

Journal of Official Statistics, Vol. 29, No. 4, 2013, pp. 609-622, http://dx.doi.org/10.2478/jos-2013-0042

# **Internet Coverage and Coverage Bias in Europe: Developments Across Countries and Over Time**

Anja Mohorko<sup>1</sup>, Edith de Leeuw<sup>2</sup>, and Joop Hox<sup>2</sup>

To estimate the coverage error for web surveys in Europe over time, we analyzed data from the Eurobarometer. The Eurobarometer collects data for the European Community across member and applicant states. Since 2005, the Eurobarometer has contained a straightforward question on Internet access. We compared respondents with and without Internet access and estimated coverage bias for demographic variables (sex, age, length of education) and sociopolitical variables (left-right position on a political scale, life satisfaction). Countries in Europe do differ in Internet penetration and resulting coverage bias. Over time, Internet penetration dramatically increases and coverage bias decreases, but the rate of change differs across countries. In addition, the countries' development significantly affects the pace of these changes.

*Key words:* Web survey; Internet; coverage; coverage bias; nonsampling error; Eurobarometer.

# 1. Introduction

Modern society relies on reliable and valid survey data, and almost every country in the world uses surveys to estimate important statistics, such as rate of unemployment, health indicators, opinions about the government and key issues in society, intention to vote in the coming elections, and people's satisfaction with services. Surveys are also one of the most common methods in the social sciences used to understand the way societies work and to test theories.

The last decennium has been marked by fast-paced technological changes that influence survey methods and survey quality. A dramatic change in survey methodology was caused by the development of Internet surveys (Bosnjak et al. 2006; Couper 2000). Internet surveys have many advantages, such as low costs, timely data, and more privacy due to self-completion. The latter is especially important when sensitive topics are being surveyed, and mode comparisons consistently show that Internet surveys give rise to less social desirability than interviews (e.g., Kreuter et al. 2008; Link and Mokdad 2005; for an

<sup>&</sup>lt;sup>1</sup> Department of Social Informatics, Faculty of Social Sciences, University of Ljubljana, Kongresni trg 12, 1000, Ljubljana, Slovenia. Email: anja.mohorko@fdv.uni-lj.si

<sup>&</sup>lt;sup>2</sup> Department of Methodology and Statistics, Utrecht University. Plantage Doklaan 40, 1018 CN Amsterdam, the Netherlands. Email: edithl@xs4all.nl and j.hox@uu.nl

Acknowledgments: The authors thank Bill Blyth for his stimulating ideas and kind support throughout this project. We also thank Mick Couper, two anonymous reviewers, and the editors of JOS for their helpful suggestions.

overview see De Leeuw and Hox 2011). In this sense, Internet surveys are indeed more like self-administered questionnaires and share their benefits, as Couper (2008) postulated.

From the onset of Internet surveys, coverage error has been a source of major concern. A main problem with Internet surveys is under-coverage resulting from the "digital divide", that is, a difference in rates of Internet access among different demographic groups (such as an unequal distribution regarding age and education for those with and without Internet access; see Couper 2000). Although Internet coverage is growing – for instance for Europe as a whole, Internet coverage increased from 15% in December 1999 to approximately 63% in June 2012 (Internet World Stats 2013) – it varies widely across countries. For example, at the beginning of the 21st century almost 15% of Europeans had Internet access, but according to the World Bank (2009) this ranged from less than 4% (e.g., Romania and Turkey) to 44% and 46% (the Netherlands and Sweden). For a more detailed overview, see Blyth (2008). This differential coverage would not be a problem if the covered part represented the general population with respect to important survey variables. However, even in countries with a high coverage a digital divide can be observed, as Internet access is unevenly distributed over the population, with highly educated and younger persons more often having an Internet connection (e.g., Bethlehem and Biffignandi 2012; Rookey et al. 2008; Couper et al. 2007). This differential coverage over countries and demographic groups may result in biased estimates of substantive variables of interest in a study. To estimate the coverage bias, one needs data on both parts of the population, that covered and that not covered.

In terms of coverage of the household population, face-to-face interviews are often viewed as the gold standard to which other modes are compared (e.g., Groves et al. 2009). Since 2005, the Eurobarometer, which is based on face-to-face interviews, contains a question about Internet access at home. This provides us with a unique data set to analyze Internet coverage and coverage bias across European countries and over time. How would substantive results change if important international studies like the Eurobarometer used Internet surveys instead of the (golden) standard face-to-face interviews? As data collection in the Eurobarometer does not depend on respondents having access to the Internet, the survey mode is held constant, and as the same battery of questions is asked over time and across countries, this data set enables us to investigate how potential *coverage* bias could influence the results if the data had been collected using Internet surveys instead of face-to-face interviews. In other words, this gives us an indication of Internet coverage and coverage bias over time and across countries.

