesis 9b: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentive) achieve lower* <u>*refusal rates*</u> *than countries that do not use these efforts.*

First, the effect of the fieldwork effort on non-contact rates was analyzed. It is assumed that more effort to contact respondents would lead to lower nonresponse rate (see Section 9 and 16.2).To test Hypothesis 9a, the mean non-contact rates for cases in which an additional fieldwork effort was used was compared to the cases in which an additional effort was not used (Table 2). For most of the variables, no significant difference was found between the mean non-contact rates. A t-test showed that the cases that used an advance letter[141] had significantly lower non-contact rates than the cases that did not use a letter (4.0% vs. 9.9%; t = -2.75; p = .008). This finding supports Hypothesis 9a because in cases with higher fieldwork efforts, lower non-contact rates are achieved. In contrast, however, in cases in which interviewers are paid per interview compared to per hour or per interview plus

[141] The number of cases in which an advance letter was used is low (n = 6).

bonus, the non-contact rates are significantly lower (6.3% vs. 3.7%; t = -3.33; p = .013). This finding does not support Hypothesis 9a; rather, it supports the contrary assumption that less effort leads to lower non-contact rates. Thus, Hypothesis 9a cannot be supported for all the variables, since different variables have different effects.

Second, the effect of fieldwork efforts on the refusal rate was tested. It can be assumed, that for example the better the interviewer training the more likely can the interviewer convince the respondent to participate. Also it may be assumed that longer fieldwork periods allow the respondents more time to answer the survey request. Thus, it can be assumed that a correlation of fieldwork effort and refusal rate exists.

This allows formulating in Hypothesis 9b. A negative correlation between fieldwork effort and refusal rate is assumed, so the higher the fieldwork effort is, the lower the refusal rate. For most of the variables, no significant difference was found between the higher fieldwork effort and the refusal rate. Thus, Hypothesis 9b cannot be supported. However, for two variables, a significant difference was found. For the variable “length of interviewer briefing,” a significant difference was found (t = -4.51; p = .008) between the mean refusal rate for countries using higher fieldwork efforts compared to those countries that did not use these higher efforts. In those cases in which the length of the interviewer briefing is longer (more than half a day vs. less than half a day), the refusal rate (28.1%) is significantly higher than in countries where the interviewer briefing is shorter (23.8%) (see Table 2). Also, with respect to the “length of fieldwork,” a significant difference (t = -3.21; p = .002) between *lower fieldwork effort*, defined as a shorter fieldwork period (89 days and less), and higher effort (90 days and more) was found. Lower fieldwork effort cases had a mean refusal rate of 19.8%, and higher fieldwork effort cases had a mean refusal rate of 27.6 Only with respect to the variables “length of interviewer briefing” and “length of fieldwork period,” the mean refusal rate is higher in the cases in which additional fieldwork effort was implemented. Also, for the other variables that were analyzed, this same pattern was present,[142] which also does not support Hypothesis 9b.

17 Fieldwork Effort Index

Differences in response rates can be caused by a variety of factors in fieldwork. Thus, to examine the relationship between fieldwork effort and response rate, an investigation of single indicators of fieldwork effort is not sufficient. Stoop et al. (2010, p. 103) have suggested constructing a compound index drawing for several dimensions of fieldwork procedures and fieldwork efforts. The aim of this fieldwork effort index is to consider the impact of several factors simultaneously and examine their combined effect on the response rate. The use of a fieldwork effort index (FEI) does not rely on individual variables; rather, it combines different aspects of fieldwork effort.

The use of an index by the present study has three positive aspects. First, a change in a response rate can be caused by different factors, and usually these factors operate in batches. As described in the Stoop et al. (2010, p. 106) study, the individual factors that contribute to a fieldwork effort cannot be analyzed separately because all other factors would need to be constant. Due to the observational nature of the study, this is not possible.[143] Therefore, it seems appropriate to consider several fieldwork factors simultaneously and evaluate their combined effect on the response rate. Second, the use of an index helps to make the

[142] For most of the analyzed variables, the analysis of the refusal rate found the same contra-intuitive pattern—a higher fieldwork effort correlated with a higher refusal rate (see Table 11). The differences between the cases with and without fieldwork effort are, in the following cases, not significant. In cases in which an incentive was used, the refusal rate was higher (27.9%) than in the countries that did not use an incentive (24.8%). The same was true for the use of brochures (when brochures were not used, the response rate was higher), the length of the interviewer briefings (a shorter briefing was associated with a higher response rate), interviewer payment (a lack of an additional bonus was associated with a higher response rate), interviewer experience (a less experienced interviewer was associated with a higher response rate), and the length of the fieldwork period (longer periods of fieldwork were associated with lower response rates).

[143] For an overview of the fieldwork efforts used in each country and round, see Analysis Part C, Section 17.

interpretation easier and more accessible. Third, in most countries, the factors making up the fieldwork effort changes between the rounds, but the number of fieldwork efforts stays the same.

17.1 Construction Fieldwork Effort Index (FEI)

Fieldwork efforts are implemented as a bundle. Stoop et al. (2010, p. 75 ff.) constructed an index of eight fieldwork factors, which includes information on interviewer features and incentives that are expected to contribute to achieving a high response rate (Dillman, 1978; Stoop et al., 2014). First, for the present study, each of these different factors was dichotomized (high efforts = 1, low efforts = 0) and then an additive fieldwork effort index (FEI) was built. The differentiation between high and low fieldwork efforts was based on Stoop et al. (2010, p. 104ff.). Often is dichotomous (e.g., use of advance letter: yes/no). For continuous variables (e.g., the percentage of briefed interviewers), as in Stoop et al. (2010), the cut point was selected based on the distribution of this variable in the ESS round 1 to 3.

In the following list, the fieldwork efforts are described. High efforts are mentioned first and low efforts second. For example, one indicator of fieldwork effort is the number of experienced interviewers working on a survey. A high fieldwork effort within a country would be when 90% or more of all interviewers have previous experience with interviewing (measured by the length of time they work for the survey company). A low fieldwork effort would when less than 90% have work experience in interviewer briefing. In the overall dissertation, for reasons of simplicity the fieldwork effort index is referred to as fieldwork index or FEI.

Interviewer:

- Experience of interviewer: 90% or more of all interviewers experienced vs. less than 90% experienced
- Payment of interviewer: paid per hour/ per interview and bonus vs. paid per interview only
- Personal briefing of interviewers: 90% or more of all interviewers personally briefed vs. less than 90% personally briefed
- Length of personal briefing sessions: more than half a day vs. half a day or less
- Interviewer trained in refusal conversion: yes vs. no

Contact to respondent:

- Use of advance letter: yes vs. no
- Use of brochure: yes vs. no
- Use of respondent incentive: yes vs. no

Three reasons exist to explain why the previous fieldwork efforts are included in the FEI. First, the index is based on a previous analysis by Stoop et al. (2010). Using the same index enables a comparison of the results over time. However, more important is the substantial reason for including these fieldwork effort indicators: the ESS and its Core Scientific Team have suggested implementing these fieldwork efforts to enhance response rates, and these indicators also are central parts of the specifications for participating countries (European Social Survey, 2009, 2011, 2013a; Koch et al., 2012; Stoop et al., 2014). Also, the practical aspect of fieldwork efforts was central. The present study analysis can be used to determine whether the suggested methods to increase the response rate, namly the fieldwork efforts, actually enhance the response rate. Thus, it seems reasonable to include the previously dicsussed variables to describe fieldwork efforts in the FEI.[144]

Following this rationale, the present study constructed an additive index. All of the fieldwork efforts suggested within the ESS specifications for participating countries were given equal weight—no indicator was considered to be more positive or to be able to enhance the response rate more than any other indicator. For this reason, all fieldwork efforts have received the same weight.

17.2 Descriptive Analysis

On average, the effort that is put into the fieldwork of the ESS is high. Descriptive analyses of the fieldwork effort index (FEI) shows that the mean fieldwork effort index value in the ESS is 6.1, which means that on average 6.1 out of eight procedures of fieldwork efforts were implemented in rounds 1 to 5.[145]

According to (Wittenberg, 1998, p. 77), if an ordinal scale variable has a minimum of five characteristic attributes and is close to a normal distribution, then for data analysis, this variable can be treated as metric. The fieldwork index has eight characteristics that are ordered according to size. One can assume that the distances between the single values is the same. For this reason, the fieldwork effort index is treated as a metric variable.

The fieldwork effort index ranges from three to eight (Table 12), which suggests that some (3 cases) countries have implemented only three out of eight procedures of high fieldwork efforts, and some (12 cases) countries have implemented all eight of the possible

[144] Due to the substantial meaning of the variables included, other options for the construction of an index, such as examining empirical relationships (cross-tabulation, correlation coefficients, factor analysis), was not applied here.

[145] As in all the analyses in the present study, this is the value based on the analysis of the 16 countries that participated in all the rounds of the ESS. The missing cases are BE1, DE4, NL3, ES2, CH4, and CH5.

factors of fieldwork efforts. These finding shows that the variance of fieldwork efforts is rather high. For further descriptive analysis see Appendix Table C1 and Figure C1.

Table 12

Number of cases at each fieldwork effort level.

FEI value	n	Countries and ESS rounds
3	3	DK2, IE1, SI2
4	7	DK1, HU3, HU4, HU5, IE2, SI1, SI3
5	13	BE2, BE3, DK4, FR1, FR3, DE5, IE3, IE4, SI4, SI5, ES1, CH1, UK1
6	19	BE4, DK3, FI2, HU1, HU2, IE5, PL1, PL2, PL3, PL4, PL5, PT1, PT2, PT4, PT5, ES3, SE2, SE3, UK4
7	20	BE5, DK5, FI1, FI3, FR4, FR5, DE2, NL1, NL4, NO1, PT3, ES4, ES5, SE1, SE4, SE5, CH2, UK2, UK3, UK5
8	12	FI4, FI5, FR3, DE1, DE3, NL2, NL5, NO2, NO3, NO4, NO5, CH3

Note: FEI = fieldwork effort index.

As can be seen in Figure 45, the mean fieldwork index in ESS 1 (2002) was 5.6, and the mean fieldwork index in ESS 5 (2010) was 6.5. Thus, one could argue that, on average, all countries increased their efforts with one additional procedure compared to ESS 1. A *t*-test of independent samples (Levene test $p = .699$, thus equal variance is assumed) showed no significant difference between the mean fieldwork effort in 2002 and 2010 ($t = -1.19$; $p = .243$).

17.3 Quantitative Analysis

As described previously, fieldwork efforts usually are implemented in bundles. A fieldwork effort index (FEI) is constructed to analyze the effect of these combined factors on response rates. Corresponding to general Hypothesis 7 (*The more fieldwork efforts—such as the use of incentives, interviewer briefings, and so on—are implemented in a survey, the higher the response rate.*), a hypothesis was constructed to test for the effect of the fieldwork effort index on response rates:

> *Hypothesis 10: The higher the fieldwork effort (the score on the compound fieldwork effort index) is, the higher the response rate.*

According to Hypothesis 10, one would expect to find that in cases with a high fieldwork effort index, the response rate should be high as well.

The fieldwork effort in the ESS increased over time from a mean fieldwork effort index of 5.7 in 2002 to an average fieldwork effort index of 6.5 in 2010 (see Figure 45). Also, the mean response rate decreased from a mean of 61.3% (ESS 1 in 2002) to 57.2% (ESS 5 in 2010). However, a *t*-test showed that neither a change in response rate nor a change in fieldwork effort was significant.

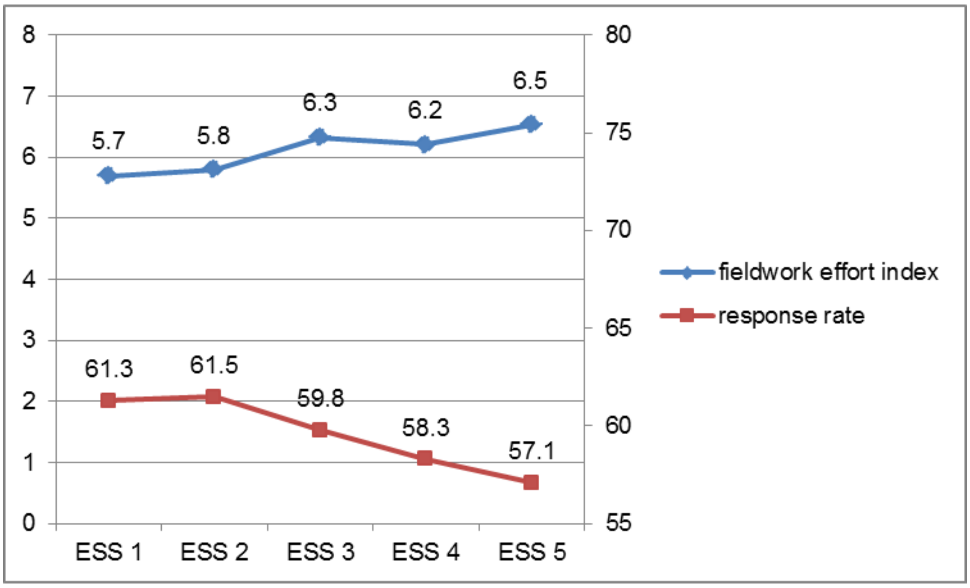

Figure 45. Mean level of fieldwork effort index (FEI) and mean response rate per round (in %).

The overall correlation between the fieldwork effort index and the response rate showed a lack of, or very low, relationship. The Pearson Correlation Coefficient (r = -.06; p = .596; n = 74) indicated that the variables of the fieldwork effort index and the response rate were not correlated. This finding showed that an increase in the mean fieldwork effort did correlate with an increase in the response rate. Thus, Hypothesis 10 cannot be supported. This result was expected, since the analyses of the separate fieldwork effort factors (Section

16.1 and Section 16.2) showed that only two out of the nine analyzed fieldwork efforts had a significant effect.[146]

The expected relationship described by Hypothesis 10 is that a high fieldwork effort index is correlated with a high response rate. By investigating the relation of fieldwork effort and response rate in more detail, a lot of variance was detected (see Figure 46 and the boxplot in Appendix Figure C1). The scatterplot (Figure 46) represents the case level, which shows the relation of the response rate and the fieldwork effort index in which the data points are widely spread. For the different levels of fieldwork, three different patterns can be described. It can be seen that in countries with low fieldwork effort index (level 0-4) (see Figure 46 marked with continuous circle) the response rate is high.[147] This is the opposite expected correlation of Hypothesis 10. In countries with a high fieldwork effort index (level 7-8) the situation differs. For some cases with a high fieldwork effort index, the response rate is high (see Figure 46, marked by a dashed circle, e.g., FI4, N02, NO3), whereas for other cases, the response rate is low (see Figure 46, marked by a dotted circle, e.g., DE1, DE3, FR3, CH3). For cases with a high level of fieldwork effort, a high variance was found in the response rates. The fact that a high variation on all levels of fieldwork effort was found supports the

[146] As described above with respect to the individual factors, the fieldwork procedures included in the index are of a non-experimental nature. This is not a problem per se, but it is important to mention that the level of the fieldwork effort index is not randomly assigned to countries—rather, it was decided by the researchers who made decisions about which fieldwork procedure should be implemented and which should not. A decision might be based on unknown or unobserved facts, such as the available survey budget, interviewer workload, general survey climate, or structure of the survey organization (e.g., are the interviewers hired by the survey organization employed full-time or are they freelancers). In Analysis Part C, Section 18, the analysis takes into account the change at the country level, but country specific factors are kept constant. Moreover, this Section 18 analysis considers the change between the rounds, but does not take into account the overall level of change.

[147] The explanation for this finding is suspected to be that in countries with high response rates, the researcher decided not to put more fieldwork efforts into the field because the response rate already was rather high. The explanation for the high response rates in those countries is suspected to be a good survey climate (among other things), rather than fieldwork efforts.

finding that the correlation of fieldwork efforts and the response rate is low. At the case level, a high fieldwork effort does not necessarily indicate a high response rate.[148]

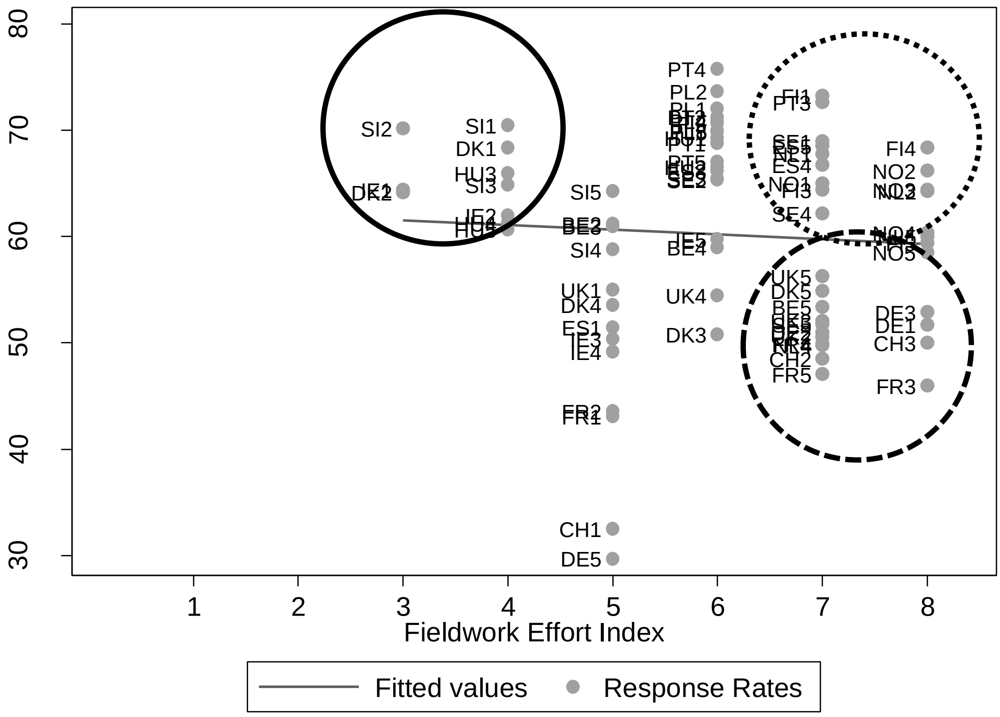

Figure 46. Scatterplot fieldwork effort index and response rate. As in other graphs in this study, the annotations are a combination of countries and round, e.g., SI2 indicates Slovenia in round 2.

Thus, the general hypothesis (Hypothesis 10) that high fieldwork efforts are correlated with high response rates cannot be supported. The analysis shows that implementing a bundle of fieldwork effort procedures does not necessarily correlate with high response rates. As can be seen in Figure 46, in some countries with a high response rate, the fieldwork effort index is low. On the other hand, in countries with high fieldwork efforts, response rates are high for some, and low for others. This finding does not mean that in some cases, fieldwork

[148] Splitting up the analysis for different points in time shows that in each round of the ESS, the variance of the response rate for the different levels of the fieldwork index is high. So, the variance of the fieldwork effort and response rate is not a time-specific issue.

procedures do not have a positive effect on a response rate within a country.[149] However, on the country level (see Figure 46), a high variance was found. The findings on the aggregate country level with high variance at the case level (countries in rounds) suggest the need to study the longitudinal development at the country level to further investigate the effects of fieldwork effort on response rate.

Due to the limited sample size other types of analysis are limited and descriptive analysis is appropriate way to gain insights. For example, a factor analysis shows that the indicators interviewer payment, length of interviewer briefing, personal briefing, interviewer experience explain 64% of variance. Comrey & Lee (1992) and Tabachnick & Fidell (2001) see problems doing a factor analysis with a limited sample size. They consider n = 80 as poor or very poor.

18 Change of Fieldwork Efforts and Change of Response Rates

Many country specific issues also may influence the response rate. These factors may be found at the macro level (e.g., society, culture, economic situation, social and cultural context), meso level of the survey itself (e.g., survey design), and micro level (e.g., respondent, interviewer) (cf. de Leeuw & de Heer, 2002; Groves & Couper, 1998). Some country specific influences cannot be kept constant, for example, the type of survey

[149] This analysis does not take into account the fact that countries also differ on other non-observed aspects of fieldwork that are not included in the analysis. Other intervening factors may operate on the micro level (such as the number of contacts of the interviewer and the respondent), macro level (such as the general survey climate), and meso level (survey level) that might help to explain more of the effects of fieldwork on the response rate. While performing the analysis of the correlation of fieldwork effort and response rate, one has to consider that it is unknown whether countries differ in other non-observed aspects of fieldwork that are not included in the analysis of the relationship of the response rate and fieldwork efforts.

organization (university, commercial, statistical office), the number of available interviewers, and the survey climate in general.[150] Thus, a longitudinal analysis is necessary.

This section analyzes the effects of the fieldwork effort on the change of response rates. The focus of this analysis is a longitudinal comparison of the changes of response rates and the changes of fieldwork efforts between rounds. The focus on the change between rounds enabled a further investigation of the contra-intuitive finding in previous sections.

By looking at the changes between the rounds the effect of the changes in the previous round on the proceeding round can be distinguished. This approach enables an analysis of whether a change in fieldwork effort has an effect on the response rate, regardless of the level of the fieldwork effort or the response rate. In addition, these analyses of the change between rounds keep country specific factors constant. By keeping country specific factors constant, the noise can be reduced. Of course, country specific factors like the survey climate and the number of contact attempts can change as well within a country. However, for this present study analysis, they are assumed to be constant.

In previous sections (Section 16 and Section 17), a negative relation between fieldwork efforts and response rate was found. The use of fieldwork efforts (such as incentives, the length of fieldwork, and so on) correlated with a lower response rate. A plausible explanation for this finding is mainly that countries, which had problems reaching a high level of response rate (e.g., in the ESS a target response rate of 70% is defined as high[151]) try to increase fieldwork efforts to stop the decrease in the response rate. Thus, it seems

---

[150] These factors not included in the model may have an effect on response rates as well. Of course, the survey climate or other country specific circumstances—such as the funding environment, the availability of high quality survey organizations, and the quality of the interviewer staff—might change. Also, the effectiveness of fieldwork procedures may be different in different countries. In the following analysis, it is assumed that these factors are constant.

[151] See European Social Survey (2011, p. 23).

plausible that in countries with low response rates, further fieldwork efforts were implemented.

18.1 Cross-Section Analysis: Overall Level

The relation between the change in fieldwork efforts and the change in response rates is analyzed in the following sections. First, different hypotheses are described and tested by conducting a regression analysis. Second, an analysis is presented of the different trends that may exist in fieldwork efforts (increasing, decreasing, constant).

This analysis is done to follow up on an issue that often arises in the discussion of the achieved response rate in a survey. When the response rate is discussed in the ESS, the national teams often say that even though the fieldwork efforts were increased, the response rate stayed constant. The national fieldwork managing team of the countries has claimed that in every round, more fieldwork efforts are necessary to keep the response rate from declining or from declining further. To empirically evaluate these country specific claims, the following analysis was conducted.

Keeping the countries constant,[152] the hypotheses on the changing relationship between fieldwork efforts and the response rate is as follows:

> *Hypothesis 11: A correlation exists between the change in fieldwork efforts (from the preceding effort to the subsequent effort) and the change in response rates (from the preceding round to the subsequent round).*

[152] As described previously, the change of fieldwork efforts and response rates across the ESS rounds allows keeping country specific factors constant, and thus enables a reduced noise analysis of the effects of these changes.

This hypothesis tests the overall effect of fieldwork efforts on the response rate. The hypothesis is further differentiated in the next sections for decreasing, increasing, and constant fieldwork efforts. These sub-hypotheses (Hypothesis 11a, 11b, 11c, see Section 18.2.1, 18.2.2, 18.2.3) focus on countries with decreasing, increasing, and stable efforts. Whereas Hypothesis 11 tests the global effort of fieldwork efforts, the following sub-hypotheses dig deeper into this effect (or the lack thereof). Hypothesis 11 is visualized, and an example is provided in Figure 47.

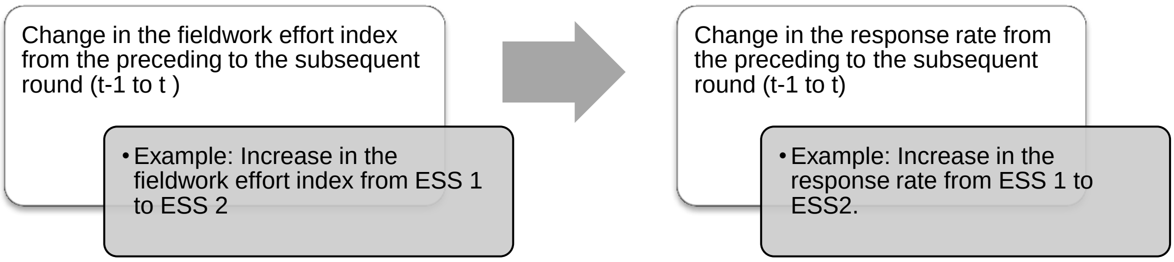

Figure 47. Hypothesis 11—change in the fieldwork effort index and change in the response rate at the country level.

To analyze the relation between the change in the fieldwork effort index (independent variable) and the change in the response rate (dependent variable), a linear regression analysis was conducted. The scatterplot, shown in Figure 48, visualizes the relation of the change in the fieldwork and the response rate from the previous to the subsequent round. The *x*-axis shows the change in the fieldwork effort index from the previous to the subsequent round (t-1 to t), and the *y*-axis shows the change in the response rate from the previous to the subsequent round (t-1 to t).[153] So, this scatterplot provides a picture of the change in the response rate and

[153] This analysis has systematic missing values. All cases in round 1 are missing, since no change to a "previous" round can be calculated. Also, a few cases are excluded. In round 3, Denmark (DK3) is excluded because the definition of the sample changed in this round. In Denmark, respondents were allowed to opt out from the register, so they are included in the gross sample, but are per definition non-respondents. This was a main change in the sample frame that caused a change in response rate. Due to the missing values for the fieldwork effort index, the following countries were not included in the analysis: BE1, DE4, NL3, ES2, CH4, and CH5.

the fieldwork efforts between rounds. For example, in Figure 8 on the *x*-axis, the value 2 can be seen, which represents Switzerland in round 2 (CH 2-point in the upper right part of the graph). Thus, by comparing the preceding (ESS 1) with the subsequent round (ESS 2), the difference in the fieldwork effort is 2, which means that the fieldwork effort has increased in ESS 2 by two additional fieldwork efforts compared to ESS 1. The comparison of the response rate (*y*-axis) of the preceding round (ESS 1) with the response rate of the subsequent round (ESS 2) can be described as an increase in the response rate of 16%. For example, in CH2, it is 16 (from 32.5% in ESS 1 to 48.5% in ESS 2).

A positive trend between the change in the fieldwork efforts and the response rate is expected (Hypothesis 11). The higher the fieldwork effort is compared to the preceding round, the higher the response rate is in the subsequent round.[154] The Pearson Correlation Coefficient shows that the variables "change in the fieldwork effort index" and "change in the response rate" are not correlated (r = -.01; p = .943; n = 55). Thus the pattern of the change of the fieldwork effort and the change of the response rate was not as expected. The regression analysis supports this finding. Figure 48 shows a positive but non-significant correlation (coef = .13; p = .361, n = 54; R^2 = .02) between the change in the fieldwork effort and the change in the response rate.[155] Nevertheless, the regression analysis shows that no correlation exists between the change in the fieldwork effort and the change in response rate.

---

[154] A high variation at every level of the fieldwork efforts (see Figure 48 and Figure C1) suggests that country specific effects may be occurring, which are followed up in Analysis Part C, Section 19. As already noted in the limitations of the fieldwork effort index, the causal relation between the fieldwork efforts and the response rate may be affected by previous results regarding the response rate in previous rounds or other surveys. This is an endogeneity problem because the causality is not clear. It may also be that the response rate in previous rounds or other surveys may affect the level of fieldwork effort, which may in turn affect the response rate in a subsequent round. The analysis above is based on the assumption that fieldwork effort affects response rates, which calls for further research.

[155] This analysis may be problematic due to the limited number of cases and also because of the low variation of the fieldwork effort. However, this is not the case for two reasons. First, this is a census of all the countries that

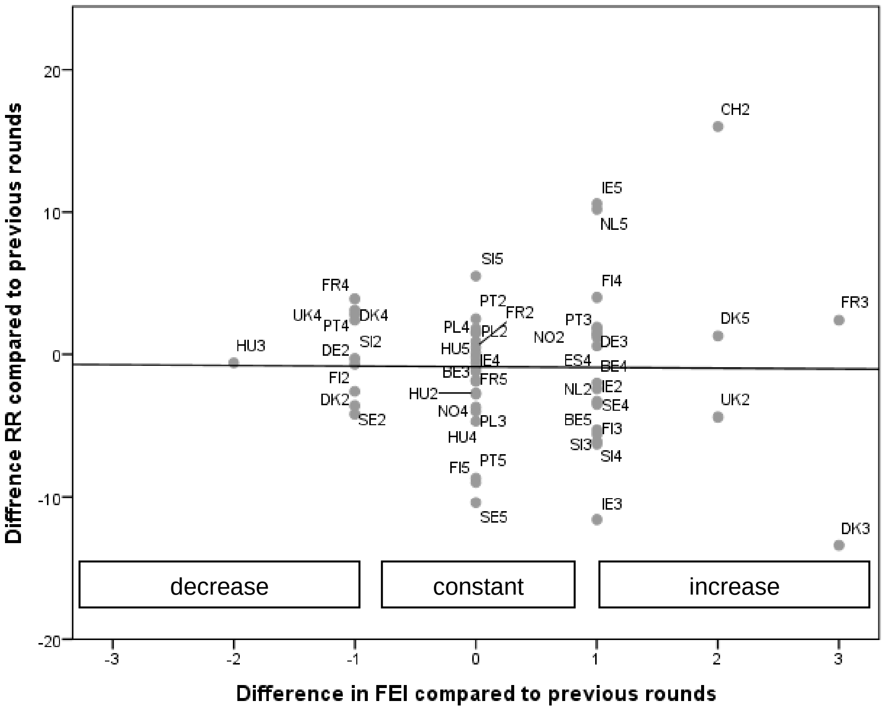

Figure 48. Change in the fieldwork effort index (FEI) and the response rate (RR) from previous rounds to subsequent round. The numbers in the labels next to the bullets mark the subsequent round (e.g., CH2 is the change between the ESS round 1 and ESS round 2 in Switzerland).

Thus, Hypothesis 11 describing the relation between the change in the fieldwork effort and the change in the response rates cannot be supported. An increase in the fieldwork effort does not correlate with an increase in the response rate. The longitudinal analysis of the change between rounds at the country level leads to the conclusion that at the country level, an increase in the fieldwork effort also does not correlate with a higher response rate. Thus, no correlation exists between the fieldwork effort and the response rate (Figure 48). This finding is in line with the finding of the cross-sectional analysis of the relation of the fieldwork effort and the response rate in Hypotheses 10, 11, 12a, 12b, and 13 (see Analysis Part C, Section 16 and 17). The cross-sectional analysis also found that an increase in the fieldwork effort had no significant effect on the response rate.

---

took part in all rounds. Second, the number of cases n = 55 and the analyzed number of variables still provide sufficient degrees of freedom (Jann, 2009).

This finding is contra-intuitive, since one would expect that more effort put into fieldwork would be related to better outcomes measured by higher response rates. Various explanations can explicate this effect, for example, fieldwork efforts only are increased when response rates are decreasing, although these efforts do not stop the decline. This argument often is used by survey agencies. Therefore, the naïve assumption that additional fieldwork efforts can stop the decline of the response rate is not reflected in the data. Another argument for the limitation of this analysis is that different effects in the change in response rates reveal unclear results.

Separate analyses of the change in fieldwork efforts may reveal more information. Thus, in the next section, further analyses are conducted on the different trends in fieldwork efforts (increase, decrease, same level of efforts) to examine whether within these different trends, the development of response rates may differ.

18.2 Separate Analysis for Different Trends in Fieldwork Efforts

This section presents a separate analysis for different trends of the response rate (decreasing, increasing, and same level of efforts), since the overall analysis (see Table 16) may mask specific information regarding the different trends in fieldwork efforts.

18.2.1 Increased fieldwork efforts

This analysis was conducted to examine the trend of increasing fieldwork efforts to determine whether in countries in which the fieldwork effort is increased, the response rate is affected. As described in the previous hypotheses (Hypothesis 7 to 10), increased fieldwork efforts are expected to increase response rates. By looking at the countries that increased fieldwork efforts from one round to the next, the effect of this increased effort on the response rate can be seen. Therefore, in this section, a further analysis is conducted only for the countries with increasing fieldwork efforts.

With respect to the effect of the increase of fieldwork effort and the change in the response rate, the following hypothesis can be formulated:

Hypothesis 11a: When countries put more effort into fieldwork than in a previous round, the response rate increases compared to the previous round.

This analysis follows up on a practical issue that often arises in the discussions about the achieved response rate in a survey. As mentioned earlier, the national teams who conduct the ESS in their country claim that more effort in fieldwork is necessary to maintain the response rate at the same level or to prevent it from declining further. So, these country teams claim that even though fieldwork efforts are increased, the response rate stays constant. The present analysis empirically tested this hypothesis.

Figure 48 shows the cases in which the fieldwork effort increased compared to the previous round (for further details, see Appendix Table C2). The data shows that in about half (52%) of the cases with an increase in fieldwork effort from the previous to the current round, an increase in the response rate occurred (in 11 out of 23 cases[156]). In the other half of the cases (12 out of 23 cases) in which the fieldwork effort increased from one round to the next, the response rates decreased. One can see that in countries with increased fieldwork efforts, the response rate can decrease: for example, in Denmark in round 3 of the ESS, the fieldwork effort was extended by three additional measures compared to round 2, but the response rate dropped by 13.4 percentage points in round 3 compared to round 2. This finding also means that in the previous round (Denmark in round 2) with less fieldwork effort, a higher response rate was achieved. To take another example, in Switzerland, in round 2 of

[156] A *case* is a country per round. The dataset included 16 countries that participated in all five rounds of the ESS—from ESS round 1 to round 5, which makes 80 cases in total. In 6 cases, the answers were missing in the fieldwork effort index (see Analysis Part C, Section 17). Thus, in total, 74 valid cases were available.

the ESS, the fieldwork was extended by two additional measures compared to the previous round, and the response rate increased by 17 percentage points.

The Pearson Correlation Coefficient (r = -.06; p = .783; n = 23) showed that the two variables—change in the fieldwork effort and change in the response rate—are only very weakly correlated. Looking at the cases in which more effort was put into fieldwork, the response rate increased in half of them and decreased in the other half. Thus, it can be argued that without the additional effort in fieldwork, the decrease in the response rate probably would have been larger. Thus, Hypothesis 11a—which states that the higher the fieldwork effort is, the higher the response rate—can be partially supported for half of the cases.

18.2.2 Decreased fieldwork efforts

Not all the countries in the ESS referred to in the present study increased their fieldwork efforts from round to round; in fact, some countries decreased their fieldwork efforts. In the next step, an analysis was conducted as to what effect this change in fieldwork effort had on the change in the response rate.

Hypothesis 11b: When countries put less effort into fieldwork than in the previous round, the response rate decreases compared to the previous round.

To analyze whether the relationship between decreasing fieldwork efforts and decreasing response rates can be detected, the cases with less fieldwork effort in subsequent rounds were examined. Cases with a decreasing response rate were rare. Only 10 out of 74 cases cut their fieldwork efforts from one round to the next (for further details, see Appendix Table C3). As can be seen in Figure 48, in about half of the cases with a fieldwork effort that was less than in the previous round (6 out of 10), the response rate decreased. In four cases, the response rate increased. However, the changes in the response rate as well as the fieldwork effort were very low.

Very few countries cut back their fieldwork efforts, and most of the time, the response rates slightly decreased. The Pearson Correlation Coefficient (r = .08; p = .838; n = 10) showed that the two variable were not correlated. This finding suggests that for those cases with decreasing fieldwork efforts, changes in the variable fieldwork effort index are not correlated with changes in the variable response rate. Thus, this finding does not support Hypothesis 11b.

18.2.3 Same level of fieldwork efforts

The third part of the present analysis concentrated on the effect of a constant fieldwork effort. With respect to the effect of a change in the fieldwork effort and the response rate, the following hypothesis was made:

Hypothesis 11c: When countries put the same effort into fieldwork as in the previous round, the response rate remains the same as in the previous round.

Looking at the cases where fieldwork efforts stay constant from one round to the next, the following pattern was detected (see Figure 48 and Appendix Table C3). In 36% of the cases (8 out of 22) the response rates increased. In 64% (16 out of 22) of cases with a constant fieldwork effort from one round to the next, the response rate decreased.

If the fieldwork effort stayed constant from one round to the next, the response rate increased in 36% of the cases, and decreased in 64% of the cases. Since one variable was constant, a Pearson correlation coefficient was not calculated. Based on these findings, Hypothesis 11c cannot be supported because when the fieldwork effort stayed constant, the response rate changed.

18.2.4 Result

The analyses of the different trends in fieldwork efforts (increasing, decreasing, and constant) from a previous to a subsequent round provided additional insights into the overall analysis of the effect of the fieldwork effort on the response rate (Figure 48). These analyses examined—at the country level—the change of fieldwork efforts from a proceeding to a subsequent round in relation to the change in the response rate from a proceeding to a subsequent round. By keeping the countries constant, the factors that might add noise to the analyses at the micro and macro levels were also kept constant.

To summarize, in cases in which the fieldwork effort was increased from the preceding to the subsequent round, the response rate increased in about half of the cases. Thus, Hypothesis 11a (*When countries put more effort into fieldwork than in a previous round, the response rate increases compared to the previous round.)* on the positive correlation of the fieldwork effort and response rate is partially supported. Hypothesis 11b (*When countries put less effort into fieldwork than in the previous round, the response rate decreases compared to the previous round.)* cannot be supported. Hypothesis 11c (*When countries put the same effort into fieldwork as in the previous round, the response rate remains the same as in the previous round.*) cannot be supported because when the change in the fieldwork effort was constant, the response rate changed.

With respect to a discussion with survey practitioners, the analysis showed that a change of fieldwork effort had a positive effect on the change of response rate (see Figure 47). Thus, Hypothesis 11a—that an increase in the fieldwork effort results in an increase in the response rate—can be supported partially.

An argument often used by survey practitioners is that extended efforts in fieldwork keep the response rate from decreasing even further. This effect can be found in the data. In cases in which more effort was put into fieldwork, the response rate increased half the time, and decreased half the time. Therefore, increasing fieldwork efforts may potentially contribute to a higher response rate. Additionally, in the cases in which the fieldwork efforts were increased but the response rate decreased, the probability exists that without these additional fieldwork efforts, the decrease in the response rate would have been larger. This would support Hypothesis 11a, which states that the higher the fieldwork effort is, the higher the response rate.

