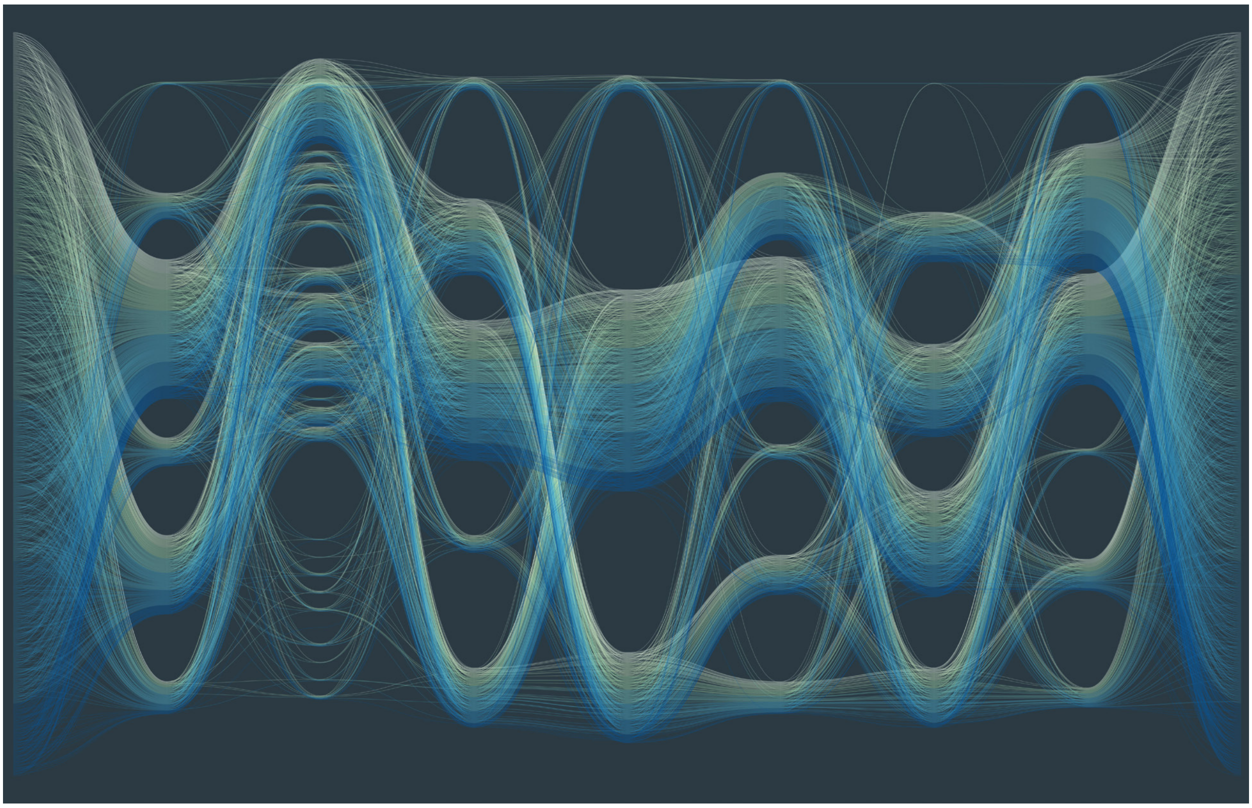
Australian National University

CENTRE FOR SOCIAL RESEARCH & METHODS

Social Research Centre

A SUBSIDIARY OF THE AUSTRALIAN NATIONAL UNVERSITY



# Panel mixed-mode effects: does switching modes in probability-based online panels influence measurement error?

**S Kocar and N Biddle**

CSRM & SRC METHODS PAPER

NO. 1/2020

## Series note

The ANU Centre for Social Research & Methods (CSRM) was established in 2015 to provide national leadership in the study of Australian society. CSRM has a strategic focus on:

- development of social research methods
- analysis of social issues and policy
- training in social science methods
- providing access to social scientific data.

CSRM publications are produced to enable widespread discussion and comment, and are available for free download from the CSRM website (http://csrm.cass.anu.edu.au/research/publications).

CSRM is located within the Research School of Social Sciences in the College of Arts & Social Sciences at the Australian National University (ANU). The centre is a joint initiative between the Social Research Centre and the ANU. Its expertise includes quantitative, qualitative and experimental research methodologies; public opinion and behaviour measurement; survey design; data collection and analysis; data archiving and management; and professional education in social research methods.

As with all CSRM publications, the views expressed in this Methods Paper are those of the authors and do not reflect any official CSRM position.

**Professor Matthew Gray**
Director, ANU Centre for Social Research & Methods
Research School of Social Sciences
College of Arts & Social Sciences
The Australian National University
January 2020

Methods Paper No. 1/2020

ISSN 2209-184X
ISBN 978-1-925715-00-2

An electronic publication downloaded from http://csrm.cass.anu.edu.au/research/publications.

For a complete list of CSRM Methods Papers, see http://csrm.cass.anu.edu.au/research/publications/methods-research-papers.

**ANU Centre for Social Research & Methods**
Research School of Social Sciences
The Australian National University

# Panel mixed-mode effects: does switching modes in probability-based online panels influence measurement error?


## S Kocar and N Biddle

**Sebastian Kocar** is a PhD candidate at the Centre for Social Research & Methods, Research School of Social Sciences, College of Arts & Social Sciences, Australian National University. He is also a Data Archivist at the Australian Data Archive. He specialises in social science statistics, research methodology and longitudinal data archiving. In his PhD research, he investigates survey errors, with a particular focus on online panel survey research.

**Nicholas Biddle** is Professor and Associate Director of the Centre for Social Research & Methods, Research School of Social Sciences, College of Arts & Social Sciences, Australian National University. He is also a Fellow at the Tax and Transfer Policy Institute, Crawford School of Public Policy, College of Arts & Social Sciences, Australian National University.


# Abstract


Online probability-based panels often apply two or more data collection modes to cover online and offline populations, and to collect data from onliners who do not respond online in time to contribute to a given wave. As a result, offline/online status can change during the life of the panel for some individuals, which can improve response rates and representativeness, but may cause increased measurement error.

In this study, we use Life in Australia™ survey data and online panel paradata to identify respondents who switched modes; almost 4% of the whole panel was interviewed using both online and offline modes in the first 2 years, and almost one-third of those 4% switched mode more than once. We selected all repeated substantive survey items, identified any relevant changes in responses that could be explained with mode effects, and determined the effect of mode switching on changes to answers, controlling for panel conditioning, panel fatigue and sociodemographic characteristics of panellists.

This study identified a limited number of panel mode effects from panellists switching modes of data collection over time. We found evidence of recency and some social desirability, and established that measurement error may be more common when the proportion of mode switchers is higher. Moreover, panel conditioning had an effect on the frequency of changing answers; respondents provided more stable answers if they were more conditioned. We conclude that combining mode effects with panel conditioning, as well as an increasing representation bias over time, may lead to less accurate estimations in longitudinal surveys.

# Acknowledgments

This research was supported by the Australian Government Research Training Program Scholarship. We thank the Social Research Centre for providing access to its Life in Australia™ data and paradata.

# Acronyms

| | |
|---|---|
| ANU | Australian National University |
| CSRM | ANU Centre for Social Research & Methods |
| OLS | ordinary least squares |

# Contents



# Tables and figures

# 1 Introduction

## 1.1 Mixing modes in online panel research

Organisations managing nonprobability-based panels tend to use only the online mode to collect data from their panellists, whereas most probability-based online panels try to allow for the fact that internet penetration is still not close to 100% (Baker at al. 2010). The aim of probability-based panels is to represent the general population, and people without computer or internet access should be included in the panel.

