



# City Research Online

## City St George's, University of London

**Citation:** Fitzgerald, R. & Aizpurúa, E. (2022). Mode and web panel experiments in the European Social Survey – lessons for EU-SILC (10.2785/18803). Publications Office of the European Union,.

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

**Permanent repository link:** <https://openaccess.city.ac.uk/id/eprint/28826/>

**Link to published version:** <https://doi.org/10.2785/18803>

**Copyright and Reuse:** Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

# 27

## Mode and web panel experiments in the European Social Survey – lessons for EU-SILC

Rory Fitzgerald and Eva Aizpurua <sup>(147)</sup>

### 27.1. Introduction

Using different modes of data collection within or between countries in a cross-national social survey has the potential to introduce methodological artefacts into the data (Martin, 2011). This means that analysts may appear to find no differences in the data or find differences that reflect the mix of data collection modes used rather than the real-world situation. The literature has shown that mode effects are likely to vary by topic, question type and country context, and are therefore hard to predict (Martin and Lynn, 2011). At the same time, methods to measure and control for mode effects are difficult and costly to implement and, in many respects, still in their infancy in terms of development (Olson et al., 2020). This means that, when modes are mixed, great care should be taken in the design and analysis stages to take account of the impact that different modes can have on the data collected.

The best way to eliminate the risk of mode effects is to ask all respondents to complete the survey using the same mode or to use a combination of modes in which all respondents answer each ques-

tion using the same mode (e.g. self-administering sensitive questions in the context of an interviewer-administered survey). If a combination of modes is to be used, efforts should be made to minimise mode effects (e.g. by ensuring that visual information is provided, rather than sometimes providing information visually and on other occasions providing it aurally). The European Social Survey (ESS) therefore decided in 2001, when it was established, to use face-to-face interviewing in all countries and has done so ever since. This decision was made because no other mode could be effectively used for interviewing in all countries and for all respondents, especially considering the length (approximately 60 minutes in English) and complexity of the survey, as well as the differences in penetration of internet and telephone technology across countries. However, it was acknowledged at the time that other modes might become more feasible in the future as a single-mode alternative and that it might become essential to combine modes for other reasons, such as increased costs of face-to-face data collection. In order to gather information to make an informed decision, the ESS established a mixed-mode methodological research programme composed of a series of six experiments. Following the conclusion of that work, it was decided on the advice of its Methods Advisory Board not to switch to mixed-mode data collection. Instead, the ESS experimented with recruiting a web panel off the back of its face-to-face survey (the Cross-national Online Survey (CRONOS)). This experimental work is of relevance to EU-SILC, in which modes are routinely mixed.

This chapter starts by providing an overview of the ESS before moving on to discuss the challenges

<sup>(146)</sup> Rory Fitzgerald is the director of the European Social Survey (ESS) European Research Infrastructure Consortium and professor of practice in survey research at City, University of London, United Kingdom. Eva Aizpurua is a research fellow at the ESS headquarters (City, University of London, United Kingdom). The authors would like to thank Peter Lynn and Lars Lyberg for their feedback. All errors and opinions are the authors' responsibility. This work was supported by Net-SILC3, funded by Eurostat and coordinated by the Luxembourg Institute of Socio-Economic Research. The European Commission bears no responsibility for the analyses and conclusions, which are solely those of the authors. Correspondence should be addressed to Rory Fitzgerald ([r.fitzgerald@city.ac.uk](mailto:r.fitzgerald@city.ac.uk)).

**Table 27.1: Overview of the ESS**

Time span	2002 to present
Frequency	Every 2 years
Management	ESS ERIC
Design	Repeated cross-sectional
Central topics	Attitudes, beliefs, values, perceptions and behaviour patterns
Target population	Residents aged 15 and over
Sampling	Probabilistic
Sample size	1 500 in each country (effective) 800 in countries with a population of less than 2 million (effective)
Survey mode	Face-to-face interviewing (CAPI)
Source questionnaire language	English
Translation	Languages spoken by 5 % of the population and more
Interview duration	60 minutes (English questionnaire)
Data access	Free of charge for non-commercial use

Source: ESS website (<https://www.europeansocialsurvey.org/>).

faced by cross-national surveys in respect of data collection, focusing on increased costs, decreasing response rates and a contraction in interviewer capacity. We then introduce different mixed-mode designs and summarise the experiments conducted to assess the feasibility and the impact of mixed-mode data collection in the ESS and more recent experiments conducted by the European Values Study (EVS). We continue with a description of the first cross-national, input-harmonised, probability-based web panel, CRONOS, which has been implemented in three countries. In the last section of the chapter, we discuss some of the lessons learned and introduce CRONOS-2, a 12-country web panel currently under construction by the ESS. The chapter concludes by considering the possible implications of this experimental work on data collection mode conducted by the ESS for EU-SILC.

## 27.2. European Social Survey

The ESS is an academically led cross-national survey that has been conducted across large parts of Europe since its establishment in 2001. From 2002 to 2019, 39 countries participated in one or more rounds of the ESS <sup>(147)</sup>. The survey is currently con-

ducted face-to-face using computer-assisted personal interviewing (CAPI) and consists of a core module, which remains largely stable from round to round, and rotating modules, which may be new or repeated from previous rounds (e.g. personal and social well-being, timing of life, welfare attitudes). Table 27.1 summarises the main characteristics of the ESS.

The aim of the ESS is to measure attitudes, beliefs, values and behaviour patterns, providing comparative data across countries and time. The ESS uses probability sampling with the aim of covering residents aged 15 and over, regardless of their nationality, citizenship or language. An effective sample size <sup>(148)</sup> of 1 500 is aimed for in each participating country (800 for countries with fewer than 2 million inhabitants). The ESS was awarded European Research Infrastructure Consortium (ERIC) status in 2013. The data are available free of charge for non-profit purposes and are widely used (as at July 2020, the ESS had 162 730 registered users from over 240 countries <sup>(149)</sup>). The number of English-language publications and presentations exceeded

<sup>(147)</sup> Further information about the survey is available on the ESS website (<https://www.europeansocialsurvey.org/>).

<sup>(148)</sup> Effective sample size refers to the actual sample size (i.e. the number of observations) divided by the design effect. This is the size of a simple random sample that would have produced the same precision.

<sup>(149)</sup> Monthly statistics on ESS data usage can be found on the ESS website ([https://www.europeansocialsurvey.org/about/user\\_statistics.html](https://www.europeansocialsurvey.org/about/user_statistics.html)).

4 400 in 2019 (Malnar, 2020). The data also have an extensive impact beyond academia in policy and third-sector work (see Technopolis Group, 2017). The survey further aims to increase cross-national comparability by using an input-harmonised approach where possible and functionally equivalent approaches where total harmonisation is not possible (Fitzgerald and Jowell, 2010).