19 Longitudinal Qualitative Within-Country Analysis

By looking in detail at the country specific implementation of fieldwork efforts, one can see whether different measures were implemented, and thus whether different effects were produced. For a more detail study of the effect of fieldwork effort procedures on response rates, four countries were selected because they had different trends in the development of their response rates (see Figure 49). This within-country analysis goes further into the details of the different variables included in the fieldwork effort index (see Section 17). This detailed, at the country level analysis showed which fieldwork efforts were implemented in which round. So, doing this analysis enabled an investigation of what fieldwork effort procedures were implemented, and what the effects of these implementations were at the country level. This analysis also describes the change of fieldwork efforts between the rounds. In addition country specific effects can be analyzed.

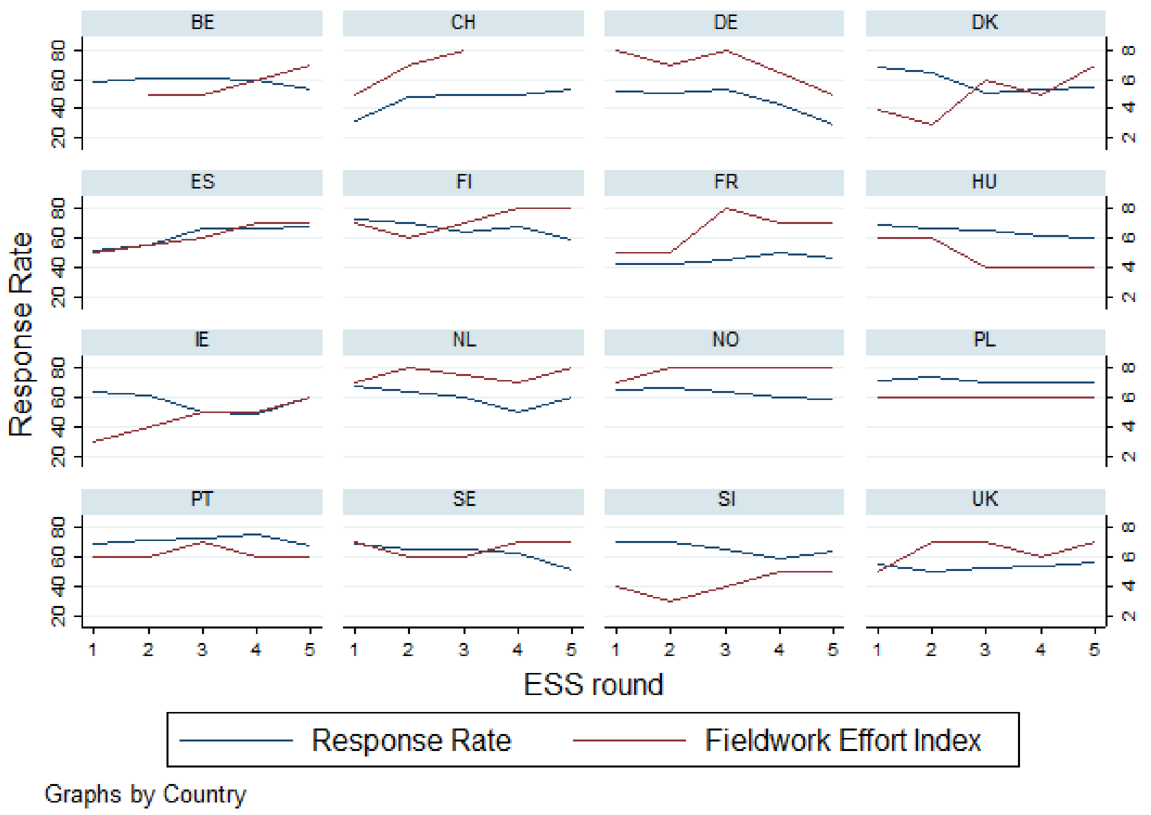

Figure 49. Mean level of fieldwork effort index and mean response rate per round (in %) per country.

After describing the change in the different fieldwork efforts, analyses of the correlation of the fieldwork efforts and response rates were conducted. In this section, the hypotheses on the change in response rates and fieldwork efforts are tested. The hypotheses put forward are (see Section 2.4):

Hypothesis 11: A correlation exists between the change in fieldwork efforts (in the preceding compared to the subsequent) and the change in response rates (in the preceding compared to the subsequent round).

For Hypothesis 11, different trends were tested and analyzed. The following hypotheses describe the different trends in fieldwork efforts more closely:

Hypothesis 11a: When countries put <u>more effort</u> into fieldwork than in the previous round, the response rate increases compared to the previous round.

Hypothesis 11b: When countries put less effort into fieldwork than in the previous round, the response rate decreases compared to the previous round.

Hypothesis 11c: When countries put the same effort into fieldwork as in the previous round, the response rate remains the same as in the previous round.

The analysis in this section provides detailed information about the country level for five rounds of the ESS (2002, 2004, 2006, 2008, and 2010). Due to this limited number of data points, the analysis performed here is more qualitative and descriptive in nature. This qualitative aspect adds important details to the quantitative analysis by providing a more detailed analysis of the different factors of the fieldwork efforts.

19.1 Germany—Decreasing Response Rate

Germany was selected for further analyses because it had a decreasing response rate. In ESS 5 (2012) the response rate in Germany (29.7%) was the lowest for all countries within the ESS. [157] Due to the decreasing response rate, it seems important to look at the fieldwork efforts.

For the analyses of the effect of fieldwork efforts on the response rate, first a list of all the implemented fieldwork efforts in Germany in each round was drawn (see Table 13). [158] Overall, the measurements that made up the fieldwork effort index stayed almost constant

---

[157] Source: European Social Survey (2016a) and Analysis Part A Table 3.

[158] For Germany in ESS round 4, the information on the percentage of experienced interviewers was missing. Personal communication (TNS Infratest, July 14, 2016).with the survey organization that conducted the fieldwork in Germany for ESS 4 revealed that this information is not available within the survey organizations. The missing value in ESS 4 for the variable “interviewer experience” was coded as “less than 90% experienced.” The reason for imputing no answer as 0 is that this is a deviation from the specification of the ESS. It is assumed that if a country complied with the specification (European Social Survey, 2009, 2011, 2013), it would be mentioned and the deviations might be omitted. An analysis of the two graphs with different versions of imputation of the missing value showed that the graphs differ slightly (see Appendix Figure C2 and Figure C3). However, there is no change in trend, and no difference in the interpretation. No substantial difference in the results with or without imputation was observed.

over the rounds. Changes over the rounds were found for the variables "interviewer experience", "personal briefing of interviewer", "length of interviewer briefing" and "refusal conversion training". For example, it was found that in ESS round 1, less than 90% of interviewers were briefed, whereas in ESS round 2, the effort was increased so more than 90% of the interviewers were briefed. This change in the efforts that were used for interviewer briefings can be described as an increase in fieldwork effort from round 1 to round 2. In round 5, a decrease in fieldwork efforts compared to the previous round was found for the variables "length of interviewer briefing" and "refusal conversion training". Overall, the fieldwork effort was at a rather high level in rounds 1 through 4, although it decreased to a medium level in round 5.

In Germany, the response rate decreased from the first round (51.7%) to the fifth round of the ESS (29.7%) (see Figure 50). At the same time, the fieldwork effort index decreased from 8 (ESS 1) to 5 (ESS 5) (see Table 13). With respect to the development between rounds, in round 5, for example, the fieldwork effort index decreased from 9 to 5. In the same round, the response rate also decreased drastically from 42.7% to 29.7%. In ESS 2, a slight decrease occurred in the response rate compared to ESS 1, and also the fieldwork index dropped from 8 points to 7 points (the change that occurred was that refusal conversion training was no longer being implemented). [159] This finding supports Hypothesis 11b (*When countries put less effort into fieldwork than in the previous round, the response rate decreases compared to the previous round.*). On the other hand, arguments can be put forward that do not support this hypothesis. In the first round of the ESS, the fieldwork effort was high with a fieldwork effort index of 8 (out of 8 possible fieldwork efforts). In ESS 3, the

[159] ESS round 4 was conducted by a different survey organization than the other rounds. TNS Infratest conducted round 4, and Infas conducted all the other rounds.

fieldwork effort index was still high (8 out of 8 points), but the response rate decreased anyway. However, in general, a pattern can be seen in that the higher the fieldwork effort is, the higher the response rate.

Table 13

Overview of fieldwork efforts in Germany

ESS round	Interviewer payment	Interviewer experience	Personal briefing of interviewer	Length interviewer briefing	Refusal conversion training	Use of advance letters	Use of brochures	Use of incentives	Fieldwork Effort Index	Response Rate
1	per hour or per interview + bonus	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	8	51.7
2	per hour or per interview + bonus	**less than 90% experienced**	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	7	51.0
3	per hour or per interview + bonus	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	8	52.9
4	per hour or per interview + bonus	90% or more of all interviewers experienced (imputed)	**less than 90% personally briefed**	more than half a day	Yes	Yes	Yes	Yes	8	42.7
5	per hour or per interview + bonus	90% or more of all interviewers experienced	**less than 90% personally briefed**	**half a day or less**	**no**	Yes	Yes	Yes	5	29.7

Note. Changes in fieldwork efforts from the proceeding to the subsequent round are bold.
[a]*Imputed value.*

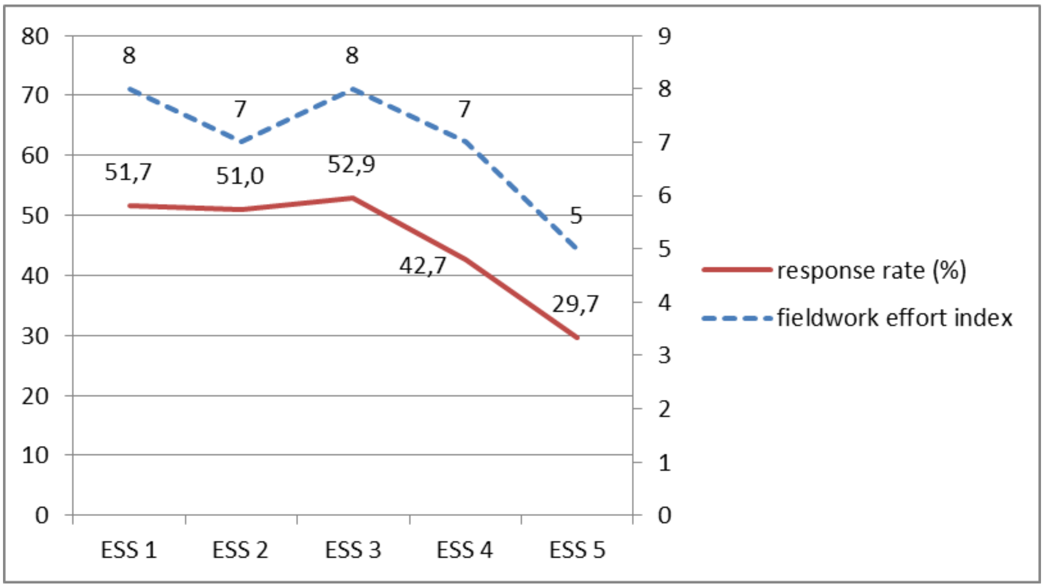

Figure 50. Germany— Mean level of fieldwork effort index and mean response rate per round (in %).

19.2 Spain—Increasing Response Rate

Spain was selected for further analysis because its response rate has a stable increase over time. Looking at the variables that constitute the fieldwork effort index, it was found that the following variables changed over the rounds in Spain (see Table 14): interviewer payment, personal briefing of interviewer, length of interviewer briefing, refusal conversion training, and use of a brochure. Over the rounds, these fieldwork efforts increased steadily from a medium level (5) to a high level (7).

Looking at the development of the fieldwork efforts and the response rate (Figure 51), it can seem that the fieldwork effort is at a medium level and increased from 5 (ESS 1) to 7 (ESS 5). At the same time, a steady increase of the response rate over the rounds (ESS 1: 51.5%, ESS 5: 68.6%) was found. As described previously, due to the limited number of data points, this analysis has a qualitative character. Although the data points showed a trend: the higher the fieldwork effort was, the higher the response rate. [160] Thus, for Spain, Hypothesis 11a can be supported. It can be assumed that the more fieldwork efforts are implemented, the higher the response rate goes.

Table 14

Overview of fieldwork efforts in Spain

ESS round	Interviewer payment	Interviewer experience	Personal briefing of interviewer	Length interviewer briefing	Refusal conversion training	Use of advance letters	Use of brochures	Use of incentives	Fieldwork Effort Index	Response Rates
1	**per interview**	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	**No**	Yes	**No**	Yes	5	51.5
2	per hour or interview + bonus	90% or more of all interviewers experienced	**less than 90% personally briefed**	more than half a day	Yes	Yes	Yes	Yes	7	54.9
3	per hour or interview + bonus	90% or more of all interviewers experienced	**less than 90% personally briefed**	**half a day or less**	Yes	Yes	Yes	Yes	6	66.2
4	per hour or interview + bonus	90% or more of all interviewers experienced	**less than 90% personally briefed**	more than half a day	Yes	Yes	Yes	Yes	7	66.8
5	per hour or interview + bonus	90% or more of all interviewers experienced	**less than 90% personally briefed**	more than half a day	Yes	Yes	Yes	Yes	7	68.6

Note. Changes in fieldwork efforts from the proceeding to the subsequent round are bold.
[a]Imputed value.

[160] In round 2, there are missing values for fieldwork effort (see Table 11). These values have been imputed. The imputation was conducted based on the assumptions that deviations from the project specifications of the ESS (European Social Survey, 2013a) were not or reluctantly reported. The low number of experienced interviewer staff is a deviation. So, if no answer was given, this was coded as 0 meaning that the fieldwork effort was not implemented. Another argument for this imputation is that the same survey organization was used as in first round (TNS Demoscopia). Assuming that the interviewer staff was constant, the same value as in ESS 1 was imputed. No difference in the results can be detected with or without imputation. For graphs, see the Appendix Figure C4 and Figure C5.

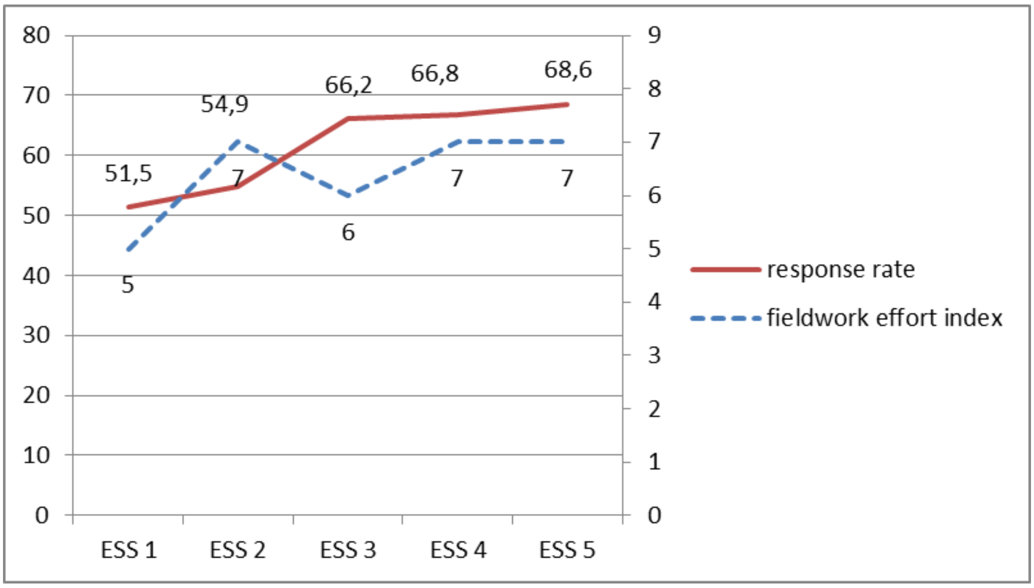

Figure 51. Spain— Mean level of fieldwork effort index and mean response rate per round (in %).

19.3 The United Kingdom—Constant Response Rate

The United Kingdom was selected for further analysis at the country level because its response rate was rather constant and at a medium level. This section also analyzes whether the same fieldwork efforts were implemented over the rounds.

Looking at the change of fieldwork effort (Table 15), it can be seen that fieldwork efforts change over the rounds, which was the case for the variables “interviewer payment”, “interviewer experience”, “length of interviewer briefing” and “use of brochures”. Even though the total number of fieldwork efforts procedures stayed the same, different procedures were implemented over the rounds. In Spain and Germany, the different procedures of fieldwork efforts were similar, or only one effort changed over the rounds (see Table 13 and Table 14). However, in the United Kingdom, the variance of combination is higher.

Table 15

Overview of fieldwork efforts in the United Kingdom

ESS round	Interviewer payment	Interviewer experience	Personal briefing of interviewer	Length interviewer briefing	Refusal conversion training	Use of advance letters	Use of brochures	Use of incentives	Fieldwork Effort Index	Response Rate
1	per interview	**less than 90% experienced**	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	**No**	Yes	5	55.0
2	**per hour or interview + bonus**	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	**half a day or less**	Yes	Yes	Yes	Yes	7	50.6
3	**per hour or interview + bonus**	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	**half a day or less**	Yes	Yes	Yes	Yes	7	52.1
4	per interview	**less than 90% experienced**	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	6	54.5
5	per interview	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	7	56.3

Note. Changes in fieldwork efforts from the proceeding to the subsequent round are bold.

An analysis of the correlation of fieldwork efforts and the response rate (see Figure 52) showed that the fieldwork effort index ranged between 5 and 7. The level of response rate is at a medium level in round 1 and increased in the second round to a high level (7) and stayed at a constant high level until round 5, with only a slight dip in that round. The response rate was rather constant between 55.0% (ESS 1) and 56.3% (ESS 5), which can be interpreted to mean that although fieldwork efforts changed, the response rate stayed constant over time. So, in the United Kingdom, no correlation existed between the fieldwork effort and the response.[161] This finding for the United Kingdom does not support Hypothesis 11a, 11b, or 11c.

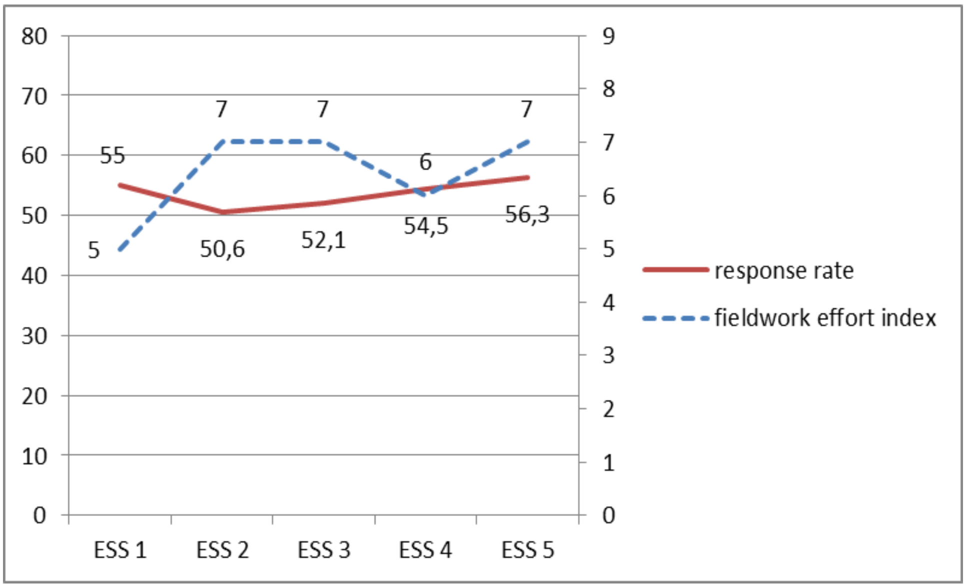

Figure 52. Mean level of fieldwork effort index and mean response rate per round (in %).

19.4 Portugal—Constant Response Rate

Portugal was selected for further research on a more qualitative analysis basis because of its rather high and constant response rate. Looking at the fieldwork efforts implemented in Portugal over the rounds (Table 16), it was found that the following variables changed over the rounds: “interviewer payment”, “interviewer experience” and “incentives”. In addition to those changes, the fieldwork efforts were rather constant and at a constant medium level of 6 (out of 8).

[161] As described previously, an interpretation of this trend is challenging due to the limited number of data points. For this reason, no further quantitative analysis was conducted, although a pattern still could be discerned.

Table 16

Overview of fieldwork efforts in Portugal

ESS round	Interviewer payment	Interviewer experience	Personal briefing of interviewer	Length interviewer briefing	Refusal conversion training	Use of advance letters	Use of brochures	Use of incentives	Fieldwork Effort Index	Response Rate
1	**per interview**	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	**No**	6	69.8
2	**per interview**	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	**No**	6	71.3
3	per hour or interview + bonus	90% or more of all interviewers experienced	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	**No**	7	72.7
4	**per interview**	**less than 90% experienced**	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	6	75.8
5	**per interview**	**less than 90% experienced**	90% or more of all interviewers personally briefed	more than half a day	Yes	Yes	Yes	Yes	6	67.1

Note. Changes in fieldwork efforts from the proceeding to the subsequent round are bold.

Not only fieldwork efforts but also response rates were at a constant high level (between 68.8% and 67.1%). The graph (Figure 53) shows a constant fieldwork effort and a constant response rate in Portugal, which can be interpreted as support for Hypothesis 11c (*When countries put the same effort into fieldwork as in the previous round, the response rate remains the same as in the previous round.*).

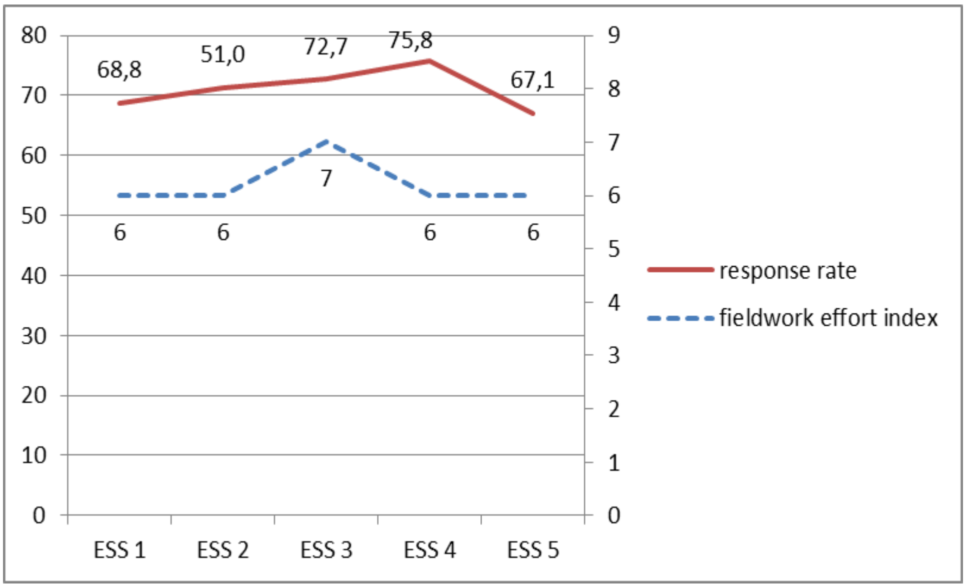

Figure 53. Portugal—Mean level of fieldwork effort index and mean response rate per round (in %).

19.5 Result

This section described, at the country level, the details of the fieldwork effort, and the relation of fieldwork effort and the response rate. It was found that in some countries (like the United Kingdom, Germany and Spain), the mix of fieldwork efforts changed over time. In other countries (like in Portugal), the fieldwork efforts remained rather constant. At the overall level, the correlation between the response rate and the fieldwork effort is low. However, at the country level, patterns were found: in Germany, where fieldwork efforts decreased, a decrease in response rates was found. This parallel development of the response rate and fieldwork effort supports Hypothesis 11b (*When countries put less effort into fieldwork than in the previous round, the response rate decreases compared to the previous round.)*. In Spain, when the fieldwork effort increased, the response rate also increased, which supports Hypothesis 11a *(When countries put more effort into fieldwork than in the previous round, the response rate increases compared to the previous round.)*. In the United Kingdom, the fieldwork effort seems to have no effect on the response rate. Portugal, the other country besides the United Kingdom with a constant response rate, supports Hypothesis

11c (*When countries put the same effort into fieldwork as in the previous round, the response rate remains the same as in the previous round.*). The analyses at the country level showed that fieldwork efforts have different effects in different countries.

20 Discussion

The limitations of these analyses are of course that the results might be affected with respect to the selection of the variables selected for the fieldwork effort index (FEI). Opportunities for further research might be may look at how other variables for measuring fieldwork efforts could be included in this analysis. For example, the effect of more information on for interviewer briefings and interviewer training might may provide as interesting results related to response rates. More and different aspects of fieldwork efforts would maybe perhaps provide information about which fieldwork efforts might show positive effects on the response rates. As a researcher, I would highly welcome if this additional information on meta- and paradata of surveys, such as the fieldwork efforts, would be published in data reports and published in datasets, preferably in a comprehensive way. Opportunities for further research would be to compare response rates and fieldwork efforts from more surveys.

A trend was found that all countries included more fieldwork effort in each subsequent round of the ESS (see Figure 49). This finding suggests a positive picture to survey researchers. The data, which provides some evidence that improvements regarding the quality of fieldwork over the rounds of the ESS is has been put into practice. The aim of the ESS as “continuous improvements over the rounds” (Jowell, 2010) is reflected in the data. As mentioned earlier, in the ESS, the countries receive feedback on compliance and deviations from the survey targets and processes (Halbherr, Koch, Kappelhof, & Stoop, 2014). Feedback on the performance in previous rounds might help to improve performance in subsequent rounds.

Comparing the results of my analysis to previous research in this area of fieldwork efforts, the results of my analysis extended to a broader data basis. Analysis from (Stoop et al., 2010, p. 105) included only a limited set of countries, and their analysis was restricted to one round of the ESS (namely the 21 countries that participated in ESS round 3 in 2006). The analysis performed in the present study done for five rounds of the ESS showed different results for the 16 countries that participated in all five rounds. In this the present analysis, a negative relationship between the index of fieldwork efforts and the response rate in ESS round 3 (r = -0.41) was detected.

Even though the index has many positive aspects, some of these might be critiqued when using an index. In this index, all the fieldwork procedures received the same weight, which means that the introduction of incentives was treated the same way as the level of experience of the interviewers. It is not known whether these factors have the same impact on response rates. Additionally, the effectiveness of procedures may differ between countries. For example, in one country, interviewer training may have a higher effect on the response rate, although the use of advance letters may be more effective in another country. Despite the fact that these assumptions have not been empirically tested, Stoop et al. (2010, p. 111) agree that using an index is a pragmatic approach to combining the various aspects of fieldwork factors into the analysis.

21 Results for Analysis Part C

This chapter analyzed the relationship between the fieldwork effort and response rate. Survey research assumes a positive correlation between high efforts in fieldwork and the achievement of high data quality and a high response rate. This hypothesis on the effect of fieldwork efforts on response rates were systematically tested in this chapter. Often information about fieldwork is not publically available or is not comparable across time and across countries. The data from the European Social Survey offers an opportunity to analyze systematically the effects of fieldwork efforts on response rates. Different hypotheses on the relation of fieldwork effort and response rate were tested in Analysis Part C. Thus, the present analysis has helped to answer questions about whether high fieldwork efforts can be associated with high response rates.

The analysis (Hypothesis 7) was conducted at four different levels. First, a cross-sectional analysis (Section 16) was conducted to examine the general relationship between separate fieldwork efforts and the response rate, non-contact rate, and refusal rate. Second, a fieldwork effort index (FEI) was constructed, and the effect of this additive index on the response rate was analyzed (Section 17). Third, the focus was directed to the longitudinal aspect and the change of fieldwork efforts and the change of response rates between the rounds (Section 18). Fourth, the correlation between fieldwork efforts and response rates was discussed from a qualitative perspective at the country level (Section 19).

The *first level of analysis at the cross-sectional level* was conducted to test the separate factors that are considered as fieldwork efforts.[162] Hypothesis 7 on the general relation between the fieldwork effort and response rate was applied to the cross-sectional level (Hypothesis 8). The analyses showed that implementing certain fieldwork efforts (such as interviewer briefing, use of advance letters, use of brochures, and so on) did not produce

[162] The included fieldwork efforts are: use of incentives (yes/no), use of brochures (yes/no), use of advance letters (yes/no), length of the interviewer briefing (more than half a day/less than half a day), personal briefing of interviewers (90% or more of briefed/less than 90% briefed), payment of interviewers (per hour or per interview + bonus/per interview), interviewer experience (90% or more experienced/less than 90% experienced), interviewer trained in refusal conversion (yes/no) and length of fieldwork (90 days and more/89 days and less) (see Table 11).

significantly higher response rates. In fact, the contrary seems to be the case. This finding was against the assumed relation of Hypothesis 8 and contra the intuitive assumption. Only for the fieldwork effort variables “incentive” and “length of fieldwork”, a significant difference between cases with and without this fieldwork effort was found. Of course, this paradoxical interpretation of the data is due to the non-experimental nature of the data. The fieldwork efforts are not randomly assigned to countries and rounds; rather, countries select fieldwork efforts on purpose. It can be assumed that countries take into account the challenges they are facing while trying to achieve a high response rate. In countries that expect a lower response rate, higher fieldwork efforts are planned. So, it is important to look at the changes in fieldwork efforts and response rates over time through a longitudinal perspective to minimize any country specific effects.

To further analyze the effect of fieldwork effort on response rates, the complementary of a response rate, the nonresponse rate, is analyzed. Following the general assumption that the high fieldwork effort leads to high a response rate, the hypothesis regarding nonresponse rates was formulated. If high fieldwork efforts are correlated with high response rates, high fieldwork efforts should be correlated with low nonresponse rates (meaning low non-contact rates and low refusal rate) as well. Based on this assumption, for the nonresponse rates (refusal and non-contact rates) the following hypotheses are formulated and tested: Hypothesis 9a and 9b (see Table 17). Regarding the non-contact rate, mixed results were found. For most of the additional fieldwork efforts, no significant effect was found. With respect to the “use of an advance letter,” Hypothesis 9a can be supported. However, for the variables “payment of interviewer,” the direction of the correlation was contrary to the hypothesis: higher fieldwork efforts—meaning better interviewer payments—were correlated to a higher non-contact rate. The same contra-intuitive finding was found when comparing the mean refusal rates in cases with and without an additional fieldwork effort of the variables “length of interviewer briefing” and “length of fieldwork”. For most of the variables, no significant difference was found between the higher fieldwork effort and the refusal rate (Hypothesis 9b). As mentioned previously, this finding can be an effect of the non-experimental nature of the experiment, which means that countries with high non-contact or refusal rates may decide to improve these rates by offering better pay to the interviewers, by lengthening the interviewer briefings, or by lengthening the time of the fieldwork.

The *second level of analysis includes an index of fieldwork efforts*. Since fieldwork efforts usually are implemented in a bundle, a fieldwork effort index that included the described fieldwork efforts was calculated. Hypothesis 10 on the positive correlation of the fieldwork effort index and the response rate also was tested (see Table 17). As for the analysis of separate fieldwork efforts, the hypothesis can be supported only partially. For some fieldwork efforts a correlation was detected. In some countries with high fieldwork efforts, response rates are high for some, while in other countries with high fieldwork efforts, response rates are low (see Figure 46). The Pearson Correlation Coefficient (r = -.06; p = .596; n = 74) showed the variable fieldwork effort index and response rate are not, or only very slightly, correlated.

The third level of analysis is the longitudinal analysis. By keeping country specific factors constant and looking at the changes between the previous and the succeeding round, the effects of the changes between rounds were analyzed. This approach enabled a focus only on the change. First, a general hypothesis was tested for the correlation of the changes of the fieldwork effort and the response rate (Hypothesis 11). Contrary to expectations, it was shown (Figure 48) that a change in fieldwork efforts did not have a positive effect on the response rate. For the different trends (increasing, decreasing, and constant) of the fieldwork efforts’ effects on the response rate (see Hypothesis 11a, 11b, and 11c), no, or only partial, support of the hypotheses was found (see also Table 17). In cases in which the fieldwork efforts increased or decreased, the response rates changed accordingly, but only in a few cases.

The argument often used by survey practitioners is that extended efforts in fieldwork keep the response rate from decreasing even further. This effect can be found in the data. In cases in which more effort was put into fieldwork, in half of the cases the response rate increased, and in the other half, it decreased. Therefore, increasing fieldwork efforts may (or may not) contribute to higher response rates. Additionally, in the cases in which fieldwork efforts increased but the response rate decreased, it is probable that without additional fieldwork efforts, the decrease in response rates may have been larger, although this was not tested.

The forth level analysis is focusing on quantitative analysis at the country level. At the country level, but not at the overall level, a pattern was found (see Table 17). For different countries, different hypotheses seemed to be supported. At the country level, an increase in fieldwork effort partially correlated with an increase in the response rate, even though the correlation showed a positive trend. Also, in a limited number of cases, a decrease in fieldwork efforts correlated with a decrease in the response rate. The results of the country specific analysis showed that fieldwork efforts are country specific. For example, in Spain, the relation of the fieldwork effort and the response rate was as expected by Hypothesis 11a. Also, in Germany, the relation is as expected by Hypothesis 11b. In Portugal, a high and constant fieldwork effort and a high and constant response rate support Hypothesis 11c. However, in the United Kingdom, no pattern of support for the above hypotheses was found.

These findings lead to the conclusion that fieldwork efforts and response rates or non-response rates are not correlated in general, but they need to be further analyzed at the variable and country levels.

For most of the tested hypotheses, above the positive relationship between fieldwork effort and response rate cannot be supported (Table 17). However, the different efforts of fieldwork still show an effect on the response rate at the country level or at the overall level. By increasing the effort, the response rate does not increase, but might be prevented from further decline.

Table 17

Results of hypotheses tests

Hypotheses	Result
General hypothesis on fieldwork effort and response rates	
Hypothesis 7: The more fieldwork efforts—such as the use of incentives, interviewer briefings, and so on—are implemented in a survey, the higher the response rate.	Partial support, variable and country specific support.
Fieldwork efforts and response rates	
Hypothesis 8: Countries that use additional fieldwork efforts (e.g., send an advance letter, or use a brochure or incentives) achieve higher *response rates* than countries that do not use these fieldwork efforts.	No support of hypothesis. Support for the variables “incentive” and “length of fieldwork.”
Fieldwork efforts and nonresponse rates	
Hypothesis 9a: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentives) achieve lower *non-contact* rates than countries that do not use these fieldwork efforts.	No support for most of the fieldwork efforts. Support for “use of an advance letter.”
Hypothesis 9b: Countries that use fieldwork efforts (e.g., send an advance letter or use a brochure or incentives) achieve lower *refusal rates* than countries that do not use these fieldwork efforts.	No support for most of the fieldwork efforts.
Fieldwork effort (as a compound “fieldwork effort index”) and response rates	
Hypothesis 10: The higher the fieldwork effort (the score on the *compound fieldwork effort index*) is, the higher the response rate.	No support of hypothesis.
Longitudinal analysis: **Change between the rounds of the fieldwork effort index and response rate**	

Hypothesis 11: A correlation exists between the *change in fieldwork efforts* (in the preceding compared to the subsequent round) and the change in response rates (in the preceding compared to the subsequent round).	No support of hypothesis.
Longitudinal analysis: **Change between the rounds of the fieldwork effort index and response rate—different trends**	
Hypothesis 11a: When countries put *more effort* into fieldwork than in the previous round, the response rate increases compared to the previous round.	No support of hypothesis. Partial support for half of the cases.
Hypothesis 11b: When countries put *less effort* into fieldwork than in the previous round, the response rate decreases compared to the previous round.	No support of hypothesis.
Hypothesis 11c: When countries *put the same effort* into fieldwork as in the previous round, the response rate remains the same as in the previous round.	No support of hypothesis.

Analysis Part D: Fieldwork Efforts and Nonresponse Bias

Most studies assumed that high fieldwork efforts lead to high response rates, and that high response rates correlate with low nonresponse biases. It is only implicitly assumed that high response rates are linked to data quality measured by a low nonresponse bias. In the methodological literature, only the link between fieldwork efforts and response rates or the link between response rates and nonresponse bias are analyzed, but less so the direct link between fieldwork efforts and the nonresponse bias (see Section 2, especially 2.3, 2.4 and 2.5).[163] Analysis Part D of the present study fills this research gap by empirically analyzing the relationship of fieldwork efforts and data quality.

An explanation for the implicit assumptions and the omission of empirical tests has been the argument that fieldwork efforts are difficult to measure. Information on fieldwork efforts is not usually publically available, and often not available in a systematic manner. The European Social Survey (ESS) provides this data, and thus enables an analysis of this often implicitly assumed link between fieldwork efforts and nonresponse bias.

Some groups are considered as a “hard to get population”, meaning that often they are less likely to take part in a survey. Groups that are more likely to take part in surveys (the “easy to get population”) are often overrepresented, and the “hard to get population” is often underrepresented (Goyder, 1987; Stoop, 2005) (see, Section 2.2).

Increasing fieldwork efforts do not mean that all respondents necessarily receive the same effort. The survey literature makes two assumptions about the relation of fieldwork

[163] Exceptions are, for example, the analysis of the relation of the number of contact attempts and nonresponse bias (Wood, White, & Hotop, 2006) or the analysis of the number of contact attempts and the coefficients of variations (Durrant, 2016).

efforts and nonresponse bias. On the one hand, increasing the fieldwork effort is said to decrease the underrepresentation. It is assumed that with more fieldwork effort, the sample is more balanced and less biased (Couper & de Leeuw, 2003, p. 165). On the other hand, it also is assumed that increasing the fieldwork effort leads only to including more of the "easy to get" respondents in the achieved sample (Stoop, 2005). Respondent groups that are already well included in the achieved sample could be even more overrepresented, and by increasing effort, "more of the same" respondents could be included in the sample. This possibility would lead to an increase in the response rate, but since only a certain group of respondents, namely "more of the same", would be included in the net sample, the nonresponse bias would be larger (see Section 2.2).

In this chapter (Analysis Part D), hypotheses on the correlation of the response rate and nonresponse bias are formulated and tested (see Table 18). The effects of processes (fieldwork efforts as an independent variable) on output quality indicators (nonresponse bias as a dependent variable) are further investigated. The research question is: "*Does more fieldwork effort lead to higher data quality measured by less nonresponse bias".* In other words, if more effort is put into fieldwork, are the respondents that are more likely to refuse or the less easy to contact respondents more likely to be included in the achieved sample?

In the present analysis, the over- and underrepresentation is systematically analyzed to assess the amount of nonresponse bias. The research question (Does more fieldwork effort lead to higher data quality [measured by less nonresponse bias]?) is further extended by asking whether higher fieldwork efforts lead to equally represented groups and low nonresponse bias or to an increased nonresponse bias with respect to the "easy to get population". Thus, it can be determined whether a correlation exists between the fieldwork effort and the nonresponse bias.