This comes with a cost, and the balance between measurement equivalence and coverage is important (Blom et al. 2016). One way to reduce noncoverage bias is to provide respondents with computer hardware and internet access. An alternative is to collect data using a mix of modes, including interviewer-administered modes (Baker at al. 2010). In Europe, four probability-based panels include the offline population. Different organisations managing probability-based online panels use different approaches to find the right balance between measurement errors, representation errors and costs. The French ELIPSS Panel focuses more on measurement equivalence by subjecting panellists to the same stimulus; consequently, all of them receive a tablet computer with internet access to fill out online questionnaires. On the other hand, the German GESIS panel is a mixed-mode panel, with the offline population participating via mailed paper questionnaires. The trade-off is that this does not guarantee measurement equivalence. More equivalence is offered by the Dutch LISS Panel and the German Internet Panel (GIP), but still less than in the case of the ELIPSS panel, because of different devices and browsers being used – only the offline population receives tablets (Blom et al. 2016).

Life in Australia™ is, similar to the GESIS panel, a mixed-mode panel. However, it uses a different offline mode – interviewer-administered telephone mode. To reduce representation and response bias, some respondents in Life in Australia™ end up being part of both online and offline populations during the panel lifecycle. Some offliners do not provide an email address at recruitment, but provide one in later waves, which means that they start by participating offline and later switch to responding online. Also, onliners are first contacted and reminded via email, then text message, but might be contacted over the phone later if they do not respond online after a certain time; a small percentage of them even respond over the phone (Kazmirek et al. 2019). Similar changes between online and offline populations happened in the LISS, GIP and ELIPSS panels. Those panels include between 7% and 10% of online panellists who were previously offline (Blom et al. 2016), which should not come as a surprise because the stability of mode preferences of longitudinal respondents is fairly low (Baghal & Kelley 2016).

## 1.2 Mode effects in mixed-mode panel research

In mixed-mode probability-based online panels, improvements in representativeness may come at the cost of measurement error. Generally speaking, there are two specific hypotheses about the possible impact of shifting from one mode to another: social desirability response bias and satisficing. Social desirability is the tendency of certain respondents to report more socially desirable, acceptable answers or those in sync with the popular opinion, rather than choosing answers reflecting their true feelings or thoughts (Grimm 2010). It is a consequence of two separate factors: self-deception and other-deception (Nederhof 1985). Satisficing occurs when a respondent generates valid but not necessarily accurate or thoughtful answers to survey questions by decreasing their cognitive effort

(Baker et al. 2010). Although social desirability should be a more significant issue in the case of interviewer surveys, self-administration might have the potential for a higher incidence of satisficing because of the ease of responding (Baker et al. 2010). Mixing interviewer-administered and self-administered modes should, therefore, result in an increased mode effect bias. In a longitudinal design, this may cause measurement inequivalence both between waves and between different modes (Cernat 2015), because change cannot be measured accurately if the respondent is presented with the same question stimulus in each wave (Dillman 2009). It has been argued (de Leeuw 2005) that mode effects should be an important survey design consideration and should be reduced as much as possible, and, in a longitudinal design, mode experiments should be carried out to help adjust for mode effects.

The concept of social desirability is made up of four nested characteristics, from large scale to small scale: cultural characteristics, personality characteristics, data collection mode and item characteristics. It is, therefore, important to control for question wording and mode of data collection, particularly in mixed-mode or cross-cultural research (Lavrakas 2008). Social desirability is more prevalent in survey modes allowing respondent identification, and in data collection in the presence of other people, and is related to questions on widely accepted attitudes, and behavioural and social norms (Grimm 2010). Bias can also be observed in interviewer-administered surveys, with respondents reporting higher satisfaction with their jobs (Kim & Kim 2016); products in market research (Albert & Tullis 2013); family, social life, health and financial troubles (Keeter et al. 2015); and democracy. Those interviewed on the phone also reported more trust in the ruling party (Zimbalist 2017) and more positive ratings of political figures (Keeter et al. 2015).

The main reason for satisficing is that some respondents tend to make the task of responding as easy as they can. This leads to using ranges or rounding values, making ratings following a few simple principles or not considering questions seriously (Tourangeau et al. 2000:254). Satisficing, which tends to have a higher incidence in self-administered surveys, can be observed in different ways: nondifferentiation, nonsubstantive responses, rapid completion (speeding), and response-order effects, such as primacy and recency (Baker et al. 2010). Primacy is a tendency to choose the first-offered answers, and recency is a tendency to choose among last categories regardless of the content (Dillman et al. 2014:104–105). It is expected that computer administration will yield the opposite response-order effect (primacy) to oral administration (recency), and so will give different distributions of responses (Baker et al. 2010).

## 1.3 Panel conditioning, panel fatigue and (in)stability of responses over time

Compared with cross-sectional research, panel research has more sources of measurement bias, such as panel conditioning and panel fatigue, which are not only interesting in themselves but should be considered when analysing mode effects specific to panels. Panel conditioning occurs when a sample unit's response is influenced by prior survey participation or contacts, and introduces so-called 'time-in-sample' bias. Because it affects responses in future waves, the estimates for certain supposedly stable concepts vary significantly over time (Lavrakas 2008). For example, in the study conducted by Halpern-Manners and Warren (2012), conditioned respondents reported lower unemployment rates and higher incidence of leaving the labour force than respondents who participated in the survey for the first time, even after controlling for attrition and mode effects. Panel fatigue, on the other hand, occurs when the quality of data from a particular respondent diminishes because they stay in the panel, providing data, for too long. It results in unit nonresponse, item nonresponse, satisficing, and other forms of lower-quality data (Lavrakas 2008).

These sources of measurement bias could be the reason for lower or higher stability of answers, or lower or higher quality of data in panel research over time. Although some authors report panel conditioning for a limited number of knowledge items only in online panel research (see Kruse et al. [2009]), other authors report significant effects of panel conditioning in longitudinal research. For example, Cernat (2015) reported an effect of panel conditioning on stability, with the results showing that stability increases with time even if no mode differences are apparent. In Australia, Wooden and Li (2013) presented similar findings – repeated participation resulted in a clear and gradual reduction in the dispersion of the target variables. Moreover, Sturgis et al. (2009) reported a reduction in the fraction of nonsubstantive answers over time as a form of panel conditioning.

## 1.4 Research questions

In this study, we answer the following research questions about the presence of respondent-level panel mode effects after controlling for sociodemographic characteristics of panellists and panel conditioning:

1. Is switching from interviewer-administered mode (telephone, offline) to self-administered mode (online) associated with changes in answers over time?
2. Does switching from interviewer-administered mode (telephone, offline) to self-administered mode (online) in probability-based online panels influence satisficing?
3. Does switching from self-administered mode (online) to interviewer-administered mode (telephone, offline) influence social desirability?
4. Does switching from self-administered mode (online) to interviewer-administered mode (telephone, offline) influence item nonresponse?

To answer these questions, the remainder of the paper is structured as follows: Section 2 presents the data, methods and survey items selected to study panel mode effects; Section 3 presents results of the mode effect analysis; and Section 4 discusses the results and practical implications of respondents switching modes in online panel research.

# 2 Methods

## 2.1 Data

The data used in this study were collected for the Australian National University by the Social Research Centre using its probability-based panel known as Life in Australia™. Five of six datasets used in this study are from ANUPolls, quarterly surveys of Australian public opinion (CSRM 2019).

Sociodemographic variables were from Life in Australia™ profile data files. Certain predictors, such as mode switching, were derived from panel participation variables from Life in Australia™ online panel paradata files.