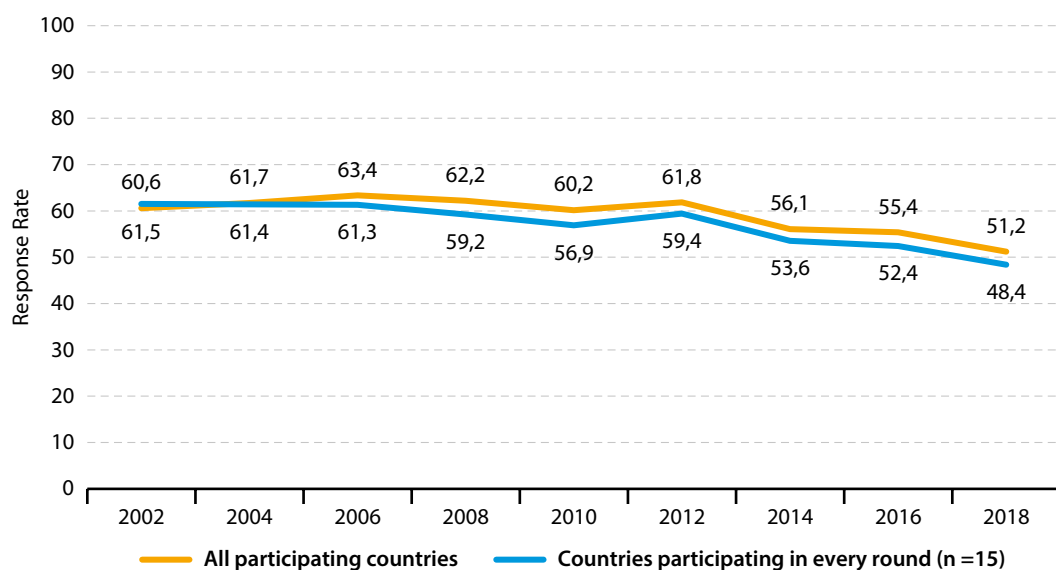
### 27.3. Challenges faced by cross-national general social surveys in terms of data collection and mode

Contemporary survey research faces important challenges, mainly related to declining response rates and increasing data collection costs (Leeper, 2019; Luiten, Hox and de Leeuw, 2020). Despite having relatively high response rates and increased fieldwork efforts to maximise these rates, the ESS

has not been immune to this trend, particularly in recent years. As shown in Figure 27.1, there has been a substantial decline in response rates, which has been attributed more to refusals than to non-contacts (Beullens et al., 2018).

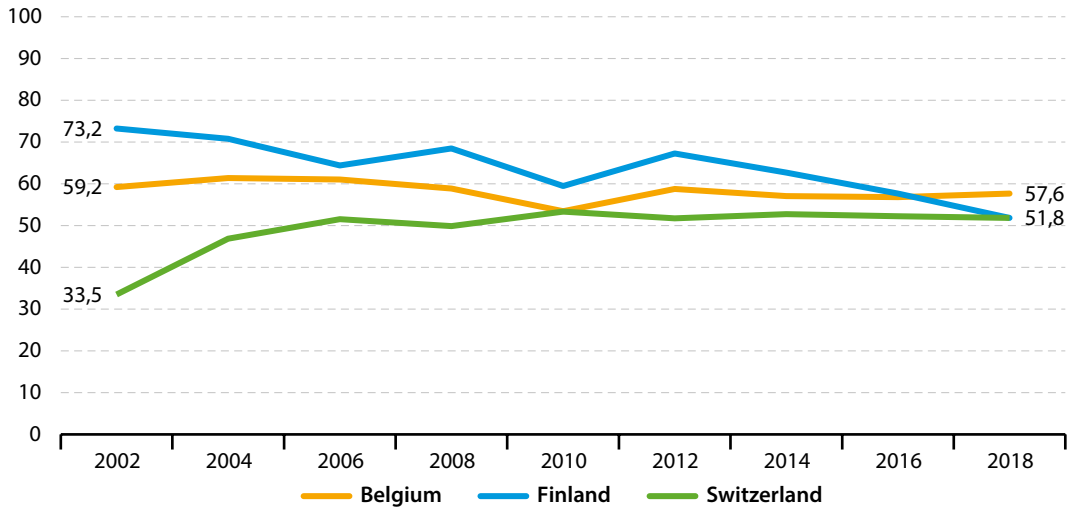
Although this trend is observable at the aggregate level, large differences exist across countries, in both the direction and the magnitude of the change in response rates over the years, because of differences in survey climate and tradition (ESS survey documentation reports, rounds 1–9 <sup>(150)</sup>). Figure 27.2 displays the changes in response rates across the rounds for three selected countries with distinct trends. Although Finland and Switzerland both achieved the same response rate (51.8 %) in round 9 (2018) of the ESS, the former exhibits a downward trend (from 73.2 % in 2002), whereas Switzerland's response rates increased by 19.9 percentage points from 2002 to 2010 and stabilised afterwards. Belgium, however, remained relatively stable over the years, with only small fluctuations in response rates from round to round.

**Figure 27.1: Average ESS response rates in rounds 1–9, 2002–2018**



Source: ESS survey documentation reports, rounds 1–9.

<sup>(150)</sup> <https://www.europeansocialsurvey.org/data/round-index.html>

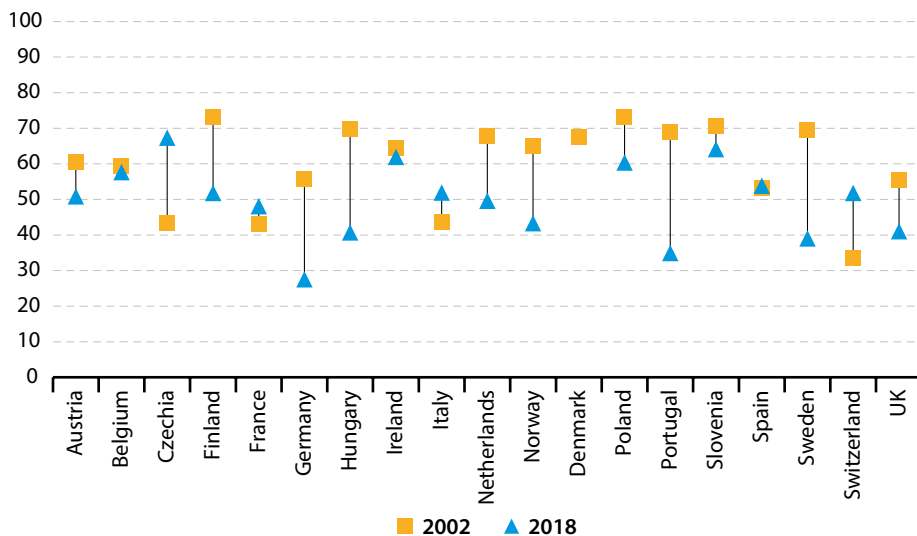
**Figure 27.2: ESS response rates in rounds 1–9 in selected countries, 2002–2018**

Source: ESS survey documentation reports, rounds 1–9.