Two aspects are analyzed (see Table 18). First at the aggregate level the relationship of the fieldwork effort index (FEI) [164] and a nonresponse bias index is analyzed. Second, a variable specific analysis is conduction. Section 23 examines the effect of the fieldwork effort index (FEI) on the nonresponse bias of separate variables. Variable specific Hypotheses are constructed and analyzed.

The systematic and empirical analysis performed by the present study extends the existing literature by including fieldwork efforts in the discussion of nonresponse bias. Previous research such as the meta-analysis by Groves and Peytcheva (2008) on the response rate and nonresponse bias found that most of the variance in the nonresponse bias is within-survey rather than across-surveys. The following section analyzes whether nonresponse bias is a general characteristic of a survey or more variable specific and country specific. By using the same survey and keeping the topic and survey set-up constant, variations can be reduced and previous findings from the benchmark study by Groves and Peytcheva (2008) can be extended.

At the variable specific level, the present analysis takes into account these two main reasons for nonresponse[165] (refusal and non-contact) (see, e.g., Groves and Couper, 1998). It is examined whether the different effects of fieldwork efforts on nonresponse bias may be due to the different causes for nonresponse. [166] The question that is asked is whether fieldwork

---

[164] For the construction of the fieldwork effort index (FEI), see Analysis Part C Section 17.1.

[165] The AAPOR definition of nonresponse includes three main categories: refusal, noncontact, and others (American Association for Public Opinion Research, 2015). AAPROR defines the *refusal rate* as: “The proportion of all cases in which a housing unit or the respondent refuses to be interviewed, or breaks-off an interview, of all potentially eligible cases.” AAPOR provides three definitions of refusal rates that differ in the way they treat cases with dispositions of unknown eligibility. The *contact rate* is defined by AAPOR as: “The proportion of all cases in which some responsible housing unit member was reached.” (American Association for Public Opinion Research, 2015). AAPOR also provides three definitions of contact rates (Stoop et al., 2012).

[166] The third category of nonresponse is “not able/ others”. This category is often described and used as a residual category. Literature and research on this category is limited. For research on the category of “not able/ other”, see Stoop et al. (2012)

efforts have different effects on different types of nonresponse. Some variables might be more related to the contactability of respondents because they relate to an at-home-pattern (e.g., employed people are less often at home). Other variables are more related to refusals (older persons are assumed to more often refuse to take part in a survey). The question that is examined is whether fieldwork efforts have different effects on persons who are difficult to contact (non-contacts) or on persons who are, in general, considered more likely to refuse (refusals).

Table 18

Hypotheses for the relationship between fieldwork efforts and nonresponse bias

General hypothesis on fieldwork efforts and nonresponse bias

Hypothesis 12: The higher the fieldwork effort is, the lower the nonresponse bias.

Fieldwork Effort and Nonresponse Bias—Variable specific application

The higher the fieldwork effort is, the lower the nonresponse bias for the variable gender.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable old persons.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable young persons.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable low education.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable high education.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable working population.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable family status (married persons).

The higher the fieldwork effort is, the lower the nonresponse bias for the variable nationality.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable one-person household.

The higher the fieldwork effort is, the lower the nonresponse bias for the variable five-or- more person household.

22 Analysis at the Aggregate Level

The following sections examine the relationship of fieldwork efforts and nonresponse bias at an aggregate level. This approach enables an examination of the correlation of fieldwork efforts and nonresponse bias to help answer the question put forward by many survey practitioners: Does high fieldwork effort correlate with better data quality? The hypotheses on fieldwork efforts and nonresponse bias are as follows:

Hypothesis 12: The higher the fieldwork effort is, the lower the nonresponse bias.

At the aggregate level in a first step (Section 22.1), the effects of fieldwork efforts on socio-demographic variables (gender, age, education, paid work, and household size) are analyzed and visualized (Figure 54). In a second step (Section 22.2), as a compound measure of nonresponse bias, an additive index of nonresponse bias for the socio-demographic variables was constructed. This nonresponse bias index provides a more general analysis of the effects of fieldwork efforts on nonresponse bias.

22.1 Effect of Fieldwork Efforts on Nonresponse Bias: Variable Specific

This section analyzes the relationship between variable specific nonresponse bias[167] and fieldwork efforts. It can be seen (Figure 54) that an over- as well as an underrepresentation of variables exists. The variance of the nonresponse bias differs between the different levels of fieldwork effort. The extent of the nonresponse bias varies for different variables. The nonresponse bias not only differs between the variables, but also between countries.

[167] In many surveys, post-stratification weights are constructed using socio-demographic variables to adjust for the nonresponse bias. The present analysis does not including a post-stratification weight only design weights.

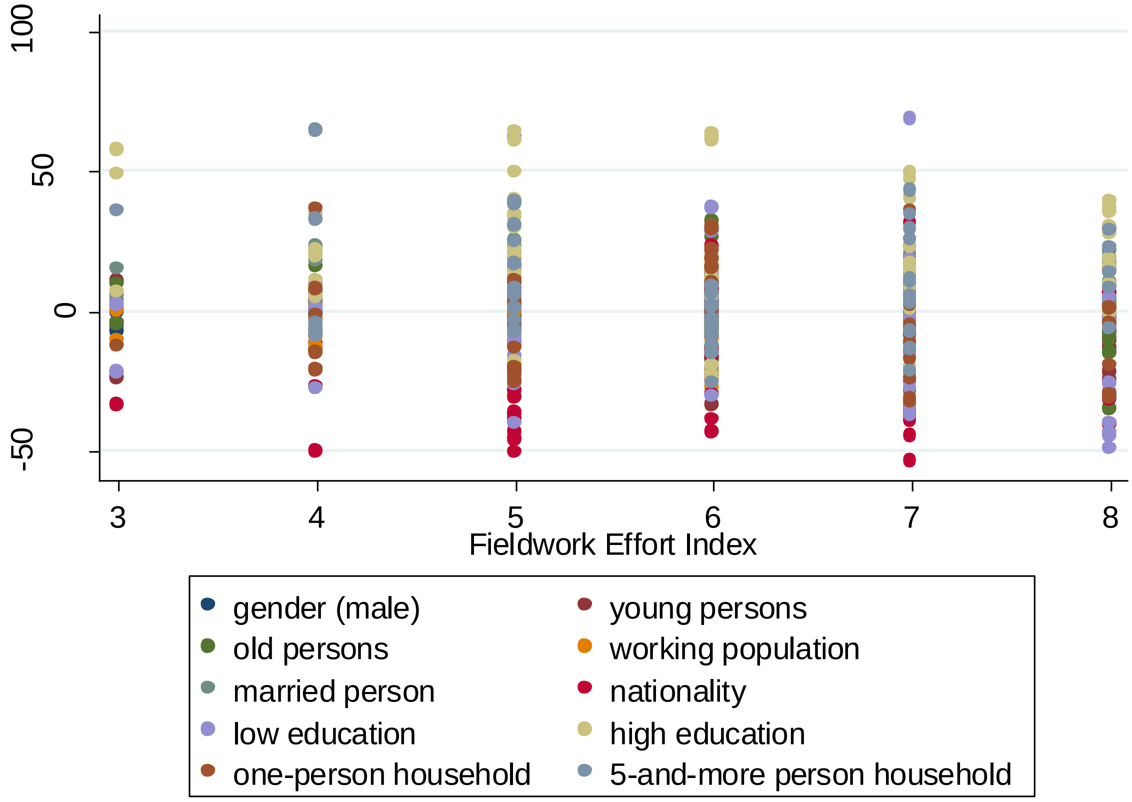

Figure 54. Fieldwork Index and relative nonresponse bias.

To empirically test the hypothesis on the relationship of fieldwork efforts and nonresponse bias the absolute values of selected socio-demographic variables was analyzed (male, younger persons [age 15-24], older persons [age 75 and older], working population, family status [married person], low education, high education, nationality, one-person household, 5-and-more person household-see Table 18). The absolute value does not take into account the over- and underrepresentation, but rather focuses on the absolute value of the bias.

It can be seen that the absolute value of the relative nonresponse bias—the level as well as the extent of nonresponse bias—varies between the variables and countries (Figure 55). A visual analysis shows a pattern—with higher fieldwork effort, the variance of nonresponse bias seems to decrease.

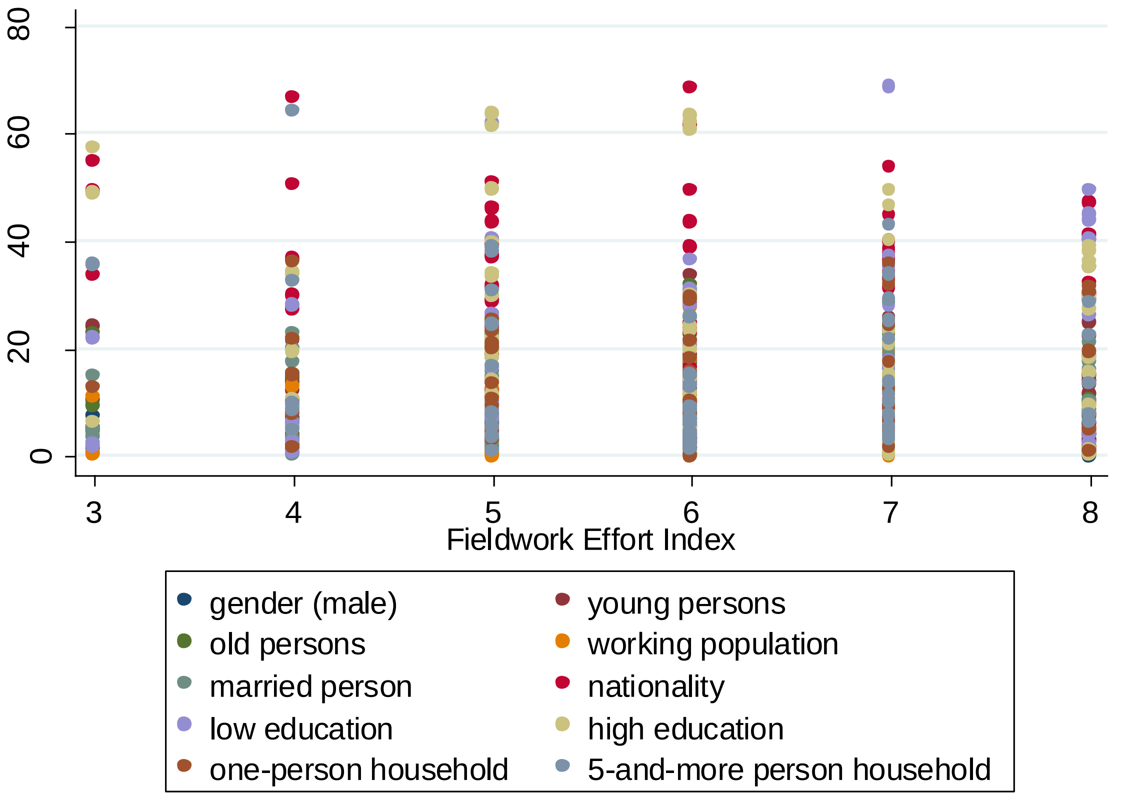

Figure 55. Fieldwork effort index and absolute relative bias.

22.2 Effect of Fieldwork Efforts on the Nonresponse Bias Index

To further empirically refine the analysis of the variables previously discussed above, an additive index of nonresponse bias was constructed[168]. The index includes the nonresponse bias (see Figure 56) of the previously analyzed socio-demographic variables (gender [male], young respondents, older respondents, working population, family status [married persons], low education, high education, nationality, one-person household, and five-or-more person household).[169] The nonresponse bias index is the sum of the absolute values of the relative

[168] A factor analysis shows that the indicators young persons, old persons, low education, gender explain 68% of variance. Comrey and Lee (1992) and Tabachnick & Fidell (2001) see problems doing a factor analysis with a limited sample size. They consider n = 80 as poor or very poor. Due to the low sample size, descriptive analysis was performed as well.

[169] Because the nonresponse bias index is based on the previous analyses other options for the construction of an index-like examining empirical relationships (cross-tabulation, correlation coefficients, factor analysis)-are not applied for the index construction.

bias of the analyzed variables divided by the number (11) of included variable.[170] In case of missing values, the standardization of the value was reduced to the valid number of variables.

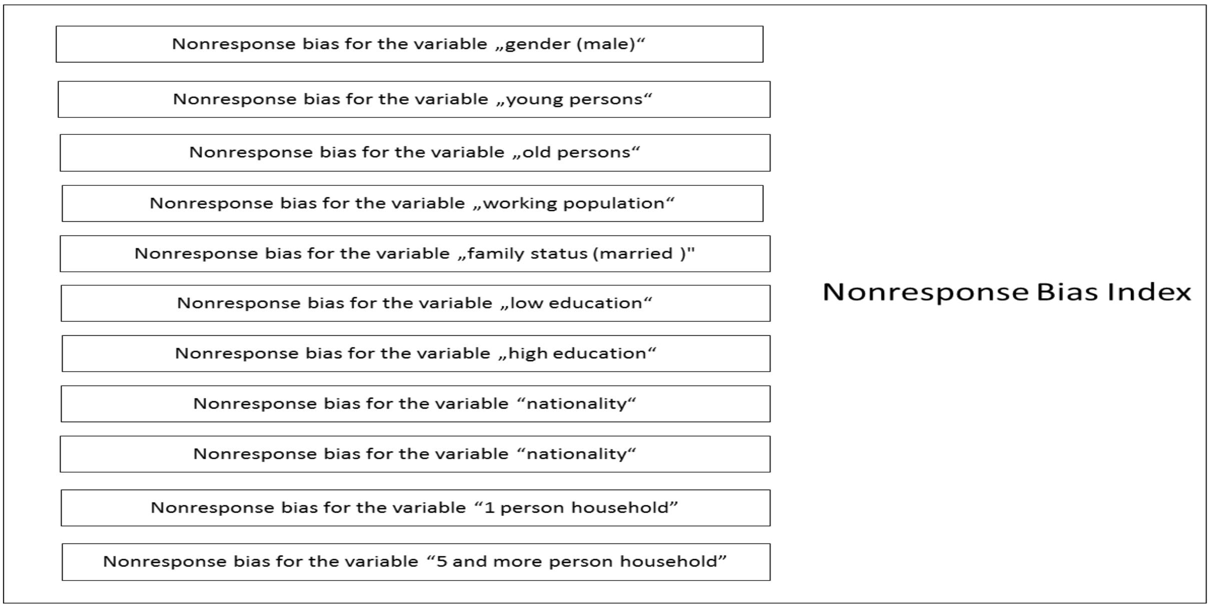

Figure 56. Visualization of the nonresponse Bias Index.

A model was constructed to analyze the relation of the fieldwork effort index (independent variable) and the nonresponse bias index (dependent variable). The Pearson Correlation Coefficient (r = -.08; p = .509; n = 74) shows a negative but not significant correlation between the variables fieldwork effort index and nonresponse bias index. A regression analysis (Figure 57) was conducted to examine the correlation and trend between the nonresponse bias variables and the fieldwork effort index. Previous analysis in the present study has shown that no assumptions for the Ordinary Least Square Regression analysis have been violated. The R Square indicates how much of the total variation in the dependent variable, the index of bias, can be explained by the independent variable, the fieldwork index. About 0.6% of the total variation can be explained, which also is very low. The ANOVA reports how well the regression equation fits the data (i.e., predicts the dependent variable).

---

[170] Extreme outliers for the variables education in the United Kingdom, nationality in Poland, and the old age group in Sweden were deleted for this analysis.

The significance level of p = .509 is well above the 0.05 significance level, which indicates that overall the regression (coef = -.24; t = -.66; p = .509) model does not predict, in a statistically significant way, the outcome variable (nonresponse bias index). Due to the limited number of data points (n = 74), the coefficient (coef = -.24) can be used to vaguely describe a negative correlation between the two variables. However, as described above, this correlation is not significant. Thus, Hypothesis 12 (*The higher the fieldwork effort is, the lower the nonresponse bias.*) cannot be supported.

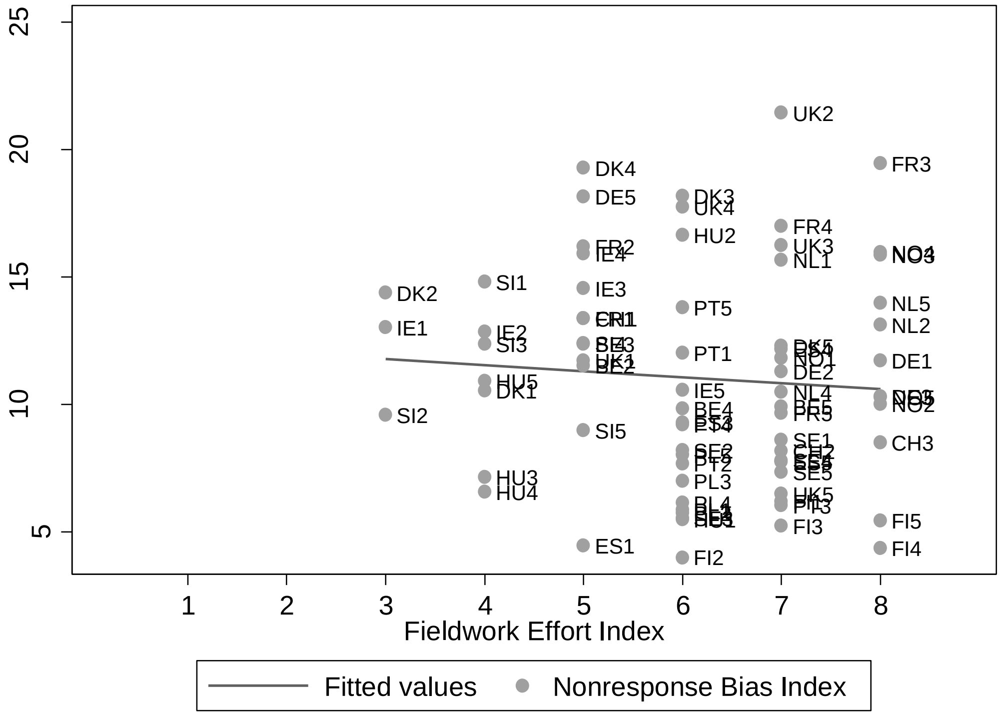

Figure 57. Fieldwork index and sum of bias (additive index of absolute relative bias).

23 Variable Specific Analysis

This section analyzes the fieldwork effort and nonresponse bias at the variable level, which enables answering the following research question: *Are there different effects for different variables? Does more fieldwork effort lead to less bias for different socio-demographic variables?* This analysis built on the one hand on the construction of the fieldwork effort index (see Analysis Part C Section 17.1), which was constructed as a compound measure of the different fieldwork procedures in the field of the survey. On the other hand the analyses were based on the calculation of the nonresponse bias analysis for the separate socio-demographic variables (see Figure 58 and Section 22). For details on the literature where the hypotheses are based upon see Section 2.6.

The general hypothesis is tested for different socio-demographic variables (variable specific):

Hypothesis 12: The higher the fieldwork effort is, the lower the nonresponse bias (for the variables gender (male), age, working population, family status, education, nationality or and household size).

.

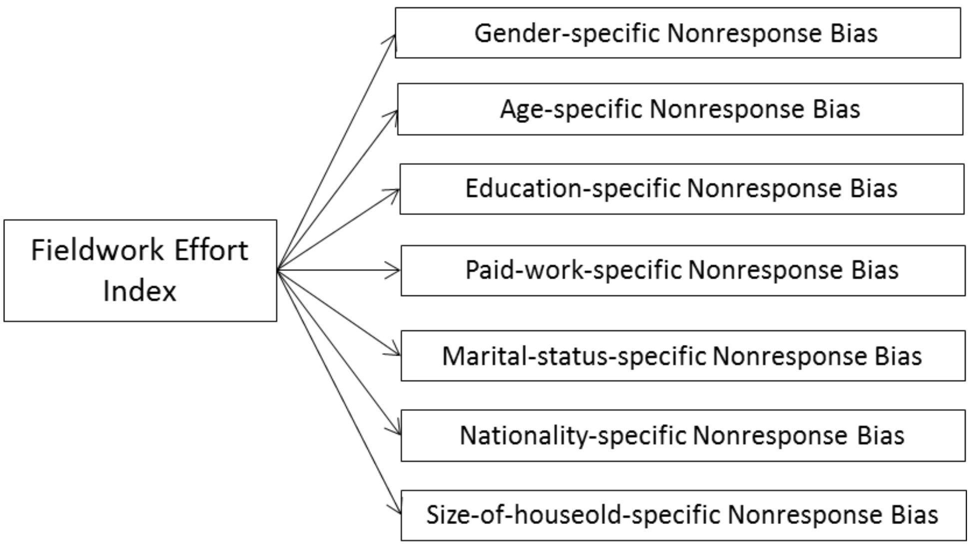

Figure 58. Hypothesis Analysis Part D, Section 23.

In the following chapters first the literature upon which the hypothesis are based upon are discussed. Than the correlation of the nonresponse bias and fieldwork efforts are analyzed. In this analyzes two steps are performed. First, the relative nonresponse bias, second the (relative value of the) absolute nonresponse bias is analyzed.

As a first step, data is analyzed using scatterplots that show the correlation of the relative bias and the fieldwork effort index. A regression analysis was performed for the separate variables for the fieldwork effort index and nonresponse bias. In addition, the Pearson Correlation Coefficient was calculated for each variable to measure the linear correlation between the variable fieldwork effort index and the different socio-demographic variables. In the second step, the same analysis was carried out with the absolute value of the nonresponse bias. The scatterplots for both kind of analysis includes regression lines, and the correlations are calculated, which was done for all variables on the variable level and on the overall level. If applicable for these variables, the reasons for nonresponse (refusal and non-contact) were discussed.

A graph with a relative bias is useful to understanding the general trend of the over- and underrepresentation of the variables. The graphs that display the (relative value of the) absolute nonresponse and show whether the bias is getting larger or smaller, independent of whether the bias decreases or increases.[171] This allows determining whether the bias is getting larger or smaller (Hypotheses 12a and 12b). So, both types of graphs (of the relative bias and the absolute value of the relative bias) are helpful for the analysis of nonresponse bias

23.1 Gender (male)

The reason for selecting the variable *male* for the nonresponse bias analysis is that male respondents are in general considered to be more difficult to contact than female respondents. Females are considered to be at home more often, since the probability that they work part-time is higher than for males. Also, female persons are more likely at home to take

[171] For advantages and disadvantages of the use of the “relative bias” and “absolute value of the relative bias” see Introduction of the Analysis Part B.

care of family, children, and older persons (Groves & Couper, 1998). Due to this at-home pattern, females are found to be easier to contact in a face-to-face survey.

Another reason for the overrepresentation of females in surveys can be due to the selection mechanism of the interviewers with respect to non-probability samples. Previous research has shown (Koch, Halbherr, Stoop, & Kappelhof, 2014; Kohler, 2007; Sodeur, 1997) that interviewers sometimes prefer to ask female respondents for an interview instead of male respondents. One reason for this might be that in the sample design—in which the interviewer has to do the household selection—she/he often commits fraud and asks the available person to do the survey instead of making an appointment with a person who is not at home (Koch et al., 2014; Kohler, 2007; Sodeur, 1997). So, an overrepresentation of females also can point to undocumented substitutions. The overrepresentation of female respondents often is interpreted as interviewer substitution of “easier cases”. Over-representation of female respondents can be considered as low interviewer compliance to the selection processes, and therefore, interviewer misbehavior.

To analyze the effect of the fieldwork effort on the nonresponse bias for the variable “gender (male)”, first a scatter plot was calculated. Figure 59 and Figure 60 show a scatterplot of the fieldwork effort index on the *x*-axis and the nonresponse bias on the *y*-axis.

The graph with the relative bias (Figure 59) shows that, in general, females were over-represented in the sample. In countries with a low level of fieldwork effort (5 and lower), females were much more overrepresented than underrepresented. Regression analysis showed a non-significant correlation (coef = -.19; p = .112; n = 74). Although, the graph suggests a negative correlation.

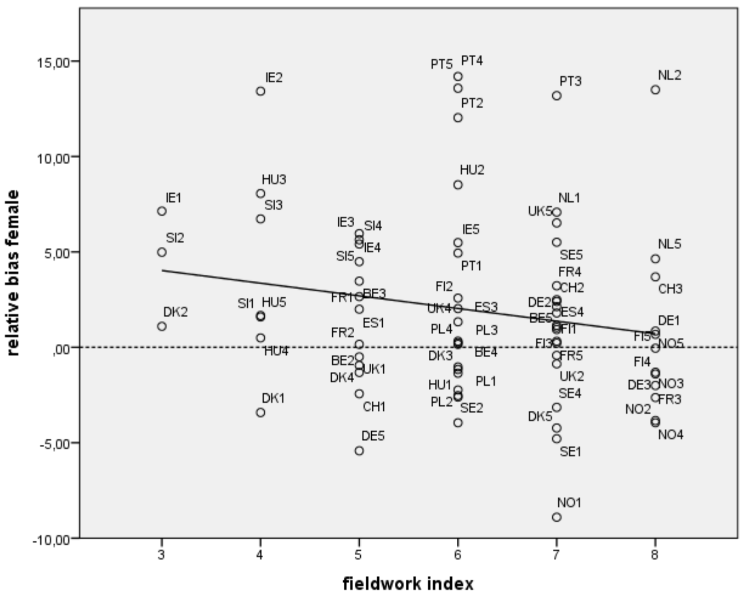

Figure 59. Relative bias for "gender (female)" and the fieldwork effort index.

Even though the absolute value of the relative bias (Figure 60) suggests that with an increasing fieldwork effort (measured by the fieldwork effort index), the nonresponse bias decreases, this correlation is not significant (r = -.10; p = .392). A regression analysis (Figure 60) indicated that the regression (coef = -.09; p = .429; n =74) model does not predict in a statistically significantly way the outcome variable (absolute value of relative bias for the nonresponse bias of the variable "gender [male]"). Thus, Hypothesis 12— which states that the higher the fieldwork effort is, the lower the bias—cannot be supported for the variable "gender (male)".

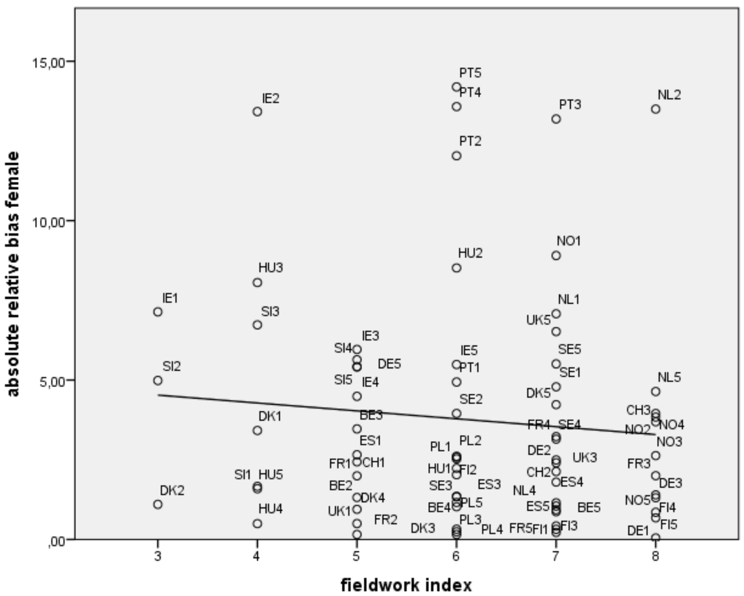

Figure 60. Absolute value for the relative bias for “gender (female)” and the fieldwork effort index.

23.2 Young Persons

The reason for selecting the variable “young persons (age 15 to 24 years)” for the nonresponse bias analysis is that younger respondents are, in general, underrepresented in surveys. This situation is in general due to the fact that young respondents are more difficult to contact (Lynn, 2003). They are at home less often, move (for schooling or universities) more often and thus their addresses are not traceable, and live more often in dorms or in one-person households. Due to this at-home pattern, they are considered as a “challenging to include group” for surveys.

The scatterplot of the relationship of fieldwork effort and nonresponse bias for the variable “young respondents aged 15 to 24 years” (Figure 61) shows an over- and underrepresentation in some cases. On average, young respondents are underrepresented. In countries with very high levels of fieldwork effort (fieldwork effort index of 8), the underrepresentation of younger respondents increases. A regression analysis showed that the correlation between fieldwork effort and the relative nonresponse bias for the variable “young respondents” was not significant (*r* = -.08; *p* = .494; *n* = 74).

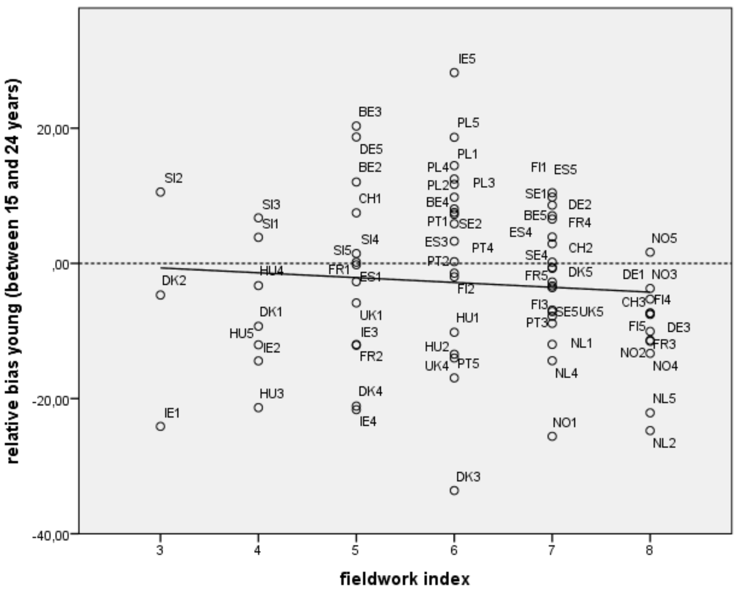

Figure 61. Relative bias for age “young persons” and the fieldwork effort index.

To analyze the effect of fieldwork effort on nonresponse bias, the scatterplot with the absolute value of the bias enables an interpretation of Hypothesis 12 (Figure 62). The visual analysis suggests that the fieldwork effort has a negative effect on the nonresponse bias, meaning that even with a higher fieldwork effort, the nonresponse bias decreased. However, a regression analysis showed that the correlation between the fieldwork effort and the absolute relative bias was a negative and not significant index (r = -.11; p = .356; n = 74). Even though the visual analysis suggested a slight tendency of lower bias with higher fieldwork efforts (as expected by the hypothesis), the empirical analysis did not support Hypothesis 12.

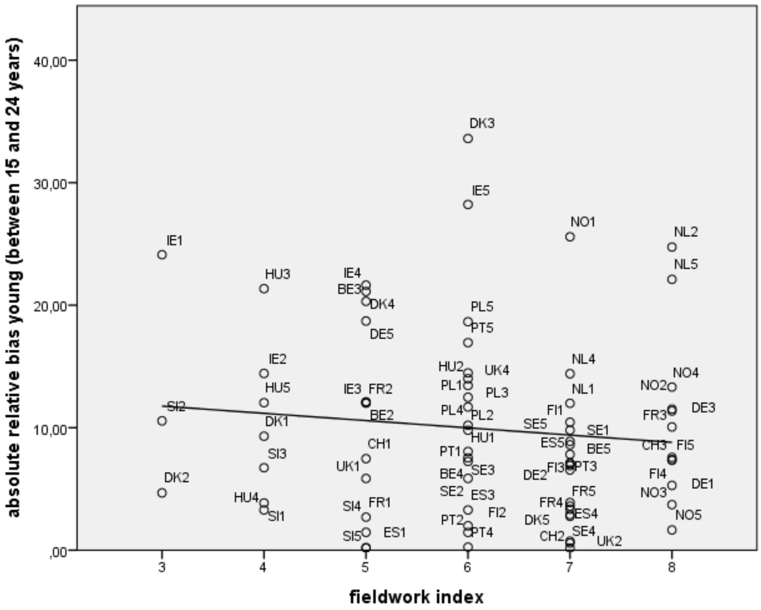

Figure 62. Absolute value for the relative bias for “young persons” and the fieldwork effort index.

23.3 Old Persons

The reason for selecting the variable “old persons (age 75 and older)” for the nonresponse bias analysis is that older respondents are considered more likely to refuse, since as Hoolbrook et. al. (2003, p. 94, p. 110) describe, they are reluctant to let a stranger in their house. Elderly persons may also be socially isolated and thus have lower cooperation rates (Groves & Couper, 1998, p. 133). Also higher incidence of health problems make elderly physically or mentally less able to participate (Cohan & Duffy, 2002, p. 21f.). Due to this fact, older respondents are considered to be more difficult to include in a sample (Kortmann and Halbherr, 2009).

The scatterplot of the effect of the fieldwork effort on the relative nonresponse bias of the variable “persons aged 75 years and older” shows, in general, an underrepresentation (Figure 63). This underrepresentation can be detected regardless of the level of the fieldwork effort.

With increasing fieldwork effort, the relative bias of persons 75 years and older seems to increase. However, a regression analysis showed that the relationship between the fieldwork effort and nonresponse bias was not significant (r = -.12; p = .350; n = 64).

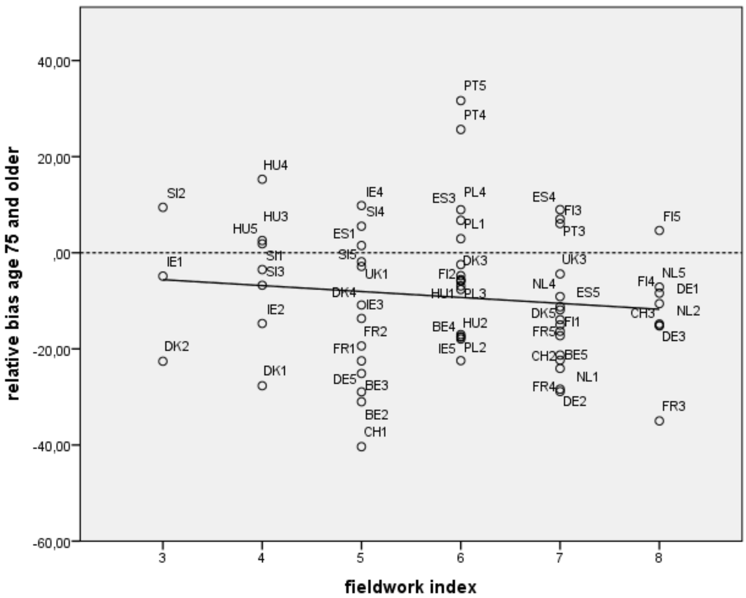

Figure 63. Relative bias for “old persons” and the fieldwork effort index.

As expected in Hypothesis 12, a high level of fieldwork effort seems to have a negative effect on the (absolute value of the relative) nonresponse bias of the population of persons 75 years and older. However, even though an empirical regression analysis showed a negative correlation (coef = -.07) between high efforts in fieldwork and high bias (absolute relative bias), the correlation was not significant (p = .595; n = 64). These analyses show that Hypothesis 12—which states that the higher the fieldwork effort is, the lower the bias—cannot be supported by the data.

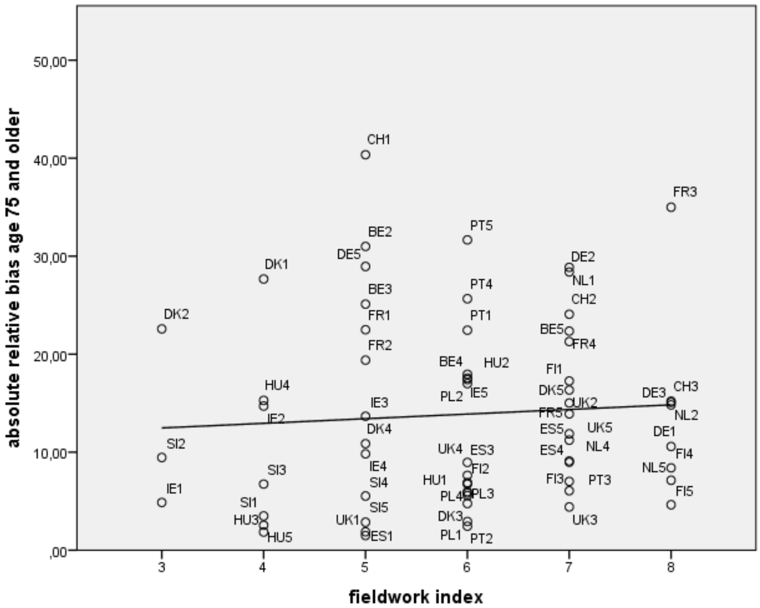

Figure 64. Absolute value for the relative bias for “old persons” and the fieldwork effort index.

23.4 Working Population

The reason for selecting the variable “working population” (definition: working for payment at least one hour in the last 7 days) for the nonresponse bias analysis is that persons in paid work are expected to be underrepresented in surveys. The reason for this underrepresentation is that they are at home less often, and thus are hard to contact (Lynn et al., 2002, p. 142). Due to this at-home pattern, persons in paid work are considered to be difficult to include surveys.[172]

The scatterplot of fieldwork efforts and the nonresponse bias for the variable “working populations” shows that persons in paid work are over- as well as underrepresented. As can be seen in Figure 65, the higher the fieldwork effort is, the better the representation of the working population. However, independently of the level of effort in general, the working population is on average overrepresented in the net sample.

---

[172] Note that here the variable “in paid work” is defined as “working for more than 1 hour in the last 7 days”. Persons working part-time are included.

This finding is against the expectation, since persons working are in general not at home and thus harder to contact for interviews. With a lower fieldwork effort, the working population is equally over- or underrepresented. With a high fieldwork effort, the working population is generally better represented with a slight tendency of overrepresentation. Thus, with increasing fieldwork effort, the nonresponse bias decreased. This trend is as we would expect according to Hypothesis 12.

An empirical analysis does support this hypothesis. A regression analysis showed a positive and significant (at the 10% significance level) correlation between the two variables (r = .20; p = .089; n = 74). The hypothesis stating a negative correlation between fieldwork effort and nonresponse bias can be supported for the relative bias.