In this study, we compare those with access to Internet at home to the whole target group of Eurobarometer face-to-face interviewees (both with and without Internet access at home). It is expected that the coverage bias between the two groups differs between countries and will decrease over time for all countries. We also expect that the rate of decrease may be different in different countries and that social and economic indicators at the country level may explain some of these differences.

In the following sections, we first describe the available data and the analysis methods used. We then present our results on trends in Internet coverage at home and the resulting coverage bias for available demographic variables and sociopolitical variables. This is followed by a multilevel analysis to model the changes over time and the influence of socioeconomic development on these trends. We end with a critical discussion and implications for research.

## 2. Method

# 2.1. Available Data

#### 2.1.1. Eurobarometer

The Eurobarometer collects data for the European Community across EU members and applicant countries four to eight times a year. The Eurobarometer has a separate data collection for East and West Germany, the Republic of Cyprus and the Turkish Republic of Northern Cyprus, and Great Britain and Northern Ireland. Therefore, the following 32 countries were included in the analyses: Austria, Belgium, Bulgaria, Croatia, Cyprus (Republic and TCC), Czech Republic, Denmark, Estonia, Finland, France, Germany (East and West), Great Britain, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Northern Ireland, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and Turkey. Since 2005, the Eurobarometer contains a yearly question about Internet access at home.

Each wave of the Eurobarometer consists of face-to-face interviews and includes a core questionnaire plus an additional questionnaire with special topics. For each standard Eurobarometer survey, new and independent samples are drawn; since October 1989, the basic sampling design has been a multi-stage probability sample. To ensure the total coverage of each country, the sampling in the first stage is based on a random selection of sampling points (PSU) after stratification by the distribution of the national, resident population in terms of metropolitan, urban, and rural areas, that is, proportional to the population size. Within the PSUs addresses are then selected using random route procedures, followed by a random selection of a person at the address (for more details on sampling and coverage, see GESIS Eurobarometer Survey series 2013).

Every household survey suffers from nonresponse (Bethlehem et al. 2011; De Leeuw and De Heer 2002; Groves and Couper 1998), and the Eurobarometer is no exception. Unfortunately, there is no detailed information on response rates made available publicly and on a regular basis by the principal investigator, the European Commission's Eurobarometer unit. Still, there is some indication that response rates vary between countries. For instance, Busse and Fuchs (2012) note that for the 2002 Eurobarometer, response rates varied between rates of around 70% for East and West Germany and 40% or less for Ireland, Denmark and the UK. No systematic nonresponse studies are available. However, the Eurobarometer data do include integrated design and poststratification weights to adjust the realized samples to EUROSTAT population data (Moschner 2012). These weights will be used in estimating the coverage bias indicators.

The core questionnaire contains trend questions about sociopolitical orientation and standard demographic questions and, since 2005, also includes a question on having an Internet connection at home, allowing us to estimate Internet access at home and the resulting coverage bias. Besides Internet access at home, interview data on the following variables were available for all countries: sex, age, length of education, political left-right self-placement and life satisfaction (see Mohorko et al. 2011 for the question wording

used); also the year of data collection was recorded. All the data were downloaded in February and March 2011, at which point the Eurobarometer data were fully available for the years 2005 to 2009. Hence, our analysis will cover this five-year period. To assess coverage bias, we analyze three demographic variables: sex, age, and length of education, and two substantive variables: political left-right self-placement and life satisfaction. The demographic variables age, sex, and education are seen as important indicators for the digital divide (e.g., Couper 2000) and correlate with many substantive variables typically assessed in academic or market research surveys (Fuchs and Busse 2009). The substantive variables political left-right self-placement and life satisfaction give us an opportunity to directly investigate the influence of undercoverage on the assessment of two major sociopolitical indicators.

# 2.1.2. Additional Country-level Variables

The data from the Eurobarometer are individual level data, collected through face-to-face interviews in each country. Apart from Internet penetration, the countries involved in the Eurobarometer also differ on socioeconomic variables, which may influence Internet coverage. To model this, we collected socioeconomic country-level data from Eurostat, the World Bank, and the Human Development Report. Contextual country-level variables are: life expectancy at birth (in years), country's educational index, duration of primary and secondary education (in years), and urbanization (the percentage of urban population). Economic indices on country level are the percentage of employed (labor force), the Gini coefficient (which measures income inequality), Gross Domestic Product growth (GDP), and inflation. For a description of these variables and the data sources including the URL, see Mohorko et al. (2011, 2013). It should be noted that these variables are measured at the country level, but they are available for each year, hence they are time-varying predictors.