Figure 27.3 shows response rates achieved in 2002 and 2018 in the 18 countries participating in rounds 1 and 9 of the ESS. Despite a general decline, which is observable for 11 countries, the differences are heterogeneous, with large decreases in Germany (from 55.7 % to 27.6 %), Hungary (from 69.9 % to 40.7 %) and Sweden (from 69.5 % to 39.0 %), and more modest decreases in others countries, including Austria (from 60.4 % to 50.9 %) and Slovenia (from 70.5 % to 64.1 %). A few countries obtained very similar response rates in both rounds (Belgium, Ireland and Spain), whereas others improved theirs. In this group are Czechia and Switzerland with large increases (24.0 and 18.4 percentage points, respectively) and France and Italy with smaller gains (5.0 and 8.2 percentage points, respectively).

In a recent assessment of response rates across surveys in Europe from 1998 to 2015, it was found that, although there were differences in the rate of decline, all types of surveys showed a downward trend in response rates (Luiten, Hox and de Leeuw, 2020). The decline in response rates has prompted survey organisations worldwide to consider alternative modes of data collection, including combinations of survey modes. In addition, as more research has moved away from using face-to-face

data collection, fewer agencies now offer this to the standard required for high-quality surveys such as the ESS and the Survey of Health, Ageing and Retirement in Europe (Sommer, 2019). In cross-national surveys, however, combining survey modes brings additional challenges associated with differences in technology penetration and disparate sampling frames (de Leeuw, Suzer-Gurtekin and Hox, 2019). The sampling frames available in some countries require contact to be made in person, although using mail or telephone is sometimes also possible. In general, if an interviewer is not to be used, then a frame of individuals is needed, so that letters can be sent to the target respondents, or alternatively a relatively complete frame of telephone numbers. These, however, are not always available. When population registers are not available, in-person contact is the best way to appropriately sample individuals within households. Unlike interviewer-administered selection procedures, which are well established, within-household selection in self-administered modes is more difficult to accomplish, with studies finding between 20 % and 30 % of selections in mail and web surveys to be inaccurate (Olson and Smyth, 2014, 2017). In addition, there are no acceptable general population frames of email addresses and no acceptable ways of drawing probability samples from them

**Figure 27.3: ESS response rates in rounds 1 (2002) and 9 (2018)**

Source: ESS survey documentation reports, rounds 1 and 9.

(Dillman, 2017). In the case of online surveys, it is therefore necessary to mix modes of contact (e.g. sending advance letters and invitations to participate in web surveys by post), which increases the costs and complexities of data collection.

Logistical demands are also amplified in cross-national surveys, as the number of actors increases along with diverging contexts. In addition, within research infrastructures levels of experience and expertise vary widely, even across European countries. Although recruiting qualified interviewers may be relatively easy in some countries, other countries – particularly those in which the chosen mode is infrequently used – may experience difficulties in hiring seasoned interviewers (De Jong, 2016). This has important implications, because inexperienced interviewers tend to produce lower response rates (West and Blom, 2017; Wuyts and Loosveldt, 2020). At the same time, the scarcity of interviewers often results in increased workloads, which lead to larger interviewer effects (West and Blom, 2017; see also Chapter 28 of this book). Furthermore, the reliability of postal systems is rather uneven across European countries, making contact by post difficult to implement consistently. This contributes to cross-national differences in the outcomes of push-to-web or postal self-completion approaches, threatening

the comparability of the data. A push-to-web (online only) cross-national design implemented by the European Union Agency for Fundamental Rights in 2017 achieved an average response rate of 18 % in countries using named person samples (Denmark, Estonia, Italy, Luxembourg, Hungary, Malta, the Netherlands, Poland, Slovenia, Finland and Sweden). However, response rates were much lower in countries not using individual registers. Specifically, address-based samples (in Austria, Belgium, Bulgaria, Croatia, France, Ireland, Latvia, Lithuania, Spain and the United Kingdom) had an average response rate of 7 %, whereas this percentage dropped to 3 % in enumeration countries (Czechia, Germany, Greece, Italy, Cyprus, Portugal, Romania and Slovakia; Smith, 2018). The ESS has recently (2021) conducted a push-to-web mixed-mode study (web and postal) in three European countries – Austria, Hungary and Serbia – with the goal of testing this approach in a cross-national setting. This three-country study has provided encouraging results, with response rates of around 40 % in all countries, two of which used address-based samples (Austria and Serbia) and only one of which used a named person sample (Hungary). It will also provide insights into the representativeness of the achieved samples. These findings should help inform survey designs as push-to-web

data collection methods gain popularity (for a review of push-to-web surveys, see Dillman, 2017).

Equivalent measurement is one of the main challenges of cross-national surveys that are comparative by design, such as the ESS. Questionnaires not only are required to be culturally relevant within and across countries but also need to accommodate multiple languages and provide invariant measures (Pennell et al., 2017). When comparing groups, functional equivalence is necessary to ensure that observed differences represent actual differences and are not the result of other factors, such as differences in data collection modes or non-equivalent translations. However, removing all sources of error from surveys is not possible. In single-country surveys, the challenge is the optimal allocation of resources to minimise total survey error. In the context of cross-national research, however, the goal is to minimise error and to make error components similar in magnitude and direction across countries (Smith, 2011). This can be promoted through planning, coordination and adoption of comparable protocols of data collection. In cross-national surveys, strong infrastructures that support and monitor the design and implementation of the survey are particularly important. The ESS infrastructure, for example, is led by a Core Scientific Team that ensures careful joined-up planning, provides support to national teams and monitors activities at the national level to maximise compliance at every stage. The Scientific Advisory Board and Methods Advisory Board ensure that the approaches used remain ‘state of the art’.

## 27.4. Mixed-mode survey designs

Mixed-mode designs are those in which respondents answer the same questions using different modes (e.g. some respondents are interviewed face-to-face while others complete the survey online). Sometimes respondents are offered a choice between multiple modes (e.g. web, telephone, face-to-face) in what is called a *concurrent mixed-mode design*. At other times potential respondents are assigned to different modes depending on the information that is available (e.g. telephone surveys

for sample members with a telephone number and face-to-face interviews for those with addresses only). On other occasions, respondents are invited to participate using a certain mode first (usually the most cost-effective mode) and offered additional modes if they are unable or unwilling to respond (*sequential mixed-mode design*; for a review of mixed-mode survey designs, see de Leeuw, 2018).

In the context of cross-national research, countries might use different modes of data collection, resulting in what has been called *across-country mixed-mode designs*. An example of this design is found in the International Social Survey Programme, which in recent years has allowed countries to choose between face-to-face interviewing, self-administered surveys and telephone surveys<sup>(15)</sup>. When one or more countries combine modes of data collection, using a concurrent or a sequential approach, the design becomes a *within-country mixed-mode design*. The third variation in cross-national time-series surveys is the *across-time mixed-mode design*, which occurs when countries transition from a single mode or combination of modes to a different mode (Martin, 2011). These three designs are not exclusive and, as shown in Chapter 24 of this book, they coexist in the context of EU-SILC.