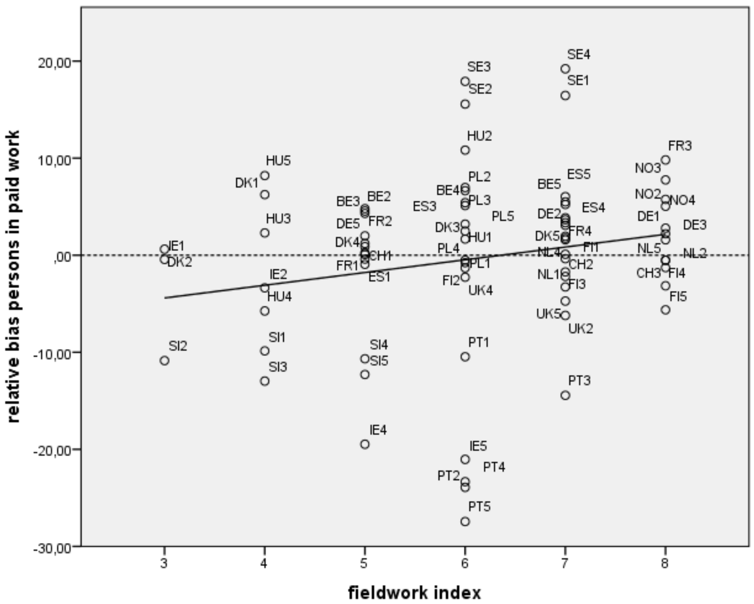

Figure 65. Relative bias for “working population” and the fieldwork effort index.

The scatterplot of the absolute value of the relative bias (Figure 66) shows that for the relative bias, the relationship was not significant (r = -.08; p = .525; n = 74). Although a visual analysis of the scatterplot (Figure 66) supports the hypothesis that the higher the fieldwork effort is, the lower the nonresponse bias. However, the empirical analysis did not support Hypothesis 12.[173]

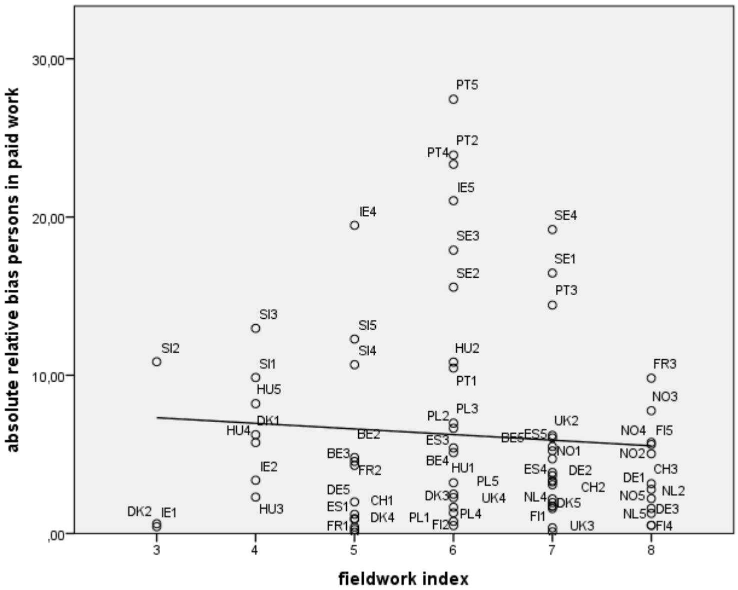

Figure 66. Absolute value for the relative bias for “working population” and the fieldwork effort index.

23.5 Family Status: Married Persons

The variable “family status: married persons” is included in the analysis of the relation of fieldwork efforts and nonresponse bias because marital status may affect contactability. Married persons are less likely to live in one-person households. Thus, the possibility to contact a married person at home may be higher if more than one person lives in the household. Since contactability is affected indirectly by marital status, the family status “married person” was included in the nonresponse bias analysis.

Figure 67 shows that, in general, married persons are mostly overrepresented in surveys. However, it can be seen that the nonresponse bias for the variable is not affected by

[173] The graphs and analysis excluding the outliers Sweden (SE), Ireland (IE), Portugal (PT), and Slovenia (SI) are displayed in Appendix Figure D1 and Figure D2.

the index of the fieldwork effort index (coef = .04, p = .756; n = 73). This shows that, independently of fieldwork effort, married persons are overrepresented in the net sample.

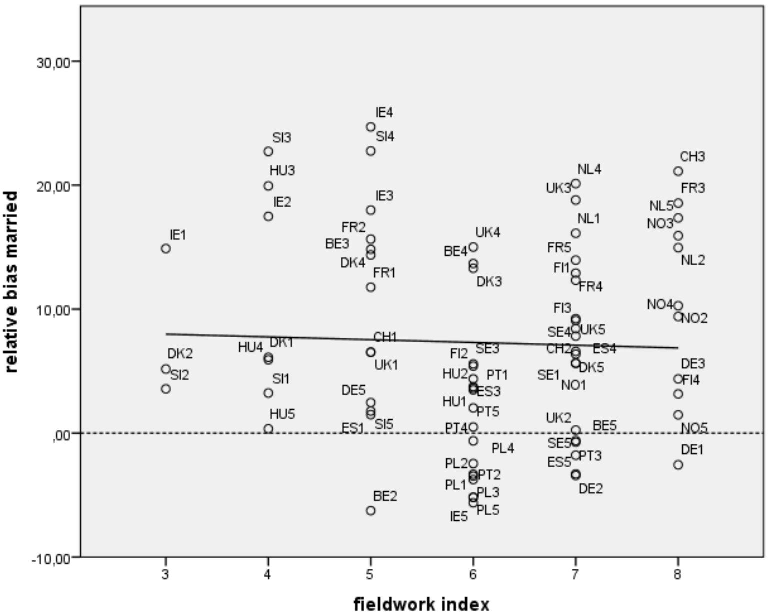

Figure 67. Relative bias for “married persons” and the fieldwork effort index.

An analysis of the effect of the fieldwork effort on the absolute value of the nonresponse bias enables a test of Hypotheses 20. A regression analysis also showed no significant correlation (coef = .03, p = .798; n = 73), which leads to the conclusion that Hypothesis 12 applied to the variable “married” cannot be supported by the data (Figure 68). No correlation between fieldwork effort and the nonresponse bias for the variable “married person” was found.

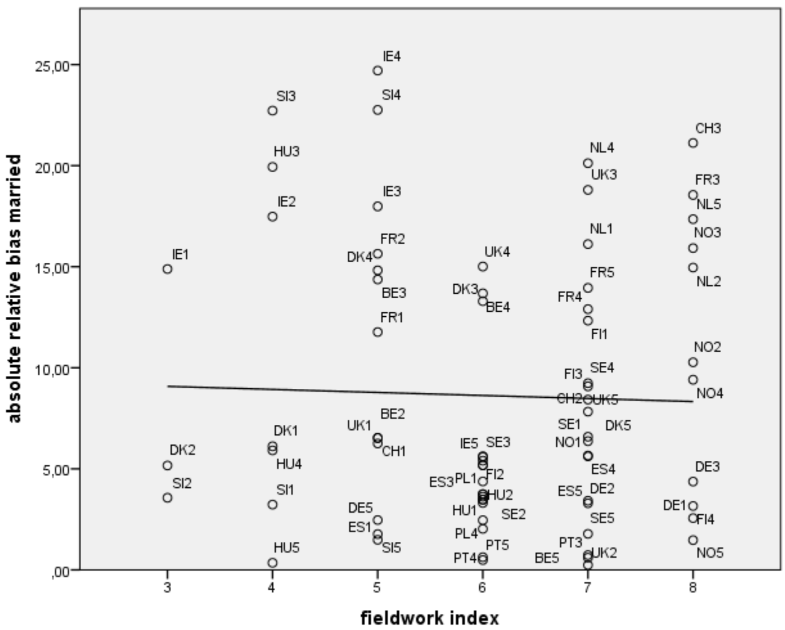

Figure 68. Absolute value for the relative bias for "married persons" and the fieldwork effort index.

23.6 Low Education

Respondents with low education (ISCED 0-2)[174] are less likely to take part in surveys, and so they often are underrepresented (Helmschrott & Martin, 2014; Koch et al., 2014). This is often is explained by the fact that persons with lower education refuse more often. Reasons for the higher nonresponse is that they often are afraid to give their opinion on a difficult survey topic (e.g., the European Union) and that the cognitive burden might be higher (Tourangeau & Smith, 1996, Holbrook et al., 2003, p. 82).

The graph with the relative bias (Figure 69) shows that, in general, persons with low education are underrepresented in the sample.[175] The graphs suggest a slight tendency of stronger underrepresentation of low-educated respondents with higher fieldwork efforts.

[174] The coding of the variable education follows the ISCED standard (Organisation for Economic Co-operation and Development, 1999). Low education is coded as ISCED 0-2, medium education as ISCED 3-4, and high education as ISCED 5-6.
[175] As described above, due to its extreme outliers (detected by calculation boxplots), the United Kingdom is excluded from the analysis of the education variable for both the relative and the absolute relative nonresponse bias.

However, an empirical regression analysis showed that a significant correlation between the two variables was not found (r = - .02; p = .855; n = 74).

A closer analysis of the countries showed that in Norway (NO) and Germany (DE), despite the very high fieldwork efforts, low-educated persons were highly underrepresented (see red circle in Figure 69). The graph suggests that the inclusion of respondents with low education could be especially relevant in those countries.

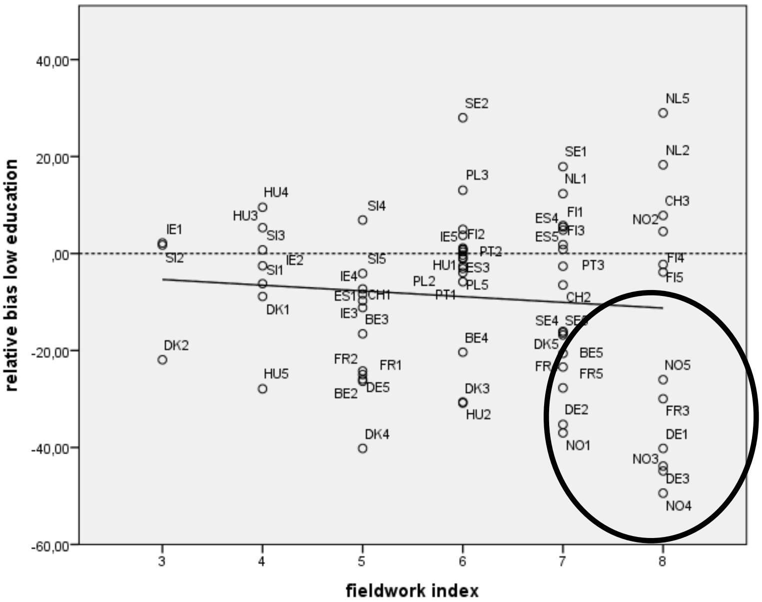

Figure 69. Relative bias for “low education” and the fieldwork effort index.

As reflected in the scatterplot of the fieldwork effort and the absolute value of nonresponse bias (Figure 70) and in the regression analysis, a positive and significant correlation (coef = .23; p = .050; n = 74) was found between high fieldwork efforts and high bias (absolute relative nonresponse bias).[176] This can be interpreted as the more effort that is put into fieldwork, the larger the nonresponse bias. This finding is against the stated relationship in Hypothesis 12.

[176] Analysis excluding the outlier United Kingdom is displayed in Appendix Figure D3 and Figure D4.

A plausible explanation of the contra-intuitive finding with respect to the relative bias is that a higher fieldwork effort and higher bias might be explained by the fact that persons with a high education are more willing to take part in a survey (for literature review and analysis of the variable "high education" see Section 23.7). So the "easy to get" persons with high education might be overrepresented, leading to an underrepresentation of persons with low education.

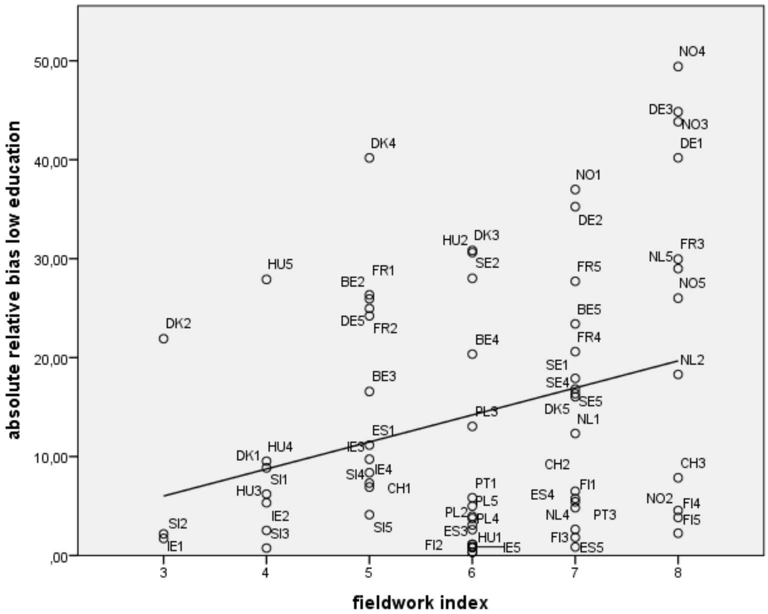

Figure 70. Absolute value for the relative bias for "low education" and the fieldwork effort index.

23.7 High Education

An overrepresentation of highly educated persons (ISCED 5-6) in surveys has been found (Helmschrott & Martin, 2014; Koch et al., 2014). This is explained by the fact that persons with higher education are more interested in different topics and not afraid of "knowing too little" to answer a survey.[177]

[177] Of course, the biases for the variables low and high education are interlinked. Mid education (ISCED 3-4) is not included in the present analysis, since it is not included in the nonresponse bias index. But the scatterplots are displayed in Appendix Figure D5 and Figure D6. A visual analysis showed that there is neither an over- nor an underrepresentation of the variables mid education. A regression analysis of the relative bias of mid

The scatterplot of fieldwork effort and the nonresponse bias for high education (Figure 71) shows that, in general, persons with high education are overrepresented in the ESS. This is the case in most of the countries, regardless of the level of fieldwork effort. However, a regression analysis did not support the visual findings of a negative trend. A significant correlation was not found between the relative bias for the variable "high education" and fieldwork efforts (r = -.16; p = .189; n = 69).

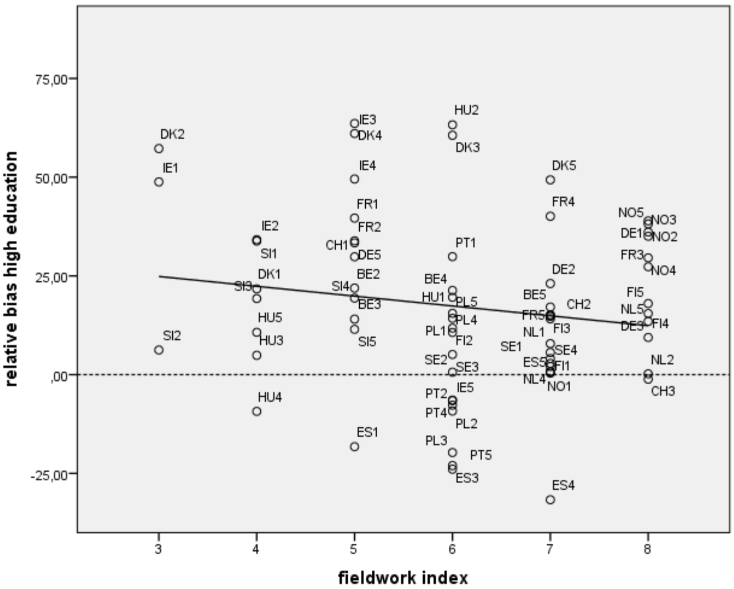

Figure 71. Relative bias for "high education" and the fieldwork effort index.

The scatterplot of the absolute value of the relative bias (Figure 72) shows the same trend for the variable "high education". A regression analysis showed a high negative correlation of coef = -.22 which is significant at the 10% level (p = .064; n = 69). This supports Hypothesis 12 that states that the higher the fieldwork effort is, the lower the nonresponse bias.

education and the fieldwork index showed no significant correlation (r = .15; p = .220; n = 68). Also, the regression analysis of the absolute value of the fieldwork effort index showed no significant correlation between the two variables (r = .03; p = .802; n = 68).

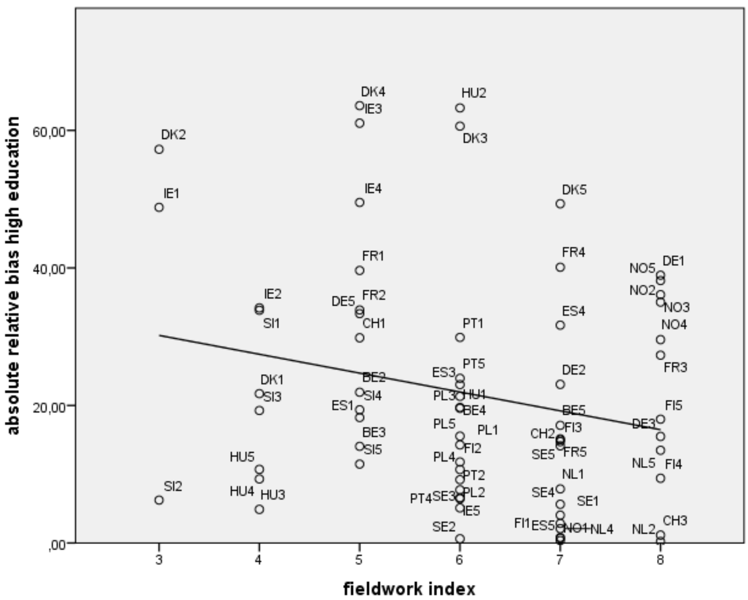

Figure 72. Absolute value for the relative bias for “high education” and the fieldwork effort index.

23.8 Nationality

Non-nationals are defined here as persons living in a country without having the nationality of that country. In survey research, non-national respondents are considered to be difficult to include in a net sample (Blohm & Diehl, 2001; Feskens et al., 2007; Helmschrott & Martin, 2014; Koch, 1997; Koch et al., 2014). In general, non-nationals are considered more mobile and less likely to be at home. Non-nationals are more likely to have language problems that may hinder participation in a survey and thus increase the chances of refusal.[178] Representation of non-nationals in a survey may be related to contactability and the refusal to participate in a survey.

Analyses were conducted to determine whether a correlation between high fieldwork efforts and better representation of non-nationals can be found. It was expected that with higher levels of fieldwork effort, the underrepresentation could be reduced considerably. Figure 73 shows that non-nationals are, in most cases, underrepresented. The extent of this

---

[178] Language problems that non-nationals might have are defined by the AAPOR outcome “eligible-non-interview” in the category “others” (American Association for Public Opinion Research, 2016).

underrepresentation is high. It was found that higher fieldwork efforts significantly reduced the underrepresentation of non-nationals (coef = -.29, p = .032; n = 55). The correlation was significant at a .05 level. However, even with the highest fieldwork effort, the group of non-nationals still is underrepresented. This analysis supports Hypothesis 12.[179]

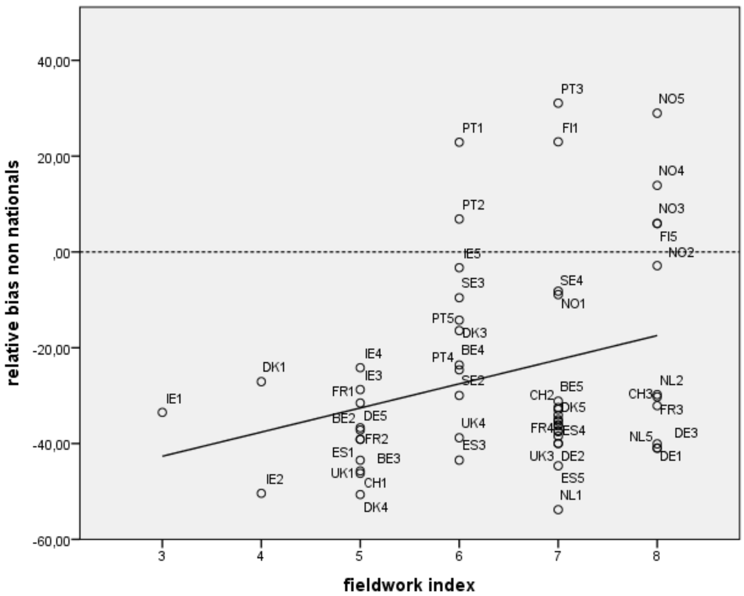

Figure 73. Relative bias for “non-nationals” and the fieldwork effort index.

An analysis of the absolute value of the nonresponse bias was used to test Hypothesis 12. The scatterplot in Figure 74 shows a negative correlation between the fieldwork effort and the absolute bias of non-nationals. The correlation is significant at the 10% level (coef = .23; p = .090; n = 55). Thus, the analysis supports the conclusion that fieldwork effort is correlated with the nonresponse bias for the variable nationality.

[179] Although in a few cases with high fieldwork effort, non-nationals are overrepresented (Portugal round 1 to 3, Finland round 1 and 5, and Norway round 3 to 5).

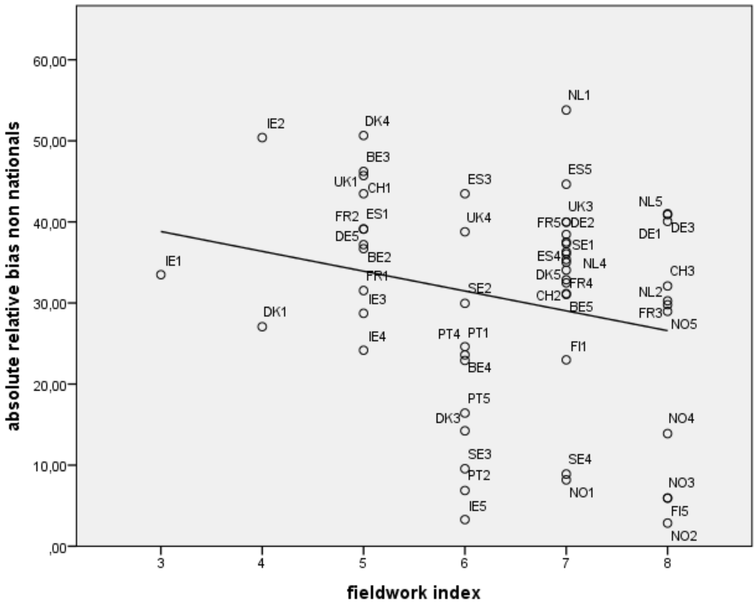

Figure 74. Absolute value for the relative bias for “non-nationals” and the fieldwork effort index.

23.9 One-Person Household

The number of persons living in a household may influence the contactability of respondents. If more people live in a household, the probability that someone opens the door is higher than if only one person lives in the household. Even though the person opening the door might not be the potential respondent, she/he can provide information about the best time and day to contact the potential respondent. Studies indicated that accessabilty to households of sampled persons is related to the probabilty that a household member is at home (Goyder (1987), Koch (1993), Lynn et al. (2002), Lynn & Clarke (2002) and Stoop, 2005, p. 70). Previous analysis showed that people living alone have higher risk of refusal and are thus more likely to be underrepresented (Durrant, & Steele (2009) Goyder (1987) and Campanelli, Sturgis, & Purdon (1997, 3ff.). Thus, household size plays a role in contactability and likelihood to refuse. Persons living in one-person households are considered to be more difficult to contact, and so more contact attempts may be necessary. [180]

For this reason, the effect of fieldwork efforts on the nonresponse bias for the variable "one-person household" was analyzed. The data shows that one-person households are both over- and underrepresented (Figure 75). A regression analysis showed no correlation between the fieldwork effort and nonresponse bias for the variable "one-person household" (coef = -.13; p = .382; n = 47).

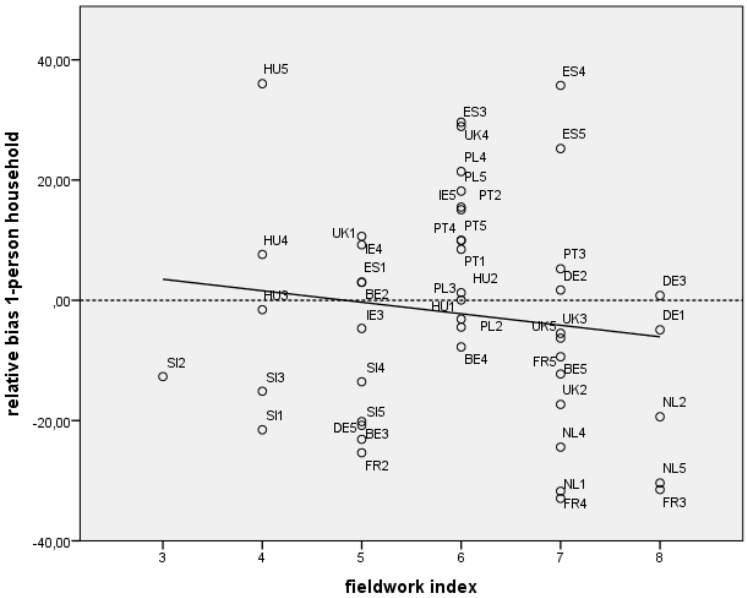

Figure 75. Relative bias for "one-person household" and the fieldwork effort index.

The absolute value of the nonresponse bias for one-person households (Figure 76) focuses on the size of the bias. It was found that the higher the fieldwork effort was, the higher the nonresponse bias in general. The correlation between the fieldwork effort and one-person households is positive but not significance (coef = -.10; p = .501; n = 47). Thus,

[180] People living in 1-person households are more likely to be young or older (Durrant & Steele, 2009; Goyder, 1987), so age may have an intervening effect.

Hypotheses 24 cannot be supported. No significant correlation between the fieldwork effort and nonresponse bias was found.

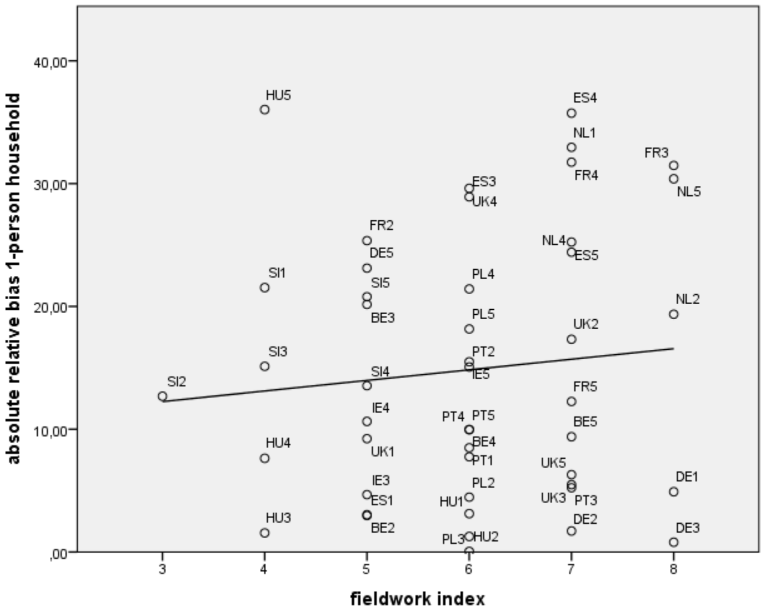

Figure 76. Absolute value for the relative bias for "one-person household" and the fieldwork effort index.

23.10 Five-or-More Person Household

As described previously (Section 23.9), the number of persons living in a household may influence the contactability of its potential respondents. The more people are living in a household, the higher the probability that someone will open the door.

The scatterplot (Figure 77) shows that, in general, persons living in five-or-more person households are overrepresented in the data. A regression analyses showed that no significant correlation between fieldwork effort and nonresponse bias for this variable was found (coef = .25; p = .111; n = 42).[181]

---

[181] The bias of five-or-more person households is large in Slovenia (SI). By deleting the data points for Slovenia, a different trend can be detected. The negative correlation for the absolute relative bias (r = -.18, p = .234) changes to a positive trend (r = .23; p = .14). The effect of one country (5 cases) in an analysis of 47 cases (for the variable household size) was rather high. For the graphs see Appendix Figure D7 and Figure D8.

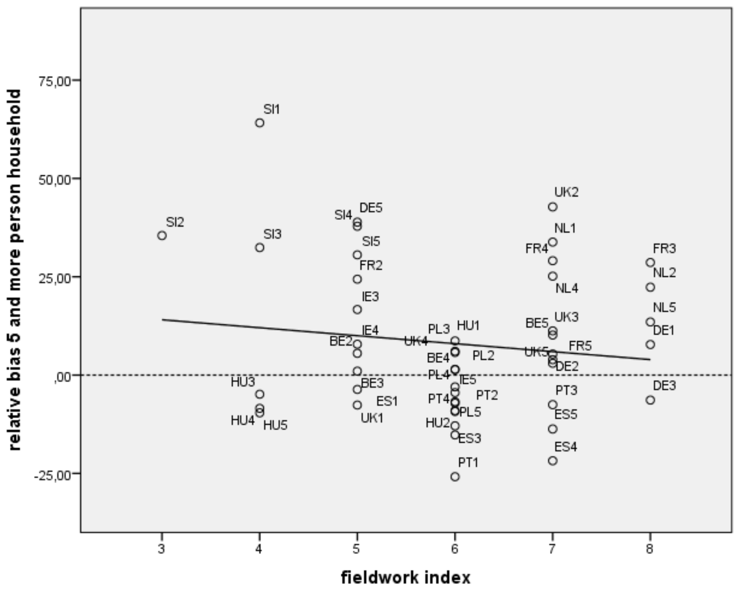

Figure 77. Relative bias for “five-or-more person household” and the fieldwork effort index.

The correlation of the absolute relative bias of five-or-more person households and fieldwork efforts is negative but not significant (coef = .18; p = .234; n = 42). Thus, a relation between the level of fieldwork effort and the nonresponse bias for the variable “five-or-more person households” was not found. Thus, Hypothesis 12 cannot be supported for the variable “five-or-more person household”.

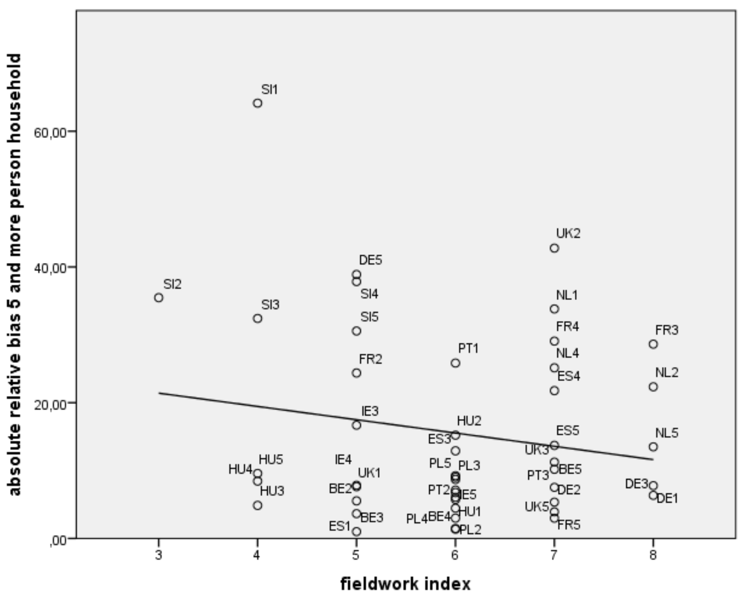

Figure 78. Absolute value for the relative bias for “five-or-more person household” and the fieldwork effort index.

23.11 Result

This section aims to answer the research question whether higher fieldwork effort leads to a lower nonresponse bias and thus higher data quality. The analysis was performed and the hypotheses were tested for 10 socio-demographic variables, such as gender, age, different levels of education, being in paid work, being married, non-nationals, and household size (see Table 19). Scatterplots were produced, and correlations were calculated.

A significant correlation was detected for four variables: “working population”, “low education”, “high education” and “nationality”.[182] For the variables “working population”, “high education”, and “nationality”, the relation is as stated in the hypotheses: More fieldwork effort is associated with less nonresponse bias.

Table 19

Results of variable specific analysis of fieldwork effort and nonresponse bias.

Result	coef	p	n	Hypotheses 12 supported
More fieldwork effort is associated with less nonresponse bias:				
Working population (relative bias)	- .20	.089+	74	Support of hypothesis.
High education (absolute bias)	- .22	.064+	69	Support of hypothesis.
Nationality (relative and absolute bias)	- .23	.090+	55	Support of hypothesis.
More fieldwork effort is associated with more bias:				
Low education (absolute bias)	- .23	.050**	74	No support of hypothesis. Contrary to of hypothesis.
No correlation between fieldwork effort and bias was found:				
Gender (female)	- .09	.429	74	No support of hypothesis.

[182] At the 10% significance level, a correlation was found for the relative nonresponse bias for the variable “working population” and for the absolute value of the nonresponse bias of the variables “low education” and “high education”.

Young persons (age 15–24 years)	- .11	.356	74	No support of hypothesis.
Old persons (75 years and older)	- .07	.595	64	No support of hypothesis.
Married persons	- .03	.798	73	No support of hypothesis.
one-person households	- .10	.501	47	No support of hypothesis.
five-or-more person households	- .18	.234	72	No support of hypothesis.

Note. Source: Fieldwork Efforts from the ESS; Nonresponse bias: comparison of the harmonized socio-demographic variable in the ESS to the LFS for the 16 countries that took part in all rounds of the ESS.
+ p < .1 *** $p < .01$. ** $p < .05$. * $p < .10$.

24 Discussion

In general, a significance test is used to draw conclusions from a sample to a population. In the present analyses, a census of all countries was examined. Thus, it can be discussed if the strategy of drawing conclusions from a sample to a population was used. However, at the same time, arguments for drawing conclusions from a sample to a population can be made because these analyses use the whole universe of countries that participated in all rounds of the ESS. So, a generalization from the present study to other surveys, countries, and rounds is possible. This argumentation allows interpreting significance tests.

Analyzing the pattern for nonresponse bias only, without only interpreting significant results Hypotheses 12 can be supported. Variables that are more related to the at-home pattern and therefor related to contactability seem to be more affected by more fieldwork effort: females, being in paid work, young respondents, non-nationals, and persons living in five-or-more person households. For these variables, the following hypothesis can be supported: The higher the fieldwork effort is, the lower the nonresponse bias. For the groups of respondents who are known to be underrepresented because they are more likely to refuse, the situation differs. For persons that are older (75 years and older) or persons with low education, the bias increased with higher fieldwork effort. However, the analysis also showed that, in general, for the variables that are more affected by contactability, more fieldwork effort decreases the nonresponse bias and improves data quality. For characteristics that are more related to refusal (such as age and low education), higher fieldwork effort does not increase data quality. But this conclusion is based on patterns and not on significant results.

Nonresponse bias is country specific because the variables that are correlated with refusal and contactability may vary with respect to countries. For example, depending on the family structure, older respondents may live in larger family households, which means that the household sizes increases and may lead to increased contactability. Further analyses may be relevant for each country to detect the variables that are related to contactability and refusal. An analysis at the country level may help to target fieldwork efforts to those cases that might have the largest nonresponse bias.

Further analysis, for example a multi-level model would be interesting in order to analyze separate the country effects. However, due to the limited sample size (five points in time and 16 countries) the number of cases is too low. To further analyze the relationship between a sample and a population, larger sample sizes would be necessary.

25 Results for Analysis Part D

The effect of process (fieldwork efforts) on output quality indicators (nonresponse bias) were further investigated in Analysis Part D. Hypotheses on the relation of fieldwork efforts and nonresponse bias were tested and analyzed.

The analysis of the relationship of fieldwork effort (as an index) and an additive index of bias of all the variables showed a negative, but not significant, correlation (see Figure 54). At the variable level only for some variables the hypotheses could be supported. For the variables “working population”, “high education” and “nationality” the relation was as stated in the hypothesis: High fieldwork efforts are correlated with a lower nonresponse bias (see Table 19).

This finding enabled a drawing of the conclusion that variables are affected differently by nonresponse bias, which may be due to the fact that different reasons exist for nonresponse. For example, more fieldwork effort may affect the nonresponse that is due to refusal differently than the nonresponse that is due to non-contact. It can be concluded that for the variables that are related in general to the at-home pattern and contactability, additional fieldwork effort did not lead to a decreased nonresponse bias (gender, young respondents, married respondents, high education and household size). This is only the case two for the variables “working population” and “nationality”. For the variables that are more related to refusal (older respondents, education), significant effects were found only for “education”. The variable “education” is especially interesting, since two different effects for the different levels of education are found. More fieldwork effort was associated with less bias for the variable “high education”. And in contrast with more fieldwork effort, the bias for “low education” increased.

To answer the general research question of whether high fieldwork efforts is correlated with lower nonresponse bias, results can be found that show the assumed negative correlation. But at the same time, results also show a positive or no correlation at all. As can be seen in the scatterplots (Figure 59 to Figure 78), the extent of the nonresponse bias differs between the variables and countries. In some cases, the nonresponse bias is rather large, whereas in other cases, it is relatively small, which leads to the conclusion that effects of nonresponse bias is not specific for one survey, but varies over variables within one survey and also varies over time and between countries.

V. Conclusion, Discussion and Outlook

This study provides fundamental research in the field of survey methodology. It makes several contributions to the field of measuring the data quality of cross-national surveys. In particular, this study will help to further an understanding of the challenges related to the quality indicators of fieldwork in face-to-face surveys. This study also provides background for researchers and practitioners interested in the process quality of surveys. The findings of large flagship methodologically rigorous cross-national surveys—such as the ESS—can provide a broad empirical background for improving data quality in cross-national and national survey contexts.

This study analyzed the implicit assumptions prevailing in survey research related to the correlation of the often entangled quality indicators—response rates, fieldwork efforts, and nonresponse bias in 16 European countries. This study also analyzed the development of response rates. A special focus was on the effects of survey-specific characteristics and fieldwork processes (meso level). This study used metadata and paradata from the European Social Survey (ESS) and the European Labour Force Survey (LFS), the latter of which was used as a reference statistic for the nonresponse bias calculation. The analyzed time span was from 2002 to 2010.

Measurements of fieldwork efforts are rare, especially in cross-national and longitudinal analyses. Often, the processes of fieldwork are unknown or not adequately documented. The present study extends the previous findings and assumptions of survey research on fieldwork processes by including relevant cross-national and cross-time data. Thus, the theoretical focus is extended and supplemented by relevant factors that have not been considered systematically before.

This analysis extends and further develops a basic model of survey research by systematically analyzing the quality indicators response rate, fieldwork effort, and nonresponse bias. This study also adds to the discussions about data quality in cross-national surveys.

Thus, this study makes four major contributions to the field of survey research. *The first is an overview of response rate development* (Analysis Part A). This analysis shows that a significant decrease in the mean response rate was not detected in the ESS between 2002 and 2010. In some countries, the response rate increased, but in most countries, it decreased significantly, and in others, it stayed constant. Also, a significant increase was not found in the mean refusal rate or the mean non-contact rate. In some countries that participated in the ESS surveys, the refusal and non-contact rates decreased or increased significantly. Although in many countries, the response rate decreased, the present study analysis of the ESS data does not support the general assumption that response rates are decreasing in general.