## 2.2 Population, sample and data collection modes

The population in this research can be defined as ‘Australian residents aged 18 years or more’, and the results from the surveys are generalisable to the Australian population. The response rate for the establishment of Life in Australia™, calculated as the product of the recruitment rate and the profile rate, was 15.5% in 2016 (*n* = 3322) and 12.2% in August 2018 (refreshment, *n* = 267). To undertake recruitment, a dual-frame random-digit dialling (RDD) sample design was employed, with a 30:70 split between landline and mobile phone sample frames in 2016; a single-frame RDD mobile sample design was employed in 2018. ‘Last birthday’ method was used to select potential panel members in landline frames and the phone answerers in the mobile sample, although only one person per household was invited to join the panel. To cover online panellists, the online web self-completion mode was used. To collect data from offline panellists, the telephone mode was used (Kazmirek et al. 2010).

Because we studied changes in responses over time, only those respondents participating in at least two waves (see Table 1) were included (*n* = 2542). About 1% (wave 21 → wave 22) and about 3% (wave 1 → wave 3) of panellists changed the survey administration mode between the analysed waves (see Table 2). Although we worked with a relatively large sample of Life in Australia™ respondents whose answers were potentially sensitive to panel conditioning, we ended up with a relatively small sample size of panellists who changed the mode of data collection (127 respondents, who switched modes a total of 172 times). That might negatively affect the reliability of results related to panel mode effects, because small samples often leave the null hypothesis unchallenged.

**Table 1 Survey data used in this study**

| SRC code | Title of survey | Month and year | Wave |
|---|---|---|---|
| 1832 | Australian Personas Survey, 2016 | December 2016 | 1 |
| 1839 | ANUPoll 2017 Housing | March 2017 | 3 |
| 2009 | ANUPoll 2017 Job Security | October 2017 | 10 |
| 2150 | ANUPoll 2018 Populism | August 2018 | 19 |
| 2165 | ANUPoll 2018 Data Governance | October 2018 | 21 |
| 2170 | ANUPoll 2018 Population | November 2018 | 22 |

ANU = Australian National University; SRC = Social Research Centre

## 2.3 Data analysis

Mode effects in cross-sectional mixed-mode studies are usually tested using binomial, ordinal and multinomial logistic regressions, partial proportional and proportional odds models, ordinary least squares (OLS) models, or structural equation modelling (SEM) (Jäckle et al. 2010). In this study, because we worked with panel data, we used:

- binary logit regression (pooled)
- multinomial logistic regression (pooled)
- multiple linear regression (OLS, pooled)
- fixed- and random-effect panel logit regression
- fixed- and random-effect panel OLS regression.

To establish what panel data model should be used in controlling for unobserved heterogeneity, fixed or random effect, we performed the Hausman test for endogeneity (Hausman 1978) each time.

## 2.4 Selection of data items to investigate panel mode effects

Because Life in Australia™ is not primarily used to measure longitudinal changes over time, few items are repeated. In ANUPolls, four variables are included in each wave to measure change over time. The items that appeared in all six waves of Life in Australia™ online panel data used in this study were:

- satisfaction with the way Australia is heading ('satisfaction')
- the most important problem facing Australia ('1st problem')
- the second most important problem facing Australia ('2nd problem')
- party support in federal election for the House of Representatives ('party support').

## 2.5 Statistical models

Panel mode effects were investigated by studying changes in responses to the same questions over time, conditional to changes in survey administration modes for panellists over time, and controlling for the extent of panel conditioning and sociodemographic characteristics of panellists.

In our models, the derived dependent variables measuring changes in responses from the same respondents over time attempted to capture certain concepts described in the literature on mode effects in cross-sectional studies – that is, those that can be observed with panel data:

- Changes in answers regardless of the change type, all four substantive items (stability): logit regression with a binary dependent variable coded as 0 – no change, 1 – any change; multiple regression analysis with a continuous dependent variable *number of answer changes in a wave* (range 0–4).
- Change from substantive to nonsubstantive answers and vice versa, all four substantive items (sensitivity): multiple regression analysis with a continuous dependent variable *number of changes between substantive and nonsubstantive answers in a particular wave* (range –4 to 4, where negative values represent increase in the total number of nonsubstantive answers).
- Change from any substantive answer to the first listed answer and vice versa, satisfaction and party support items (primacy effect): multinomial regression with a nominal dependent variable coded as 0 – no change, 1 – other answer to the first one, 2 – first answer to other.
- Change from any substantive answer to the last listed answer and vice versa, satisfaction (recency effect): multinomial regression with a nominal dependent variable coded as 0 – no change, 1 – other answer to the last one, 2 – last answer to other.
- Increase of satisfaction, change from less popular answers to more popular answers (1st problem, party support), change from other answers to 'environment' (1st problem), vice versa (social desirability): multinomial regression with a nominal dependent variable coded as 0 – no change, 1 – change to socially desirable answer, 2 – change to socially less desirable answer.

We also derived a number of regressors for our regression models. The following independent variables are included, with information about which model they are included in.

**Mode effects**

- Change of mode, binary regressor coded as 0 – no change, 1 – any change (panel mode effects), model with *any change of answers* as the dependent variable only.
- Change of mode, nominal regressor coded as 0 – no change, 1 – online to telephone, 2 – telephone to online (panel mode effects), all other models.

**Panel conditioning**

- Number of times a respondent was asked the same question before responding in a particular wave (extent of panel conditioning) – we assume that panel conditioning will have a greater effect if certain questions are asked more times.
- Time in months since previously asked the question (extent of panel conditioning) – we assume that panel conditioning is less severe if the gap in time between asking the respondents the same question is greater; this item could also have an effect on the propensity of changing answers that cannot be consistent over time.

**Panel fatigue**

- The total number of waves a respondent participated in over the first 22 waves of Life in Australia™ data collection. It should be kept in mind that we also include non-ANU-based surveys in this calculation, which made up 16 of the first 22 waves of data collection. This control variable had to be added, because the literature reports that sensitivity of answers is related to item nonresponse, but so is panel fatigue, which is further associated with nonsubstantive responses, panel nonresponse and attrition (see Lavrakas [2008] for more detail).

**Demographic controls**

- Age group.
- Gender.
- Education.
- State.
- Country-of-birth group.

# 3 Results

## 3.1 Descriptive statistics

First, we present the results of the descriptive analysis of the composition of the sample in this study (Table 2). We have calculated and added propensities (predictive margins from logistic regression models) for the panel phenomena empirically investigated in the next sections of this paper – propensity of a particular sociodemographic subgroup to participate in a panel wave, propensity to change answers to the same questions over time, and propensity to change mode (online to telephone or vice versa) over time (holding other factors constant).

In the full Life in Australia™ sample, females; people between 55 and 75 years of age; the more educated (bachelor degree or higher and certificate/diploma/trade); people living in the ACT, Tasmania or South Australia; and Australian-born people are overrepresented compared with Australian 2016 Census results. The youngest (18–24 years of age), people living in New South Wales and the least educated are the most underrepresented (ABS 2016). In the seven waves of data that we use in this study, females, respondents 55 years of age and older, people living in South Australia, the most educated and Australian-born people have a higher propensity to participate in surveys. The propensity to answer questions consistently over time is higher for females and people with certificate/diploma/trade and year 11 or less education levels. The propensities to change modes are very low and are fairly consistent across all sociodemographic groups.

## 3.2 Stability of answers over time

Further to the propensity to change any answer in consecutive waves participated in, Figure 1 presents propensities for changing answers to any of the four specific questions repeated in ANU-commissioned Life in Australia™ surveys, over time. Different types of changes in answers to individual questions are presented as well.

The results show that party support is the item with the greatest consistency of answers over time, and 2nd problem is the item with the least consistency, which is why it was not included in the analysis presented in Table 2. Overall, there seem to be two factors associated with the propensity to change answers: the number of waves participated in and the time between waves participated in. The propensity for consistency generally increases over time and when the time between waves decreases. The same conclusion can be drawn for changes between substantive and nonsubstantive answers. The relationship between the factors is to be further explained using statistical modelling.