## 2.2. Analysis

# 2.2.1. Coverage and Indicators of Coverage Bias

Coverage is defined as the percentage of the population of interest that is included in the sampling frame; ideally the coverage should be 100%. Furthermore, there should be a one-to-one correspondence between the population of interest or target population and the (sampling) frame population. If this is not the case, *and* if those missing in the frame differ from the target population on a key variable of interest in the study, coverage error occurs (Biemer and Lyberg 2003; Groves et al. 2009). Groves (1989, p. 11) describes coverage error as follows: "Coverage error exists because some persons are not part of the list or frame (or equivalent materials) used to identify members of the population. Because of this they *never* can be measured whether a complete census of the frame is attempted or a sample studied."

Undercoverage is one of the main concerns for the validity of conclusions based on Internet surveys (Couper 2000). Although Internet access is growing, there are still many individuals who are not covered, and if those without Internet access differ on key measures from those with Internet access, the resulting estimators will be biased. For example, if wealthier households are more likely to have Internet access, then a survey about household assets that is based exclusively on the Internet will produce income estimates that are too high (Lohr 2008).

To investigate *coverage* problems in Internet-based surveys, we compare the responses of the subgroup of Internet-at-home with those of the total group of Eurobarometer respondents. Since the Eurobarometer was conducted face-to-face in all countries and face-to-face surveys have the least coverage problems (Groves et al. 2009, p. 163; De Leeuw 2008, p. 125), the total Eurobarometer group in this study is regarded as a proxy for the target population. Differences between those with an Internet connection at home and the total Eurobarometer group give an indication of the bias due to *undercoverage* if an Internet survey had been implemented instead of a face-to-face survey.

The net coverage bias is defined by Lessler and Kalsbeek (1992, p. 59-60) as

$$\bar{y}_{\text{covered}} - \bar{y}_{\text{target}} = \frac{N_{\text{not covered}}}{N_{\text{target}}} (\bar{y}_{\text{covered}} - \bar{y}_{\text{not covered}})$$
(1)

which is used by Bethlehem and Biffignandi (2012, p. 289) to define bias due to the non-Internet population. Based on Equation (1), we use two indices to assess the amount of coverage bias: the relative bias (Lessler and Kalsbeek 1992, p. 60) and the absolute relative bias (Groves and Peytcheva 2008). The relative coverage bias is used for descriptive purposes, as the sign of this estimate indicates the over- or undercoverage of specific groups (e.g., if more men than women have Internet access at home in a certain year and in a certain country). However, when modeling changes occur over time and across countries, positive and negative values for relative coverage can cancel each other out and the resulting regression coefficients may falsely give the impression that the overall coverage error is close to zero. Therefore, we use the absolute relative coverage bias in our multilevel analyses.

The relative and absolute relative coverage bias due to lack of Internet access are defined as

relative coverage bias 
$$= \frac{\bar{y}_{Int} - \bar{y}_{EB}}{\bar{y}_{EB}}$$
 (2)

and

absolute relative coverage bias = 
$$\left| \frac{\bar{y}_{Int} - \bar{y}_{EB}}{\bar{y}_{EB}} \right|$$
 (3)

where *EB* represents the total achieved Eurobarometer sample, which is viewed as our target population, and *Int* represents the covered Internet subsample. Analogous  $\bar{y}_{EB}$  and  $\bar{y}_{Int}$  represent the means of the Eurobarometer target population and the Internet subsample on the variable *y*.

#### 2.2.2. Statistical Analyses

The relative coverage bias is used for descriptive analyses over countries and time. Positive values indicate that surveys, which are exclusively conducted through the Internet, will result in estimates that are too high, whereas negative values indicate that these will result in estimates that are too low. Multilevel analysis on the absolute relative coverage bias is used to model and explain trends over time and country for all bias indicators (sex, age, length of education, political left-right self-placement and life satisfaction). For ease of interpretation, the absolute relative coverage bias is expressed as percentage points. In the multilevel model, the lowest level represents the years, indicated by a time variable coded 2005 = 0, 2006 = 1, et cetera. To estimate change over time, we analyze a null model that always includes the linear effect of time and tests whether the variance component for the slope of time is significant. If this random component is not significant using a likelihood ratio test, it is removed from the null model. Since the plots for the effect of time in Figure 1 indicate possible nonlinearity, we test for nonlinear effects by analyzing the quadratic effect of time. If the quadratic term is not significant at the conventional 5% level, it is removed from this model; the linear term for time is always retained in the null model.

In a second step, we add country-level socioeconomic variables. Country-level variables model initial differences in bias between countries in the starting year 2005. Since the country-level variables vary across time, they may also explain change over time. Because the country-level variables are correlated with time, adding them to the model may replace (part of) the explanatory power of the time variable as estimated in the null model.