This comparative analysis narrows the research gap because the analysis is based on the same survey with its constant topics and survey set-up over time and across countries. Therefore, in this analysis, other factors that might have an effect on the response rate are kept constant. This analysis also provides comparative information with up-to-date data for the European context, whereas most existing comparative studies have focused mainly on the US context.

The second contribution of this study is a systematic analysis using comparative data that compares the relationship of response rates and nonresponse bias (Analysis Part B) *at a general and a country-specific level*. Thus, a conclusion can be drawn with respect to the prevailing hypotheses in the literature as to whether high response rates lower the risk of nonresponse bias. The present study analysis was conducted on a variety of socio-demographic variables (*gender, education, occupation, nationality, household size*, and *family status*) across time (between 2002 and 2010) and for 16 countries. A pooled cross-sectional analysis as well as a regression showed that, in general, lower response rate levels are associated with a higher nonresponse bias level. The Pearson Correlation Coefficient of the variable response rate and the (absolute value of the) nonresponse bias showed a small negative but significant strength of association. Thus, the hypothesis found in the literature that higher response rates are associated with lower nonresponse bias in general is supported by the data.

Further analyses at the variable level showed that the relationship of the response rate and nonresponse bias did not apply to all variables. For the majority of the variables (*old persons*, *married persons*, *persons with low education*, *persons with high education, nationals of the country,* and *persons living in one-person households*), the relation of response rates and nonresponse bias was as assumed—a higher response rate is indeed correlated with a lower nonresponse bias. However, for the variables *gender (male)* and *persons living in a five-and-more person household*, a significant, reversed correlation was detected. Against the assumed relation for these two variables, a higher response rate was correlated with a higher nonresponse bias. Also, the level of nonresponse bias differed between the variables. For the variables *gender (male)* and *nationals of the country,* the level of nonresponse bias was rather low. For the variables *older persons*, *education (low and high),* and *household size (one-person household and five-and-more person household,* the overall nonresponse bias was very large.

Descriptive analyses of different countries showed different patterns in the development of response rates and nonresponse bias. It was found that nonresponse bias at the variable level differed between the countries. Therefore, the present study showed that country-specific circumstances need to be considered when analyzing the nonresponse bias.

These findings contribute to the existing literature. The level as well as the pattern of nonresponse bias was found to be variable-specific and differed across variables within in a survey. In addition, the nonresponse bias of different variables differs not only within a survey, but also between the different countries in a survey. In other words, nonresponse bias is a variable and also a country-specific issue.

The third contribution to the existing literature is the analysis of the relationship between fieldwork efforts and response rates (Analysis Part C). The general hypothesis prevailing in the survey literature is that the higher the fieldwork efforts of a survey are, the higher the response rates. The present study provides an empirical background to test this theoretical assumption of survey methodology. Due to the lack of data available to previous studies in the literature, an empirical testing of this hypothesis was very limited. The ESS provides the missing data and therefore the opportunity for a systematic analysis of the effects of fieldwork efforts on response rates.

To comprehensively analyze the relationship between fieldwork efforts and response rates, an analysis was conducted at four different levels. First, the present study used a cross-sectional analysis to examine the general relationship between separate fieldwork efforts (the use of advance letters, brochures, and incentives; length of the fieldwork; length of the interviewer briefing; personal briefing of interviewers; payment of interviewers; interviewer training; and interviewer experience) and the response rate. This analysis found that the hypothesis that single fieldwork efforts in general increase the response rate was not supported by the data. Specifically, the data shows that the uses of certain fieldwork efforts per se do not increase the response rate or decrease the nonresponse rate. Second, a fieldwork effort index (FEI) was constructed, and the effects of this additive index on the response rate was analyzed. This analysis showed the same results—no significant correlation. The third level of analysis utilized a longitudinal perspective. The focus of this analysis was on the change in the fieldwork effort and the change in the response rate between the rounds of the ESS within countries. Contrary to expectations, a change in fieldwork efforts did not have a positive effect on the response rate. These three levels of analysis showed only very limited support for the assumed hypothesis on the relation between fieldwork efforts and response rates. However, a qualitative analysis (forth level of analysis) at the country level showed positive correlations between the response rate and fieldwork efforts for some countries and some variables. Even though high fieldwork efforts are not correlated to high response rates in general, descriptive analysis at the country level showed that implementing certain fieldwork efforts positively affects the response rate.

Although these findings lead to a conclusion that fieldwork efforts and response rates or nonresponse rates are not correlated in general, positive effects can be detected and need to be further analyzed at the variable and country levels. This finding is especially relevant for implementing a survey or cross-national survey. Thus, the findings of this study suggest that future researchers should carefully collect and analyze information about fieldwork processes at the country level. The implementation of a simple “one-size-fits-all-strategy” for fieldwork with the goal to enhance response rates in all countries or surveys does not seem adequate.

The fourth contribution to the existing literature is the analysis of the relation of fieldwork efforts and nonresponse bias (Analysis Part D). The current literature implicitly assumes that a high fieldwork effort is correlated to low nonresponse bias. However, these assumptions have not been tested empirically, mainly because information on fieldwork efforts usually is not available publicly, and often is not available in a systematic and comparable manner. The ESS data on fieldwork efforts provides such an opportunity. Also, the possibility to calculate the *nonresponse bias*, defined as the deviation of the ESS from the “gold standard” LFS, provides an opportunity to narrow this research gap. The analysis of fieldwork efforts and nonresponse bias showed a negative, non-significant correlation between the fieldwork effort index and the nonresponse bias index. Thus, more fieldwork efforts are not correlated with higher data quality (measured by a lower nonresponse bias) for all socio-demographic variables in general. However the analysis at the variables level showed significant positive effects of some variables, such as *working population, high education,* and *nationality*. For these variables, a lower nonresponse bias—and therefore a better representation of survey respondents in relation to population characteristics—was correlated with higher fieldwork efforts.

Taking into account the pattern of variables, including the non-significant results, showed that for the variables related to a better contactability of potential respondents (*females, being in paid work, young respondents, non-nationals,* and *persons living in five-or-more person households*), more fieldwork efforts decrease the nonresponse bias and thus improve data quality. For the groups of respondents who are known to be underrepresented because they are more likely to refuse, the situation differs. For persons who are older (75 years and older) or persons with low education, the nonresponse bias increased with higher fieldwork efforts. Reasons for nonresponse (refusal and non-contact) play an important role. The findings of the present study suggest that fieldwork efforts have country-specific and variable-specific effects on nonresponse bias. Therefore, the effects of fieldwork efforts on nonresponse bias need to be considered and discussed at the country and variable levels, and thus, fieldwork efforts need to be tailored for each country.

Knowledge about *country specific* underrepresentation and overrepresentation regarding different socio-demographic and substantial variables is essential. Also knowledge about the effects of fieldwork efforts and nonresponse bias at the country level is important. This tailoring should be done using quantitative analyses and qualitative discussions with fieldwork experts for different countries. On this basis, fieldwork efforts need to be tailored according to the survey topic and country-specific circumstances.

Lessons learned from this analysis are that more data on fieldwork is needed. This kind of information about the fieldwork process should be made publicly available to enable further analysis. The provision of metadata on fieldwork processes and the analysis and monitoring of fieldwork processes need to be central concerns in future studies. This approach could involve cooperation between survey organizations, survey methodologists, and substantive specialists as a way to learn more about the country-specific aspects of data collection, and to harmonize information about the data collection process. For reasons related to comparability, the documentation of fieldwork efforts needs to be standardized (similar to the AAPOR response rate definition).

A further contribution of the present study, in addition to its contribution to the survey literature, is the provision of a theoretical background for *survey practitioners* working in a survey agency or conducting a survey. Findings from PIAAC (2010) and ESS have shown that high quality surveys cannot be created using a single strategy; rather, a bundle of different efforts are required.

Challenges for the future include reaching out to the users and producers of cross-national survey data to provide them with information about process quality, especially with respect to fieldwork. The analysis and use of fieldwork efforts should be more prominent in further discussions about response rates, nonresponse bias, and data quality. Making fieldwork efforts measurable and including the quality of fieldwork in discussions about nonresponse bias and response rates are important advances for the survey community. The data collection process is a black box for researchers, and so a further emphasis on the process of data collection is important when analyzing data quality. With respect to the comparability of countries, the metadata analysis of response rates, fieldwork efforts, and nonresponse bias will become increasingly important. As the demand increases for high-quality surveys, discussions about data quality will become more central—for example, when researchers have to justify the use of high cost and high quality face-to-face probability surveys in contrast to low cost online non-probability access panel surveys.

The *limitations of this study* are related to the limited sample size, namely 16 countries that participated in all rounds of the ESS 1 to 5. Thus, the sample size of 80 cases limits the power of the analysis, may cause problems when calculating a factor analysis, and is problematic with respect to further analysis such as a multi-level analysis. Also the interdependency of the cases of the longitudinal analysis may limit the findings. The interdependency across time may bias the results, since we cannot rule out a correlation between the countries in each round (e.g., Germany in Round 1 and Round 2). Second, the cross-sectional interdependency may be less of a problem, although latent cluster effects may exist (Southern vs. Northern Europe). Both problems lead to deflated standard errors. The robust standard errors are a conservative (high) standard estimation that enables a regression analysis when the OLS assumption of autocorrelation and homoscedasticity are violated. Since the observations are independent, this calculated estimator is robust for these types of misspecifications. The present study calculations take these issues into account, and therefore provide robust standard errors or a sandwich estimator of variance. However, due to the limited sample size and the selectivity of the sample, the validity of the results may be limited, and so their application to other surveys needs to be carefully interpreted.

It can be concluded that the fieldwork process needs to become a central aspect within the survey lifecycle and the discussions of data quality. Additional monitoring tools and the use of paradata and process data with respect to fieldwork processes need to become more prominent when planning or evaluating of a study. This focus on fieldwork processes as central aspects of data quality would lead to the following general aim of survey methodology—to ensure and provide high quality quantitative data that enables representative and meaningful substantive analyses.

VI. Literature

Aday, L. (1996). *Designing and conducting health surveys.* San Francisco, CA: Jossey-Bass.

Alreck, P. L., & Settle, R. B. (1995). *The survey research handbook: Guidelines and strategies for conducting a survey* (2nd ed.). Chicago, IL: Chicago : Irwin.

Allen, M. P. J. (1997). *Understanding regression analysis.* New York: Plenum Press.

American Association for Public Opinion Research. (2015). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (8th ed.). Retrieved from https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/Standard-Definitions2015_8thEd.pdf.

American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (9th ed.). Retrieved from http://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf.

Andridge, R. R., & Little, R. J. A. (2011). Proxy pattern-mixture analysis for survey nonresponse. *Journal of Official Statistics*, *27*(2), 153–180.

Babbie, E. R. (1990). *Survey research methods.* Belmont, CA: Wadsworth.

Babbie, E. R. (2007). *The Practice of Social Research.* Belmont, CA: *Wadsworth.*

Backstrom, C. H., & Hursh, G. (1963). *Survey research*. Evanston, IL: Northwestern University Press.

Barbier, S., Loosveldt, G., & Carton, A. (2016). Measuring the survey climate: The Flemish case. Retrieved from http://www.surveyinsights.org/?p=7430.

Bethlehem, J. (2002). Weighting nonresponse adjustments based on auxiliary information. In M. R. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 275-288). New York, NY: Wiley.

Biemer, P. P. (2010). Total survey error: Design, implementation, and evaluation. *Public Opinion Quarterly, 74*(5), 817–848. doi: 10.1093/poq/nfq058

Biemer, P. P., & Lyberg, L. E. (2003). *Introduction to survey quality*. Hoboken, NJ: John Wiley & Sons, Inc.

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

Blohm, M. (2006). Datenqualität durch Stichprobenverfahren bei der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften-ALLBUS. In F. Faulbaum & C. Wolf (Eds.), *Stichprobenqualität in Bevölkerungsumfragen*, Vol. 12 (pp. 37–54). Bonn: IZ InformationsZentrum Sozialwissenschaften.

Blohm, M., & Diehl, C. (2001). Wenn Migranten Migranten befragen. Zum Teilnahmeverhalten von Einwanderern bei Bevölkerungsbefragungen. *Zeitschrift für Soziologie, 30*(3), 223–424.

Blohm, M., Hox, J., & Koch, A. (2007). The influence of interviewers' contact behavior on the contact and cooperation rate in face-to-face household surveys. *International Journal of Public Opinion Research, 19*, 97–111.

Blom, A. G. (2009). *Measuring, explaining and adjusting for cross-country differences in unit nonresponse: What can process data contribute?* (Doctoral dissertation). Retrieved from http://survex.org/fileadmin/user_upload/PhD_thesis_Annelies_Blom.pdf.

Blom, A. G., Jäckle, A., & Lynn, P. (2010). The use of contact data in understanding cross-national differences in unit nonresponse. In J. A. Harkness, M. Braun, B. Edwards, T. P. Johnson, L. E. Lyberg, P. Ph. Mohler, B-E. Pennell, T. W. Smith (Eds.), *Survey methods in multinational, multiregional, and multicultural context* (pp. 335–372): John Wiley & Sons, Inc.

Blom, A. G., Lynn, P., & Jäckle, A. (2008). Understanding cross-national differences in unit non-response: The role of contact data. *ISER Working Paper Series*, *2008-01*. Retrieved from https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2008-01.pdf.

Brehm, J. (1993). *The phantom respondents: Opinion surveys and political representation.* Ann Harbor, MI: University of Michigan Press.

Brehm, J. (1994). Stubbing our toes for a foot in the door? Prior contact, incentives and survey response improving response to surveys. *International Public Opinion Research*, *6*(1), 45–64.

Brick, M. J., & Williams, D. (2013). Explaining rising nonresponse rates in cross-sectional surveys. *The ANNALS of the American Academy of Political and Social Science*, *645*, 36–59. doi:10.1177/0002716212456834

Bronner, A. E. (1988). Surveying ethnic minorities. In W. E. Saris & I. N. Gallhofer (Eds.), *Sociometric research,* Vol. 1 (pp. 36–47). London, UK: Macmillian.

Brüderl, J., & Ludwig, V. (2015). Fixed-effects panel regression. In H. Best & C. Wolf (Eds.), *Regression analysis in causal inference* (pp. 327–357). London, UK: Sage.

Callegaro, M.; Baker, R.P.; Bethlehem, J.G. Göritz, A.S. Krosnick, J.A. & Lavrakas, P.J. (2014). *Online Panel Reserch: A Data Quality Perspective*. West Sussex: Wiley.

Campanelli,P.; Sturgis, P. & Purdon, S. (1997). Can you hear me knocking: An investigation into the impoact of interviewers on survey response rates. London: The Survey Method Centre at SCPR. Retrieved from http://eprints.soton.ac.uk/80198/.

CASRO. (1982). *On the definition of response rates*. Port Jefferson, NY: CASRO. Retrieved from https://c.ymcdn.com/sites/www.casro.org/resource/resmgr/docs/casro_on_definitions_of_resp.pdf.

Cohan, G. & Duffy, J.C. (2002). Are nonrespondents to health surveys less healthy than respondents? *Journal of Official Statistics*, *18*(1), 13-23.

Comrey, A. L., & Lee, H. B. (1992). *A First course in factor analysis*, Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Couper, M.P. (2013). Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys. *Survey Research Methods*, 7(3), 145:56. doi: http://dx.doi.org/10.18148/srm/2013.v7i3.5751

Couper, M. P., & de Leeuw, E. (2003). Nonresponse in cross-cultural and cross-national surveys. In J. A. Harkness, F. J. R. Van de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 157–177). Hoboken, NJ: John Wiley & Sons.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly, 64*(4), 413–428. doi: 10.1086/318638.

Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly, 69*(1), 87–98. doi: 10.1093/poq/nfi002

de Heer, W. (1999). International response trends: Results of an international survey. *Journal of Official Statistics*, *15*(2), 129–144.

de Leeuw, E., & de Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillmann, J. L. Sltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 41–54). Hoboken, NJ: John Wiley & Sons.

de Leeuw, E., Callegaro, M; Hox; Korendijk, E.; Lensvelt-Mulders, G. (2007). The influence of advance letters on response in telepone survey: A meta-analysis. *Public Opinion Quarterly, 71*(3), 414–443. doi: 10.1093/poq/nfm014

Demographic and Health Surveys. (2016). Malaria indicator survey. Guidelines for the malaria indicator survey interviewer training. In ICF International (Ed.). Rockville, MD: ICF International. Retrieved from http://malariasurveys.org/toolkitfiles/08%20%20MIS%20Guidelines%20for%20Interviewer%20Training.pdf.

Dillman, D. A. (1978). *Mail and telephone surveys: The total design method.* New York, NY: Wiley.

Dixon, J., & Tucker, C. (2010). Survey nonresponse. In P.V. Marsden & J. D. Wright (Eds.), *Handbook of survey research* (pp. 593–630). Bingley, UK: Emerald.

Duncan, O. D., & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review, 20*(2), 210–217.

Durrant, G.B. (2016, July). *Assessing risk of nonresponse bias and dataset representativeness during survey data collection.* Paper presented at the ESRC Method Festival, Bath, United Kingdom. Retrieved from: http://www.ncrm.ac.uk/RMF2016/uploads/A1/86/77/Durrant_presentation_RMF_2016_3.pdf.

Durrant, G. B., & Steele, F. (2009). Multilevel modelling of refusal and noncontact in house-hold surveys: Evidence from six UK government surveys. *Journal of the Royal Statistical Society, 172*(2), 361–381.

Dutwin, D.; Loft, J. D.; Darling, J.; Holbrook, A.; Johnson, T.; Langley, R.E.; Lavrakas, P. J.; Olson, K.; Peytcheva, E.; Stec, J.; Triplett, T.; Zukerberg, A. (2014). AAPOR: *Current Knowledge and Considerations Regarding Survey Refusals*. Retrieved from: https://www.aapor.org/AAPOR_Main/media/MainSiteFiles/RefusalTF_FINAL090814.pdf

Elfff, M. (2015). Estimation techniques: Ordinary least square and maximum likelyhood. In H. Best & C. Wolf (Eds.), Regression analysis and causal inference (pp. 7-30). London: Sage.

European Social Survey. (2009). *Round 5 Specification for participating countries*. London, UK: Centre for Comparative Social Surveys, City University London.

European Social Survey. (2011). *Round 6 specification for participating countries*. London, UK: Centre for Comparative Social Surveys, City University London.

European Social Survey. (2013a). *Round 7 specification for participating countries*. London, UK: Centre of Comparative Social Surveys, City University London.

European Social Survey. (2013). *European Social Survey* [data set]. Retrieved from http://www.europeansocialsurvey.org/data/.

European Social Survey. (2014a). *ESS1 - 2002 documentation report* (6.4 ed.). Retrieved from http://www.europeansocialsurvey.org/docs/round1/survey/ESS1_data_documentation_report_e06_4.pdf.

European Social Survey. (2014b). *ESS2 - 2004 documentation report* (3.5 ed.). Retrieved from http://www.europeansocialsurvey.org/docs/round2/survey/ESS2_data_documentation_report_e03_5.pdf.

European Social Survey. (2014c). *ESS3 - 2006 documentation report* (3.5 ed.). Retrieved from http://www.europeansocialsurvey.org/docs/round3/survey/ESS3_data_documentation_report_e03_5.pdf.

European Social Survey. (2014d). *ESS4 - 2008 documenation report* (5.3 ed.). Retrieved from http://www.europeansocialsurvey.org/docs/round4/survey/ESS4_data_documentation_report_e05_3.pdf.

European Social Survey. (2014e). *ESS5 - 2010 documenation report* (4.0 ed.). Retrieved from http://www.europeansocialsurvey.org/docs/round5/survey/ESS5_data_documentation_report_e04_0.pdf.

European Social Survey. (2016a). *ESS6 - 2012 documenation report* (2.2 ed.). Retrieved from http://www.europeansocialsurvey.org/docs/round6/survey/ESS6_data_documentation_report_e02_2.pdf.

European Social Survey. (2017). About the European Social Survey European Research Infrastructure – ESS ERIC. Retrieved from http://www.europeansocialsurvey.org/about/.

European Union Labour Force Survey. (2016). European Union Labour Force Survey. Retrieved from http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey.

Feskens, R., Hox, J. J., Lensvelt-Mulders, G., & Schmeets, H. (2007). Nonresponse among ethnic minorities: A multivariate analysis. *Journal of Official Statistics, 23*(3), 387–408.

Fortin-Rittberger, J. (2015). Time-series cross-section. In H. Best & C. Wolf (Eds.), *Regression analysis and causal inference* (pp. 387-408). London: Sage.

Foster, K., & Bushnell, D. (1994, September). *Non-response bias on government surveys in Great Britan*. Paper presented at the 5th International Workshop on Household Survey Nonresponse, Ottawa, Ontario.

Freund, R. J. & Wilson, W.J. (1998). *Regression analysis: Statistical Modeling of a response variable*. San Diego, CA: Academic Press.

Fricker, S., & Tourangeau, R. (2010). Examining the realtionship between nonresponse propensity and data quality in two national household surveys. *Public Opinion Quarterly, 74*(5), 934–955. doi: 10.1093/poq/nfq064

Gengler, J. J. (2016, May). *Assessing survey attitudes in the Middle East and Arab Gulf*. Paper presented at the WAPOR World Association for Public Opinion Reserach, Austin, Texas, USA.

Goyder, H. (1987). *The silent minority: Nonrespondetnts on sample survey*. Bolder, CO: Westview Press.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70*(5), 646–675. doi: 10.1093/poq/nfl033

Groves, R. M. (2011). Three Eras of Survey Research. *Public Opinion Quarterly, 75*(5), 167-189. doi: https://doi.org/10.1093/poq/nfr057

Groves, R. M., Brick, J. M., Couper, M. P., Kalsbeek, W., Harris-Kojetin, B., Kreuter, F., Pennell, B. E., Raghunathan, T., Schouten, B.; Smith, T., Tourangeau, R., Bowers, A.; Jans, M., Kennedy, C., Levenstein, R., Olson, K., Peytcheva, E., Ziniel, S., Wagner, J. (2008). Issues facing the field: Alternative practical measures of representativeness of survey respondent pools. *Survey Practice*, *1*(3). Retrieved from http://www.surveypractice.org/index.php/SurveyPractice/article/view/221.

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly, 56*(4), 475–495. doi: 10.1086/269338

Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. Hoboken, NJ: John Wiley & Sons.

Groves, R. M., Couper, M. P., Presser, S., Singer, E., Tourangeau, R., Acosta, G. P., & Nelson, L. (2006). Experiments in producing nonresponse bias. *Public Opinion Quarterly, 70*(5), 720–736. doi: 10.1093/poq/nfl036

Groves, R. M., Dillman, D. A., Eltinge, J. L., & Little, R. J. A. (Eds.). (2002). *Survey nonresponse*. New York, NY: John Wiley & Sons.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., & Singer, E. (2004). *Survey methodology*. Hoboken, NJ: Wiley.

Groves, R. M., & Heeringa, S. G. (2006). Responsive design for household surveys: Tools for actively controlling survey errors and costs. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 169*(3), 439–457. doi: 10.1111/j.1467-985X.2006.00423

Groves, R. M., & Lyberg, L. (Eds.). (1988). *An overview of nonresponse issues in telephone surveys, in telephone survey methodology*. New York, NY: John Wiley & Sons.

Groves, R. M., & Lyberg, L. (2010). Total survey error: Past, present, and future. *Public Opinion Quarterly, 74*(5), 849–879. doi: 10.1093/poq/nfq065

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly, 72(2),* 167–189. doi: 10.1093/poq/nfn011

Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly, 68*(1), 2–31. doi: 10.1093/poq/nfh002

Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation. *Public Opinion Quarterly, 64*(3), 299–308.

Halbherr, V. (2017). *Harmonization of Labour Force Survey (LFS) and the European Social Survey (ESS). A guide for weighing and nonrepsonse bias analysis*. Manuscript in preparation.

Halbherr, V., & Koch, A. (2010). *Report on deviations from specifications (ESS 5)*. Mannheim: GESIS Leibniz Institute for the Social Science, European Social Survey.

Halbherr, V., Koch, A., Ederle, F., & Mayn, C. (2014). *Report on deviations from specifications (ESS 6) - 29 countries in second data release*. London, UK: European Social Survey ERIC.

Halbherr, V., Koch, A., Kappelhof, J., & Stoop, I. (2014). *Report on planned deviations for ESS 7*. London, UK: European Social Survey ERIC.

Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P ., Pennell B.-E., & Smith, T. W. (Eds.). (2010). *Survey methods in multinational, multiregional, and multicultural contexts*. Hobocken, NJ: John Wiley & Sons, Inc.

Harris-Kojetin, B., & Tucker, C. (1999). Exploring the relation of economic and political conditions with refusal rate to a government survey. *Journal of Official Statistics, 15(2),* 167–184.

Hartmann, P., & Schimpl-Neimanns, B. (1992). Sind Sozialstrukturanalysen mit Umfragedaten möglich? Analysen zur Repräsentativität einer Sozialforschungsumfage. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 44*, 315–346. Retrieved from http://www.gesis.org/fileadmin/upload/forschung/programme_projekte/sozialwissenschaften/Amtliche_Mikrodaten/AB92_01.pdf.

Heeringa, S.G, West, B. & Berglund P. A. 2015). Regression with complex samples. In H. Best & C. Wolf (Eds.), *Regression analysis and causal inference* (pp. 225-248). London: Sage.

Helmschrott, S., & Martin, S. (2014). Nonresponse in PIAAC Germany. *Methods, Data, Analyses, 8*(2), 243–266. doi: 10.12758/mda.2014.010

Holbrook, A. L., Green, M., & Krosnick, J. A. (2003). Telephone vs. face-to-face interviewing of national probabilty samples with long questionnaires: Comparison of respondents satisficing and social desirability response bias. *Public Opinion Quarterly*, *67*(1), 79-125. doi: 10.1086/346010

Holbrook, A. L., & Krosnick, J. A. (2010). Measuring voter turnout by using the randomiszed response technique evidence calling into question the method's validity. *Public Opinion Quarterly, 74*(2), 328–343. doi: 10.1093/poq/nfq012

Holbrook, A. L., Krosnick, J. A., & Pfent, A. (2008). The causes and consequences of response rates in surveys by the news media and government contractor survey research firms. In J. M. Lepkowski, N. C. Tucker, J. M. Brick, E. D. De Leeuw, L. Japec, P. J. Lavrakas, M. W. Link, & R. L. Sangster (Eds.), *Advances in telephone survey methodology* (pp. 499-528). New York, NY: Wiley.

Hox, J. J., de Leeuw, E. D., & Vorst, H. (1996). A reasoned action explanation for survey nonresponse. In S. Laaksonen (Ed.), *International perspectives on nonresponse* (pp. 101–110). Helsinki: Statistics Finland.

Huffington, A. (1998, October 14). Margin of arrogance is huge for pollsters. *Chicago Sun-Times*, p. 43.

International Workshop on Household Survey Nonresponse. (2017). International Workshop on Household Survey Nonresponse. Retrieved from http://www.nonresponse.org/.

Jann, B. (2009). Diagnostik von Regressionsschätzungen bei kleinen Stichproben (mit einem Exkurs zu logistischer Regression). In P. Kriwy & C. Gross (Eds.), *Klein aber fein!: Quantitative empirische Sozialforschung mit kleinen Fallzahlen* (pp. 93–125). Wiesbaden: VS Verlag für Sozialwissenschaften.

Johnson, T. P., & Owens, L. (2003, August). *Survey response rates reporting in the professional literature*. Paper presented at the Annual Meeting of the American Statistical Association ASA, San Franciso, California.

Jowell, R. (2010). *ESS5 – European social survey round 5 year 1 central coordination-monitoring social and political change in Europe.* Retrieved from http://www.europeansocialsurvey.org/docs/about/ESS5_end_of_grant_report.pdf.

Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64*(2), 125–148. doi: 10.1086/317759

Koch, A. (1997). Teilnahmeverhalten beim Allbus 1994: Soziodemographische Determinanten von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 49*(1), 99–122.

Koch, A. (1993). Sozialer Wandel als Artefakt unterschiedlicher Ausschöpfungen? Zum Einfluss von Veränderungen der Ausschöpfungsquoten auf die Zeitreihen des ALLBUS*, ZUMA-Nachrichten, 33, 43-113.*

Koch, A. (1998). “Wenn Mehr“ nicht gleichbedeutend mit “ Besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrung in Allgemeinen Bevölkerungsumfragen. *ZUMA-Nachrichten, 42,* 66–90.

Koch, A. (2006). *Deviations from the specifications for participating countries in round 1 of the ESS.* Mannheim: GESIS Leibniz Institute for the Social Science, European Social Survey.

Koch, A. (2007). *Deviations from the specifications for participating countries in round 2 of the ESS*. Mannheim: GESIS Leibniz Institute for the Social Science, European Social Survey.

Koch, A. (2010). *Compliance with the "specification for participating countries" in round 4 of the ESS*. Mannheim: GESIS Leibniz Institute for the Social Science, European Social Survey.

Koch, A., Blom, A. G., Stoop, I., & Kappelhof, J. (2009). Data collection quality assurance in cross-national surveys: The example of the ESS. *Methoden - Daten - Analysen*, *3*(2), 219–247.

Koch, A., Fitzgerald, R., Stoop, I., Widdop, S., & Halbherr, V. (2012). *Field procedures in the European Social Survey round 6: Enhancing response rates*. Mannheim: European Social Survey.

Koch, A., Halbherr, V., Stoop, I., & Kappelhof, J. (2014). *Assessing ESS sample quality by using external and internal criteria.* Mannheim: GESIS Leibniz Institute for the Social Science, European Social Survey.

Kohler, U. (2007). Surveys from inside: An assessment of unit nonresponse bias with internal criteria. *Survey Research Methods*, *1*(2), 55–67.

Kortmann, K., & Halbherr, V. (2009). *Alterssicherung in Deutschland 2007 (ASID '07)*. Bundesministerium für Arbeit und Soziales (Ed.), *Forschungsbericht* (Vol. 431/1). München: Bundesministerium für Arbeit und Soziales. Retrieved from https://www.tns-infratest.com/sofo/_pdf/ASID_2007_ZB_2009_02_16.pdf.

Kreuter, F. (2013a). Facing the nonresponse challenge. *The ANNALS of the American Academy of Political and Social Science*, *645*(1), 23–35 doi: doi: 10.1177/0002716212456815

Kreuter, F. (2013a). *Improving surveys with paradata: Analytic uses of process information*. Hoboken, NJ: John Wiley & Sons.

Kreuter, F., & Casas-Cordero, C. (2010). Paradata. *RatSWD Working Paper Series, No. 136*.

Kreuter, F.; Müller, G. & Trappmann, M. (2010). Nonresponse and Measurement Error in Employment Research: Making Use of Adminstrative Data. *Public Opinion Quarterly, 74*(5), 880-906. doi: https://doi.org/10.1093/poq/nfq060

Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T. M., Casas-Cordero, C., Lemay, M., Peytchev, A., Groves, R. M., & Raghunathan, T. E. (2010). Using proxy measures and other correlates of survey outcomes to adjust for non-response: Examples from multiple surveys. *Journal of the Royal Statistical Society Series a—Statistics in Society, 173,* 389–407.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*(3), 213–236. doi: 10.1002/acp.2350050305

Krosnick, J. A., Holbrook, A. L., Berent, M. K., Carson, R. T., Hanemann, W. M., Kopp, R. J., Mitchell, R.C., Presser, S., Ruud, P. A., Smith, K., Moody, W. R., Green, M. C., & Conaway, M. (2002). The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quarterly, 66*(3), 371–403. doi:10.1086/341394

Krueger, B. S., & West, B. T. (2014). Assessing the potential of paradata and other auxiliary data for nonresponse adjustments. *Public Opinion Quarterly, 78*(4), 795–831. doi:10.1093/poq/nfu040

Lacy, M. G. (1997). Efficiently studying rare events: Case-control methods for sociologists. *Sociological Perspectives, 40*(1), 129–154.

Lelkes, Y., Krosnick, J. A., Marx, D. M., Judd, C. M., & Park, B. (2012). Complete anonymity compromises the accuracy of self-reports. *Journal of Experimental Social Psychology, 48*(6), 1291–1299. doi: 10.1016/j.jesp.2012.07.002

Levy, P. S., & Lemeshow, S. (2013). *Sampling of populations: Methods and applications,* Vol. 4. New York, NY: Wiley.

Little, R. J. (1986). Survey nonresponse adjustments for estimates of means. *International Statistical Review/Revue Internationale de Statistique, 54*(2), 139–157. doi:10.2307/1403140

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Little, R. J., & Vartivarian, S. (2005). Does weighting for nonresponse increase the variance of survey means? *Survey Methodolgy, 31*(2), 3897-3904.

Luiten, A. (2013). *Improving survey fieldwork with paradata* (Doctoral dissertation). Retrieved from https://www.cbs.nl/nl-nl/achtergrond/2013/44/improving-survey-fieldwork-with-paradata.

Lyberg, L., & Dean, P. (1992). *Automated coding for survey responses: An international review*. Stockholm: Statistics Sweden.

Lynn, P. (2003). PEDAKSI: Methodology for collecting data about survey non-respondents. *Quality and Quantity, 37*(1), 239–261.

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

Lynn, P., Clarke, P., Martin, J., & Sturgis, P. (2002). The effects of extended interviewer efforts on nonresponse bias. In R. M. Groves, D. Dillman, J. Eltinge, & R. J. Little (Eds.), *Survey nonresponse* (pp. 135–148). New York, NY: Wiley.

Massey, D., & Roger, T. (2013). Introductions: New challenges to social measurement. *The ANNALS of the American Academy of Political and Social Science*, *645*(1).

Matsuo, H., Billiet, J. L., Loosveldt, G. & Malnar, B. (2010). *Response based quality assessment of ESS round 4: Results for 30 countries based on contact files*. Leuven: European Social Survey, University of Leuven.

Matsuo, H., & Loosveldt, G. (2013). *Report on quality assessment of contact data files in round 5: Final report 27 countries*. Leuven: European Social Survey, University of Leuven.

Merkle, D. M., & Edelman, M. (2002). Nonresponse in exit polls: A comprehensive analysis. In R. M. Groves, D. Dillman, J. Eltinge, & R. J. Little (Eds.), *Survey Nonresponse* (pp. 243–257). New York, NY: Wiley.

Meulemann, B, Loosveldt, G. & Emonds, V. (2015). Assumptions and diagnostics. In H. Best & C. Wolf (Eds.), *Regression analysis and causal inference* (pp. 83-110). London: Sage.

Organisation for Economic Co-operation and Development. (1999). *Classifying educational programmes: Manual for ISCED-97 implementation in OECD countries* (1999 ed.). Retrieved from http://www.oecd.org/edu/1841854.pdf.

Organisation for Economic Co-operation and Development. (2010). The survey of adult skills (PIAAC). Retrieved from http://www.oecd.org/site/piaac/surveyofadultskills.htm.

Peytchev, A. (2013). Consequences of survey nonresponse. *The ANNALS of the American Academy of Political and Social Science, 645*(1), 88–111.

Peytcheva, E., & Groves, R. M. (2009). Using variation in response rates of demographic subgroups s evidence of nonresponse bias in survey estimates. *Journal of Official Statistics, 25*, 193–201.

Pforr, K., Blohm, M., Blom, A. G., Erdel, B., Felderer, B., Fräßdorf, M., Hajek, K., Helmschrott, S., Kleinert, C., Koch, A., Krieger, U., Kroh, M., Martin, S., Saßenroth D., Schmiedeberg, C., Trüdinger, E.-M., & Rammstedt, B. (2015). Are incentive effects on response rates and nonresponse bias in large-scale, face-to-face surveys generalizable to Germany? Evidence from ten experiments. *Public Opinion Quarterly, 79*(3), 740–768. doi: 10.1093/poq/nfv014

Programme for the Assessment of Adult Competencies. (2010). Reducing nonresponse bias and preliminary nonresponse bias analysis. *NRB Paper, 10*. Retrieved from https://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0ahUKEwjd8c_RhaPNAhXBDiwKHUNRD3wQFggoMAA&url=http%3A%2F%2Fwww.reginfo.gov%2Fpublic%2Fdo%2FDownloadDocument%3FobjectID%3D24511501&usg=AFQjCNFjsMlChWoscCanqKtYNy_EEakflQ&cad=rja.

Rea, L. M., & Parker, R. A. (1997). *Designing and conducting survey research: A comprehensive guide*. San Francisco, CA: Jossey–Bass.

Representativity Indicators for Survey Quality. (2017). Risq Project. Retrieved from http://www.risq-project.eu/indicators.html.

Sakshaug, J. W., & Kreuter, F. (2011). Using paradata and other auxiliary data to examine mode switch nonresponse in a "recruit-and-switch" telephone survey. *Journal of Official Statistics, 27*(2), 339–357.

Sakshaug, J. W., Yan, T., & Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly, 74*(5), 907–933. doi:10.1093/poq/nfq057

Schneider, S. L. (2008). The application of the ISCED-97 to the UK's educational qualifications. *The international standard classification of education (ISCED-97): An evaluation of content and criterion validity for 15 European Countries*. pp. 281–300.

Schneider, S. L., & Kogan, I. (2008). The international standard classification of education 1997: Challenges in the application to national data and the implementation in cross-national surveys. *The international standard classification of education (ISCED-97): An evaluation of content and criterion validity for 15 European countries*. pp. 13–46.

Schnell, R. (1991). Wer ist das Volk?: Zur faktischen Grundgesamtheit bei "allgemeinen Bevölkerungsumfragen"; Undercoverage, Schwererreichbare und Nichtbefragbare. *Kölner Zeitschrift für Soziologie und Sozialpsychologie, 43*(1), 106–154.

Schnell, R. (1997). *Nonresponse in Bevölkerungsumfragen: Ausmaß, Entwicklung und Ursachen*. Opladen: Leske + Budrich.

Schouten, B. (2004). *Adjustment for bias in the integrated survey on household living conditions (POLS) 1998*. Discussion paper 04001. Voorburg/Heelen: Statistics Netherlands.

Schouten, B., Bethlehem, J., Beullens, K., Kleven, Ø., Loosveldt, G., Luiten, A., Rutar, K., Shlomo, N., & Skinner, C. (2012). Evaluating, comparing, monitoring, and improving representativeness of survey response through R-Indicators and Partial R-Indicators. *International Statistical Review, 80*(3), 382–399. doi:10.1111/j.1751-5823.2012.00189

Schouten, B., & Cobben, F. (2007). *R-indexes for the comparison of different fieldword strategies and data collection modes*. Discussion paper 07002. Voorburg/Heerlen: Statistics Netherlands.

Schouten, B., Shlomo, N., Da Silva, D. & Skinner, C. (2016). *Estimation of response propensities and indicators of representative response using population-level information*. Machester, UK: The University of Manchester.

Schouten, B., Shlomo, N., & Skinner, C. (2011). Indicators for monitoring and improving representativeness of response. *Journal of Official Statistics, 27*(2), 231–253.