## 3.3 Panel mode effects controlled for panel conditioning

Because the focus of this research is on predictors of changes in answers over time, respondents who participated in at least two waves were included in the studied sample. In practice, this means that there was a certain level of panel conditioning present for all respondents. The results are presented by the different types of mode effects related to mixed-mode research and described in the literature:

- any change in answers (instability of answers)
- change between substantive and nonsubstantive answers (sensitivity)
- change from a substantive answer to the first answer (primacy)
- change from a substantive answer to the last answer (recency)
- change to potentially socially desirable answers (social desirability).

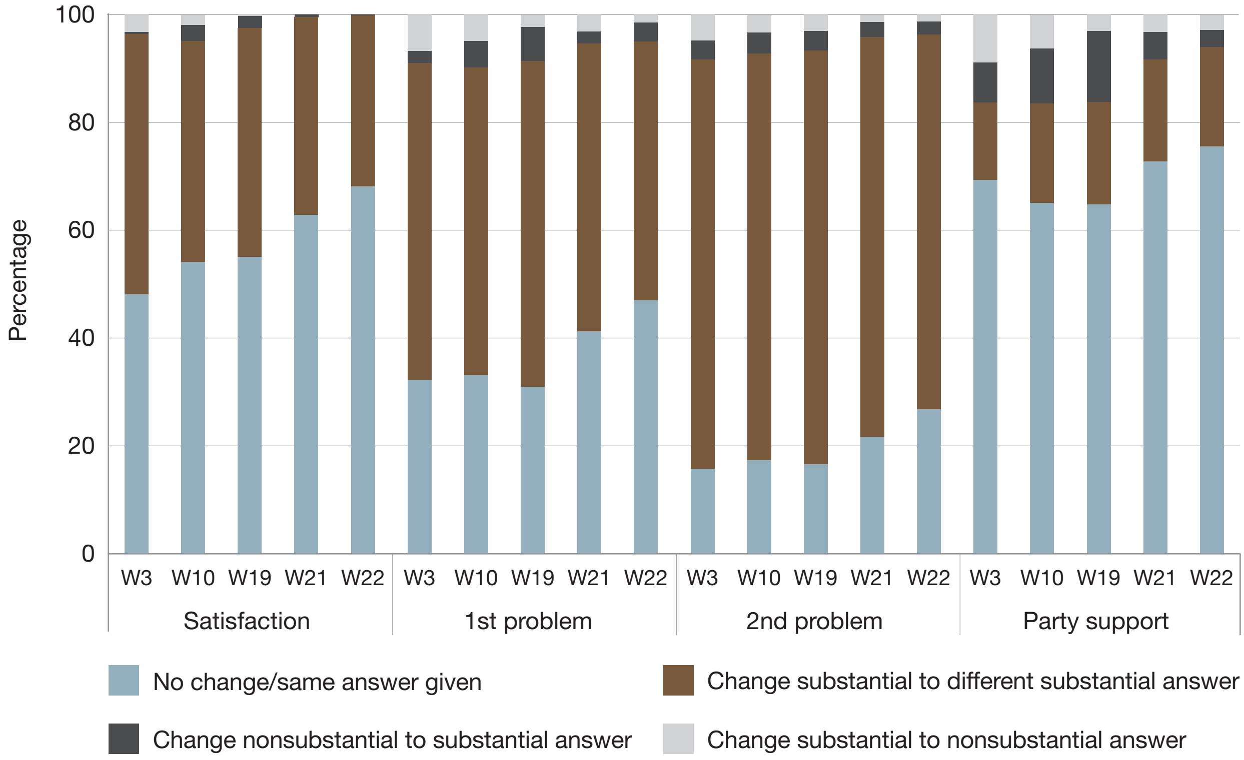
**Table 2** **Distribution of sociodemographic variables and calculated propensities for participation, changing answers and changing modes between waves**

| Control variable | Fraction of total (%) | Propensity to participate in wave | Propensity to change answer[a] | Propensity to change mode |
|---|---|---|---|---|
| | | Margin (95% CI) | Margin (95% CI) | Margin (95% CI) |
| Sex | | | | |
| Male | 48.0 | 0.666 (0.657, 0.675) | 0.827 (0.816, 0.837) | 0.014 (0.010, 0.017) |
| Female | 52.0 | 0.693 (0.684, 0.701) | 0.854 (0.845, 0.863) | 0.019 (0.015, 0.022) |
| Age group | | | | |
| 18–24 years | 9.2 | 0.512 (0.488, 0.536) | 0.864 (0.837, 0.891) | 0.009 (0.002, 0.016) |
| 25–34 years | 14.9 | 0.609 (0.591, 0.626) | 0.814 (0.793, 0.835) | 0.015 (0.008, 0.022) |
| 35–44 years | 14.9 | 0.640 (0.622, 0.657) | 0.826 (0.806, 0.845) | 0.012 (0.006, 0.018) |
| 45–54 years | 17.9 | 0.684 (0.669, 0.699) | 0.840 (0.823, 0.856) | 0.018 (0.012, 0.023) |
| 55–64 years | 19.6 | 0.729 (0.715, 0.742) | 0.845 (0.830, 0.859) | 0.015 (0.010, 0.020) |
| 65–74 years | 16.0 | 0.787 (0.773, 0.801) | 0.856 (0.840, 0.871) | 0.022 (0.015, 0.028) |
| 75 or more years | 7.4 | 0.707 (0.684, 0.729) | 0.851 (0.827, 0.875) | 0.021 (0.011, 0.030) |
| Education | | | | |
| Bachelor degree or higher | 37.2 | 0.742 (0.732, 0.751) | 0.822 (0.810, 0.833) | 0.011 (0.008, 0.014) |
| Certificate/diploma/ trade | 35.8 | 0.643 (0.632, 0.654) | 0.858 (0.846, 0.869) | 0.024 (0.019, 0.029) |
| Year 12 or equivalent | 12.3 | 0.682 (0.663, 0.700) | 0.841 (0.819, 0.862) | 0.017 (0.009, 0.025) |
| Year 11 or less | 14.8 | 0.605 (0.587, 0.623) | 0.858 (0.840, 0.876) | 0.012 (0.007, 0.017) |
| State | | | | |
| NSW | 30.0 | 0.663 (0.651, 0.675) | 0.838 (0.825, 0.851) | 0.015 (0.010, 0.019) |
| Vic | 25.2 | 0.680 (0.667, 0.692) | 0.843 (0.829, 0.857) | 0.015 (0.010, 0.020) |
| Qld | 19.4 | 0.688 (0.674, 0.702) | 0.852 (0.837, 0.867) | 0.015 (0.010, 0.020) |
| SA | 8.3 | 0.748 (0.727, 0.769) | 0.839 (0.816, 0.862) | 0.018 (0.010, 0.026) |
| WA | 11.4 | 0.676 (0.657, 0.694) | 0.829 (0.807, 0.850) | 0.019 (0.011, 0.027) |
| Tas | 2.6 | 0.65 (0.609, 0.690) | 0.846 (0.802, 0.889) | 0.029 (0.009, 0.048) |
| NT | 1.0 | 0.585 (0.517, 0.653) | 0.888 (0.822, 0.953) | 0.061 (0.009, 0.112) |
| ACT | 2.3 | 0.673 (0.629, 0.716) | 0.816 (0.768, 0.863) | 0.017 (0.000, 0.034) |
| Country of birth | | | | |
| Australia | 71.8 | 0.701 (0.693, 0.708) | 0.838 (0.830, 0.846) | 0.017 (0.014, 0.020) |
| Mainly NESB background | 16.1 | 0.604 (0.587, 0.621) | 0.856 (0.838, 0.873) | 0.015 (0.008, 0.022) |
| Mainly ESB background | 12.2 | 0.657 (0.638, 0.675) | 0.842 (0.822, 0.861) | 0.013 (0.007, 0.019) |

CI = confidence interval; ESB = English-speaking background; NESB = non-English-speaking background

a The propensity to change answer to any of satisfaction with the country heading, problem no. 1 in Australia, or party preference (problem no. 2 excluded) questions between consecutive waves completed.