Finally, differences between countries in the rate of change over the years, as indicated by variation in the slopes of the time variable, are modeled as interactions of country-level



Fig. 1. Internet access at home across Europe 2005–2009, based on the Eurobarometer's weighted data. The lines represent the 32 countries/regions distinguished in the Eurobarometer

variables with the time variable. Again, effects that are not significant are removed from the model. A two-sided significance level of alpha = 0.05 is used throughout.

# 3. Results

#### 3.1. Coverage Bias in European Countries

Internet access at home increases over time across Europe, but the rate of increase differs across countries (see Figure 1). The actual proportions per country and per year are presented in Appendix A. These numbers show that for countries with an initial low Internet penetration, for example Bulgaria and Romania, the proportions increase rapidly, while for countries with an initial high penetration, for example Sweden and the Netherlands, the growth is less steep.

But even with an Internet penetration above 80%, there still may be considerable differences between those with and without Internet access. This is indicated by the relative coverage bias, which is based on the standardized difference between the subgroup of those who do have Internet at home compared to the total (Internet and non-Internet at home) group. Full descriptive tables with the values of the relative coverage bias for each country in the Eurobarometer and each year are available in Mohorko et al. (2011).

For the demographic variables sex, age, and length of education, the descriptive tables indicate a digital divide. In Europe, those with Internet at home are more often male, younger, and highly educated (Mohorko et al. 2011, Appendix D, Tables D1-D3); similar patterns have been found in the USA (cf. Couper 2008). The bias for sex is relatively low and decreases strongly over time. The highest value was found for Greece with 8.5% more men than women having Internet access in 2006, which decreased to 5.5% in 2009. The lowest values (less than 1% more men) were found for countries like Sweden, Slovenia, Ireland, and the Netherlands in 2009. In general, the gender gap is closing very fast over time. Furthermore, the age difference is becoming smaller over time; younger people are still overrepresented, but for some countries (e.g., Sweden and the Netherlands) the age bias is really low (around -0.04) in 2009, while for others (e.g., Bulgaria) it is still rather high (-0.22 in 2009). The same can be seen for length of education. It should be noted that countries with the smallest digital divide regarding the demographics of age, sex, and education are also the countries with the highest Internet penetration. This gives an optimistic outlook for the future that as Internet penetration increases, the digital divide will decrease.

When we take a closer look at the descriptive tables for the substantive variables political left-right self-placement and life satisfaction (for the detailed tables per country over the years, see Mohorko et al. 2011, Table D4 and D5), we again note that the differences are becoming smaller over time. On average, the coverage bias is very low for political left-right self-placement, where its bias decreases towards zero over time with the largest differences found in Bulgaria (from 0.23 in 2005 to 0.075 in 2009). It should be noted that the coverage bias for this variable does not take the same direction in all countries. For some countries, those with an Internet connection at home place themselves more on the left (e.g., Austria, West-Germany, Great Britain), for other countries they place themselves more on the right (e.g., Bulgaria). For the second substantive variable life

satisfaction, we see that in every country and every year there is a positive bias, indicating that those with Internet at home are more satisfied with life than the Eurobarometer population in general. This bias decreases slightly over time.

# 3.2. Changes in Coverage Bias Over Time

The change in coverage bias over time is analyzed using multilevel analysis, with years (coded  $2005 = 0, \ldots, 2009 = 4$ ) nested within countries. This allows us to test whether the change over time is significant and to test if country-level variables can predict changes over time. The analysis showed that the effect of time squared was never significant, and therefore only the linear trend of time is included in the model. Table 1 presents the parameter estimates for each dependent variable for two models: a model with only the linear time indicator and a model with the time indicator and the significant country variables.

When we examine the effect of time in the first model, the results show a steady decrease in absolute relative coverage bias across time, as indicated by a negative value for the regression coefficient of time, except for political left-right self-placement where the overall effect of time is not significant. For all five bias indicators, Table 1 shows a significant and sometimes large country-level variance, which means that there were clear differences in overall bias between countries in 2005. For three out of five bias indicators, the time variable has a significant slope variation (indicated in Table 1 under "time slope variance"), which means that the biases for "age", "political left-right self-placement" and "life satisfaction" decrease at different rates across countries. Compared to the size of the regression coefficient for the time variable itself, these variances are relatively large. This indicates large differences in the rate of decline between countries for these bias indicators.

## 3.3. Coverage Bias and Country Differences

There are differences between countries in the size of the coverage bias and, for some variables, in the rate of the decrease of this bias over time. These differences are modeled by the direct effects of the available country-level variables: life expectancy, educational index, duration of primary and secondary education, urbanization, employment, Gini index, GDP growth rate, and inflation. The differences in rate of decrease are modeled by the interactions of these variables with the time indicator.

The explanatory variables secondary education, GDP growth rate, and inflation were never significant and are omitted from the model. Table 1 shows the estimated multilevel model and the significant regression coefficients for each of the five coverage bias indicators. The bias for political left-right placement could not be predicted by any of the available country variables. The other four coverage bias indicators can be predicted by different subsets of country-level variables. Thus differences between persons with and without Internet across countries can be predicted using different country-level variables.