Schupp, J., & Wolf, C. (Eds.). (2015). *Nonresponse bias: Qualitätssicherung sozialwissenschaftlicher Umfragen*. Wiesbaden: Springer-Verlag.

Singer, E. (2006). Introduction—Nonresponse bias in household surveys. *Public Opinion Quarterly, 70*(5), 637–645. doi: 10.1093/poq/nfl034

Singer, E., & Ye, C. (2013). The use and effects of incentives in surveys *The ANNALS of the American Academy of Political and Social Science, 645*(1), 112–141.

Singleton, R. A., & Straits, B. C. (2005). *Approaches to social research* (4th ed.). New York, NY: Oxford University Press.

Sinibaldi, J. (2014). *Evaluating the quality of interviewer observed paradata for nonresponse applications* (Doctoral dissertation). Retrieved from https://edoc.ub.uni-muenchen.de/18656/1/Sinibaldi_Jennifer.pdf.

Smith, T. W. (1983). The hidden 25 percent: An analysis of nonresponse on the 1980 General Social Survey. *Public Opinion Quarterly, Vol. 47(3),* 386-404. doi: 10.1086/268797

Smith, T. W. (2002). Developing nonresponse standards. In R. M. Groves, D. Dillman, J. Eltinge, & R. J. Little (Eds.), *Survey nonresponse* (pp. 27–44). New York, NY: Wiley & Sons.

Sodeur, W. (1997). Interne Kriterien zur Beurteilung von Wahrscheinlichkeitsauswahlen. *ZA-Information/Zentralarchiv für Empirische Sozialforschung, 41*, 58–82.

Statistical Office of the European Communities (Eurostat). (2012). *Labour force survey in the EU, candidate and EFTA countries: Main characteristics of national surveys, 2012*. Retrieved from http://ec.europa.eu/eurostat/documents/3888793/5856085/KS-TC-13-003-EN.PDF/15feeb32-717d-4326-8e7b-da6e2530ed4c.

Statistical Office of the European Communities (Eurostat). (2014). *Quality report of the European Union Labour Force Survey, 2014*. Retrieved from http://ec.europa.eu/eurostat/documents/3888793/7018036/KS-TC-15-004-EN-N.pdf.

Statistical Office of the European Communities (Eurostat). (2017a). Description of the dataset. Retrieved from http://ec.europa.eu/eurostat/web/microdata/european-union-labour-force-survey.

Statistical Office of the European Communities (Eurostat). (2017b). Statistics explained. Retrieved from http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Glossary:Labour_force_survey_%28LFS%29.

Stoop, I. (2005). *The hunt for the last respondent. Nonresponse in sample surveys* (Doctoral dissertation). Retrieved from https://www.scp.nl/english/Publications/Publications_by_year/Publications_2005/The_Hunt_for_the_Last_Respondent.

Stoop, I. (2015). *Nonresponse and fieldwork problems*. Paper presented at the 26th International Workshop on Household Survey Nonresponse, Leuven.

Stoop, I., Billiet, J., Koch, A., & Fitzgerald, R. (2010). *Improving survey response: Lessons learned from the European Social Survey*. Hoboken, NJ: John Wiley & Sons.

Stoop, I., Devacht, S., Billiet, J., Loosveldt, G., & Philippens, M. (2003). *The development of a uniform contact description form in ESS*. Paper presented at the 14th International Workshop on Household Survey Nonresponse, Leuven.

Stoop, I., Koch, A., Halbherr, V., Fitzgerald, R., & Widdop, S. (2014). *Field Procedures in the European Social Survey round 7: Enhancing Response Rates.* The Hague: European Social Survey, SCP.

Stoop, I., Koch, A., & Matsuo, H. (2012). *Not able and other reasons for survey nonresponse*. Paper presented at the International Workshop on Household Survey Nonresponse, Ottawa, Ontario.

Tabachnick, B. G. & Fidell L. S. (2001). *Using Multivariate Statistics*. Essex: Pearson Education Limited.

Tourangeau, R. & Smith T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly, 60*(2), 275-304. doi: 10.1086/297751

TNS Infratest Sozialforschung. (2011). *SOEP 1990/91–Methodenbericht Ostdeutschland zu den Befragungsjahren 1990/91 (Welle 1/2 – Ost) des Sozio-oekonomischen Panels*. Berlin: DIW Berlin, SOEP. Retrieved from http://panel.gsoep.de/soep-docs/surveypapers/diw_ssp0014.pdf.

Vehovar, V., & Slavec, A. (2012). Post-stratification weighting of the ESS: Methodology development and data preparation. DACE project, activity of WP12 London: European Social Survey.

Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. (1996). Mail surveys for election forecasting? An evaluation of the Colombia Dispatch Poll. *Public Opinion Quarterly, 60*(2), 181–227. doi:10.1086/297748

Voogt, R. (2004). “I’m not interested”. Nonresponse bias, response bias ans stimulus effect in election research. (Doctoral dissertation). Retrieved from http://dare.uva.nl/document/2/49548.

Wagner, J. (2010). The fraction of missing information as a tool for monitoring the quality of survey data. *Public Opinion Quarterly, 74*(2), 223–243. doi: 10.1093/poq/nfq007

Wagner, J. R. (2012). A comparison of alternative indicators for the risk of nonresponse bias. *Public Opinion Quarterly*, 76(3), 555–575. doi: 10.1093/poq/nfs032

Wagner, J. R., & Couper, M. P. (2011). *Using paradata and responsive design to manage survey nonresponse*. Paper presented at the 58th World Statistics Congress (ISI 2011), Dublin. Retrieved from http://2011.isiproceedings.org/papers/450080.pdf.

Wasmer, M., Blohm, M., Walter, J., Scholz, E. & Jutz, R. (2014). *Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2012.* GESIS Technical Report 2014/22. Retrieved from http://www.gesis.org/allbus/inhalte-suche/methodenberichte/.

Wolf, C. & Best, H. (2015). Linear regression. In H. Best & C. Wolf (Eds.), *Regression analysis and causal inference* (pp. 57-83). London: Sage.

Wittenberg, R. (1998). *Grundlagen computerunterstützter Datenanalyse*. Stuttgart: Lucius & Lucius.

Wood, A. M, White, I. R., & Hotopf, M. (2006). Using number of failed contact attempts to adjust for non-ignorable non-response. *Journal of the Royal Statistical Society, 169(3),* 525–542.

.

VII. Appendices

Appendix A: Analysis Part A

In the Appendix A of Analysis Part A, additional analyses and graphs are described and displayed. This section provides further details on the analyses of Section 8 and Section 9. It provides additional tables and figures of response rates in the ESS. The first section provides an analysis of the trend of response rates at different response rate levels according to an internal criterion, namely the quartiles of response rates in ESS 1.

Trends in Response Rate Levels According to Quartiles

The definition of countries with low or high response rates is based on the calculation of the response rate quartile of the response rate of the ESS in 2002 (ESS 1). Countries with low response rates (ranging from 33% to 53%) are Switzerland, Germany, Spain, and France. Countries within the 25%–50% (ranging from 55% to 65%) response rate quartile are Belgium, Ireland, Norway, and the United Kingdom. Countries within the third quartile (ranging from 68% to 69%) are Denmark, the Netherlands, Portugal, and Sweden. Countries within the fourth quartile (ranging from over 69% to 73.3%) are Finland, Hungary, Poland and Slovenia.

Analysis carried out by the present study at the country level of the four countries in the lowest response rate quartile (ranging from 33% to 53%) in 2002 found that different trends were present for different countries. In Germany, Spain, and Switzerland, large changes in response rates were found when the first round of the ESS (ESS1) was compared to the last round (ESS5).[183] In countries with a low response rate level in 2002, response rates increased in three out of four cases. Spain, France, and Switzerland had an upward trend. In Germany, the response rate started at a low level and has decreased since 2006 (ESS 3). As shown in Figure A1, most of the variation for Switzerland and France is within the confidence interval around the regression lines. Regression analysis over time for the countries with low response rates shows a positive coefficient for the response rates (b = .01; t = .69; p = .501; R^2 = .03), but the trend is not significant. Visual analysis[184] of the trend can be described as a regression to the mean level of response rate, since in the country with low response rate level the response rate increases.

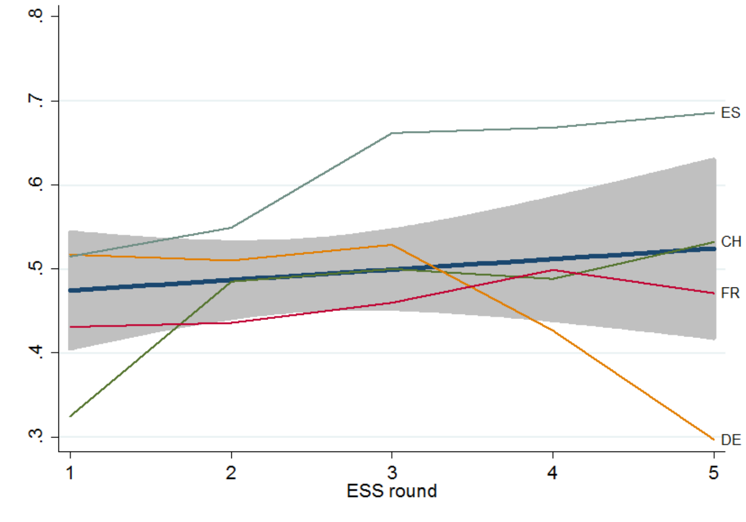

Figure A1. Development of response rates in countries with low response rates in ESS 1 (first quartile).

For the four countries in the second response rate quartile (ranging from 55% to 65%) in 2002, the regression line showed a slight negative trend (Figure A2). However, although

---

[183] Germany had a decrease of 22.0 percentage points; Spain had an increase of 17.1 percentage points; and Switzerland had an increase of 20.7 percentage points. In France, an upward trend in response rates was steady from 2002 (ESS1) to 2010 (ESS5). In Switzerland, the largest proportion of the increase in response rates occurred during the first to the second round of the ESS; after that period, Switzerland's response rates stayed stable until they increased again between ESS4 and ESS5. Out of all these four countries, Germany had a clear trend of declining response rates.

[184] Due to the limited number of cases (n = 4), a visual analysis was used.

this trend is negative (*b* = -.01), it was not significant (*t* =-1.78; *p* = .092; R^2 = .11). In Norway, Belgium, and Ireland, the response rate trend was negative and in the United Kingdom it was rather constant (an increase of 1.3% between 2002 and 2010). Thus, three out of the four countries in the second quartile had decreases in their response rates.

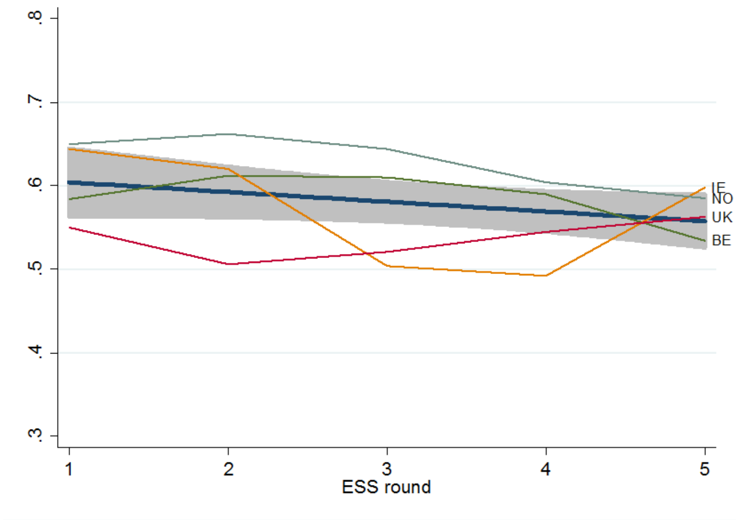

Figure A2. Development of response rates in countries with low-mean response rates in ESS 1 (second quartile).

The countries in the third quartile (ranging from 68% to 69%) and fourth quartile (ranging from over 69% to 73%) had a negative trend in their response rates (Figure A3 and Figure A4). In the third quartile, response rates decreased in all countries (Portugal, Netherlands, Denmark, and Sweden), which also is reflected in a negative regression line that showed a significant negative trend (*b* = -.03; *t* = -3.16; *p* = .005; R^2 = .25).

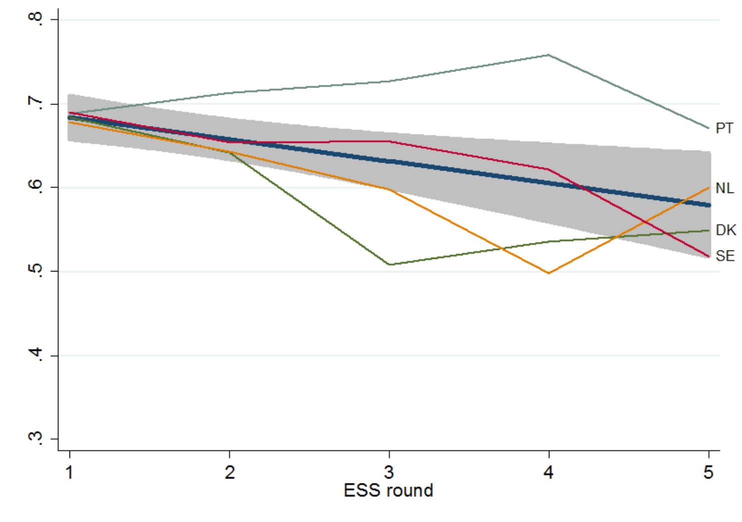

Figure A3. Development of response rates in countries with mean-high response rates in ESS 1 (third quartile).

Regression analysis of the response rate in the four countries in the highest response rate quartile (response rate between over 69%-73%) shows also a significant negative effect (b = - .02; t = -3.81; p = .001; R^2 = .44). As can be seen in Figure A4 in all of countries (Poland, Slovenia, Hungary, Finland) within the fourth quartile the response rates decreases from the 2002 to 2010. The trend in the countries in the third and fourth quartile, which are the countries with above the mean response rate in 2002, can be described as a regression to the mean.

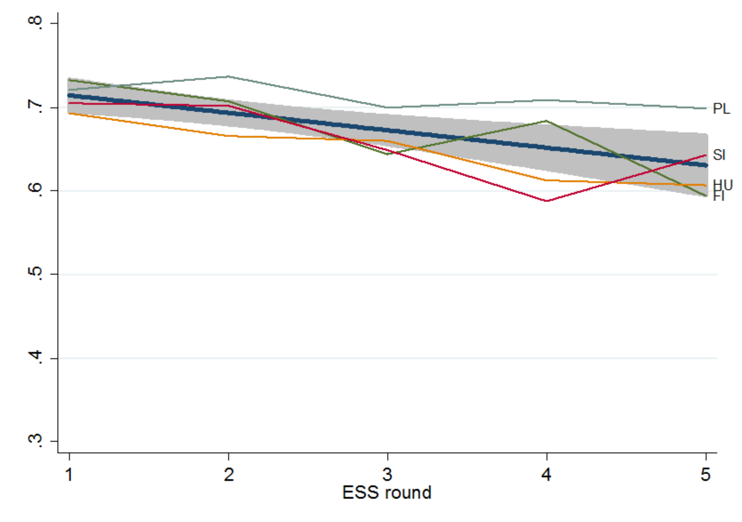

Figure A4. Development of response rates in countries with high response rates in ESS 1 (fourth quartile).

Tables and Figure of Response Rates

Table A1

Changes of response rates (in percentages) between ESS rounds

Country	ESS 1	Δ ESS1 to ESS2	ESS 2	ΔESS2 to ESS3	ESS 3	Δ ESS3 to ESS4	ESS 4	Δ ESS4 to ESS	ESS 5	Δ ESS1 to ESS5
Belgium	58.4	**-2.8**	61.2	**0.2**	61.0	**2.0**	59.0	**5.6**	53.4	**5.0**
Switzerland	32.5	**-16.0**	48.5	**-1.5**	50.0	**1.2**	48.8	**-4.4**	53.2	**-20.7**
Germany	51.7	**0.7**	51.0	**-1.9**	52.9	**10.2**	42.7	**-10.5**	53.2	**-1.5**
Denmark	68.4	**4.2**	64.2	**13.4**	50.8	**-2.8**	53.6	**-1.3**	54.9	**13.5**
Spain	51.5	**-3.4**	54.9	**-11.3**	66.2	**-0.6**	66.8	**-1.8**	68.6	**-17.1**
Finland	73.3	**2.6**	70.7	**6.3**	64.4	**-4.0**	68.4	**9.0**	59.4	**13.9**
France	43.1	**-0.5**	43.6	**-2.4**	46.0	**-3.9**	49.9	**2.8**	47.1	**-4.0**
Hungary	69.3	**2.7**	66.6	**0.6**	66.0	**4.7**	61.3	**0.6**	60.7	**8.6**
Ireland	64.4	**2.4**	62.0	**11.6**	50.4	**1.2**	49.2	**-10.6**	59.8	**4.6**
Netherlands	67.8	**3.5**	64.3	**4.5**	59.8	**10.0**	49.8	**-10.2**	60.0	**7.8**
Norway	65.0	**-1.2**	66.2	**1.8**	64.4	**4.0**	60.4	**1.9**	58.5	**6.5**
Poland	72.1	**-1.6**	73.7	**3.7**	70.0	**-0.9**	70.9	**0.9**	70.0	**2.1**
Portugal	68.8	**-2.5**	71.3	**-1.4**	72.7	**-3.1**	75.8	**8.7**	67.1	**1.7**
Sweden	69.0	**3.6**	65.4	**-0.1**	65.5	**3.3**	62.2	**10.4**	51.8	**17.2**
Slovenia	70.5	**0.3**	70.2	**5.3**	64.9	**6.1**	58.8	**-5.6**	64.4	**6.1**
United Kingdom	55.0	**4.4**	50.6	**-1.5**	52.1	**-2.4**	54.5	**-1.8**	56.3	**-1.3**
Mean	61.3	**-0.2**	61.5	**1.7**	59.8	**1.5**	58.3	**1.1**	57.2	**4.1**

Note. Δ = change. The ESS rounds and corresponding years are as follows: ESS 1: 2002; ESS 2: 2004; ESS 3: 2006; ESS 4: 2008; ESS 5: 2010.

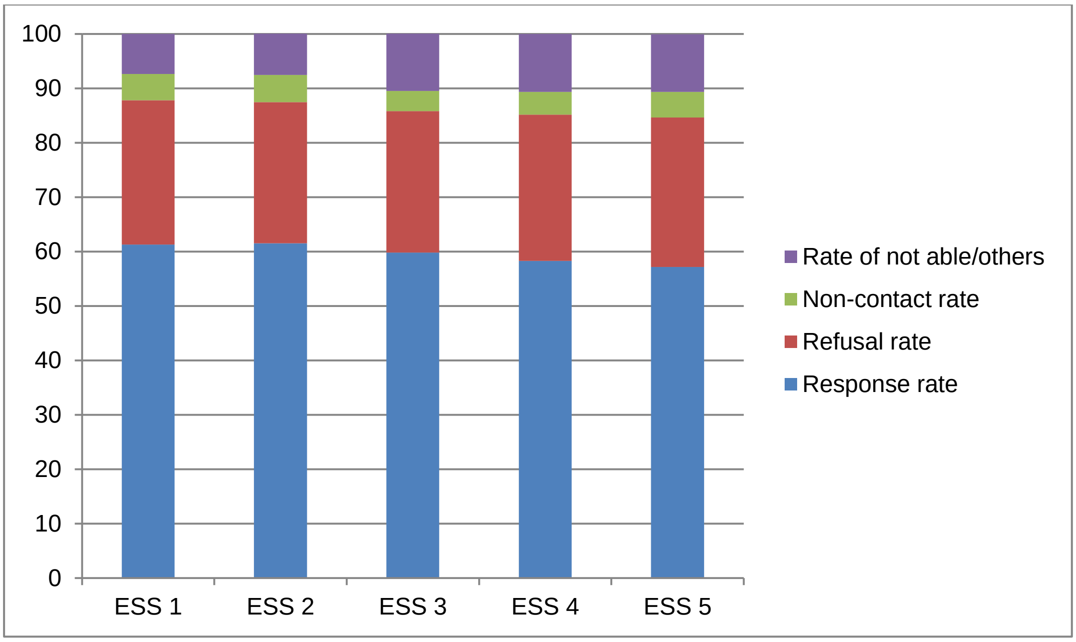

Figure A5. Stacked bar chart of response rate, refusal rate, non-contact rate, and rate of not able/others (in percentages).

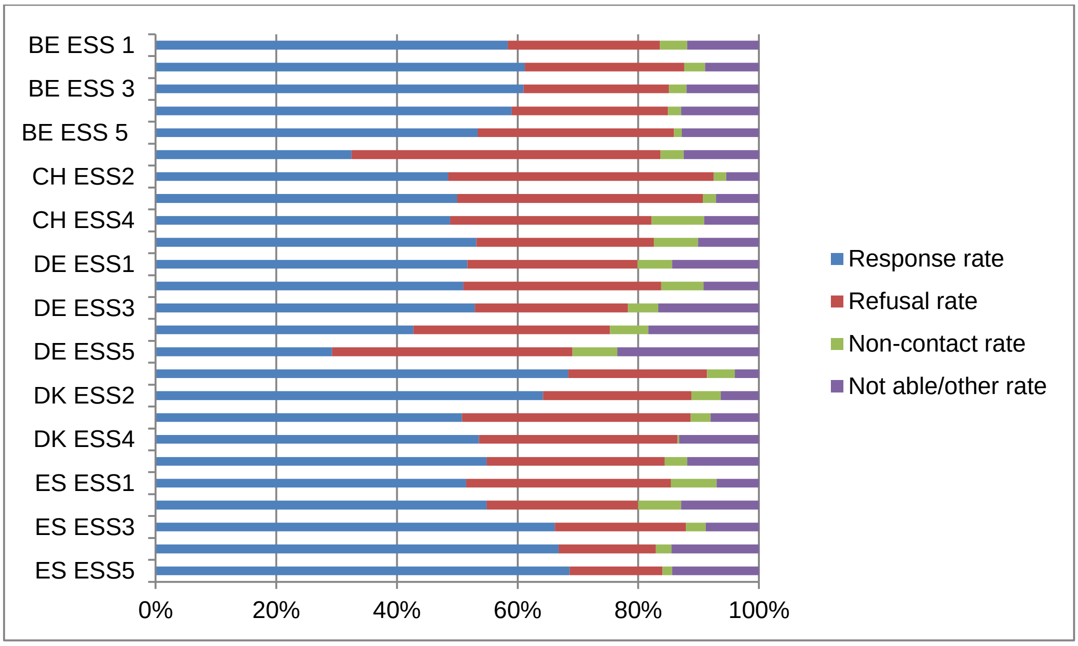

Figure A6. Chart of response rate, refusal rate, non-contact rate, and rate of not able/others (in percentages).

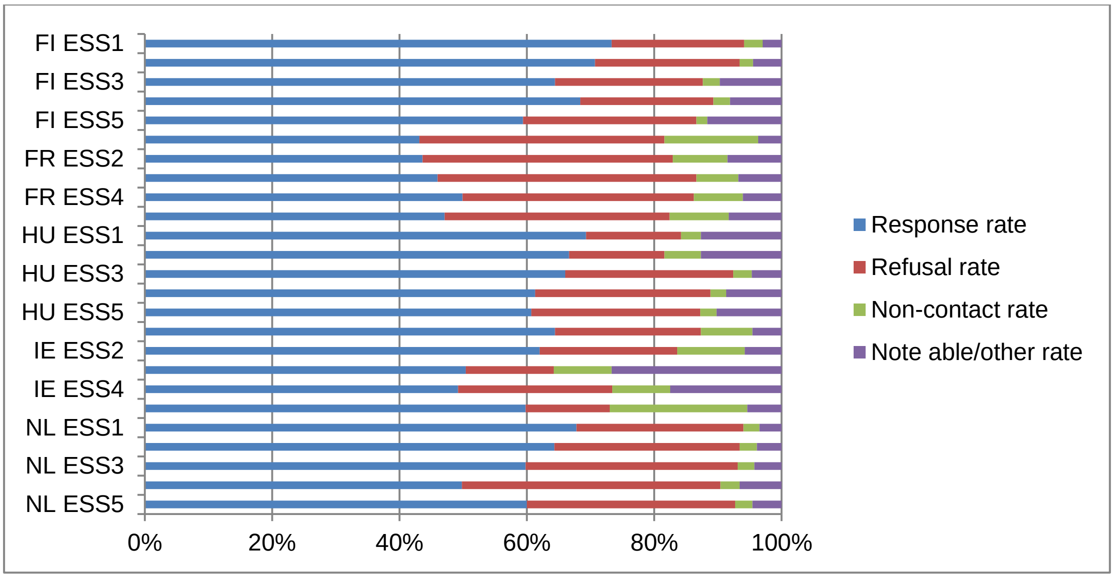

Figure A7. Chart of response rate, refusal rate, non-contact rate, and rate of not able/others (in percentages).

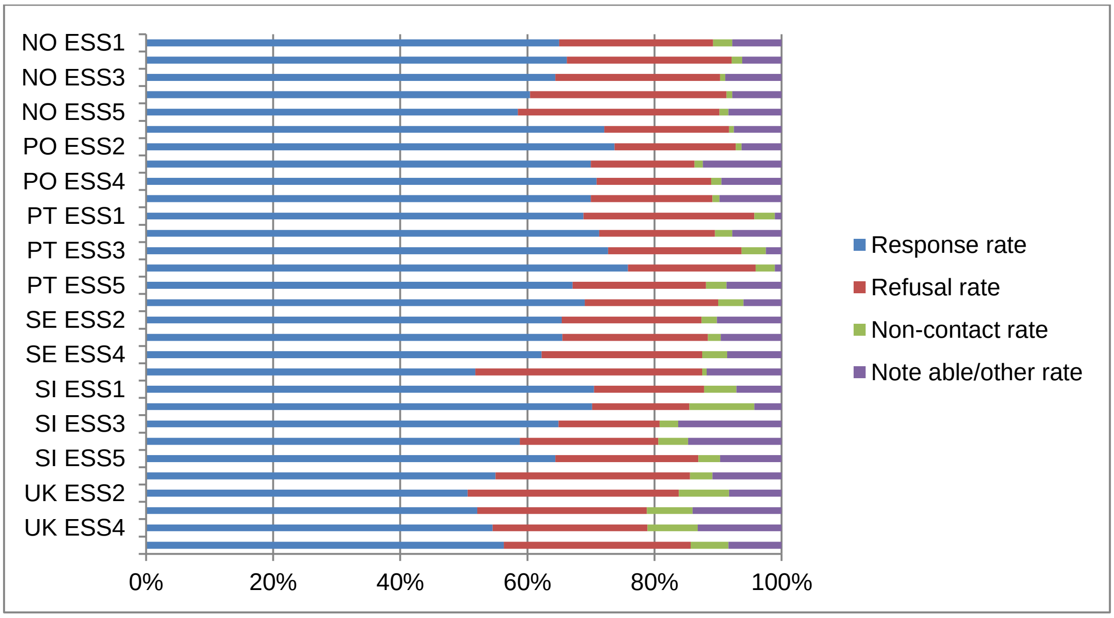

Figure A8. Response rates, refusal rates, non-contact rates, and rates of not able/others (in percentages) per country and round.

Appendix B: Analysis Part B

Appendix B provides additional information for the analysis of the effects of response rates on nonresponse bias that was performed in Analysis Part B. As a supplement to Analysis Part B Section 13, this Appendix contains additional graphs. In Figure B1 and Figure B2, the separate nonresponse bias variables are displayed in different colors with a regression line in the same color. In these figures, the nonresponse bias varies between the variables, and the slope of the regression line also differs. In Figure B3 to Figure B22 the graphs from Section 13 are included without comparable scales. Comparable scales in the figures allow comparison between the variable on the overall extent of nonresponse bias. Different scales (see Appendix) allow identification of countries and rounds.

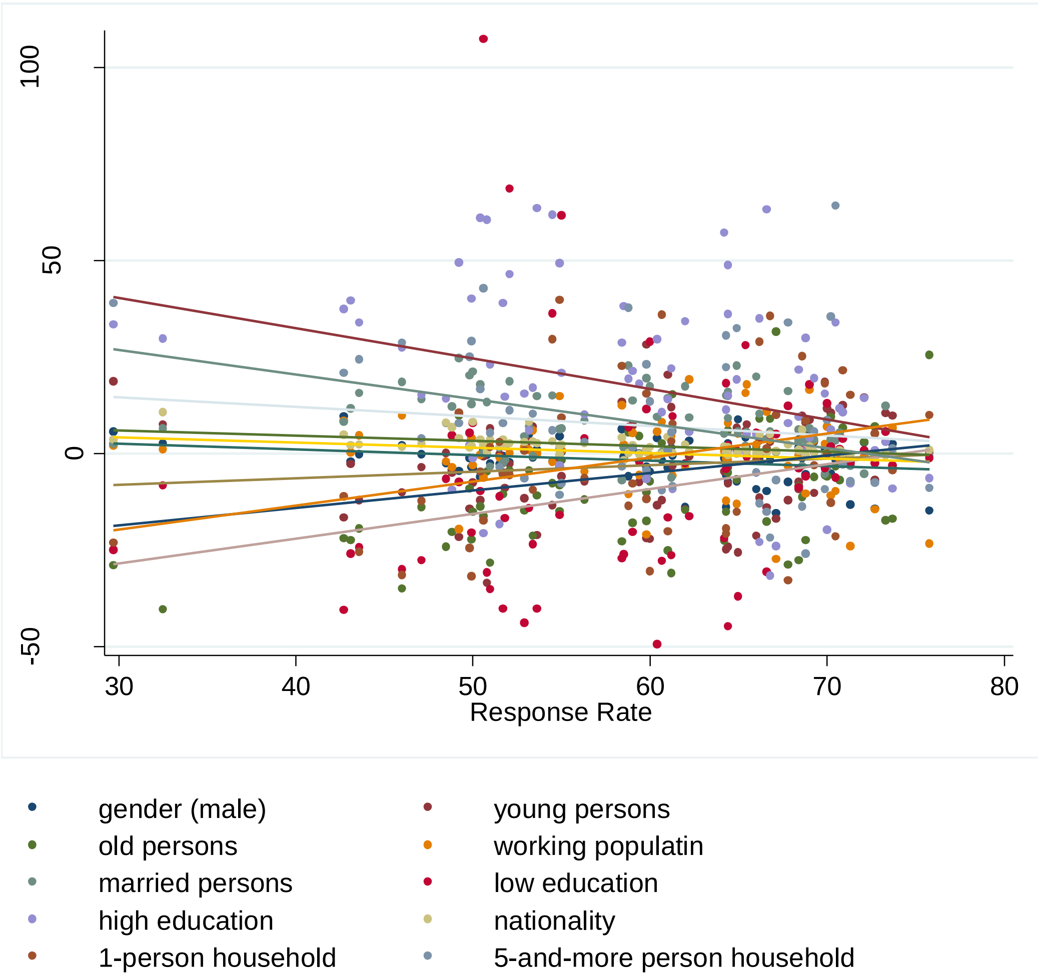

Figure B1. Nonresponse bias (relative bias) for all nonresponse bias variables and the response rates (in percentages).

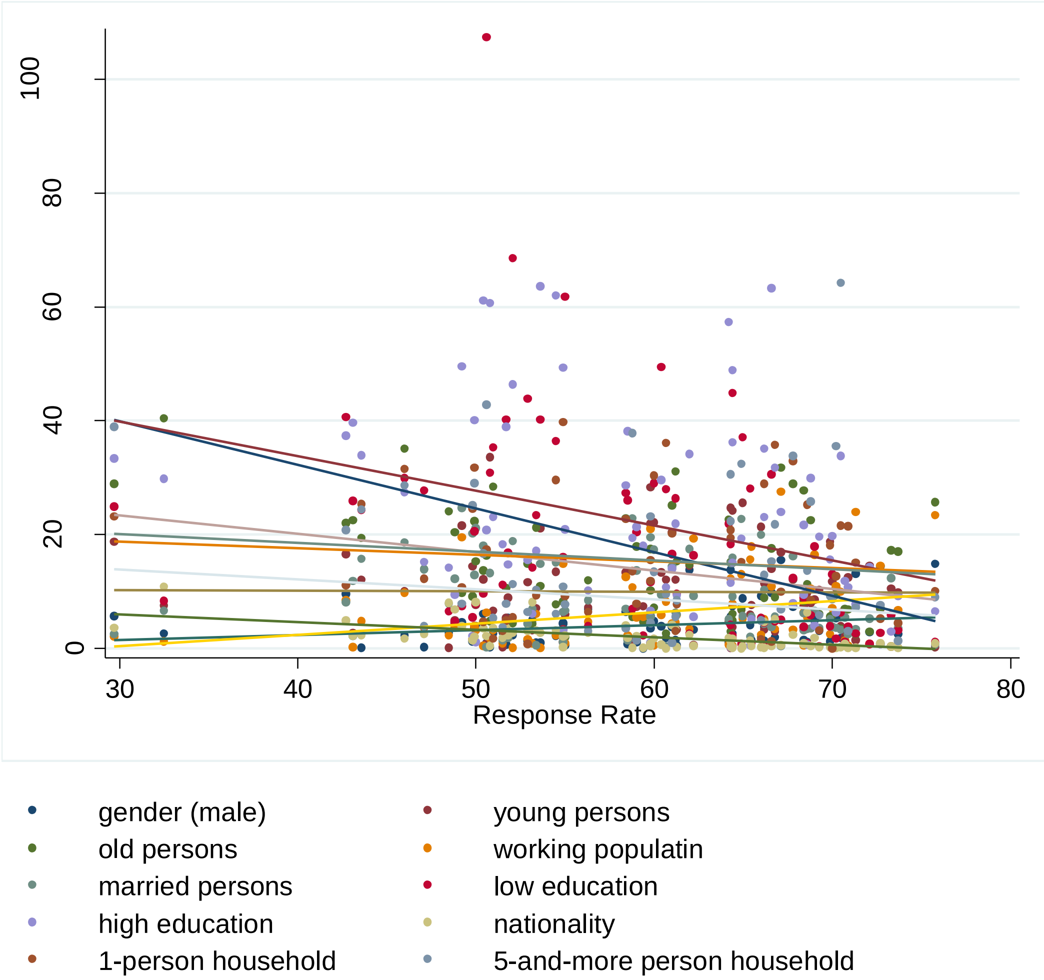

Figure B2. Nonresponse bias (absolute value of relative bias) for all nonresponse bias variables and the response rates (in percentages).

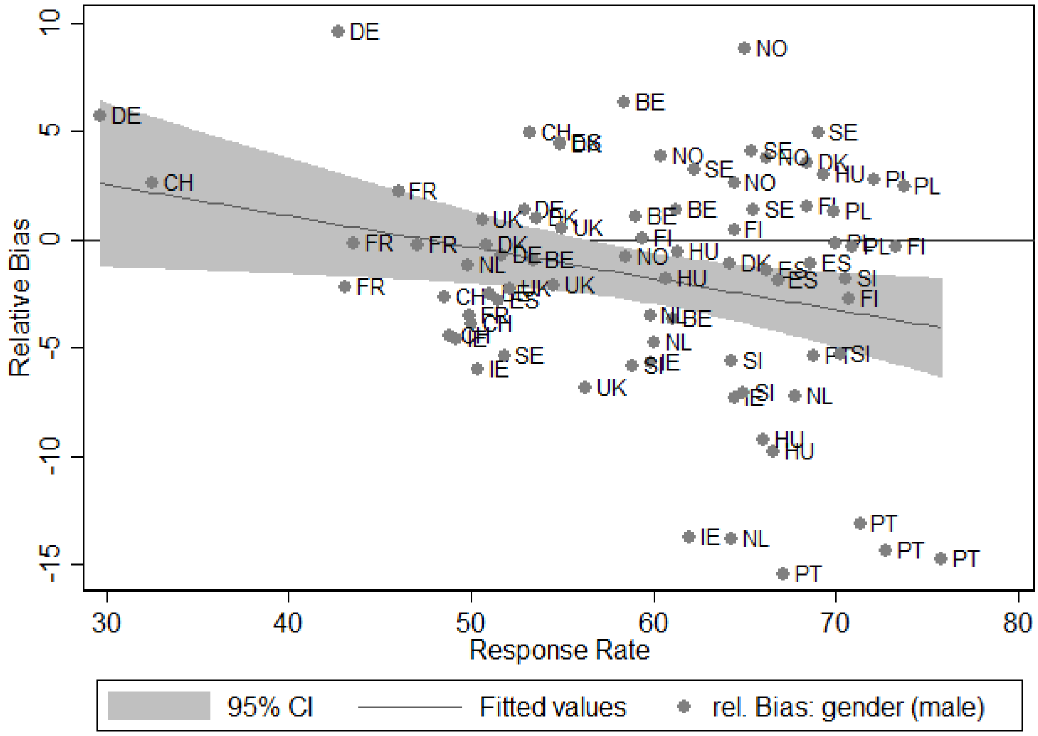

Figure B3. Relative bias for “gender (male)” and the response rate (in percentages).

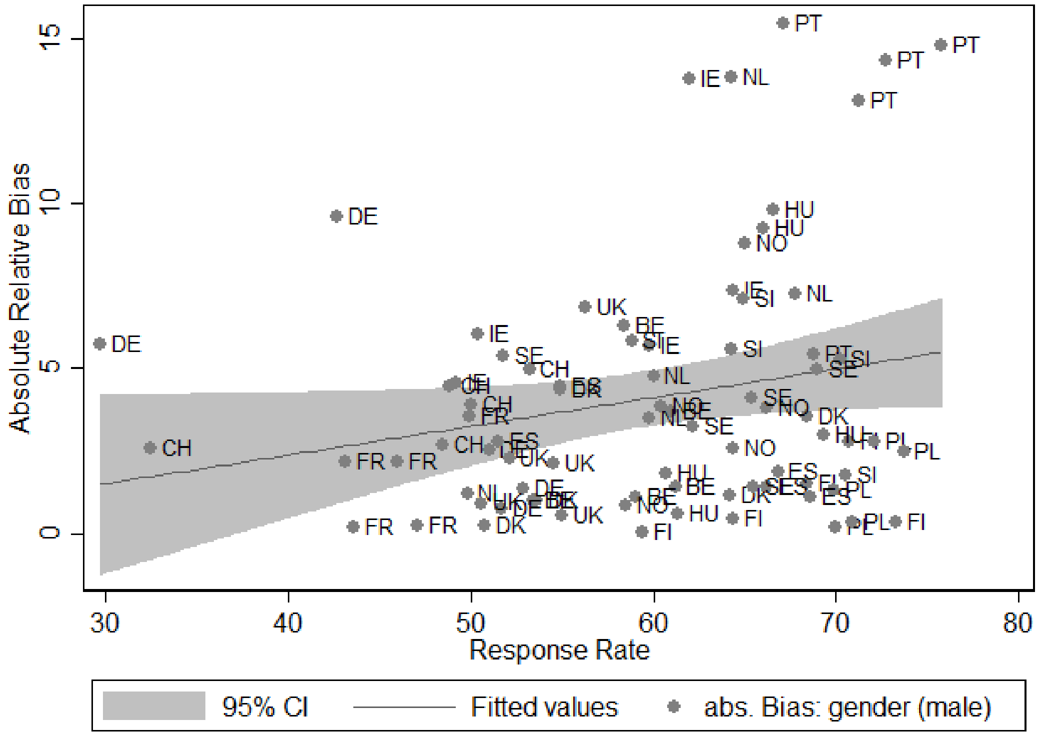

Figure B4. Absolute value for the relative bias for “gender (male)” and the response rate (in percentages).