**Figure 1** Types of changes in answers from the same respondents over time




W = wave

While satisficing and social desirability were controlled for by demographics and panel conditioning, sensitivity-associated item nonresponse was also controlled for by panel fatigue.

### 3.3.1 Instability of answers

The analysis of the types of changes in answers from the same respondents over time, results of which are shown in Figure 1, are here extended with multivariate analysis of instability of answers over time as an indicator of general mode effects and panel conditioning effects. The results of logit regression analysis (a pooled regression model) and dynamic logit regression analysis are presented in Tables 3 and 4. We carried out the Hausman test to look for a correlation between errors and regressors in the models, so we could choose between using fixed-effect and random-effect models. The results showed that fixed-effect models were the appropriate solution for controlling for unobserved time-invariant heterogeneity in all models except for the one with party support changes in answers as the dependent variable.

Regression analysis shows that mode changes, either online → telephone or telephone → online, are not predictors of changes in answers in any of the eight models for four response variables. For items with the lowest propensity for changing answers – satisfaction and party support – mode change predictors are almost statistically significant in the pooled models (satisfaction at $P = 0.1$, party support at $P = 0.15$). Further, the results show that both indicators of panel conditioning in all models were statistically significant predictors of changes in answers, although there was very little difference between the coefficients of pooled and dynamic logit models. Panel conditioning generally affects the changes in two different ways: the more times a question is asked, the lower the probability of change; and the longer the gap (measured in months) between a question being asked and then repeated, the higher the probability of change. The changes in answers to the satisfaction question seem to be more affected by how many times the question was asked than the changes in answers to the party support question.

**Table 3** **Logit regression, random-effect and fixed-effect within-person logistic regression results; dependent variable: *any change of answers***

| Substantive repeated item | Predictor of any change in answers over time | Logit regression model (pooled) | | | | Fixed-effect logit regression model | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| Satisfaction | Mode change (any) | 0.28 | −0.03 | 0.58 | 0.073 | 0.19 | −0.28 | 0.67 | 0.430 |
| | No. times question asked | −0.15 | −0.18 | −0.12 | <0.001** | −0.23 | −0.26 | −0.19 | <0.001** |
| | Months since question asked | 0.03 | 0.02 | 0.04 | <0.001** | 0.03 | 0.01 | 0.04 | <0.001** |
| 1st problem | Mode change (any) | 0.14 | −0.19 | 0.46 | 0.416 | −0.19 | −0.67 | 0.29 | 0.430 |
| | No. times question asked | −0.09 | −0.12 | −0.06 | <0.001** | −0.16 | −0.19 | −0.12 | <0.001** |
| | Months since question asked | 0.05 | 0.03 | 0.06 | <0.001** | 0.05 | 0.04 | 0.07 | <0.001** |
| 2nd problem | Mode change (any) | −0.08 | −0.48 | 0.32 | 0.696 | −0.23 | −0.81 | 0.35 | 0.440 |
| | No. times question asked | −0.11 | −0.15 | −0.08 | <0.001** | −0.17 | −0.21 | −0.12 | <0.001** |
| | Months since question asked | 0.04 | 0.03 | 0.05 | <0.001** | 0.04 | 0.02 | 0.05 | <0.001** |
| | | | | | | Random-effect logit regression model | | | |
| Party support | Mode change (any) | 0.26 | −0.06 | 0.57 | 0.108 | 0.21 | −0.21 | 0.63 | 0.325 |
| | No. times question asked | −0.06 | −0.09 | −0.03 | <0.001** | −0.09 | −0.12 | −0.05 | <0.001** |
| | Months since question asked | 0.05 | 0.04 | 0.06 | <0.001** | 0.06 | 0.05 | 0.08 | <0.001** |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

Note: Sociodemographic controls in the models were sex, age group, education, state, and country of birth.

**Table 4** **Multiple linear regression and fixed-effect within-person regression results; dependent variable: *number of changes of answers in a particular wave***

| Derived variable | Predictor of any changes in answers over time | Multiple linear regression model (pooled) | | | | Fixed-effect regression model | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| Number of any changes in answers | Mode change (any) | 0.15 | 0.01 | 0.30 | 0.039* | −0.01 | −0.19 | 0.16 | 0.881 |
| | No. times questions asked | −0.08 | −0.09 | −0.06 | <0.001** | −0.10 | −0.12 | −0.09 | <0.001** |
| | Months since questions asked | 0.03 | 0.03 | 0.04 | <0.001** | 0.03 | 0.03 | 0.04 | <0.001** |
| | Constant | 2.06 | 1.96 | 2.16 | <0.001** | 2.22 | 2.17 | 2.27 | <0.001** |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

Note: Sociodemographic controls in the models were sex, age group, education, state, and country of birth.

The results from Table 4 explain the relationship between changes in answers, mode changes and panel conditioning from a slightly different perspective. This time, the dependent variable in the models is the number of changes in answers between two consecutive waves (range 0–4). The results fully support the findings of modelling with individual survey items. However, this time, the mode change is a statistically significant predictor of the number of changes in answers in the pooled OLS model – mode changes increase the propensity to change answers.

### 3.3.2 Sensitivity

Sensitivity as a result of interviewer administration – the mode effect concept that can result in item nonresponse or nonsubstantive answer selection – was studied with static and dynamic regression models. Multiple linear regression analysis results (a pooled regression model) and dynamic regression analysis results with a continuous dependent variable are presented in Table 5. Based on the results of the Hausman test, we decided to carry out fixed-effect modelling. In these two regression models, the predictor variable *any mode change* from the previous models is split into *online to telephone* and *telephone to online* mode changes. Also, the *number of changes between substantive and nonsubstantive answers in a particular wave* (range 0–4) is this time controlled for panel fatigue as well (see Section 2.5 for more details).

The results show that the mode changes and the number of times questions were asked were not statistically significant predictors of the number of changes between substantive and nonsubstantive answers in a particular wave. On the other hand, both the number of months since the questions were asked and the panel fatigue indicator had an effect on the number of changes between substantive and nonsubstantive answers. The longer the gap in months between a question being asked and then repeated, and the more times respondents participated in Life in Australia™ research, the higher the probability of changes in the substantive answers direction. Panel fatigue, which was highly correlated with the number of times questions were asked, had a fairly small effect on answer changes.

**Table 5 Multiple linear regression and fixed-effect within-person regression results; dependent variable: *number of changes between substantive and nonsubstantive answers in a particular wave***

| Derived variable | Predictor of changes between substantive and nonsubstantive answers | Multiple linear regression model (pooled) | | | | Fixed-effect regression model | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| Number of changes between substantive and nonsubstantive answers in particular wave | Mode change: online to telephone | 0.06 | –0.06 | 0.17 | 0.336 | 0.10 | –0.09 | 0.29 | 0.309 |
| | Mode change: telephone to online | 0.01 | –0.10 | 0.12 | 0.872 | 0.06 | –0.09 | 0.21 | 0.448 |
| | No. times questions asked | –0.01 | –0.02 | 0.00 | 0.088 | 0.01 | –0.02 | 0.04 | 0.414 |
| | Months since questions asked | 0.02 | 0.01 | 0.02 | <0.001** | 0.02 | 0.02 | 0.02 | <0.001** |
| | Panel fatigue indicator | 0.010 | 0.007 | 0.013 | <0.001** | 0.01 | 0.00 | 0.01 | 0.026* |
| | Constant | –0.11 | –0.16 | –0.05 | <0.001** | –0.17 | –0.20 | –0.13 | <0.001** |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

### 3.3.3 Primacy

Primacy as a response-order effect was studied with static and dynamic regression models. The results of logit regression and random-effect regression models (after performing the Hausman test) with *change in satisfaction* and *change in party support* answers as the dependent variables are presented in Table 6. This time, the type of answer changes – both substantive answer to first-offered answer and first-offered answer to other substantive answer – were considered as well.