Table 1 shows that coverage bias for age is higher in countries with a high income inequality as indicated by the Gini-coefficient, while coverage bias for age is lower in countries with a higher educational index, a higher life expectancy, longer duration of primary school education, and high urbanicity. In contrast, coverage bias for sex is only

	Bias Sex Composi	ition	Bias Age Composit	tion	Bias Length of Edu	cation	Bias Political Left- Self-Placemer	Right 1t	Bias Life Satisfactio	u
Model	Year (2005–09)	Country predictors	Year (2005–09)	Country predictors	Year (2005–09)	Country predictors	Year (2005–09)	Country predictors	Year (2005–09)	Country predictors
Fixed part Intercept Time	b (se) 2.87 (.26) - 0.27 (.08)	b (se) 0.10 (1.34) -0.27 (.08)	b (se) 16.10 (.83) - 0.60 (.11)	b (se) 82.69 (16.07) -0.34 (.11)	b (se) 10.11 (.97) -0.78 (.11)	b (se) 62.91 (12.03) - 0.69 (.17)	b (se) 2.83 (.74) 0.05 (.16) <sup>ns</sup>	b (se) 3.02 (.41) -	b (se) 5.55 (.92) -0.35 (.13)	b (se) 84.60 (12.46) -
Country variances Gini coefficient Educational index Employment Life expectancy Primary		0.09 (.04) - - -		0.23 (.08) -0.21 (.08) - -0.59 (.19) -0.92 (.44)		- - 0.30 (.12) - 3.82 (1.58) -		1 1 1 1 1		- 4.26 (1.23) - 0.66 (.17)
school dur. Urbanicity <b>Random part</b> Residual variance Country variance variance	s <sup>2</sup> (se) 1.96 (.25) 0.98 (.35)	- s <sup>2</sup> (se) 1.96 (.25) 0.80 (.32)	s <sup>2</sup> (se) 2.13 (.31) 20.98 (5.66) 0.15 (.09)	$\begin{array}{c} -0.10 \ (.05) \\ \mathrm{s}^2 \ (\mathrm{se}) \\ 2.45 \ (.32) \\ 7.64 \ (2.26) \end{array}$	s <sup>2</sup> (se) 3.98 (.50) 27.91 (7.29)	$\begin{array}{c} -0.15 \ (.06) \\ \mathrm{s}^2 \ (\mathrm{se}) \\ 1.84 \ (.33) \\ 22.69 \ (6.68) \\ 0.45 \ (.23) \end{array}$	s <sup>2</sup> (se) 2.80 (.40) 15.64 (4.40) 0.51 (.21)	- s <sup>2</sup> (se) 2.80 (.40) 15.30 (4.26) 0.49 (.20)	s <sup>2</sup> (se) 2.77 (.40) 25.39 (6.89) 0.26 (.15)	-0.1 (.04) s <sup>2</sup> (se) 3.57 (.53) 6.67 (2.15)

Mohorko, De Leeuw, and Hox: Internet Coverage and Coverage Bias in Europe

"-" indicates parameter tested but removed because parameter was not significant at a two-sided alpha of 0.05.

<sup>ns</sup> indicates non-significant coefficient for time.

associated with the Gini coefficient; coverage bias for sex is higher in countries with high income inequality (high Gini). Coverage bias in length of education is lower in countries with a higher educational index, a higher employment level, and a higher urbanicity. Coverage bias in life satisfaction is lower in countries with a higher employment rate, higher life expectancy and high urbanicity.

There were no significant interactions with time, meaning that the available countrylevel variables do not predict the differences in the rate of bias decrease. When we compare the model with country variables added to the model with only time as predictor, an interesting pattern emerges. For all four bias indicators with a significant effect of time, Table 1 shows that adding country-level variables to the model decreases the size of both the regression coefficient for time and the variance across countries. Thus part of the effect of time is the result of changes over time in country-level variables. The signs of the regression coefficients for the country variables suggest that, in general, coverage bias decreases when education, employment, life expectancy, and urbanicity increase. In other words, differences between persons with and without Internet access decrease when the value of these variables increase. In contrast, the differences between persons with and without Internet access increase when the income distribution is more unequal.

## 4. Conclusion and Discussion

As expected, Internet penetration has increased over time in all countries included in this study. As a result, the absolute relative bias in the estimates of four out of five variables has also decreased; only political left-right self-placement does not show this trend. In other words, differences in age, sex, education, and life satisfaction between those with and without Internet access are diminishing. Multilevel analyses show that for those four bias indicators, the decrease in coverage bias over time differs across countries and that the countries' development affects the pace of this decrease. For age and life satisfaction, the variation in decrease is fully explained by the country-level variables in the model, albeit only partially for sex and education.