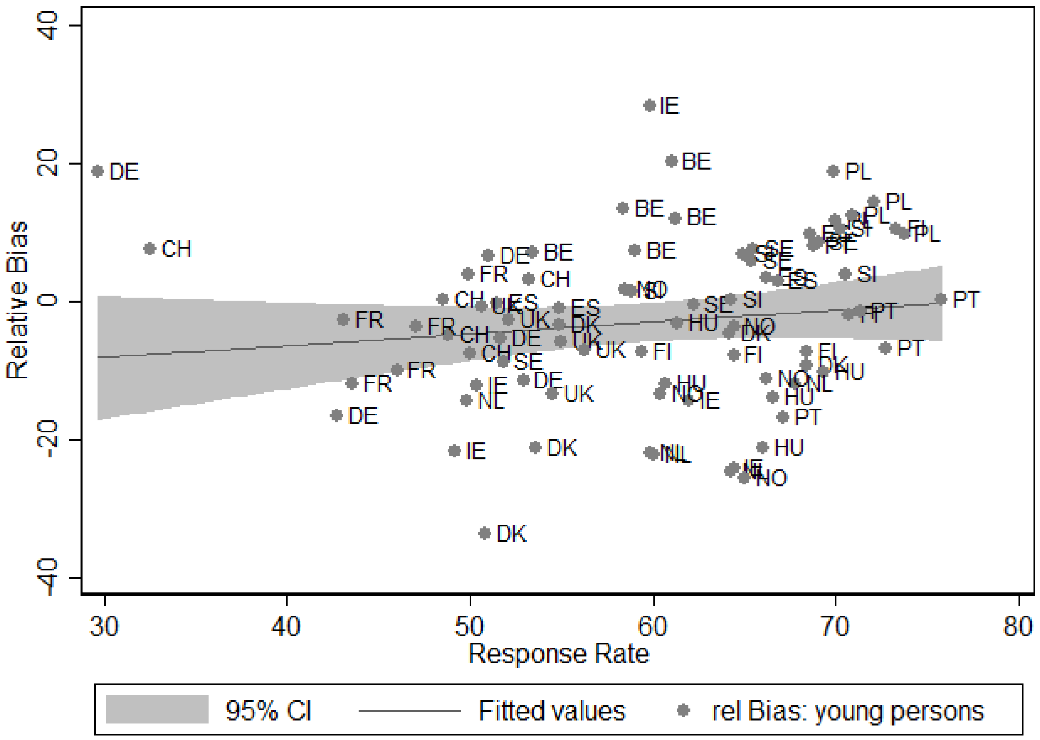

*Figure B5. R*elative bias for “young persons” and the response rate (in percentages).

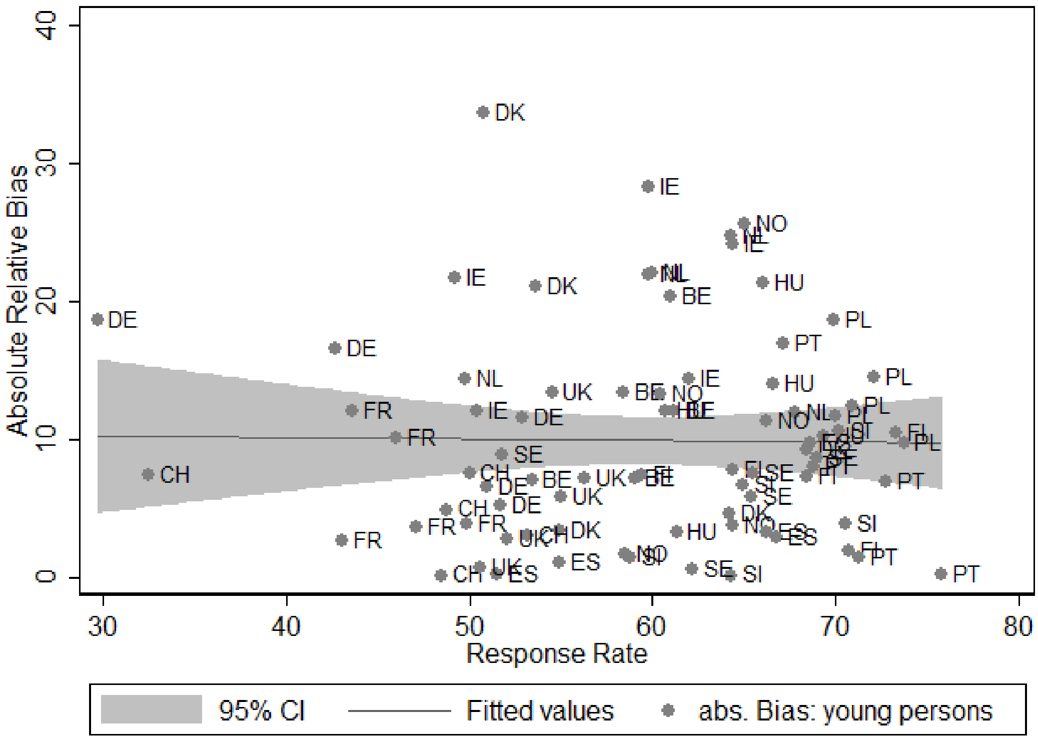

Figure B6. Absolute value for the relative bias for “young persons” and the response rate (in percentages).

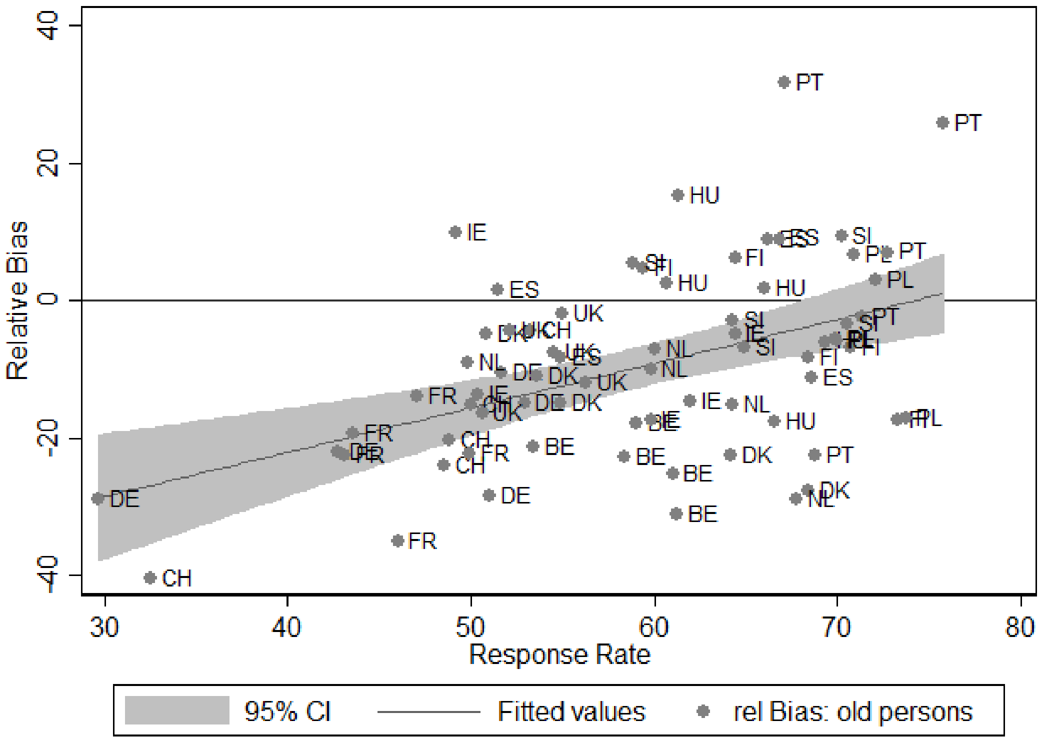

Figure B7. Relative bias for “old persons” and the response rate (in percentages).

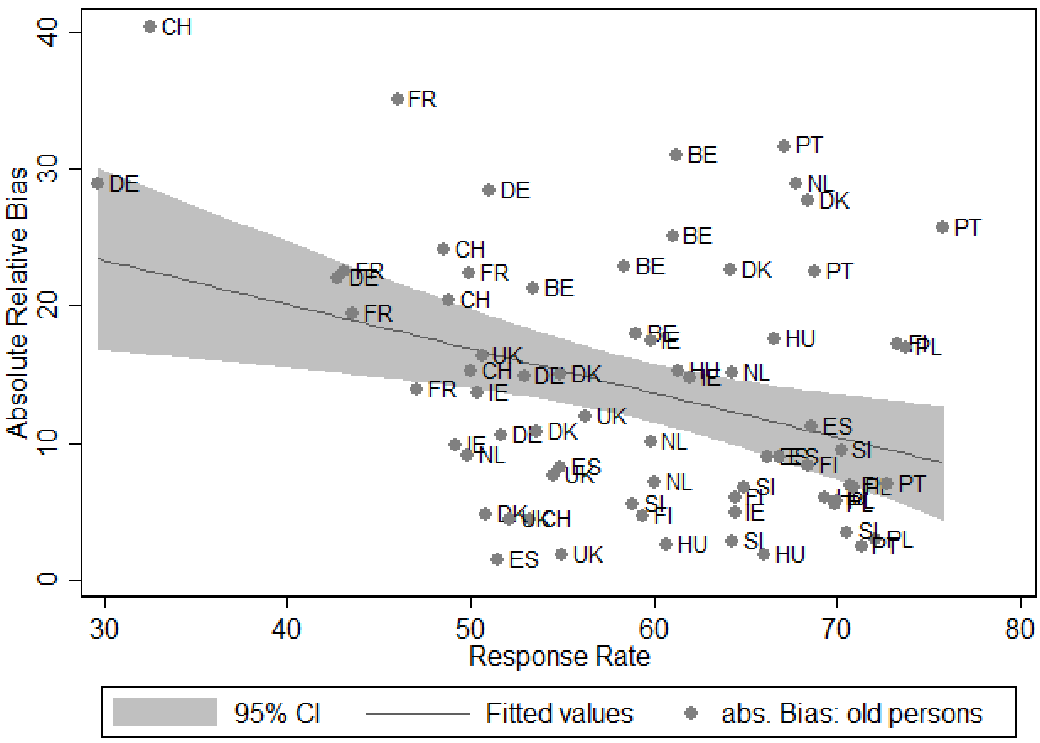

Figure B8. Absolute value for the relative bias for “old persons” and the response rate (in percentages).

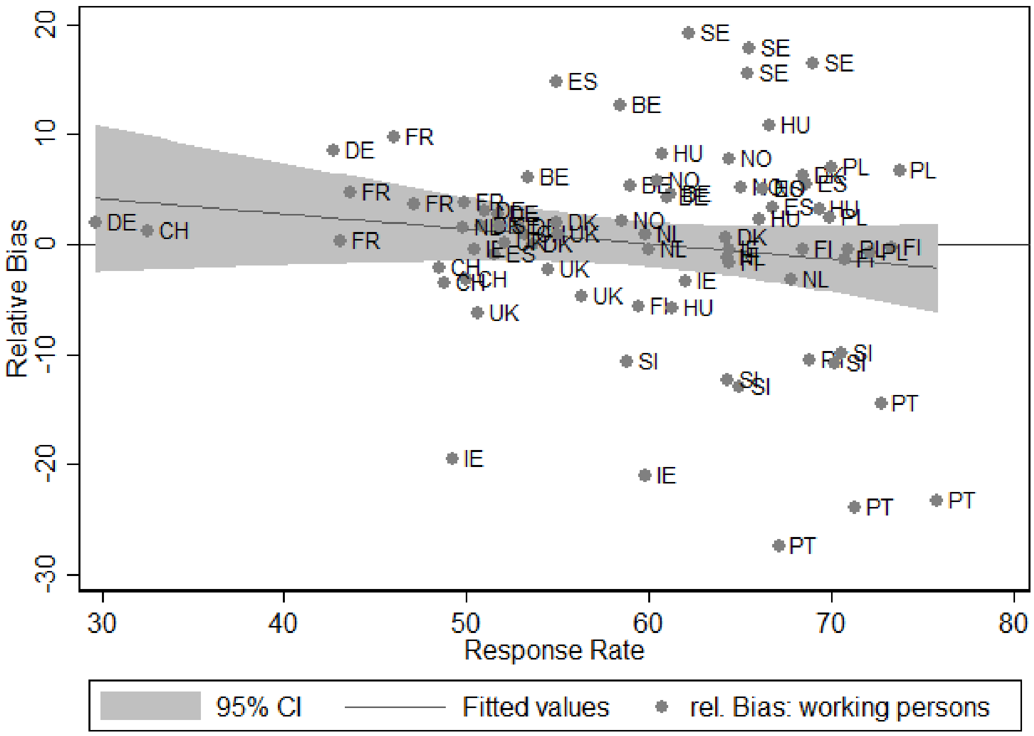

Figure B9. Relative bias for “working population” and the response rate (in percentages).

Figure B10. Absolute value for the relative bias for “working population” and the response rate (in percentages).

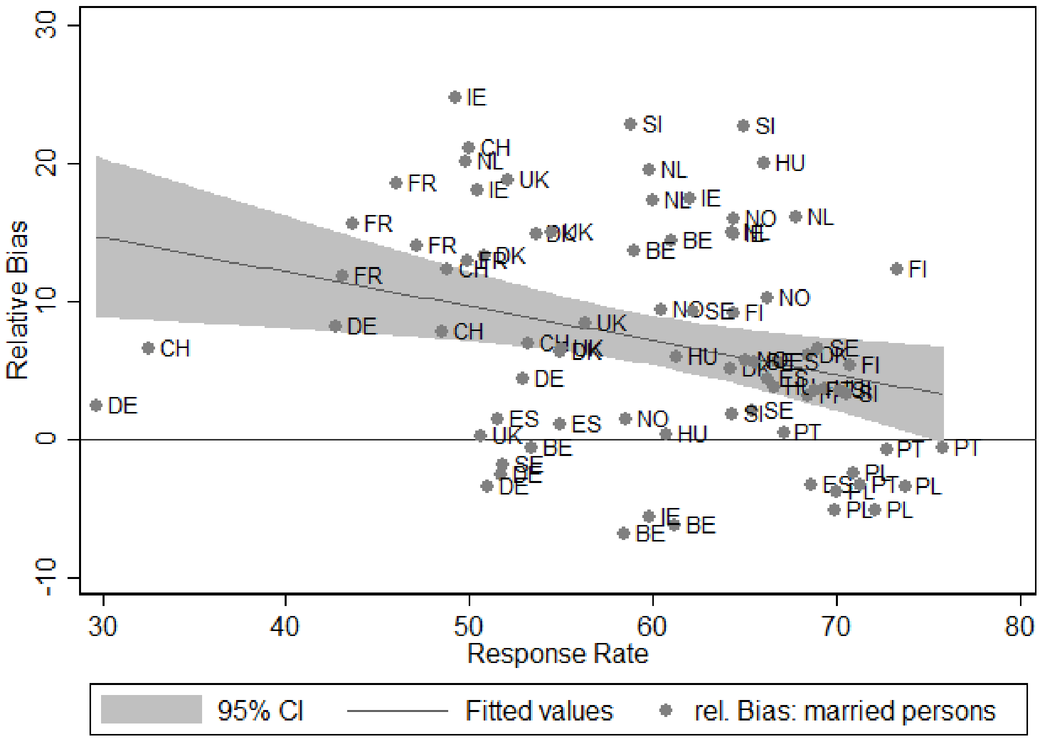

Figure B11. Relative bias for “married persons” and the response rate (in percentages).

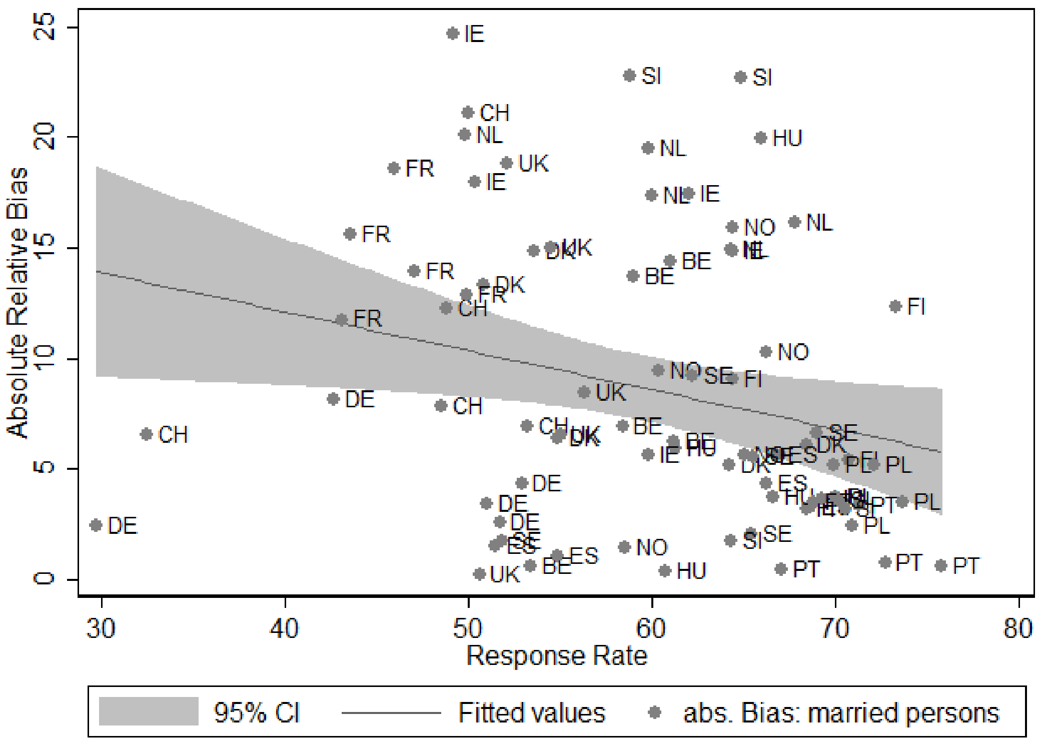

Figure B12. Absolute value for the relative bias for “married persons” and the response rate (in percentages).

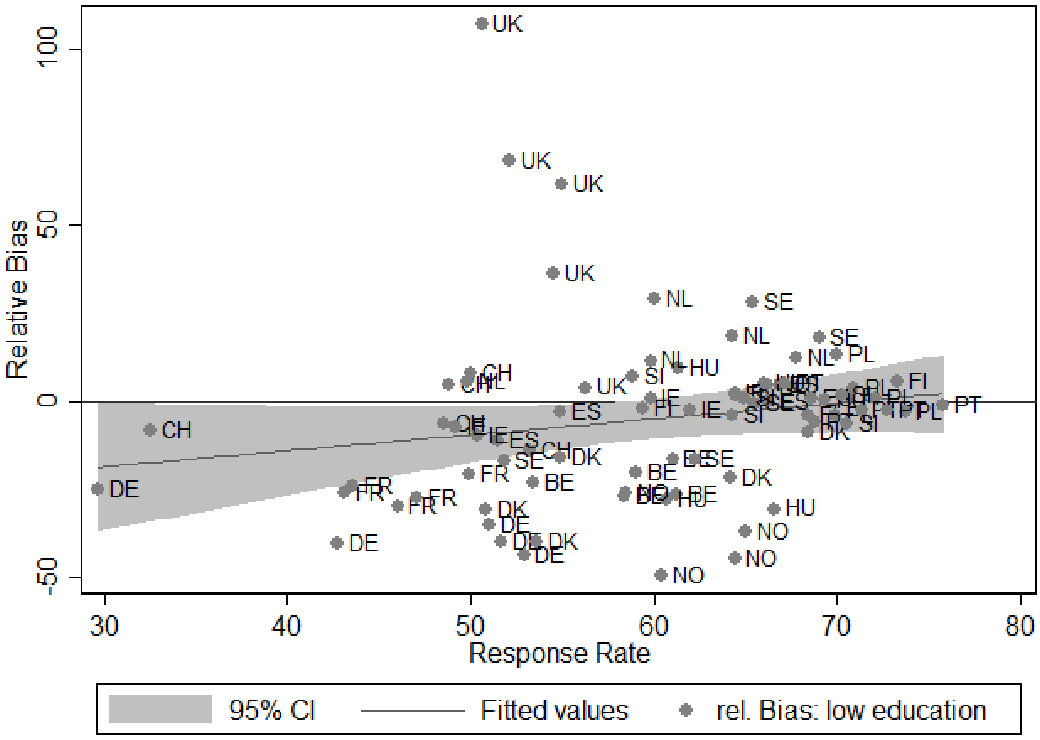

Figure B13. Relative bias for “low education” and the response rate (in percentages).

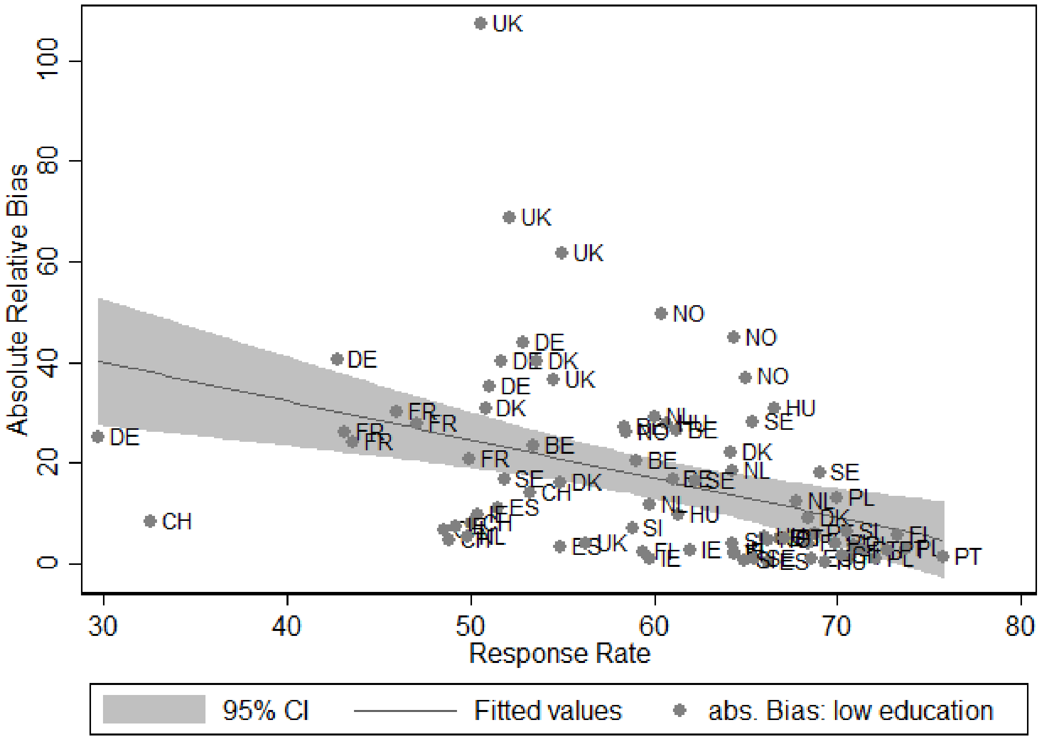

Figure B14. Absolute value for the relative bias for “low education” and the response rate (in percentages).

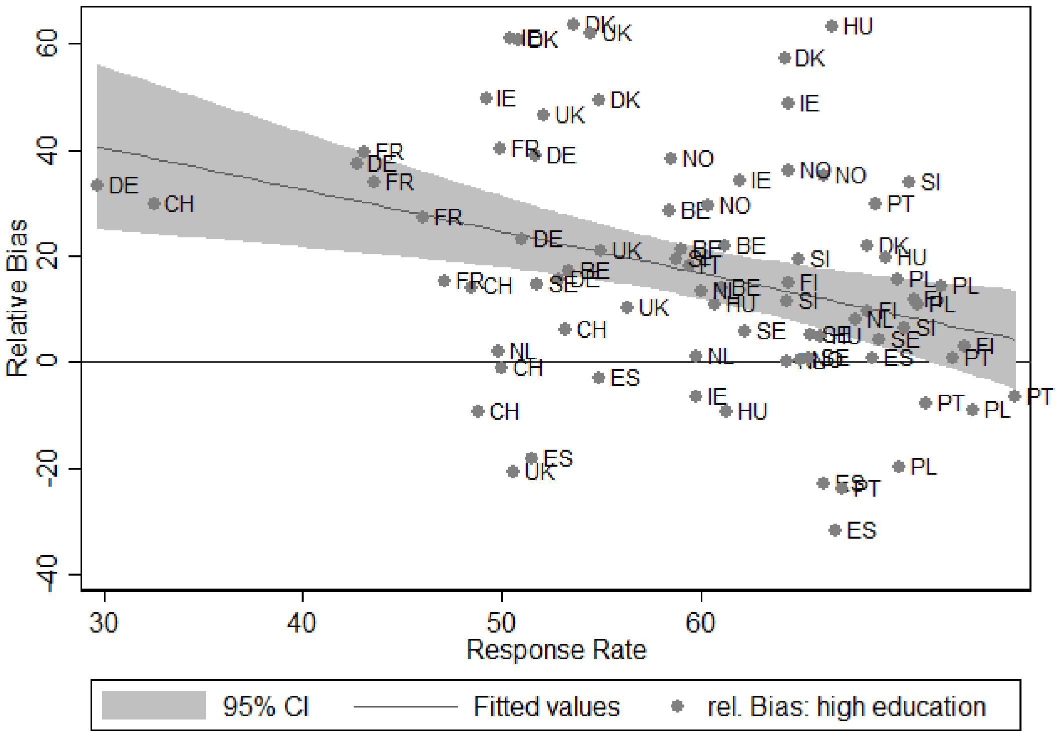

Figure B15. Relative bias for “high education” and the response rate (in percentages).

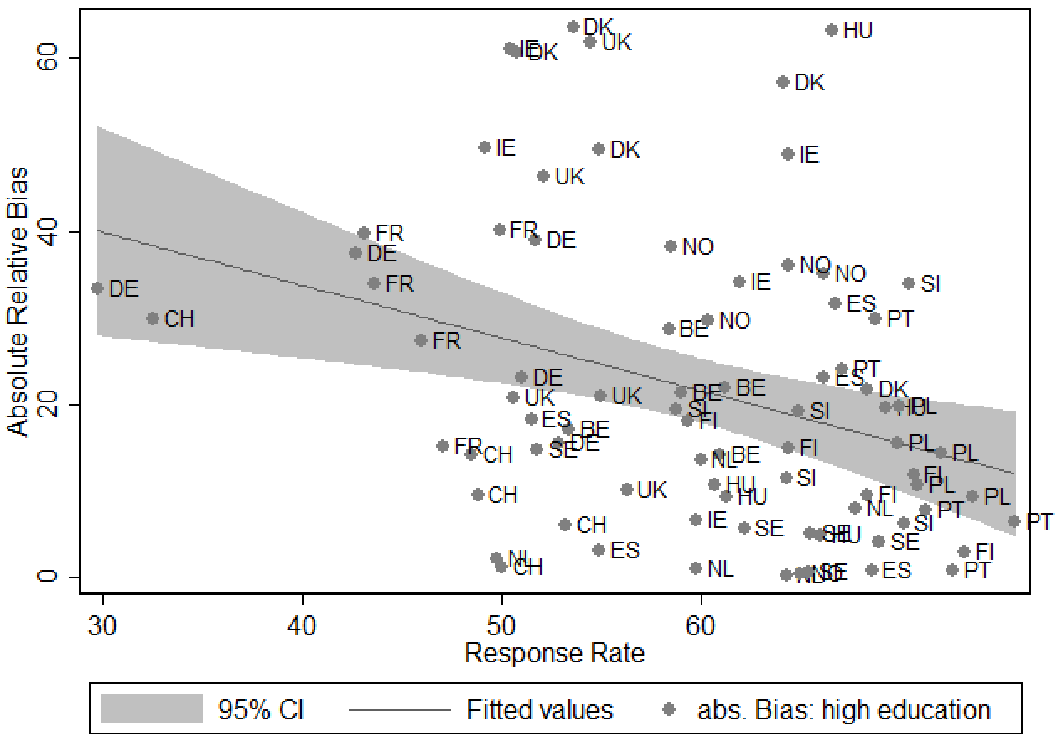

Figure B16. Absolute value for the relative bias for “high education” and the response rate (in percentages).

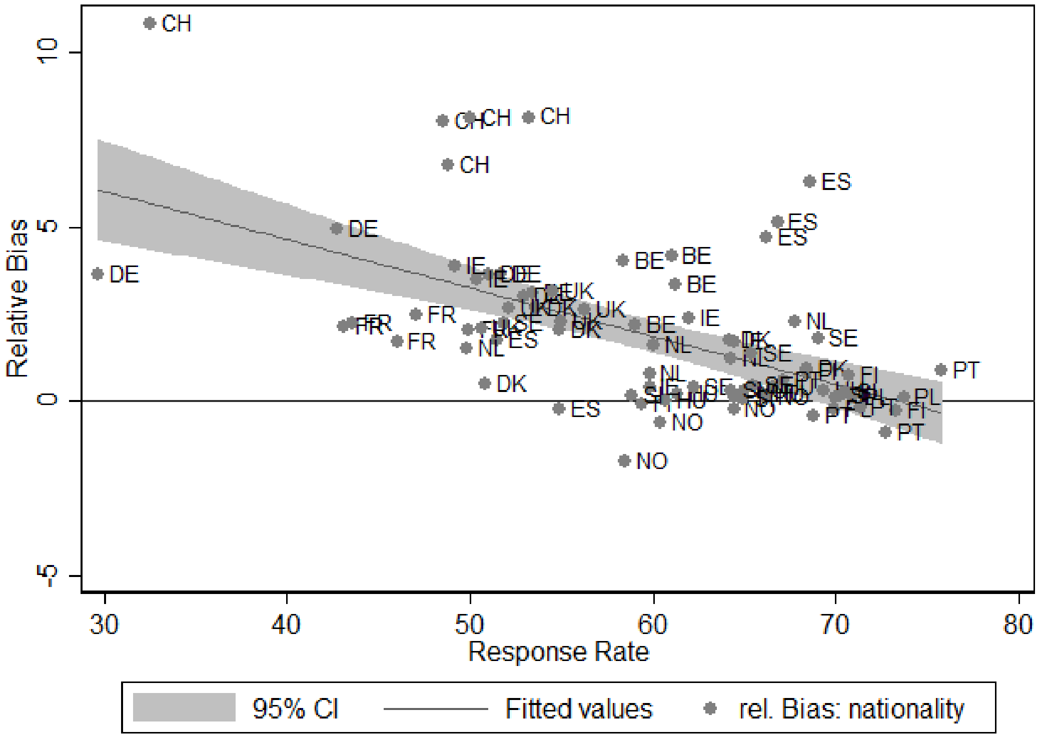

Figure B17. Relative bias for “nationality“ and the response rate (in percentages).

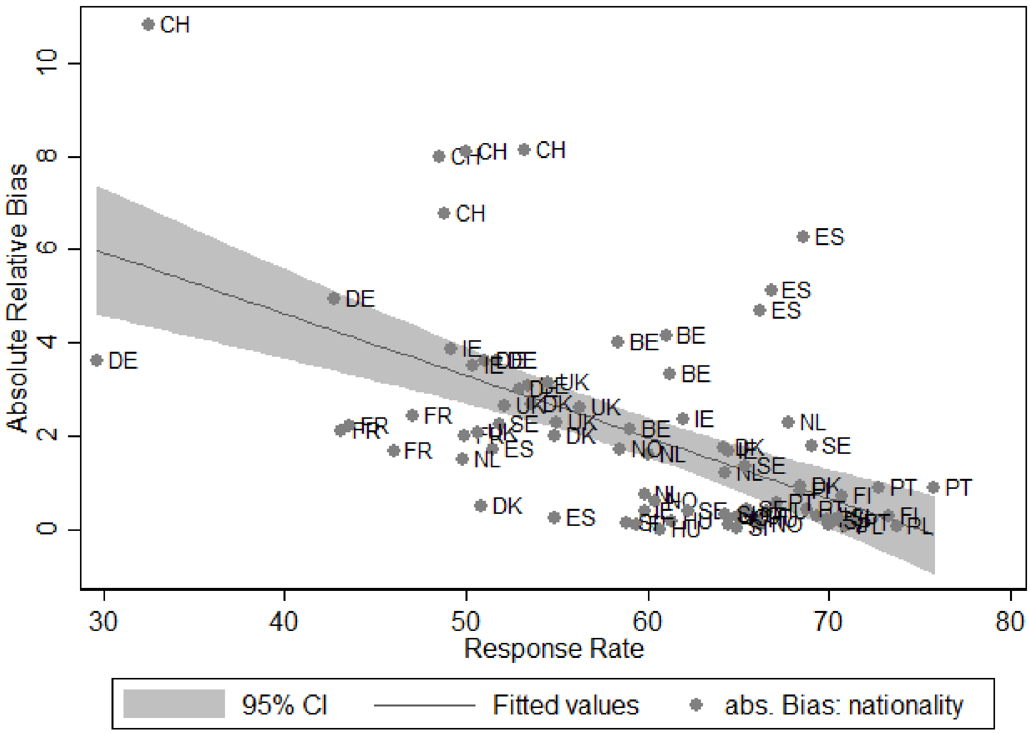

Figure B18. Absolute value for the relative bias for “nationality” and the response rate (in percentages).

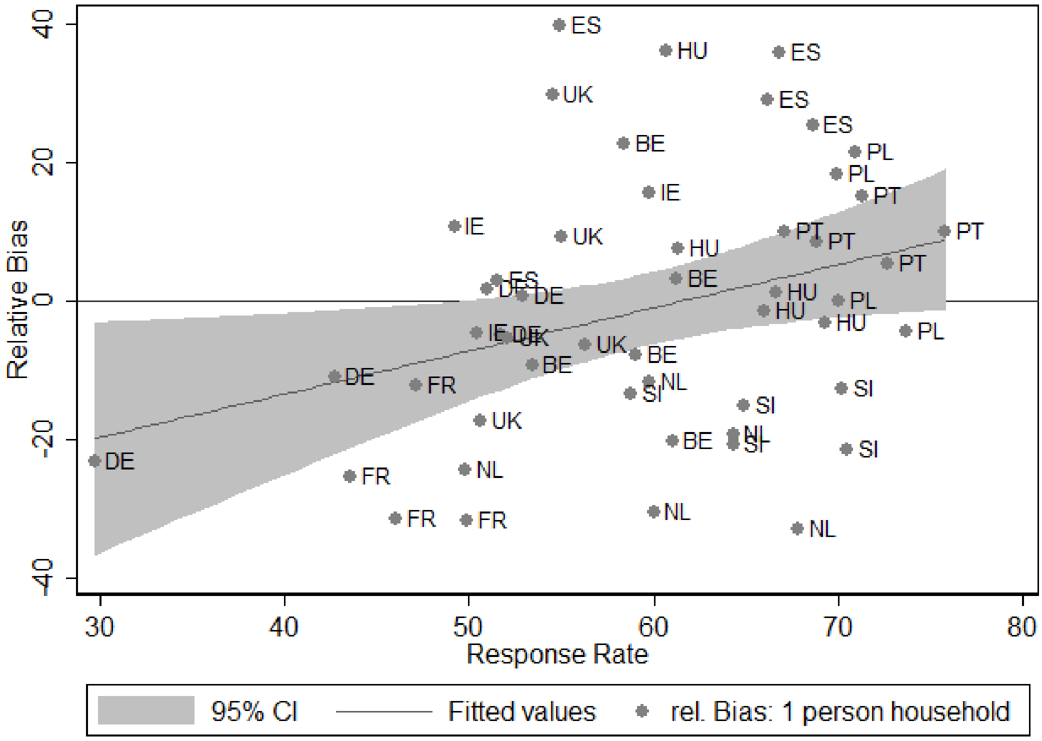

Figure B19. Relative bias for “one-person household” and the response rate (in percentages).

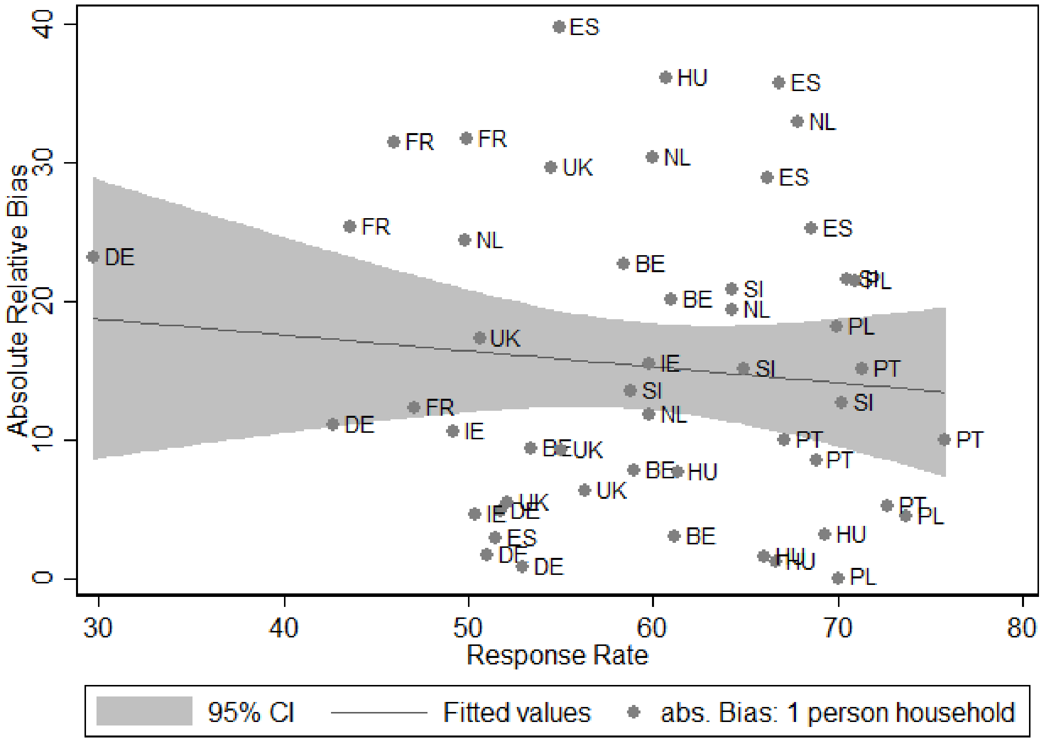

Figure B20. Absolute value for the relative bias for “one- person household” and the response rate (in percentages).