The results show that the mode changes *online to telephone* and *telephone to online* are not statistically significant predictors of primacy-related changes in answers. However, the change in the first-offered answer to the satisfaction question after switching to the online mode had a significant effect at the $P$ = 0.1 level in both static and dynamic models. Moreover, both panel conditioning indicators are statistically significant predictors of primacy-related satisfaction answer changes. Although *number of times question asked* had a positive effect on both types of answer changes (answer change effect), *months since question asked* showed only a primacy effect in this particular case for the satisfaction item. On the other hand, we did not observe any effects of predictor variables on the party support primacy-related answer changes.

### 3.3.4 Recency

Recency as a response-order effect was studied with static and dynamic regression models. The results of logit regression and random-effect regression models (after performing the Hausman test) with *change in satisfaction* answers as the dependent variable are presented in Table 7. The type of answer changes – both substantive answer to last-offered answer and last-offered answer to other substantive answer – were considered as well.

The results show that the *online to telephone* mode change predictor had a statistically significant effect on recency-related *substantive answer to last-offered answer* change of answers in both logit and random-effect logit models. The other mode change, *telephone to online,* had a significant positive effect on 'change away from recency' at $P$ = 0.1 in both models. The predictor *number of times question asked* had a statistically significant negative effect on recency (answer change effect), and *months since question asked* had a positive effect on any recency-related answer changes in dynamic models and in one of the two logit regression models.

### 3.3.5 Social desirability

Social desirability as a type of response bias, related to reporting more socially desirable, acceptable answers or those in sync with the popular opinion, was studied with static and dynamic regression models. The results of logit regression and fixed-effect regression models (after performing the Hausman test), with changes to socially desirable answers as dependent variables, are presented in Tables 8 and 9.

The results show that the mode change *online to telephone* is not a statistically significant predictor of the social desirability–related changes of satisfaction or 1st problem answers. On the other hand, the mode change *telephone to online* positively affects decreased satisfaction (pooled model only). Those respondents reported lower satisfaction in the self-administered mode. On the other hand, the mode change *online to telephone* positively affects changing party support answers to the 'popular opinion' answers, but at $P$ = 0.1 and in the pooled model only. Those respondents supported the two biggest Australian parties, with a slightly higher propensity in the interviewer-administered mode. On the other hand, *number of times questions asked* negatively affected any satisfaction and other changes to 'popular opinion' answers (party support, 1st problem), and positively affected changes between 'environment' and other answers and vice versa (answer change effect). Variable *months since question asked* positively affected changes to socially desirable answers: increased satisfaction and selecting 'popular opinion' answers to party support and 1st problem 'environment' answer (social desirability effect).

**Table 6** **Logit regression and random-effect within-person regression results; dependent variable: *change of answers from any to first-offered answer* (and vice versa)**

| Substantive repeated survey item | Type of change | Predictor of primacy change over time | Logit regression model (pooled) | | | | Random-effect logit regression model | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| Satisfaction | Substantive answer to first-offered answer | Mode change: online to telephone | −0.43 | −1.85 | 0.98 | 0.549 | −0.31 | −1.83 | 1.21 | 0.691 |
| | | Mode change: telephone to online | 0.59 | −0.25 | 1.43 | 0.170 | 0.66 | −0.30 | 1.61 | 0.177 |
| | | No. times question asked | 0.18 | 0.11 | 0.26 | 0.000** | 0.23 | 0.15 | 0.31 | 0.000** |
| | | Months since question asked | 0.06 | 0.04 | 0.09 | 0.000** | 0.07 | 0.04 | 0.09 | 0.000** |
| | First-offered answer to other substantive answer | Mode change: online to telephone | 0.33 | −0.84 | 1.50 | 0.583 | 0.42 | −0.84 | 1.68 | 0.512 |
| | | Mode change: telephone to online | 0.83 | −0.02 | 1.68 | 0.055 | 0.84 | −0.09 | 1.78 | 0.078 |
| | | No. times question asked | 0.09 | 0.02 | 0.16 | 0.016* | 0.10 | 0.02 | 0.17 | 0.010* |
| | | Months since question asked | −0.09 | −0.12 | −0.05 | 0.000** | −0.09 | −0.13 | −0.05 | 0.000** |
| Party support | Substantive answer to first-offered answer | Mode change: online to telephone | 0.11 | −0.92 | 1.13 | 0.836 | 0.16 | −0.99 | 1.32 | 0.783 |
| | | Mode change: telephone to online | 0.39 | −0.45 | 1.24 | 0.362 | 0.44 | −0.51 | 1.40 | 0.362 |
| | | No. times question asked | −0.05 | −0.11 | 0.02 | 0.135 | −0.05 | −0.12 | 0.02 | 0.181 |
| | | Months since question asked | 0.02 | −0.01 | 0.04 | 0.149 | 0.02 | 0.00 | 0.05 | 0.107 |
| | First-offered answer to other substantive answer | Mode change: online to telephone | −0.17 | −1.35 | 1.00 | 0.772 | −0.12 | −1.42 | 1.17 | 0.851 |
| | | Mode change: telephone to online | −0.72 | −2.14 | 0.69 | 0.315 | −0.90 | −2.42 | 0.62 | 0.245 |
| | | No. times question asked | 0.03 | −0.03 | 0.09 | 0.335 | 0.03 | −0.03 | 0.10 | 0.330 |
| | | Months since question asked | 0.00 | −0.03 | 0.02 | 0.709 | 0.00 | −0.03 | 0.02 | 0.793 |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

**Table 7** **Logit regression and random-effect within-person regression results; dependent variable: *change of answers from any to last-offered answer* (and vice versa)**

| Substantive repeated survey item | Type of change | Predictor of recency change over time | Logit regression model (pooled) | | | | Random-effect logit regression model | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| Satisfaction | Substantive answer to last-offered answer | Mode change: online to telephone | 1.06 | 0.35 | 1.78 | 0.004** | 1.05 | 0.17 | 1.92 | 0.019* |
| | | Mode change: telephone to online | −0.60 | −2.01 | 0.82 | 0.408 | −0.78 | −2.30 | 0.75 | 0.319 |
| | | No. times question asked | −0.13 | −0.20 | −0.06 | 0.000** | −0.13 | −0.21 | −0.06 | 0.001** |
| | | Months since question asked | 0.02 | 0.00 | 0.05 | 0.086 | 0.03 | 0.00 | 0.06 | 0.033* |
| | Last-offered answer to other substantive answer | Mode change: online to telephone | 0.33 | −0.70 | 1.35 | 0.533 | 0.26 | −0.88 | 1.40 | 0.659 |
| | | Mode change: telephone to online | 0.72 | −0.06 | 1.51 | 0.072 | 0.75 | −0.13 | 1.64 | 0.096 |
| | | No. times question asked | 0.04 | −0.03 | 0.11 | 0.278 | 0.04 | −0.03 | 0.12 | 0.288 |
| | | Months since question asked | 0.06 | 0.04 | 0.09 | 0.000** | 0.07 | 0.04 | 0.10 | 0.000** |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