The general trend is that higher levels of economic development, education, and health are associated with lower coverage bias, whereas higher income inequality is associated with higher levels of bias. Given the general economic and demographic trends, one conclusion of our study is that coverage bias due to low Internet penetration is disappearing across countries in Europe. The multilevel analyses also show variation across countries in both the initial level and rate of decrease of coverage biases for demographic variables. This shows that the "digital divide" (Couper 2000) not only differs between countries, but also is diminishing at different rates over time in these countries.

Our measure of Internet penetration and coverage bias is based on a question in the Eurobarometer that inquires specifically about Internet access at home. However, there are alternative ways to access the Internet, for instance at work, in libraries, or on mobile devices. For this reason, our analyses are based on the assumption that for surveys that consist of more than a couple of pop-up questions, respondents will prefer to answer in an environment where they have time, feel comfortable, and have privacy. Although mobile Internet is promising, only one third of the population was covered by mobile Internet in Europe in 2007. Furthermore, coverage biases for demographic variables for the mobile

web were larger than for landline Internet (Fuchs and Busse 2009). The use of mobile Internet on telephones and tablet devices is likely to increase further in the near future, which will necessitate a change in the measurement of Internet access. Provided that survey methodologists adapt their surveys to these new devices (e.g., Callegaro 2010), this will not change our conclusion that coverage bias for Internet surveys is decreasing over time.

This study focuses on coverage bias. Good coverage is a necessary but not a sufficient condition for high quality survey data. Other error sources exist, such as nonresponse error or mode effects. Meta-analyses (Cook et al. 2000; Lozar Manfreda et al. 2008) show that Internet surveys yield on average 11% lower response rate than other modes. Clearly, measures should be taken to increase this response rate. For a discussion of such measures we refer to Dillman et al. (2009). Compared to face-to-face interviews, responses to Internet surveys may differ due to mode effects, especially when sensitive topics are addressed. For a discussion, we refer to De Leeuw and Hox (2011), Dillman et al. (2009), and Kreuter et al. (2008).

In our study we treat the data from the face-to-face Eurobarometer samples as a representative sample of the total target population, and our results are conditional on the selection and nonresponse processes in the Eurobarometer. Therefore, in estimating the bias indicators, we used the design and post-stratification weights included in the Eurobarometer data. Nevertheless, nonresponse in the Eurobarometer samples can still affect our results. The use of adjustment weights amounts to treating nonresponse as missing at random (MAR, cf. Little and Rubin 2002). However, if the nonresponse in the Eurobarometer were related to Internet access itself (and were therefore missing not at random or MNAR), there is a potential for nonresponse bias. Hence we view our findings as an indication of a generally decreasing coverage bias in the countries studied, but not as precise estimates of this bias.

A potential alternative data source for a future follow-up study would be the European Social Survey (ESS), which recently added a question on Internet access to the core module. Like all surveys, the ESS also has differential nonresponse across countries, but the ESS response rates and sources of nonresponse are well documented and available for more in-depth analyses (Stoop et al. 2010). Ideally, in some countries it may be possible to validate survey-based information on Internet access with registry data.

In conclusion, even if Internet coverage is not complete, Internet surveys may still compete with other survey modes. For instance, in 2008 the Netherlands had an 86% Internet coverage, while the landline telephone coverage was around 60-70% (Bethlehem et al. 2011, p. 100 and p. 102). The same trend can be seen in other countries; for instance, Smyth and Pearson (2011, p. 16 and p. 17) report that in 2008 the US had an Internet coverage of just over 70%, and random digit dialing landline telephones had a coverage of about 78%. However, landline telephone coverage is decreasing (cf. Busse and Fuchs 2012; Mohorko et al. 2013), while Internet coverage is rapidly increasing over time – as this study shows.