Mixed-mode designs have increased in recent years due to their potential to lower financial costs and reduce coverage and unit non-response errors. However, mixing modes of data collection is not without drawbacks, as this practice may threaten the comparability of the data between groups and, in the case of time-series surveys, across time (de Leeuw, Hox and Scherpenzeel, 2019). In this context, measurement differences attributed to survey mode (e.g. social desirability bias) may be confounded with substantive differences, undermining the comparisons made. Mixed-mode designs also increase the logistic complexity of the survey, requiring additional work at the design (e.g. adapting the questionnaires to different modes), implementation (e.g. following up in different modes, additional coordination) and analysis (e.g. data cleaning, adjustment and harmonisation) stages (Martin, 2011).

<sup>(15)</sup> Information on modes of data collection is available on the website of the Leibniz Institute for the Social Sciences (<https://zacat.gesis.org/webview/index/en/ZACAT/ZACAT.c.ZACAT/ISSP.d.58/by-Year/fCatalog/Catalog69>).

**Table 27.2: Summary of the mixed-mode experiments conducted as part of the ESS methodological programme, 2003–2012**

Study	Year (round)	Country	Mode	Sampling	Research design
Study 1: Measurement differences across four modes	2003 (R1)	Hungary	Face-to-face, telephone, PAPI and web surveys	Convenience	Within-subjects reinterview design
Study 2: Causes of measurement differences between face-to-face surveys and telephone surveys	2005 (R2)	Hungary (Budapest) and Portugal (Lisbon)	Face-to-face (with and without showcards) and telephone surveys	Probabilistic	Between-subjects design
Study 3: Measurement differences between face-to-face surveys and web surveys	2010–2011 (R5)	United Kingdom	Face-to-face and web surveys	Probabilistic	Within-subjects reinterview design
Study 4: Feasibility of conducting the ESS using telephone surveys	2006 (R3)	Cyprus, Germany, Hungary, Poland and Switzerland	Face-to-face and telephone surveys	Probabilistic	Between-subjects design
Study 5: Comparing concurrent and sequential mixed-mode designs	2008 (R4)	Netherlands	Face-to-face, telephone and online surveys	Probabilistic	Between-subjects design
Study 6: Feasibility of mixed-mode designs across countries	2012 (R6)	Estonia, Sweden and the United Kingdom	Face-to-face, telephone and web surveys	Probabilistic	Between-subjects design

NB: PAPI, paper and pencil interviewing.

Source: Adapted from Villar and Fitzgerald (2017).

## 27.5. Mode experiments in the European Social Survey

To assess the feasibility and implications of transitioning from the face-to-face mode to a different mode, or a combination of modes, the ESS implemented a methodological programme to assess the impact of mixing data collection modes on the quality of survey estimates (for a detailed review of the experiments and findings, see Villar and Fitzgerald, 2017). This programme included six experiments conducted in 10 European countries between 2003 and 2012. A summary of the experiments can be found in Table 27.2. The goal of the programme was to inform the implementation of future rounds of the ESS, providing evidence to support or discard the adoption of a mixed-mode design. The first three experiments focused on measurement equivalence, whereas the following three examined the feasibility of conducting the

ESS using a different mode or a combination of modes. The remainder of this section summarises the findings of this methodological programme with regard to survey participation (response rates and sample composition) and measurement effects. This is supplemented by the results of mixed-mode experiments conducted by the EVS in six countries (Denmark, Finland, Germany, Iceland, the Netherlands and Switzerland) during wave 5 (2017).

### 27.5.1. Survey participation

One of the premises of mixed-mode designs is that each survey mode may attract different types of respondents. As a result, combining data collection modes has the potential to reduce coverage and non-response errors (de Leeuw, 2018). For this reason, selection effects (i.e. different types of respondents selecting different modes) are desired, although later adjustments may be needed, particularly in across-country mixed-mode designs

in which selection effects may differ between the countries. The findings from the ESS methodological programme did not show improvements in response rates compared with the face-to-face main survey; they pointed, instead, to a deterioration that varied in magnitude depending on the country and the mode(s). For example, in the study conducted in the Netherlands comparing face-to-face interviewing with a concurrent mixed-mode design and a sequential mixed-mode design (online, telephone and face-to-face interviews), response rates were lower for the mixed-mode designs (46 % in the concurrent design and 45 % in the sequential design, compared with 52 % in the face-to-face single-mode design). In an earlier study comparing telephone and face-to-face surveys in five European countries, it was found that, when administering the full ESS (which takes approximately 60 minutes), response rates were consistently lower in the telephone mode. Differences varied widely across the countries, ranging from a relatively small difference in Switzerland (38 % versus 46 %) to a very large difference in Hungary (18 % versus 66 %; Villar and Fitzgerald, 2017). The lack of improvement in response rates was expected, given that face-to-face surveys tend to have the highest coverage and response rates. The more recent experiments conducted during the last wave of the 2017 EVS (Luijckx et al., 2020) show that, in three countries, response rates were lower in the self-administered modes than for face-to-face interviewing (15 % lower in Switzerland, 17 % lower in Denmark and 29 % lower in Finland), whereas in Germany and Iceland response rates were actually lower in the face-to-face mode (28 % versus 35 % and 41 % versus 45 %, respectively) <sup>(152)</sup> (Christmann et al., 2019).

Another dimension explored in the ESS experiments was the demographic composition of the samples achieved. Different modes are linked to different levels of non-coverage and non-response errors, and they are likely to attract different groups of respondents. This is, in fact, one of the benefits of mixed-mode designs: because of selection effects, respondents who would not or could not participate in a single-mode design may participate if

multiple modes are offered. The major problem is that selection effects and measurement effects (e.g. social desirability, acquiescence) are often confounded, making it difficult to ascertain the extent to which differences (or similarities) between the modes are the result of differences in respondents or differences in measurement error (Vannieuwenhuyze, Loosveldt and Molenberghs, 2010). The findings of the ESS programme pointed to small differences in sample composition between the single-mode design and the mixed-mode designs. When differences were found, the composition of the face-to-face survey sample was generally closer to the population estimates than the composition of mixed-mode surveys (Villar and Fitzgerald, 2017). For example, two of the studies comparing telephone surveys with the standard face-to-face mode indicated that telephone interviews tended to over-represent those with higher educational levels. However, the results were not consistent, revealing cross-national differences in how mode affected sample composition.