**Table 8** **Logit regression and fixed-effect within-person regression results; dependent variable: *increased satisfaction and changed support of a 'popular' party* (and vice versa)**

| Substantive repeated survey item | Type of change | Predictor of change associated with social desirability over time | Logit regression model (pooled) | | | | Fixed-effect logit regression model | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| Satisfaction | Increased satisfaction | Mode change: online to telephone | −0.10 | −0.67 | 0.48 | 0.742 | −0.70 | −1.83 | 0.44 | 0.232 |
| | | Mode change: telephone to online | 0.37 | −0.14 | 0.88 | 0.155 | 0.24 | −0.47 | 0.94 | 0.512 |
| | | No. times question asked | −0.08 | −0.12 | −0.05 | 0.000** | −0.10 | −0.15 | −0.06 | 0.000** |
| | | Months since question asked | 0.06 | 0.05 | 0.07 | 0.000** | 0.07 | 0.05 | 0.08 | 0.000 |
| | Decreased satisfaction | Mode change: online to telephone | 0.25 | −0.26 | 0.77 | 0.336 | 0.25 | −0.77 | 1.28 | 0.629 |
| | | Mode change: telephone to online | 0.51 | 0.01 | 1.00 | 0.044* | 0.26 | −0.45 | 0.96 | 0.479 |
| | | No. times question asked | −0.22 | −0.25 | −0.18 | 0.000** | −0.31 | −0.36 | −0.27 | 0.000** |
| | | Months since question asked | −0.01 | −0.02 | 0.01 | 0.249 | −0.02 | −0.04 | 0.00 | 0.028 |
| Party support | Any other answer to 'popular opinion' answer | Mode change: online to telephone | 0.66 | −0.06 | 1.38 | 0.074 | 0.00 | −1.19 | 1.20 | 0.997 |
| | | Mode change: telephone to online | 0.19 | −0.60 | 0.98 | 0.634 | 0.34 | −0.75 | 1.43 | 0.545 |
| | | No. times question asked | −0.05 | −0.11 | 0.00 | 0.057 | −0.15 | −0.22 | −0.07 | 0.000** |
| | | Months since question asked | 0.03 | 0.01 | 0.05 | 0.004** | 0.03 | 0.01 | 0.06 | 0.007** |
| | 'Popular opinion' answer to any other answer | Mode change: online to telephone | 0.56 | −0.25 | 1.36 | 0.174 | 0.29 | −1.17 | 1.75 | 0.694 |
| | | Mode change: telephone to online | 0.13 | −0.71 | 0.98 | 0.758 | −0.30 | −1.52 | 0.91 | 0.628 |
| | | No. times question asked | −0.03 | −0.08 | 0.03 | 0.317 | −0.07 | −0.15 | 0.00 | 0.064 |
| | | Months since question asked | −0.02 | −0.04 | 0.01 | 0.186 | 0.00 | −0.03 | 0.03 | 0.856 |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

**Table 9** **Logit regression and fixed-effect within-person regression results; dependent variable: *changing answers to popular opinion about most important problems in the country* (and vice versa)**

| Substantive repeated survey item | Type of change | Predictor of change associated with social desirability over time | Logit regression model (pooled) | | | | Fixed-effect logit regression model | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Coef | L 95% CI | U 95% CI | *P* value | Coef | L 95% CI | U 95% CI | *P* value |
| 1st problem | Any other answer to 'popular opinion' answer | Mode change: online to telephone | 0.44 | −0.09 | 0.97 | 0.106 | 0.41 | −0.41 | 1.23 | 0.330 |
| | | Mode change: telephone to online | 0.09 | −0.47 | 0.65 | 0.747 | −0.13 | −0.85 | 0.59 | 0.732 |
| | | No. times question asked | −0.06 | −0.09 | −0.02 | 0.002** | −0.10 | −0.14 | −0.05 | 0.000** |
| | | Months since question asked | 0.00 | −0.02 | 0.01 | 0.955 | 0.01 | −0.01 | 0.02 | 0.550 |
| | 'Popular opinion' answer to any other answer | Mode change: online to telephone | −0.24 | −0.93 | 0.45 | 0.497 | −1.05 | −2.23 | 0.13 | 0.080 |
| | | Mode change: telephone to online | 0.05 | −0.52 | 0.63 | 0.856 | −0.01 | −0.82 | 0.81 | 0.990 |
| | | No. times question asked | −0.03 | −0.07 | 0.00 | 0.083 | −0.03 | −0.08 | 0.02 | 0.209 |
| | | Months since question asked | 0.02 | 0.00 | 0.03 | 0.027* | 0.04 | 0.02 | 0.05 | 0.000 |
| 1st problem | Any other answer to 'environment' answer | Mode change: online to telephone | 0.36 | −0.50 | 1.22 | 0.409 | 0.18 | −1.05 | 1.41 | 0.778 |
| | | Mode change: telephone to online | 0.19 | −0.66 | 1.04 | 0.661 | −0.28 | −1.50 | 0.94 | 0.654 |
| | | No. times question asked | 0.06 | 0.00 | 0.12 | 0.039* | 0.16 | 0.08 | 0.24 | 0.000** |
| | | Months since question asked | 0.06 | 0.04 | 0.08 | 0.000** | 0.09 | 0.06 | 0.12 | 0.000** |
| | 'Environment' answer to any other answer | Mode change: online to telephone | 0.48 | −0.45 | 1.41 | 0.312 | −0.78 | −2.47 | 0.91 | 0.365 |
| | | Mode change: telephone to online | −0.26 | −1.42 | 0.90 | 0.665 | −0.41 | −1.98 | 1.17 | 0.612 |
| | | No. times question asked | 0.09 | 0.03 | 0.15 | 0.004** | 0.18 | 0.10 | 0.27 | 0.000** |
| | | Months since question asked | 0.00 | −0.03 | 0.02 | 0.716 | 0.03 | 0.00 | 0.06 | 0.037 |

* = significant at the 0.05 level; ** = significant at the 0.01 level; Coef = model regression coefficient; L 95% CI = lower limit of 95% confidence interval; U 95% CI = upper limit of 95% confidence interval

# 4 Discussion and recommendations

The existing literature on mode effects in probability-based mixed-mode research is limited. One of the limitations to understanding mode effects better is the inability to fully disentangle mode effects from subsample composition effects. To achieve this, an optimal randomised design would have to be applied, which means that all onliners and offlliners in treatment and control groups would have to have an equal nonzero probability of being assigned to either the online or the offline mode. However, this kind of randomisation is almost impossible, because most offline respondents cannot, or refuse to, respond online, and the cost of administering telephone compared with online delivery mode means that survey companies are unlikely to significantly (and randomly) increase the number of offline repondents. In this study, we instead used the fact that certain respondents, although the percentage is small and nonrandom, appear in both modes over time and respond to a limited number of repeated questions. Consequently, we could not only study mode effects related to questionnaire administration, we could also assess how much of a measurement error may be introduced by allowing respondents to respond in different modes, especially if we would like to measure changes over time in a quasi-longitudinal design.

An important finding of this study was that answers from the same respondents vary greatly over time, even for items for which a slightly higher consistency would be expected, such as party preference, and for short time gaps between survey interviews. Stability of answers differed significantly between different political attitude items, but very little instability could be explained by sociodemographic characteristics. At the same time, respondents switching modes affected stability to a smaller extent than we expected based on the relevant literature. This might be a result of there being only a small subsample of mode switchers. We observed several coefficients that indicated an impact of switching modes on changing answers, consistent with the mode effect literature, but the effects were often significant at the $P = 0.1$ and not the $P = 0.05$ level. Bigger samples – for example, with additional panel survey data with mode switchers or people with a higher propensity to change modes – would increase the statistical power and may show the effects to be more statistically significant.