Country\Year	2005	2006	2007	2008	2009	Grand Total
Austria	0.46	0.53	0.56	0.55	0.56	0.53
Belgium	0.55	0.60	0.63	0.63	0.62	0.61
Bulgaria	0.09	0.16	0.21	0.29	0.32	0.21
Croatia	0.32	0.42	0.43	0.41	0.41	0.40
Cyprus Rep.	0.33	0.37	0.43	0.48	0.48	0.42
Cyprus (TCC)	0.27	0.34	0.41	0.55	0.53	0.42
Czech Rep.	0.33	0.42	0.45	0.49	0.53	0.44
Denmark	0.77	0.77	0.81	0.83	0.84	0.80
Estonia	0.44	0.52	0.61	0.63	0.62	0.56
Finland	0.63	0.69	0.75	0.75	0.79	0.72
France	0.42	0.50	0.59	0.57	0.62	0.54
Germany East	0.46	0.44	0.51	0.57	0.61	0.52
Germany West	0.56	0.62	0.60	0.65	0.69	0.62
Great Britain	0.57	0.59	0.63	0.63	0.65	0.61
Greece	0.22	0.26	0.27	0.24	0.30	0.26
Hungary	0.17	0.21	0.27	0.34	0.37	0.27
Ireland	0.50	0.50	0.51	0.64	0.65	0.56
Italy	0.37	0.35	0.38	0.41	0.39	0.38
Latvia	0.16	0.30	0.42	0.50	0.50	0.38
Lithuania	0.17	0.26	0.36	0.37	0.38	0.31
Luxembourg	0.65	0.71	0.69	0.72	0.76	0.71
Malta	0.50	0.52	0.63	0.66	0.66	0.59
Northern Ireland	0.54	0.54	0.56	0.62	0.65	0.58
Poland	0.25	0.31	0.41	0.42	0.48	0.37
Portugal	0.21	0.24	0.33	0.35	0.33	0.29
Romania	0.11	0.19	0.22	0.32	0.33	0.23
Slovakia	0.18	0.26	0.33	0.40	0.43	0.32
Slovenia	0.54	0.54	0.61	0.63	0.64	0.59
Spain	0.33	0.39	0.38	0.43	0.41	0.39
Sweden	0.78	0.85	0.84	0.87	0.88	0.84
The Netherlands	0.83	0.83	0.86	0.88	0.90	0.86
Turkey	0.08	0.15	0.17	0.22	0.25	0.17
Grand Total	0.40	0.45	0.50	0.53	0.55	0.49

# Appendix A: Growth of internet access at home across Europe: 2005–2009 based on the Eurobarometer weighted data for that time period

# 5. References

Bethlehem, J. and Biffignandi, S. (2012). Handbook of Web Surveys. Hoboken, NJ: Wiley.

Bethlehem, J., Cobben, F., and Schouten, B. (2011). Handbook of Nonresponse in Household Surveys. New York: Wiley, Wiley Series in Survey Methodology.

- Biemer, P.P. and Lyberg, L.E. (2003). Introduction to Survey Quality. New York: Wiley, Wiley Series in Survey Methodology.
- Blyth, B. (2008). Mixed-Mode: The Only "Fitness" Regime? International Journal of Market Research, 50, 241–266.

- Bosnjak, M., Forsman, G., Isaksson, A., Lozar Manfreda, K., Schonlau, M., and Tuten, T. (2006). Preface to JOS Special Issue on Web Surveys. Journal of Official Statistics, 22, iii.
- Busse, B. and Fuchs, M. (2012). The Components of Landline Telephone Survey Coverage Bias. The Relative Importance of No-Phone and Mobile-Only Populations. Quality and Quantity, 46, 1209–1225. DOI: http://www.dx.doi.org/10.1007/s11135-011-9431-3
- Callegaro, M. (2010). Do You Know Which Device Your Respondent Has Used to Take Your Online Survey? Survey Practice, December: www.surveypractice.org. Available at http://surveypractice.org/2010/12/08/device-respondent-has-used/ (accessed August 2012).
- Cook, C., Heath, F., and Thompson, R.L. (2000). A Meta-Analysis of Response Rates in Web- and Internet-Based Surveys. Educational and Psychological Measurement, 60, 821–836. DOI: http://www.dx.doi.org/10.1177/00131640021970934
- Couper, M.P. (2000). Web Surveys: A Review of Issues and Approaches. Public Opinion Quarterly, 64, 464–494. DOI: http://www.dx.doi.org/10.1086/318641
- Couper, M.P. (2008). Designing Effective Web Surveys. New York: Cambridge University Press.
- Couper, M.P., Kapteyn, A., Schonlau, M., and Winter, J. (2007). Noncoverage and Nonresponse in an Internet Survey. Social Science Research, 36, 131–148. DOI: http://www.dx.doi.org/10.1016/j.ssresearch.2005.10.002
- De Leeuw, E.D. (2008). Choosing the Method of Data Collection. International Handbook of Survey Methodology, E.D. de Leeuw, J.J. Hox, and D.A. Dillman (eds). New York: Routledge, Taylor & Francis, European Association of Methodology (EAM) Methodology Series.
- De Leeuw, E.D. and De Heer, W. (2002). Trends in Household Survey Nonresponse: A Longitudinal and International Comparison. In Survey Nonresponse, R.M. Groves, D.A. Dillman, J.L. Eltinge, and R.J.A. Little (eds). New York: Wiley, Wiley Series in Survey Methodology.
- De Leeuw, E.D. and Hox, J.J. (2011). Internet Surveys as Part of a Mixed Mode Design. Social Research and the Internet. In Advances in Applied Methods and New Research Strategies, M. Das, P. Ester, and L. Kaczmirek (eds). New York: Routledge, Taylor & Francis, European Association of Methodology (EAM) Methodology Series.
- Dillman, D.A., Smyth, J.D., and Christian, L.M. (2009). Internet, Mail, and Mixed-Mode Surveys; The Tailored Design Method. New York: Wiley, Wiley Series in Survey Methodology.
- Fuchs, M. and Busse, B. (2009). The Coverage Bias of Mobile Web Surveys Across European Countries. International Journal of Internet Science, 4, 21–33. Available at: http://www.ijis.net/ijis4\_1/ijis4\_1\_fuchs\_pre.html (accessed July 2012).
- GESIS, Eurobarometer Survey Series (2013). Available at: http://www.gesis.org/ eurobarometer-data-service/survey-series (accessed July 2013).
- Groves, R.M. (1989). Survey Errors and Survey Costs. New York: Wiley, Wiley Series in Survey Methodology.
- Groves, R.M. and Couper, M.P. (1998). Nonresponse in Household Interview Surveys. New York: Wiley, Wiley Series in Survey Methodology.
- Groves, R.M. and Peytcheva, E. (2008). The Impact of Nonresponse Rates on Nonresponse Bias A Meta-Analysis. Public Opinion Quarterly, 72, 167–189. DOI: http://www.dx.doi.org/10.1093/poq/nfn011