Results from the 2017 EVS experiments are consistent with these findings, suggesting that samples obtained through face-to-face interviewing tend to be more similar to the overall population, although differences were generally small (Christmann et al., 2019). In the case of Germany, where the face-to-face mode was compared with two self-administered mixed-mode (computer-assisted web interviewing and paper self-completion) designs – one a matrix design <sup>(153)</sup> and one with a full-length questionnaire – differences were found in some variables (e.g. age, nationality, household size) but not in others (e.g. gender), and the size of the differences fluctuated. For instance, the samples achieved under-represented foreigners in all cases, but differences were larger in the two mixed-mode designs. Although nearly one in eight individuals in the population was foreign, this fraction dropped to one in nine for the face-to-face interviews and to 1 in 15 for the mixed-mode designs (Christmann et al., 2019). In terms of education, the samples achieved over-represented the group with the highest educational level, with the largest difference being between the mixed-mode

<sup>(152)</sup> As part of the EVS experiments, full-length and matrix questionnaires were used. The figures reported here refer to the comparisons between the full length (approximately 60 minutes) face-to-face and self-administered questionnaires.

<sup>(153)</sup> In a matrix design, the questionnaire is split into shorter versions to which respondents are randomly assigned. For a full description of the experiments, please see Luijckx et al. (2020).

designs and the population (40.3 % for the matrix design and 39.0 % for the design with the full questionnaire, compared with 24.3 % among the population) and a smaller difference between the face-to-face single-mode design (34.8 %) and the population (Christmann et al., 2019).

### 27.5.2. Measurement effects

Mode effects are the result of both mode self-selection effects, which are produced when assignment to the modes is not randomised, and measurement effects, which are attributed to mode differences (e.g. interviewer effects, questionnaire design). Unlike mode selection effects, mode measurement effects represent a source of measurement error and, in a mixed-mode survey, constitute undesired effects. They arise when respondents' answers depend on the mode of data collection (e.g. more honest responses to the same question when it is self-administered). The results from the ESS experiments revealed differential mode measurement effects that threatened the equivalence and the comparability of the data. For example, the study conducted in Hungary and Portugal indicated that telephone respondents were more likely to provide socially desirable responses across a range of indicators than face-to-face respondents (Jäckle, Roberts and Lynn, 2006). In general, attitudinal questions, which are dominant in the ESS, showed a lower level of consistency across modes than behavioural questions. Among these attitudinal questions, the largest differences were found for estimates of personal well-being, political attitudes and participation, and attitudes towards immigrants. In addition, large differences were found for self-reported income, a variable that, in general surveys, tends to yield low-quality data in terms of high item non-response and inconsistencies with administrative data (Moore, Stinson and Welniak, 2000) – which is particularly important for EU-SILC. One of the experiments used a reinterview design, randomly assigning respondents to complete the survey in a different mode, finding that answers to the income question were different in 48 % of the cases (Villar and Fitzgerald, 2017). A later study also revealed that telephone respondents were less likely to report lower household income than those interviewed face-to-face (Jäckle, Roberts and Lynn, 2006).

Although some measurement effects are inherent in the mode, others can be avoided or minimised. Changes in questionnaire design between modes can exacerbate the unwanted effects, threatening the comparability of the data (de Leeuw, Hox and Scherpenzeel, 2019). In mixed-mode surveys, designing and implementing questionnaires that are equivalent is particularly important to prevent avoidable mode effects. For instance, a study comparing face-to-face and web responses to the 2008 Dutch EVS found that the responses to 64 % of the items differed between the modes (Bennink, Moors and Gelissen, 2013). The differences were attributed to changes in question wording (e.g. definitions provided by interviewers in face-to-face interviewing that had to be included as part of the question in the online survey) and the ways in which non-substantive responses (refusals, 'don't know') were presented (visible versus non-visible) and navigated (possibility of leaving a question unanswered).

Major challenges associated with differential measurement error include the existence of heterogeneous effects across variables and the lack of a single method that could be used to adjust for these differences in all types of analyses (Martin and Lynn, 2011). Based on the findings from the ESS programme, the Core Scientific Team decided, on the advice of the Methods Advisory Board, not to adopt a mixed-mode strategy, continuing instead with the face-to-face mode. It was agreed, however, that implementing a cross-national probability-based web panel, to be recruited off the back of the ESS, would be trialled. In a sense, this still leads to a mixed-mode design, as analysts can combine answers from the face-to-face survey with those from web follow-ups at the individual level, but the main ESS remains, at least for now, in face-to-face mode<sup>(154)</sup>.

## 27.6. Cross-national Online Survey

Internet use continues to increase, with 85 % of Europeans using the internet at least once a week

<sup>(154)</sup> The Core Scientific Team of the ESS is currently reviewing whether to recommend a change to the mode of data collection in the future, including the possibility of using a combination of modes.

**Table 27.3: Characteristics of CRONOS**

Participating countries	Estonia Great Britain Slovenia
Data collection years	2016–2018
Recruitment approach	Piggyback sampling (ESS round 8)
Population	All ESS respondents aged 18 and over
Incentives	Unconditional (GBP 5 / EUR 5 with each survey invitation)
Number of waves	Six waves plus a welcome survey
Periodicity of waves	Bimonthly
Survey duration	20 minutes
Data access	Free of charge for non-commercial use (CRONOS data can be linked to ESS round 8 data)

Source: Adapted from Jessop et al. (2019).

in 2019 (DESI, 2020). This, along with the reduced costs and fieldwork times associated with online surveys, has resulted in a very rapid increase in this mode of data collection. In addition, online surveys are associated with reduced social desirability biases, one form of measurement error that occurs when respondents provide inaccurate responses to comply with social norms (Krumpal, 2013). Because interviewers are absent, acquiescent or agreeable responses, in which individuals tend to agree or provide affirmative answers to questions, are also reduced (Liu, Conrad and Lee, 2017). Despite this, online surveys have important shortcomings, including low response rates and self-selection biases. A recent meta-analysis, for example, revealed that web surveys still yield lower response rates than other modes. Daikeler, Bošnjak and Manfreda (2020) found that response rates were 12 percentage points lower for online surveys than for other modes. Online surveys have also been found to be less representative than other single-mode surveys (Cornesse and Bošnjak, 2018). Although the digital divide has lessened, access to the internet still varies widely across and within countries, with large differences in the percentage of people who regularly use it. In the United Kingdom, for instance, 99 % of adults aged 16–44 were recent internet users in 2019, whereas this percentage dropped to 47 % in the case of those aged 75 and older (ONS, 2019). As at June 2020, just 67 % of adults in Bulgaria had access to the internet, compared with 96 % of those in the Netherlands <sup>(155)</sup>. Therefore, under-cov-

erage and non-response are still serious threats to the validity of online surveys, especially at the cross-national level.