Nevertheless, we found a few notable measurement biases after switching modes. The mode change was a statistically significant predictor of the number of changes in answers – switching decreases the stability of answers, has a positive effect on recency when switching to interviewer-administered telephone mode, and has a negative effect on social desirability in the self-administered mode for a limited number of items. These findings are in line with the theory on mode effects in online panels (Baker et al. 2010). On the other hand, many different changes in answers could be better explained by panel conditioning. Generally, the more times the same questions are asked over time, the lower the probability of changes, and the longer the gap (measured in months) between asking questions, the lower the stability of answers. This is consistent with findings of Cernat (2015) and Wooden and Li (2013) on the effect of panel conditioning on reliability and stability of answers. In our study, the analysis of individual types of changes, normally attributed to mode effects, offered mixed evidence for both indicators of the extent of panel conditioning. Both regressors were associated with something we called 'answer change effect', but in some cases in a positive and in other cases in a negative way. The number of months since the question was asked was slightly more strongly associated with phenomena normally attributed to mode effects than the other indicator of panel conditioning. The contribution of this study is to present evidence on the severity of panel conditioning effects when respondents are conditioned repeatedly with short time intervals, sometimes being asked the same question in consecutive months. The existing literature on panel conditioning (e.g. Sturgis et al. 2009, Wooden & Li 2013, Cernat 2015) mostly studied this source of measurement error in longitudinal studies, where the time between data collection waves is much longer. We can conclude that panel conditioning seems to play an important role in the stability of answers; researchers should pay extra attention if the same question is asked several times in a short period of time, which might prevent respondents from reporting naturally changed attitudes over time.

In this study, we faced a number of limitations. Because we investigated mode effects as changes in responses to the same questions from the same respondents over time, we had to control the effect of switching modes with the other sources of measurement error specific to panel research. The reason for this is that, in this study design, all respondents were conditioned in at least one wave before providing the same or different answers in the next wave. The subsample used in this study therefore consisted of respondents who participated in at least two waves out of six for which we could find repeated items measuring political attitudes. Infrequent respondents were not included in the sample, which might have introduced some representation bias. Moreover, panel conditioning had to be controlled in the models in a slightly different way. We did not compare distributions of the selected response variables between those who answered the question for the first time and those who had been conditioned by being asked the same question in the past. With this study design, we instead controlled mode effects with the effect of the extent/severity of panel conditioning. Also, we note that some respondents, who are panellists in Life in Australia™, regularly participate in other cross-sectional and/or nonprobability-based panel research. Unfortunately, we could not control for potential panel conditioning as a result of survey participation outside of Life in Australia™ research. One out of five different concepts related to measurement bias – sensitivity – was investigated through item nonresponse and nonsubstantive answers. Because these concepts are associated with another source of measurement error specific to panel research – panel fatigue – we included an indicator of panel fatigue in the models investigating factors affecting sensitivity. Last but not least, our findings might be less generalisable in fields outside political attitudes research, because all of the survey items used in this study were from ongoing political poll research.

The contribution of this study is, first and foremost, in identifying certain mode effects, such as recency or social desirability, as a result of panellists switching modes of data collection. We also noted that switching modes might induce more measurement error due to satisficing and social desirability if the proportion of mode switchers was higher. Although mode effects themselves do not seem to affect the accuracy of estimates in a very negative way, combining them with panel conditioning, as well as opt-out attrition and nonresponse, may lead to less accurate estimations. The future research on the accuracy of estimation of attitudinal changes over time in probability-based panels should, therefore, focus on studying concurrent sources of survey errors specific to online panels and their effects on accuracy.

# References


ABS (Australian Bureau of Statistics) (2016). 2016 Census – cultural diversity. Findings based on use of ABS TableBuilder data.

Albert W & Tullis T (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*, 2nd edn, Morgan Kaufmann, Waltham, Massachusetts.

Baghal TA & Kelley J (2016). The stability of mode preferences: implications for tailoring in longitudinal surveys. *Methods, Data, Analyses* 10(2):143–166.

Baker R, Blumberg SJ, Brick JM and 17 other authors (2010). AAPOR report on online panels. *Public Opinion Quarterly* 74(4):711–781.

Blom AG, Bosnjak M, Cornilleau A, Cousteaux AS, Das M, Douhou S & Krieger U (2016). A comparison of four probability-based online and mixed-mode panels in Europe. *Social Science Computer Review* 34(1):8–25.

Cernat A (2015). The impact of mixing modes on reliability in longitudinal studies. *Sociological Methods & Research* 44:427–457.

CSRM (Centre for Social Research and Methods) (2019). *ANUPoll*, https://csrm.cass.anu.edu.au/research/surveys/anupoll.

de Leeuw ED (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics* 21(5):233–255.

Dillman DA (2009). Some consequences of survey mode changes in longitudinal surveys. In: Lynn P (ed.), *Methodology of longitudinal surveys*, Wiley, Chichester, 127–140.

——— , Smyth JD & Christian LM (2014). *Internet, phone, mail, and mixed-mode surveys: the tailored design method*, John Wiley & Sons.

Grimm P (2010). Social desirability bias. In: Kamakura W (ed.), *Wiley international encyclopedia of marketing – vol 2: Marketing research*, Wiley, Chichester.

Halpern-Manners A & Warren JR (2012). Panel conditioning in longitudinal studies: evidence from labor force items in the current population survey. *Demography* 49(4):1499–1519.

Hausman JA (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society* 46(6):1251–1271.

Jäckle A, Roberts C & Lynn P (2010). Assessing the effect of data collection mode on measurement. *International Statistical Review* 78(1):3–20.

Kazmirek L, Phillips B, Pennay DW, Lavrakas PJ & Neiger D (2019). *Building a probability-based online panel: Life in Australia™*, CSRM & SRC Methods Paper 2/2019, ANU Centre for Social Research & Methods, Canberra.

Keeter S, McGeeney K, Igielnik R, Mercer A & Mathiowetz N (2015). *From telephone to the web: the challenge of mode of interview effects in public opinion polls*, Pew Research Center, Washington, DC.

Kim SH & Kim S (2016). Social desirability bias in measuring public service motivation. *International Public Management Journal* 19(3):293–319.

Kruse Y, Callegaro M, Dennis J, DiSogra C, Subias S, Lawrence M & Tompson T (2009). Panel conditioning and attrition in the AP-Yahoo! news election panel study. In: *Proceedings of the 64th Conference of the American Association for Public Opinion Research*, Hollywood (Florida), 14–17 May 2009.

Lavrakas PJ (2008). *Encyclopedia of survey research methods*, Sage Publications, Thousand Oaks, California.

Nederhof AJ (1985). Methods of coping with social desirability bias: a review. *European Journal of Social Psychology* 15(3):263–280.

Sturgis P, Allum N & Brunton-Smith I (2009). Attitudes over time: the psychology of panel conditioning. In: Lynn P (ed.), *Methodology of longitudinal surveys*, Wiley, Chichester, 113–126.

Tourangeau R, Rips LJ & Rasinski K (2000). *The psychology of survey response*, Cambridge University Press.

Wooden M & Li N (2013). Panel conditioning and subjective well-being. *Social Indicators Research* 117(1):235–255.

Zimbalist Z (2017). 'Fear-of-the-state bias' in survey data. *International Journal of Public Opinion Research* 30(4):631–651.

CENTRE FOR SOCIAL
RESEARCH & METHODS

+61 2 6125 1279
csrm.comms@anu.edu.au

The Australian National University
Canberra ACT 2601 Australia

www.anu.edu.au

CRICOS PROVIDER NO. 00120C