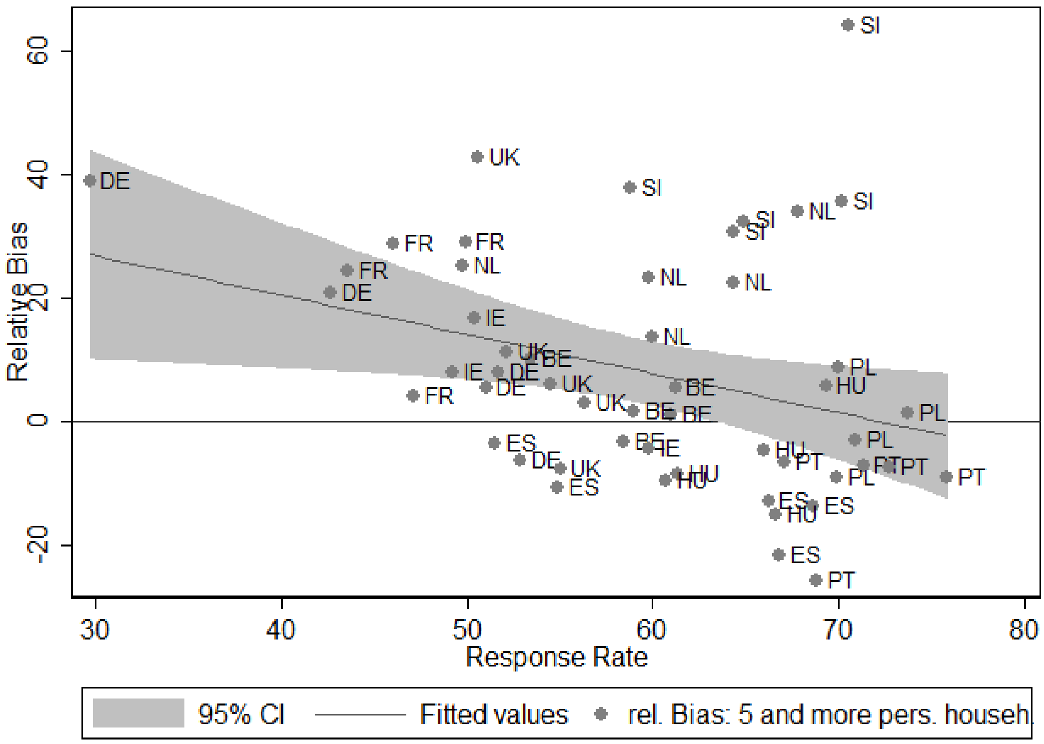

Figure B21. Relative bias for “five-or- more person household” and the response rate (in percentages).

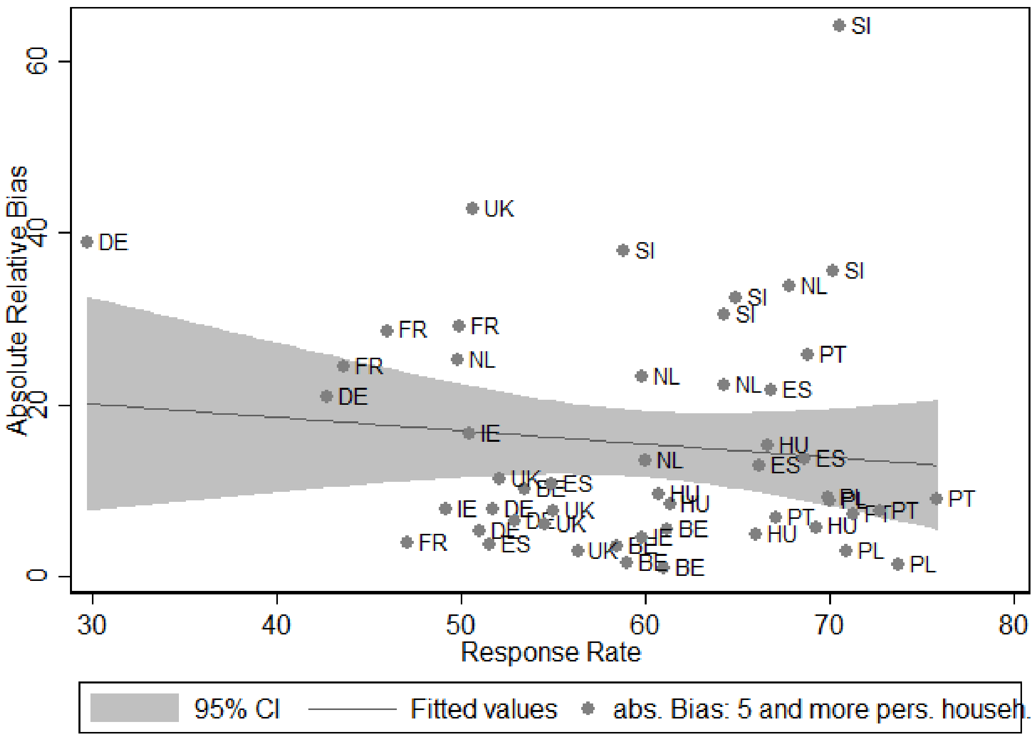

Figure B22. Absolute value for the relative bias for “five-or-more person household” and the response rate (in percentages).

Appendix C: Analysis Part C

Appendix C provides additional information for the analysis of the effects of fieldwork efforts on response rates that was performed in Analysis Part C. It is divided into three parts. First descriptive statistics of the variable fieldwork effort index (Table C1) and boxplots for the fieldwork efforts and response rates are shown (Figure C1). These analyses supplement the analysis conducted in Analysis Part C Section 17. In the second part of the analysis the focus is on longitudinal aspects, and changes between the rounds are further described in Table C2 and Table C3, Figure C4 and Figure C5.

These provide further material related to the analysis in Analysis Part C, Section 18. In the third part, different kind of imputation methods for missing values for the longitudinal qualitative within-country analysis are provided (Figure C3 and Figure C4 and Figure C5). This provides additional information to Analysis Part C Section 19.

Table C1

Descriptive statistics of the variable fieldwork effort index

Round	n	Min.	Max.	Mean	Median	Mode	SD	Variance
ESS 1	15	3	8	**5.67**	6	5	1,397	1,952
ESS 2	15	3	8	**5.80**	6	6	1,568	2,457
ESS 3	15	4	8	**6.33**	6	6	1,397	1,952
ESS 4	14	4	8	**6.21**	6	6	1,188	1,412
ESS 5	15	4	8	**6.53**	7	7	1,187	1,410
Total	74	3	8	**6.11**	6	7	1,361	1,851

Note. SD = standard deviation.

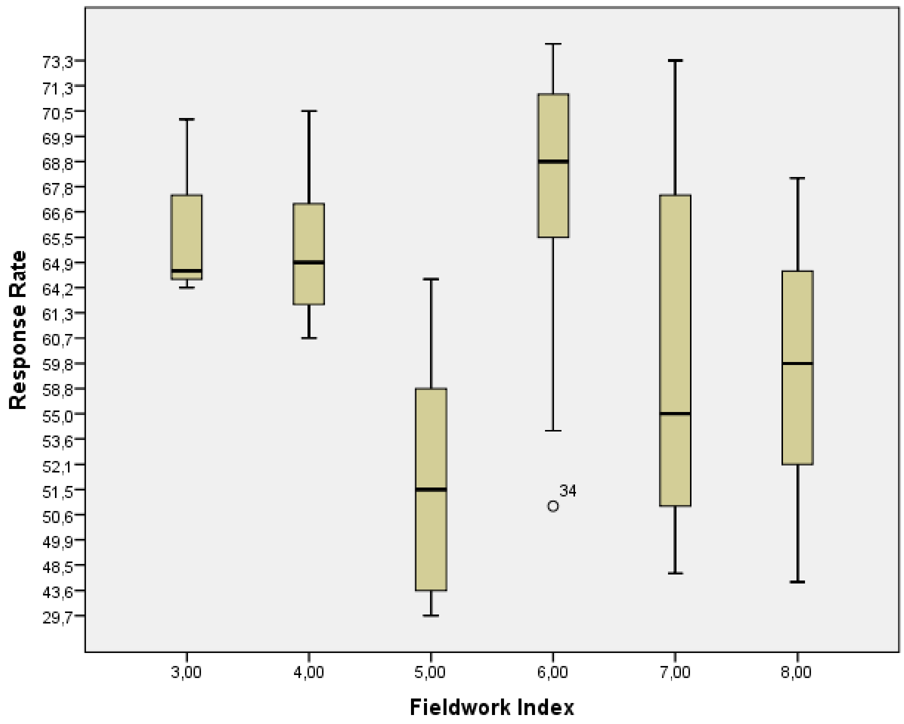

Figure C1. Boxplot response rates (in percentages) and the fieldwork effort index.

Longitudinal Analysis

Table C2

Change of response rate for countries that increased their fieldwork efforts

Country	Round	ΔFEI	ΔRR
Denmark	3	3	-13.4
Ireland	3	1	-11.6
Finland	3	1	-6.3
Slovenia	4	1	-6.1
Belgium	5	1	-5.6
Slovenia	3	1	-5.3
United Kingdom	2	2	-4.4
Netherlands	2	1	-3.5
Sweden	4	1	-3.3
Ireland	2	1	-2.4
Belgium	4	1	-2.0
Spain	4	1	0.6
Norway	2	1	1.2
Denmark	5	2	1.3
Portugal	3	1	1.4
Switzerland	3	1	1.5
United Kingdom	5	1	1.8
Germany	3	1	1.9
France	3	3	2.4
Finland	4	1	4.0
Netherlands	5	1	10.2
Ireland	5	1	10.6
Switzerland	2	2	16.0

Note. ΔFEI = Change in fieldwork effort index; ΔRR = Change in response rate.

Table C3

Change of response rate for countries that decreased their fieldwork efforts

Country	Round	ΔFEI	ΔRR
Denmark	2	-1	-4.2
Sweden	2	-1	-3.6
Finland	2	-1	-2.6
Germany	2	-1	-0.7
Hungary	3	-2	-0.6
Slovenia	2	-1	-0.3
United Kingdom	4	-1	2.4
Denmark	4	-1	2.8
Portugal	4	-1	3.1
France	4	-1	3.9

Note. ΔFEI = Change in fieldwork effort index from the previous to the last round; ΔRR = Change in response rate from the previous to the last round.

Table C4

Change of response rate for countries with constant fieldwork efforts

Country	Round	ΔFEI	ΔRR
Sweden	5	0	-10.4
Finland	5	0	-9.0
Portugal	5	0	-8.7
Hungary	4	0	-4.7
Norway	4	0	-4.0
Poland	3	0	-3.7
France	5	0	-2.8
Hungary	2	0	-2.7
Norway	5	0	-1.9
Norway	3	0	-1.8
Ireland	4	0	-1.2
Poland	5	0	-1.0
Hungary	5	0	-0.6
Belgium	3	0	-0.2
Sweden	3	0	0.1
France	2	0	0.5
Poland	4	0	0.9
United Kingdom	3	0	1.5
Poland	2	0	1.6
Spain	5	0	1.8
Portugal	2	0	2.5
Slovenia	5	0	5.5

Note. ΔFEI = Change in fieldwork effort index from the previous to the last round; ΔRR = Change in response rate from the previous to the last round.

Different Imputation for the Longitudinal Qualitative Within-Country Analysis

Germany

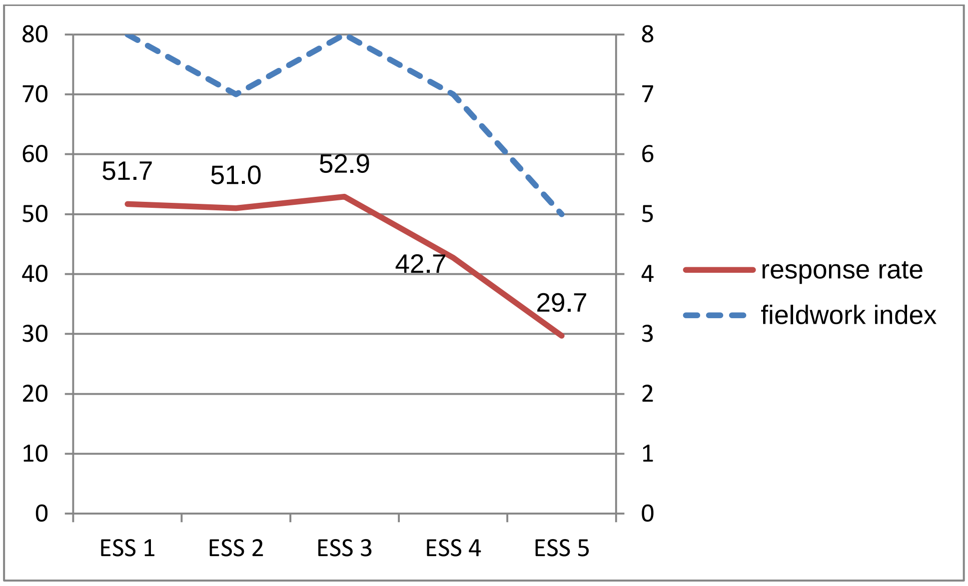

Figure C2. Germany— Mean level of fieldwork effort index and mean response rate per round (in %). Imputation of missing value in round 4 for the variable “interviewer experience”.

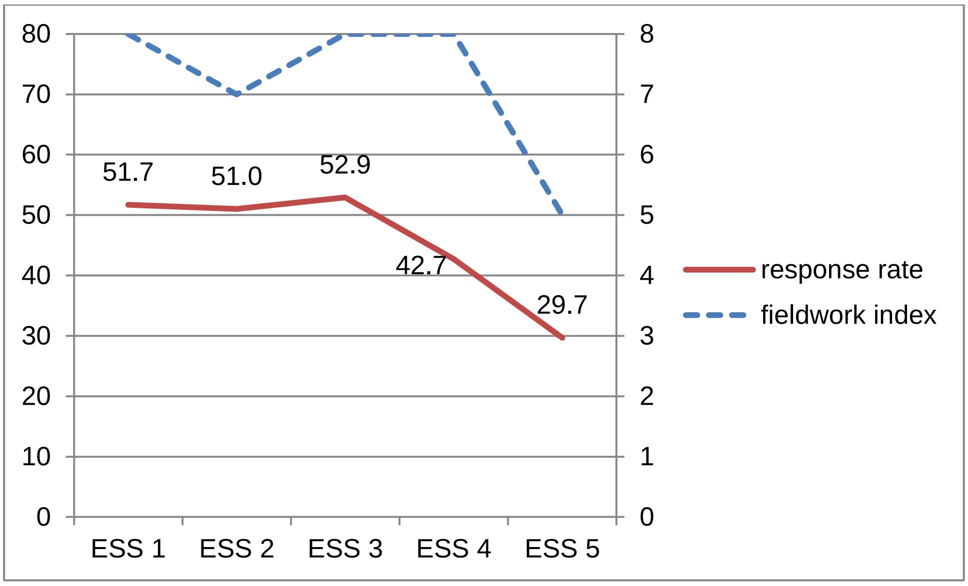

Figure C3. Germany—Mean level of fieldwork effort index and mean response rate per round (in %). Imputation of missing value in round 4 for the variable “interviewer experience”.

Spain

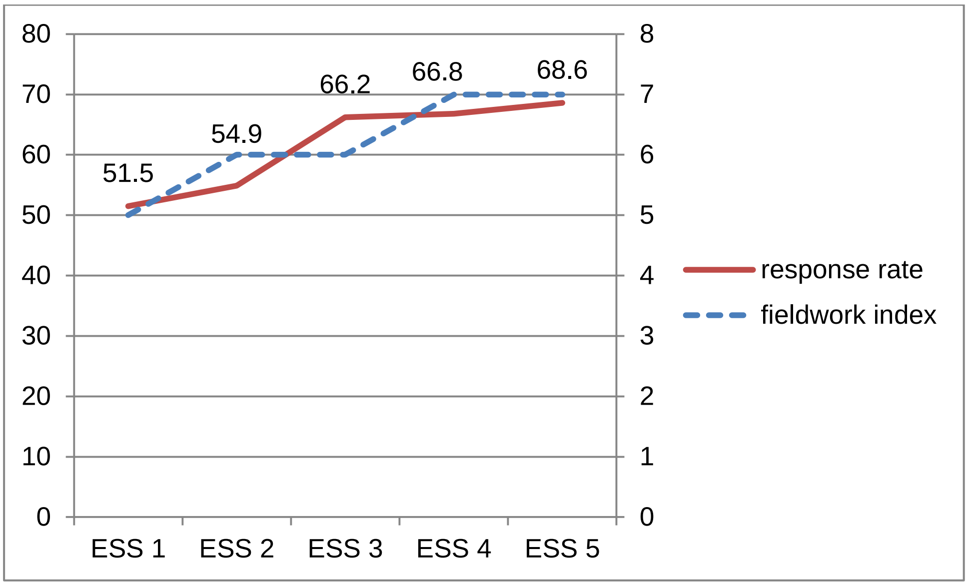

Figure C4. Spain—Mean level of fieldwork effort index and mean response rate per round (in %). Imputation of missing value in round 2 for the variable “interviewer experience

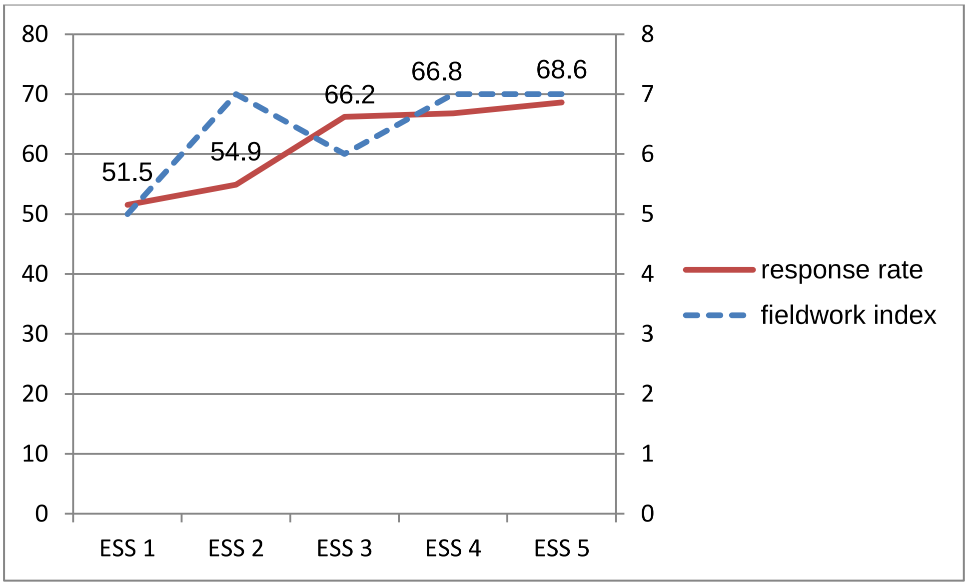

Figure C5. Spain—Mean level of fieldwork effort index and mean response rate per round (in %). Imputation of missing value in round 2 for the variable “interviewer experience”.

Appendix D: Analysis Part D

The Appendix D of Analysis Part D describes and displays additional analyses and graphs. Overall, these sections provide additional information for the analyses carried out in Section 23 for additional analysis of relative bias and fieldwork effort index. The first three sections of Appendix D include outliers which are not included in the analysis in chapter “Analysis Part D”. In addition, a further analysis of the education variable is provided in the last section of Appendix D.

Working Population

In the following graphs, the outliers Sweden (SE), Ireland (IE), Portugal (PT), and Slovenia (SI) were excluded The inclusion of these countries does not change the interpretation.

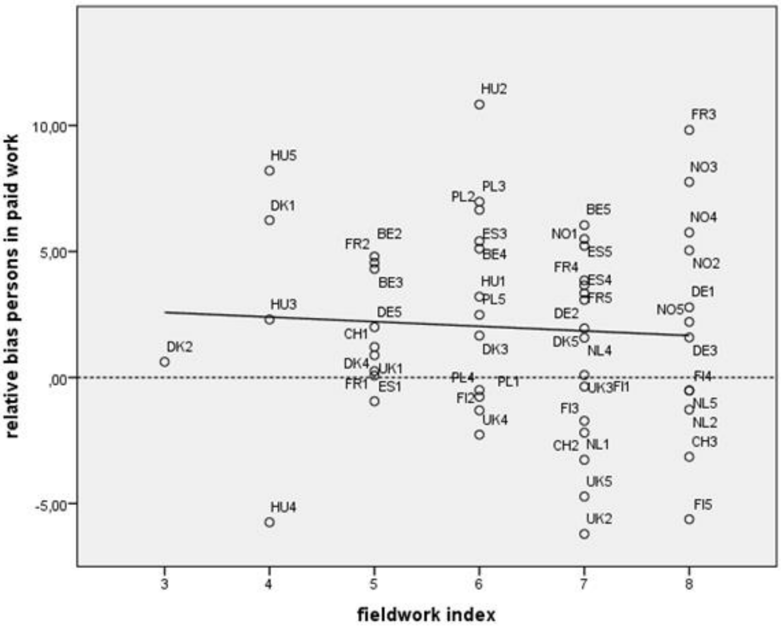

Figure D1. Relative bias for “working population” and the fieldwork index, excluding Portugal, Sweden, Ireland, and Slovenia.

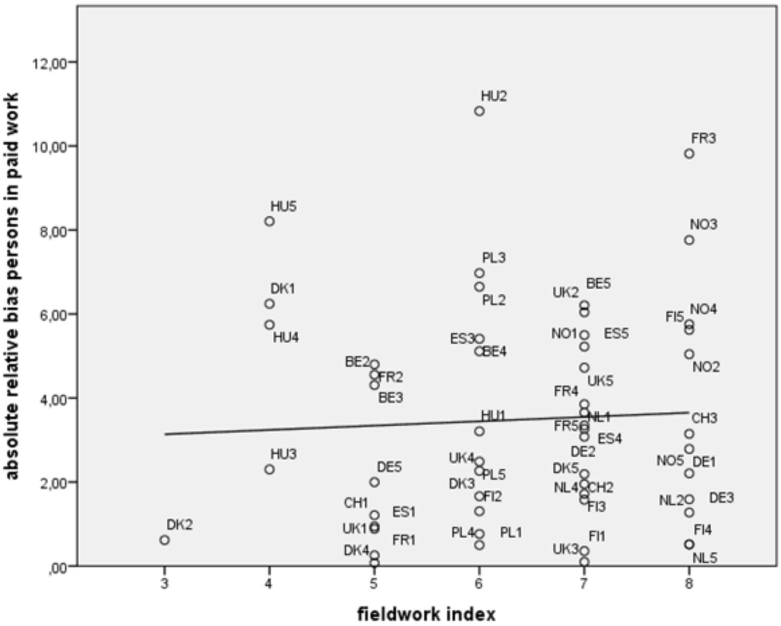

Figure D2. Absolute value for the relative bias for "working population" and the fieldwork index, excluding Portugal, Sweden, Ireland, and Slovenia.

Low Education

In Figure D3 and Figure D4 the relative bias of the variable "low education" is displayed including the outlier United Kingdom (UK).

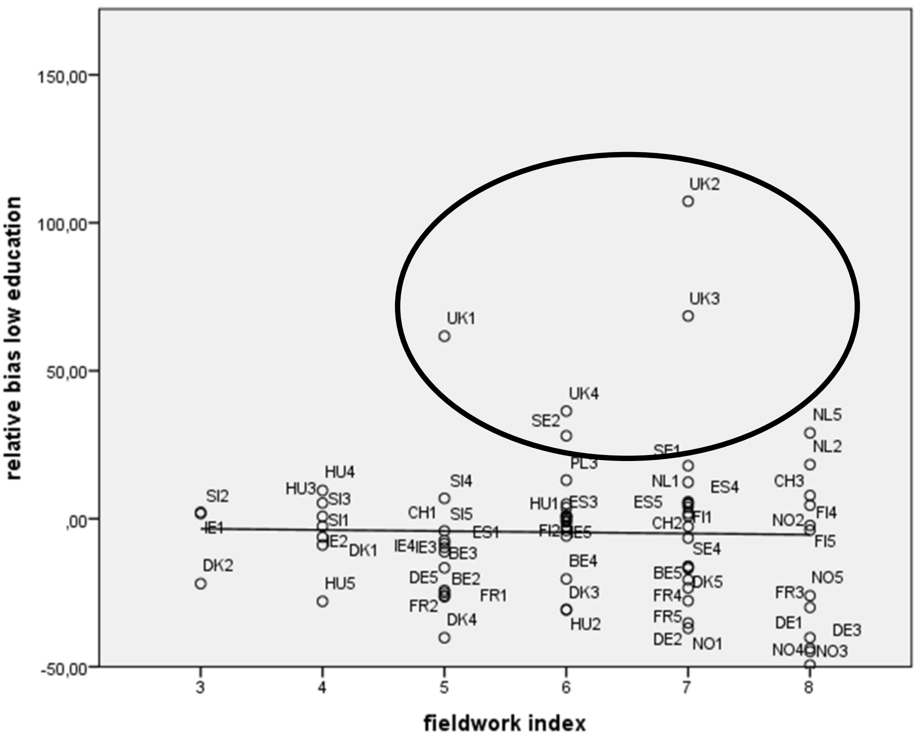

Figure D3. Relative bias for "low education" and the fieldwork index, including the United Kingdom.

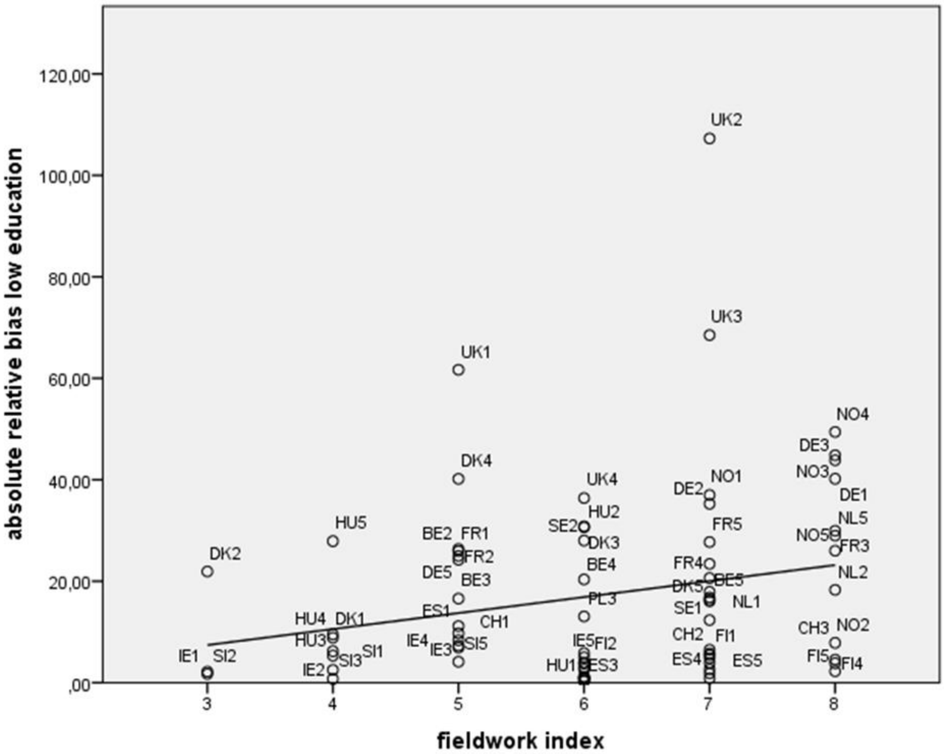

Figure D4. Absolute value for the relative bias for “low education” and the fieldwork index, including the United Kingdom.

Mid Education

In Figure D5 and Figure D6 the relative bias of the variable “mid education” is analyzed. In Analysis Part D, the variable “low education” and “high education” are analyzed. The category of “mid education” supplements this analysis.

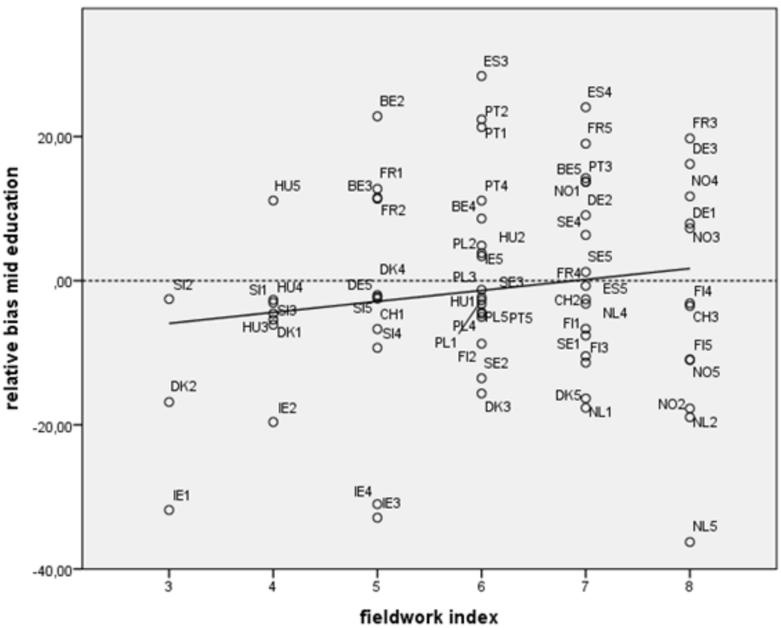

Figure D5. Relative bias for “mid education” and the fieldwork index (r = .15; p = .220; n = 68).

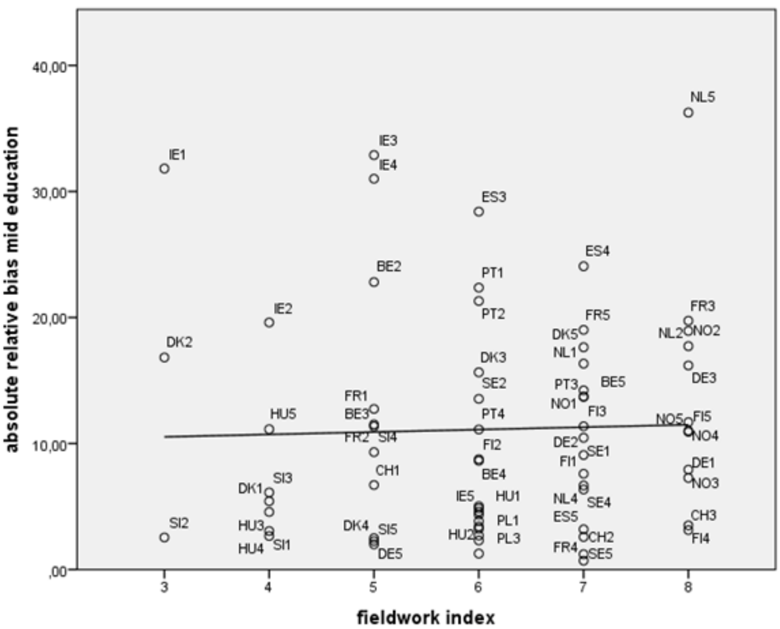

Figure D6. Absolute value for the relative bias for “mid education” and the fieldwork index (r = .03; p = .802; n = 68).

Five-or-More Person Households

The following graphs show the relative and absolute bias for the variable “5-and-more person households” and fieldwork efforts index (Figure D7 and Figure D8). In comparison to the analyses in Analysis Part D, Section 23, the outlier Slovenia (SI) is excluded.

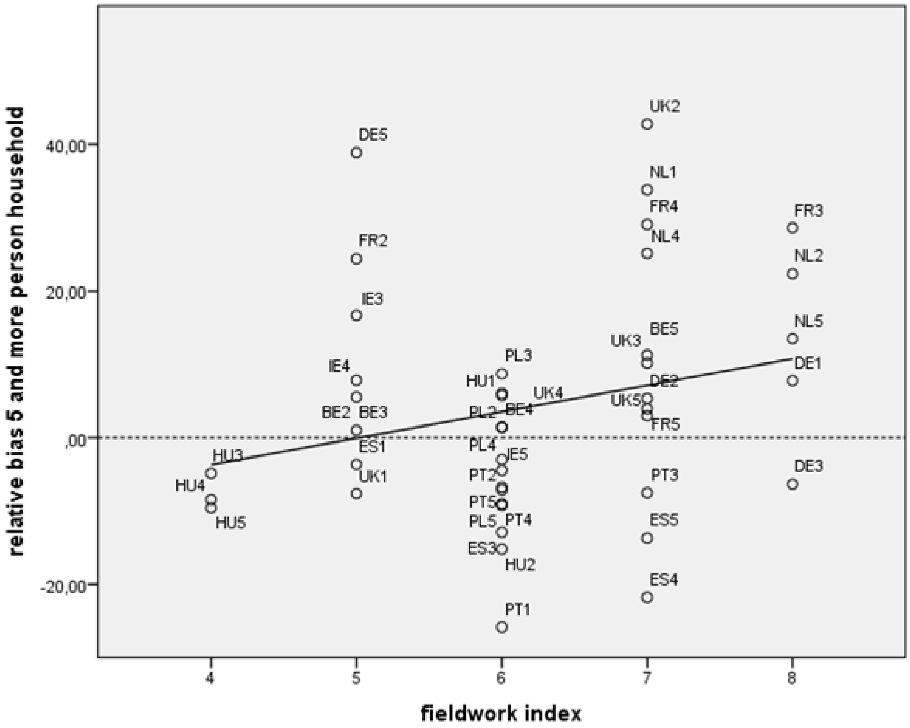

Figure D7. Relative bias for “five-or-more person households” and the fieldwork index, not including Slovenia (r = .25; p = .111; n = 42).

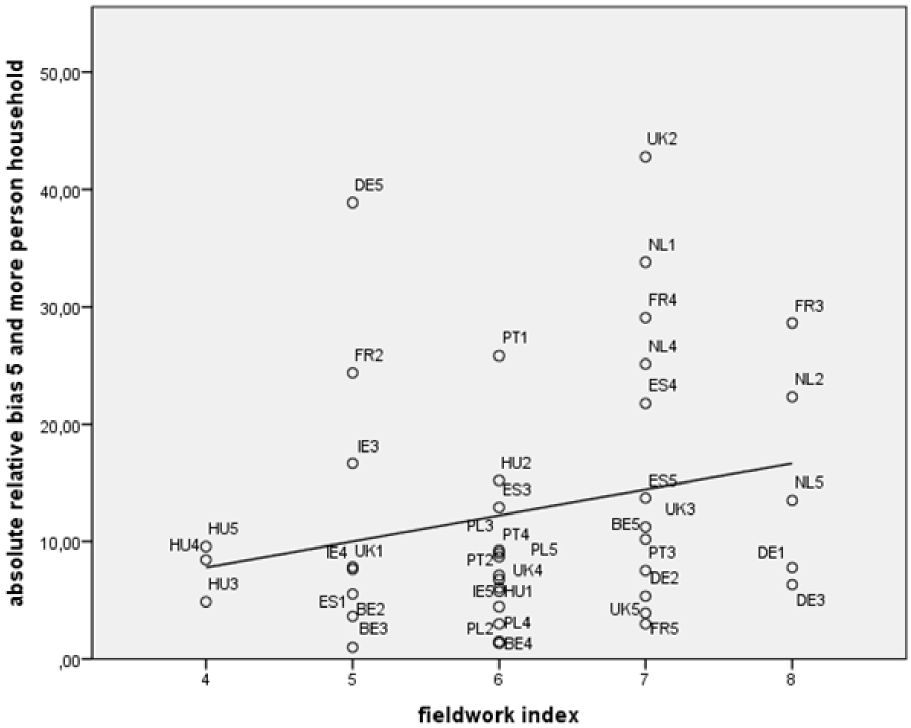

Figure D8. Absolute value for the relative bias for “five-or-more person households” and the fieldwork index, not including Slovenia (r = .23; p = .137; n = 42).

Appendix E: Literature Review

$$\text{REF1} = \frac{\text{R}}{(\text{I} + \text{P}) + (\text{R} + \text{NC} + \text{O}) + (\text{UH} + \text{UO})}$$

Refusal Rate 1 (REF1) is the number of refusals divided by the interviews (complete and partial) plus the non-respondents (refusals, non-contacts, and others) plus the cases of unknown eligibility.

$$\text{REF2} = \frac{\text{R}}{(\text{I} + \text{P}) + (\text{R} + \text{NC} + \text{O}) + \text{e}(\text{UH} + \text{UO})}$$

Refusal Rate 2 (REF2) includes estimated eligible cases among the unknown cases similar to Response Rate 3 (RR3) and Response Rate 4 (RR4) above.

$$\text{REF3} = \frac{\text{R}}{(\text{I} + \text{P}) + (\text{R} + \text{NC} + \text{O})}$$

Refusal Rate 3 is analogous to Response Rate 5 (RR5) and Response Rate 6 (RR6) above. As in those cases the elimination of the unknowns from the equation must be fully justified by the actual situation. Non-contact and other rates can be calculated in a manner similar to refusal rates. Refusal, non-contact, and other rates will sum to equal the non-response rate.

Figure E1. Refusal rates definitions by the American Association for Public Opinion Research (2015, p. 55).

$$\text{CON1} = \frac{(\text{I} + \text{P}) + \text{R} + \text{O}}{(\text{I} + \text{P}) + \text{R} + \text{O} + \text{NC} + (\text{UH} + \text{UO})}$$

Contact Rate 1 (CON1) assumes that all cases of indeterminate eligibility are actually eligible.

$$\text{CON2} = \frac{(\text{I} + \text{P}) + \text{R} + \text{O}}{(\text{I} + \text{P}) + \text{R} + \text{O} + \text{NC} + \text{e}(\text{UH} + \text{UO})}$$

Contact Rate 2 (CON2) includes in the base only the estimated eligible cases among the undetermined cases.

$$\text{CON3} = \frac{(\text{I} + \text{P}) + \text{R} + \text{O}}{(\text{I} + \text{P}) + \text{R} + \text{O} + \text{NC}}$$

Contact Rate 3 (CON3) includes in the base only known eligible cases.

Figure E2. Contact rates definitions by the American Association for Public Opinion Research (2015, p. 56).

Eidesstattliche Erklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig angefertigt und mich keiner anderen als der in ihr angegebenen Hilfsmittel bedient zu haben. Insbesondere sind sämtliche Zitate aus anderen Quellen als solche gekennzeichnet und mit Quellenangaben versehen.

Mannheim, den 27. Juli 2016

Verena Halbherr