CRONOS was implemented during round 8 of the ESS <sup>(156)</sup>. The objective of CRONOS was to assess the feasibility of establishing a cross-national probability-based panel following a harmonised approach. This assessment was used to create a blueprint intended to guide the development of such a panel in the future (Jessop et al., 2019). As shown in Table 27.3, CRONOS was piloted in three countries – Estonia, Great Britain and Slovenia – acting as proof of concept for the viability of a European online panel. Because CRONOS used a ‘piggy-back’ recruiting approach, in which all ESS adult respondents were invited – at the end of their ESS interview – to join the panel, fieldwork costs were significantly reduced (e.g. there was no need to source a new sampling frame or to hire additional interviewers). CRONOS followed a centralised management approach, with a high level of standardisation of procedures across countries while allowing adaptations if needed. A panel design such as that used in CRONOS provides important advantages, making it possible to capture individual-level variation across time. For this reason, CRONOS was seen as a very valuable complement to the main ESS, although the sample size at country level remained rather small, largely due to the orig-

<sup>(156)</sup> The CRONOS panel work was developed under the Synergies for Europe’s Research Infrastructures in the Social Sciences project, which was funded by the EU’s Horizon 2020 research and innovation programme under grant agreement No 654221. The CRONOS initiative was also supported by the 2015–2017 and 2017–2019 ESS ERIC work programmes.

<sup>(155)</sup> <https://www.internetworldstats.com/stats9.htm>

inal sample size in the face-to-face study and the limited cooperation in joining the panel.

The results of CRONOS highlighted the feasibility of implementing a cross-national online panel. Participation rates <sup>(157)</sup> were reasonable, ranging from 56 % in Great Britain to 78 % in Estonia. However, comparisons between the sample composition of CRONOS and the target population revealed multiple discrepancies. For example, CRONOS over-represented females, citizens and married individuals while under-representing older and the least educated groups (Bottoni and Fitzgerald, 2021). Although internet-enabled tablets were provided to potential respondents who had no internet access, the propensity to join the panel increased with the frequency of internet use. In addition, when comparing individuals who participated in the panel with those who did not, some differences emerged in attitudinal and behavioural indicators. For instance, it was found that CRONOS respondents had higher levels of social and institutional trust, greater life satisfaction and more tolerant attitudes towards the lesbian, gay, bisexual and transgender community than non-participants. They also reported higher levels of political participation and better perceived health.

Differences in measurement quality between questions included in the main ESS (round 8) and questions included in CRONOS have been found to be small. Non-differentiation (i.e. variance in the respondents' answers to a given topic) was equivalent across modes, whereas item non-response, although generally low, was higher in the online panel. Primacy effects (i.e. tendency to select the first answer categories) were larger in CRONOS, whereas recency effects (i.e. tendency to select the last answer categories) were generally comparable. There was also evidence of metric equivalence, providing support to the comparison of unstandardised relationships across the ESS and CRONOS. The results for scalar equivalence were less robust, suggesting some caution when comparing means (Cernat and Revilla, 2020).

CRONOS was the first attempt to establish a cross-national probability-based panel under an

<sup>(157)</sup> Calculated by dividing the number of actual participants by the number of people invited to participate in the panel.

input harmonisation framework, in which panel design and maintenance followed the same principles in all participating countries. The results of this experience showed the feasibility of developing such a panel in terms of costs, response rates and data quality. Web panels such as CRONOS are a viable complement to ongoing cross-national surveys, providing the opportunity to further explore certain topics and evaluate individual-level differences. However, they present important challenges associated with non-response bias that require further attention. Following the successful implementation of CRONOS and taking into consideration the challenges encountered during the project, a blueprint for a comparative probability-based online survey was developed (Jessop et al., 2019). The main recommendations included in the blueprint, grouped by stage of the survey cycle, are summarised in Table 27.4.

Building on the knowledge acquired from CRONOS, the ESS is working on the implementation of a larger-scale probability-based panel. During round 10 of the ESS, adult participants will be recruited at the end of the interviews. CRONOS-2 will cover 12 European countries and will comprise six waves, allowing the study of individual- and country-level differences. It is anticipated that CRONOS-2 will help to build expertise and infrastructure so that the field is prepared for a large-scale switch to the online mode in the future. This includes the development of a sample management system for cross-national surveys <sup>(158)</sup> that is linked to the Qualtrics survey platform. Procedures for management of translation and for centralised communication with the panel are also being trialled. Most notably, CRONOS-2 aims to introduce web-based interviewing in a comparative format to new countries where probability-based national panels have not been established. Although piloting and capacity building are the focus now, the longer-term 'dream' is pan-European coverage, with online interviews being dominant, and a complementary mode designed to include those without internet access gradually being phased out over time.

<sup>(158)</sup> This tool is being developed under the Horizon 2020 Social Sciences & Humanities Open Cloud project under grant agreement No 823782.

**Table 27.4: Recommendations for the design and implementation of a cross-national panel**

<b>Sampling and sample management</b>
<ul style="list-style-type: none"> <li>• The parent survey recruits participants using probability sampling.</li> <li>• Equivalent sampling approaches are used in all participating countries, using the best random sampling practice in each case.</li> <li>• The sample size achieved is sufficiently large for the effective statistical analysis of country-level data.</li> <li>• Participants' contact details are updated throughout the duration of the panel.</li> <li>• Targeted approaches to fieldwork and data collection are used based on available data.</li> </ul>
<b>Recruitment</b>
<ul style="list-style-type: none"> <li>• All eligible people who complete the parent survey are invited to participate in the panel, including those who do not have internet access.</li> <li>• The recruitment approach is standardised across countries.</li> <li>• Panel members recruited early in the parent survey's fieldwork receive a 'welcome mailing' and a 'welcome survey' to prevent disengagement.</li> <li>• Incentives are used in recognition of the time and effort of panel members.</li> </ul>
<b>Questionnaire development</b>
<ul style="list-style-type: none"> <li>• Questionnaire content is carefully developed taking into consideration comparability issues.</li> <li>• The questionnaire is translated and pretested. At a minimum, a cross-national expert review and advance translation efforts are used before the questions are fielded.</li> <li>• Questionnaires are adapted to be displayed on multiple devices (smartphones, tablets, PCs).</li> <li>• Questionnaire length ranges between 15 and 20 minutes, to prevent data-quality issues and break-offs.</li> </ul>
<b>Fieldwork</b>
<ul style="list-style-type: none"> <li>• Between 4 and 12 waves of data are collected per year.</li> <li>• For each wave, fieldwork periods of around 4 weeks are recommended.</li> <li>• Fieldwork protocols (e.g. incentives used) are adapted to the countries to optimise response rates and sample representativeness.</li> <li>• Panel members are sent multiple communications to keep them engaged and informed.</li> <li>• The primary mode of communication is email, supplemented by other modes (e.g. postal, text messaging).</li> <li>• Reminders are sent at different times and on different days, not exceeding more than one in any given week.</li> <li>• Between-wave mailing is used to maintain the engagement of panel members.</li> </ul>
<b>Management and data security</b>
<ul style="list-style-type: none"> <li>• A centralised survey management approach is used to achieve high input harmonisation.</li> <li>• Data reduction is practised to minimise the risk of harm.</li> <li>• Only those who need it, and are trained, have access to identifiable information.</li> <li>• All data outputs are reviewed for disclosure risk.</li> </ul>

Source: Adapted from Jessop et al. (2019).