- Groves, R.M., Fowler, F.J., Couper, M.P., Lepkowski, J.M., Singer, E., and Tourangeau, R. (2009). Survey Methodology. New York: Wiley, Wiley Series in Survey Methodology.
- Internet World Stats. (2013). Available at: http://www.internetworldstats.com/stats.htm (accessed June 2013).
- Kreuter, F., Presser, S., and Tourangeau, R. (2008). Social Desirability Bias in CATI, IVR, and Web Surveys. The Effect of Mode and Question Sensitivity. Public Opinion Quarterly, 72(5), 847–865. DOI: http://www.dx.doi.org/10.1093/poq/nfn063
- Lessler, J.T. and Kalsbeek, W.D. (1992). Nonsampling Error in Surveys. New York: Wiley.
- Link, M.W., and Mokdad, A.H. (2005). Effects of Survey Mode on Self-Reports of Adult Alcohol Consumption: A Comparison of Mail, Web, and Telephone Approaches. Journal of Studies on Alcohol, March 2005, 239–245.
- Little, R.J.A. and Rubin, D.B. (2002). Statistical Analysis with Missing Data. New York: Wiley.
- Lohr, S.L. (2008). Coverage and Sampling. In International Handbook of Survey Methodology, E.D. de Leeuw, J.J. Hox, and D.A. Dillman (eds). New York: Routledge, Taylor & Francis, European Association of Methodology (EAM) Methodology Series.
- Lozar Manfreda, K., Bosnjak, M., Berzelak, J., Haas, I., and Vehovar, V. (2008). Web Surveys versus Other Survey Modes – A Meta-Analysis Comparing Response Rates. International Journal of Marketing Research, 50, 79–104.
- Mohorko, A., De Leeuw, E., and Hox, J. (2011). Internet Coverage and Coverage Bias in Countries Across Europe and over Time: Background, Methods, Question Wording and Bias Tables. Available at: www.joophox.net (accessed June 2013).
- Mohorko, A., De Leeuw, E., and Hox, J. (2013). Coverage Bias in European Telephone Surveys: Developments of Landline and Mobile Phone Coverage across Countries and over Time. Survey Methods: Insights from the Field. Available at: http://surveyinsights. org/?p = 828 (accessed February 2013).
- Moschner, M. (2012). GESIS, Weighting overview. Available at: http://www.gesis.org/ eurobarometer-data-service/survey-series/candidate-countries-eb/weighting-overview/ (accessed February 2013).
- Rookey, B.D., Hanway, S., and Dillman, D.A. (2008) Does a Probability-based Household Panel Benefit from Assignment to Postal Response as an Alternative to Internet-only? Public Opinion Quarterly, 72, 962–984. DOI: http://www.dx.doi.org/10.1093/ poq/nfn061
- Smyth, J.D. and Pearson, J.E. (2011). Internet Survey Methods: A Review of Strengths, Weaknesses, and Innovations. In Advances in Applied Methods and New Research Strategies, M. Das, P. Ester, and L. Kaczmirek (eds). New York: Routledge, Taylor & Francis, European Association of Methodology (EAM) Methodology Series.
- Stoop, I., Billiet, J., Koch, A., and Fitzgerald, R. (2010). Improving Survey Response. Lessons Learned from the European Social Survey. New York: Wiley.

World Bank (2009). Available at: http://data.worldbank.org/ (accessed May 2011).

Received December 2011 Revised March 2013 Accepted August 2013