## 27.7. Conclusions and lessons from the European Social Survey for EU-SILC

Cross-national general social surveys have important challenges ahead, resulting from differences in technology penetration, survey tradition and climate across countries, as well as declining response rates, rising costs and decreasing face-to-face capacity. In this environment, mixed-mode and on-line surveys have become increasingly popular. Although mixed-mode surveys have the potential to reduce coverage and non-response error, the

results obtained under the ESS programme show that these designs do not always lead to smaller errors or better data quality and can actually introduce other forms of error (Villar and Fitzgerald, 2017). Isolating and correcting for mode-specific measurement error is still a complex task for which no universally accepted procedure exists (Martin and Lynn, 2011). This has implications for comparative surveys such as the ESS and EU-SILC, in which achieving measurement invariance is essential to study differences across countries and over time. In addition, the impact of mixed-mode approaches on the planning and management of the fieldwork requires consideration, as well as the adjustments needed after data collection, to prevent changes in

data collection modes from threatening the comparability of the estimates.

The use of online surveys has also grown considerably in the past decade, allowing for rapid data collection. Despite this, response rates remain lower than in other modes (Daikeler, Bošnjak and Manfreda, 2020), and the absence of general sampling frames of internet users requires the use of alternative modes of contact, which increases fieldwork costs. Other challenges of online surveys, such as variations in internet penetration and differences in technology access and use, are likely to lessen over time. All things considered, the evidence today suggests that face-to-face interviewing is still needed in the short term, although the use of online surveys and mixed-mode approaches is on the rise and is likely to continue to increase (Schober, 2017).

The experience of the ESS with CRONOS indicates that building a cross-national probability-based panel off the back of an established survey (or perhaps recruiting directly) is feasible and provides important opportunities for the research community. CRONOS was successfully implemented in three European countries and recruited participants off the back of the main ESS (round 8). The characteristics of CRONOS panellists were not very different from those of the target population, although older respondents and those who used the internet less often were under-represented (Bottoni and Fitzgerald, 2021). Building on the pioneering experience of CRONOS, the ESS is currently planning CRONOS-2 to test the implementation of an online panel across a larger and more diverse set of countries. Its results will contribute to further developing the methodology for a cross-national web panel and will provide further open access data for researchers and the general public.

For EU-SILC, these experiments underline the importance of designing questionnaires across modes to minimise mode effects. Although mode-inherent factors, such as interviewers being present or absent, cannot always be avoided, mode measurement effects can be reduced by the design of the questionnaire. For this, the adoption of a unified mode design, in which equivalent questionnaires (e.g. question structures, wording) are developed for each mode, is recommended (Dillman, 2017;

de Leeuw, Suzer-Gurtekin and Hox, 2019). This approach precludes design differences across modes (e.g. the use of grids in online/paper questionnaires versus sequential questions in face-to-face/telephone interviews) that may lead to unintended mode differences and, ultimately, threaten the validity of comparisons across groups. In addition to adopting a unified mode design, mixing modes that are most similar will restrict mode-specific errors (de Leeuw, 2018). Two characteristics are often considered when comparing modes: the degree of interviewer involvement (e.g. self-administered surveys versus interviewer-administered surveys) and the channel of communication used to present questions and provide answers (e.g. aural communication versus visual communication).

Because questionnaire design cannot reduce mode-inherent errors (e.g. how people answer sensitive questions), estimating and adjusting for unwanted mode effects is necessary (de Leeuw, 2018). In addition, treating differences found in data with some caution would be advisable, especially for more subjective measures and sensitive topics. In the longer run, if quality is to be improved, efforts to reduce the variety of modes used within and between countries should be a priority, especially with a shift towards greater use of online interviewing.

## References

- Bennink, M., Moors, G. and Gelissen, J. (2013), 'Exploring response differences between face-to-face and web surveys: a qualitative comparative analysis of the Dutch European Values Survey 2008', *Field Methods*, Vol. 25, No 4, pp. 319–338.
- Beullens, K., Loosveldt, G., Vandenplas, C. and Stoop, I. (2018), 'Response rates in the European Social Survey: increasing, decreasing, or a matter of fieldwork efforts?', *Survey Methods: Insights from the Field* (<https://surveyinsights.org/?p=9673>).
- Bottoni, G. and Fitzgerald, R. (2021), 'Establishing a baseline: bringing innovation to the evaluation of cross-national probability-based online panels', *Survey Research Methods*, Vol. 15, No 2, pp. 115–133.

- Cernat, A. and Revilla, M. (2020), 'Moving from face-to-face to a web panel: impacts on measurement quality', *Journal of Survey Statistics and Methodology*, Vol. 9, No 4, pp. 745–763.
- Christmann, P., Gummer, T., Hähnel, S. and Wolf, C. (2019), 'Does the mode matter? An experimental comparison of survey responses between face-to-face and mixed-mode surveys', paper presented at the 8th Conference of the European Survey Research Association, 15–19 July, Zagreb, Croatia.
- Cornesse, C. and Bošnjak, M. (2018), 'Is there an association between survey characteristics and representativeness? A meta-analysis', *Survey Research Methods*, Vol. 12, No 1, pp. 1–13.
- Daikeler, J., Bošnjak, M. and Manfreda, K. L. (2020), 'Web versus other survey modes: an updated and extended meta-analysis comparing response rates', *Journal of Survey Statistics and Methodology*, Vol. 8, No 3, pp. 513–539.
- De Jong, J. (2016), 'Data collection: face-to-face surveys', in *Guidelines for Best Practice in Cross-cultural Surveys*, Survey Research Centre, Institute for Social Research, University of Michigan, Ann Arbor, MI (<https://ccsg.isr.umich.edu/index.php/chapters/data-collection-chapter/face-to-face-surveys>).
- de Leeuw, E. D. (2018), 'Mixed-mode: past, present, and future', *Survey Research Methods*, Vol. 12, No 2, pp. 75–89.
- de Leeuw, E. D., Hox, J. and Scherpenzeel, A. (2019), 'Mode effects versus question format effects: an experimental investigation of measurement error implemented in a probability-based online panel', in Lavrakas, P., Traugott, M., Kennedy, C., Holbrook, A., de Leeuw E. and West, B. (eds), *Experimental Methods in Survey Research: Techniques that combine random sampling with random assignment*, Wiley, Hoboken, NJ, pp. 151–165.
- de Leeuw, E. D., Suzer-Gurtekin, Z. T. and Hox, J. J. (2019), 'The design and implementation of mixed-mode surveys', in Johnson, T. P., Pennell, B. E., Stoop, I. A. L. and Dorer, B. (eds), *Advances in Comparative Survey Methods*, Wiley, Hoboken, NJ, pp. 387–408.
- DESI (Digital Economy and Society Index) (2020), *Use of Internet Services*, European Commission, Brussels (<https://digital-strategy.ec.europa.eu/en/policies/desi-use-internet>).
- Dillman, D. A. (2017), 'The promise and challenge of pushing respondents to the web in mixed-mode surveys', *Survey Methodology*, Vol. 43, No 1, pp. 3–30.
- Fitzgerald, R. and Jowell, R. (2010), 'Measurement equivalence in comparative surveys: the European Social Survey – from design to implementation and beyond', in Harkness, J. A., Braun, M., Edwards, B., Johnson, T. P., Lyberg, L., Mohler, P. P. et al. (eds), *Survey Methods in Multinational, Multiregional, and Multicultural Contexts*, Wiley, Hoboken, NJ, pp. 485–495.
- Jäckle, A., Roberts, C. and Lynn, P. (2006), 'Telephone versus face-to-face interviewing: mode effects on data quality and likely causes – report on Phase II of the ESS-Gallup Mixed Mode Methodology Project', *Institute for Social and Economic Research Working Papers*, No 2006-41, Institute for Social and Economic Research, University of Essex, Colchester.
- Jessop, C., Bottoni G., Sommer E., Sibley E. and Fitzgerald R. (2019), 'Blueprint for comparative web panel', deliverable 7.7 of the Synergies for Europe's Research Infrastructures in the Social Sciences project funded under the European Union's Horizon 2020 research and innovation programme under grant agreement No 654221 ([www.seriss.eu/resources/deliverables](http://www.seriss.eu/resources/deliverables)).
- Krumpal, I. (2013), 'Determinants of social desirability bias in sensitive surveys: a literature review', *Quality and Quantity*, Vol. 47, pp. 2025–2047.
- Leeper, T. J. (2019), 'Where have the respondents gone? Perhaps we ate them all', *Public Opinion Quarterly*, Vol. 83, pp. 280–288.
- Liu, M., Conrad, F. G. and Lee, S. (2017), 'Comparing acquiescent and extreme response styles in face-to-face and web surveys', *Quality and Quantity*, Vol. 51, pp. 941–958.
- Luijckx, R., Jónsdóttir, G. A., Gummer, T., Ernst Stähli, M., Frederiksen, M., Ketola, K. et al. (2020), 'The European Values Study 2017: on the way to the future using mixed-modes', *European Sociological Review*, Vol. 37, No 2, pp. 330–347.
- Luiten, A, Hox J. and de Leeuw, E. (2020), 'Non-response trends and fieldwork effort in the 21st century: results of an international study across coun-

- tries and surveys', *Journal of Official Statistics*, Vol. 36, No 3, pp. 469–487.
- Malnar, B. (2020), *European Social Survey Academic Impact Monitoring – Annual report 2019*, University of Ljubljana, Ljubljana.
- Martin, P. (2011), 'A good mix? Mixed mode data collection and cross-national surveys', *Research & Methods*, Vol. 20, No 1, pp. 5–26.
- Martin, P. and Lynn, P. (2011), 'The effect of mixed mode survey designs on simple and complex analyses', *Centre for Comparative Social Surveys Working Paper Series*, No 4, Centre for Comparative Social Surveys, City, University of London, London.
- Moore, J. C., Stinson, L. and Welniak, E. J. (2000), 'Income measurement error in surveys: a review', *Journal of Official Statistics*, Vol. 16, No 4, pp. 331–361.
- Olson, K. and Smyth, J. (2014), 'Accuracy of within-household selection in web and mail surveys of the general population', *Field Methods*, Vol. 25, No 1, pp. 56–69.
- Olson, K. and Smyth, J. (2017), 'Within-household selection in mail surveys: explicit questions are better than cover letter instructions', *Public Opinion Quarterly*, Vol. 81, No 3, pp. 688–713.
- Olson, K., Smyth, J., Horwitz, R., Keeter, S., Lesser, V., Marken, S. et al. (2020), 'Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR task force report', *Journal of Survey Statistics and Methodology*, pp. 1–31.
- ONS (Office for National Statistics) (2019), 'Internet users in the UK: 2019' (<https://www.ons.gov.uk/releases/internetusersintheuk2019>).
- Pennell, B. E., Cibelli Hibben, K., Lyberg, L. E., Mohler, P. and Worku, G. (2017), 'A total survey error perspective on surveys in multinational, multiregional, and multicultural contexts', in Biemer, P. P., de Leeuw, E., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L. E. et al. (eds), *Total Survey Error in Practice*, Wiley, Hoboken, NJ, pp. 179–201.
- Schober, M. F. (2017), 'The future of face-to-face interviewing', *Quality Assurance in Education*, Vol. 26, No 2, pp. 290–302.
- Smith, P. (2018), 'The feasibility of a cross-European push to web survey', in Cleary, A., Stannard, J., Smith, P., de Luis Iglesias, R., Horton, S., Clemens, S. et al., *Ipsos Research Methods Centre Update: November 2018*, pp. 7–10 (<https://www.ipsos.com/sites/default/files/ct/publication/documents/2018-11/ipsos-research-methods-newsletter-november-2018.pdf>).
- Smith, T. (2011), 'Refining the total survey error perspective', *International Journal of Public Opinion Research*, Vol. 29, No 4, pp. 464–484.
- Sommer, E. (2019), 'Survey experts network meeting 4: survey fieldwork cost', deliverable 5.12 of the Synergies for Europe's Research Infrastructures in the Social Sciences project funded under the European Union's Horizon 2020 research and innovation programme under grant agreement No 65422 ([https://seriss.eu/wp-content/uploads/2020/08/SERISS-Deliverable-5.12-fieldwork-cost-workshop-report\\_final.pdf](https://seriss.eu/wp-content/uploads/2020/08/SERISS-Deliverable-5.12-fieldwork-cost-workshop-report_final.pdf)).
- Technopolis Group (2017), *Comparative impact study of the European Social Survey (ESS) ERIC – Final report* (<https://www.europeansocialsurvey.org/docs/findings/ESS-Impact-study-Final-report.pdf>).
- Vannieuwenhuyze, J., Loosveldt, G. and Molenberghs, G. (2010), 'A method for evaluating mode effects in mixed-mode surveys', *Public Opinion Quarterly*, Vol. 74, No 5, pp. 1027–1045.
- Villar, A. and Fitzgerald, R. (2017), 'Using mixed modes in survey research: evidence from six experiments in the ESS', in Breen, M. J. (ed.), *Values and Identities in Europe – Evidence from the European Social Survey*, Routledge, New York, pp. 259–293.
- West, B. and Blom, A. G. (2017), 'Explaining interviewer effects: a research synthesis', *Journal of Survey Statistics and Methodology*, Vol. 5, No 2, pp. 175–211.
- Wuyts, C. and Loosveldt, G. (2020), 'Measurement of interviewer workload within the survey and an exploration of workload effects on interviewers' field efforts and performance', *Journal of Official Statistics*, Vol. 36, No 3, pp. 561–588.

