Survey Research Methods (2022)
Vol. 16, No. 3, pp. 283-299
doi:10.18148/srm/2022.v16i3.7846
European Survey Research Association

Survey Response in RDD-Sampling SMS-Invitation Web-Push Study

Sebastian Kocar
University of Tasmania
Australia

In this study, I tested a fairly new survey data collection approach, using random digit dialing (RDD) selection of mobile phone numbers combined with SMS invitations, with respondents asked to complete the survey online. The SMS included a URL that directed recipients to an online questionnaire consisting of primary and secondary socio-demographic questions, as well as items on the use of Internet, health, technology, and life satisfaction.

The main aims of the study were to calculate and forecast response rate in a mobile web-push survey, and, through random assignment, to causally identify practices affecting response rates in survey research using this sampling type and the online survey mode. For this reason, a number of data collection characteristics were randomized and later used as predictors of survey (non)response.

The results showed that response rates in a survey using the proposed approach are low, i.e., below 2% in these survey experiments. They remain lower than most other survey modes, probably below 5%, even when using an optimized survey design with pre-notifications, reminders, and the sending of texts to mobile numbers with appended geo-demographic information. These were the effective maximization methods and techniques identified in this study. The benefits of this approach to sample recruitment are its simplicity and cost-effectiveness, and its potential for use in the future by students, academics, and social and market research companies, as the nonresponse bias and coverage bias did not seem to exceed representation bias in similar probability-based surveys. Traditionally, cross-sectional general population surveys use many other recruitment approaches, for example, mail outs, telephone calls, or face-to-face contacts. Text messaging proved to be more than just an additional communication channel, or a medium for sending survey reminders.

Keywords: nonresponse; probability sampling; random digit dialing; SMS recruitment; web-push; mobile web survey

1 Introduction

Survey data collection underpins a large proportion of social science research across multiple disciplines, but is increasingly difficult. Surveys are facing unprecedented challenges, response rates have declined steadily over the years, and methods to sample national populations are growing more expensive and complex (Couper, Antoun, & Mavletova, 2017). Generally speaking, there are three things more valued than anything else in survey research in practice: low cost, data quality, and time-efficiency, but you can only have two of them at the same time (Keeter, 2019). Many surveys that could normally be conducted time-efficiently and for relatively low costs, are generally considered of lesser quality compared to large-scale nationally representative surveys; an example of lower-quality surveys could be convenience samples or volunteer opt-in panel surveys, which are not based on probabilistic principles (Baker et al., 2010). In this study, I will test a survey data collection approach that provides an option for quick data collection at a relatively low cost. In contrast to volunteer panels, it is a probability-based online survey. However, due to its simplistic design, it might be conducted with some loss of quality, similarly to nonprobability surveys. The issue of data quality will be briefly addressed at the end of this article.

Primarily, this article focuses on the response dimension of the approach to a particular kind of data collection. Nonresponse is an issue in both probability and nonprobability surveys, and this project is also meant to provide some evidence on the connection between nonresponse and bias in a survey with low response rates. More specifically, I will study response rates in an RDD-sampling SMS-invitation web-push survey. This study is one of the first to combine these three approaches to sampling (RDD), recruitment (SMS), and data collection methodology (web-push/text-to-web[1]). To the best of my knowledge, it was previously tested

Contact information: Sebastian Kocar, Private Bag 51, Hobart TAS 7001, Australia (E-Mail: sebastian.kocar@utas.edu.au)

[1]Web-push surveys, also known as push-to-web surveys, are sur-

only in Germany (Bucher & Sand, 2021). Applications of this approach in practice would not be possible without the rise of mobile internet devices, such as smartphones, and/or an increase of internet coverage (Couper et al., 2017). For example, in 2019 in Australia, internet penetration rate was approximately 88% (Datareportal, 2020), 91% of adult Australians used mobiles to go online (Australian Communications and Media Authority, 2022), and the smartphone penetration rate was approximately 91% (Deloitte, 2019). Smartphone surveys might be a promising tool to collect data, but more work should be done to improve response and decrease nonresponse bias in this, to some degree, intrusive task (Elevelt, Lugtig, & Toepoel, 2019).

Thus, I will investigate how different data collection solutions and maximization approaches affect response, and how researchers can reduce nonresponse to mitigate potential representation bias. We have to have in mind that low response rates might not result in representation bias, as the link between response rates and nonresponse bias has been reported as weak at best (Groves & Peytcheva, 2008). However, the decline in response to RDD surveys increased the potential for bias in estimates (Brick, 2008). I can argue that there has to be a threshold below which surveys with extremely low response rates fail to sufficiently capture the socio-demographic, attitudinal, behavioral, or factual variability of the population; that is, in combination with coverage bias.

2 Literature review

2.1 Probabilistic sampling and undercoverage in web and smartphone surveys

Probability-based sampling requires each unit of the population to have a known non-zero chance of being selected to the sample (Neyman, 1938). Generally speaking, there is a missing link between probabilistic sampling and web surveys, as there are no general population sampling frames of email addresses. To address this issue, either offline recruitment (F2F, CATI, or postal) combined with a web-push approach, or non-probabilistic approaches such as river sampling are used in practice (Callegaro, Manfreda, & Vehovar, 2015).

In CATI surveys, random digit dialing (RDD) is often used as a set of techniques. The advantage of RDD is the probabilistic nature of selection into the sample. It is a method for generating telephone numbers randomly, either landline or mobile numbers. In survey methodology, it has been predominantly used in telephone surveys (Brick, 2008). With a decreasing percentage of landlines in developed countries like Australia, RDD methodology has had to adjust to these changes and sample more mobile numbers than landline numbers. This approach to sampling has previously been effectively used in recruitment to online surveys, such as probability-based online panels. In 2016, the Social Research Centre (Melbourne, Australia), managing the only national probability-based online panel, included 30% of landline numbers and 70% of mobile numbers in their sampling design. In 2018, the refreshment sample was recruited exclusively via mobile phones (Kaczmirek, Phillips, Pennay, Lavrakas, & Neiger, 2019), showing a trend towards mobile-only recruitment in the future. However, sampling of only mobile phone numbers for cross-sectional general population surveys has been quite rare and not well documented in the literature.

It is still unclear whether smartphone web surveys are a promising alternative to web surveys on desktop and laptop computers (PCs); they might suffer from coverage and nonresponse, possibly more than web surveys on PCs (Antoun, Conrad, Couper, & West, 2019) that are generally known for undercoverage of people without access to the internet (Couper, 2000). This might result in undercoverage bias for a number of non-demographic items (Hsia, Zhao, & Town, 2020). However, web surveys are not the only survey mode subject to coverage error. For example, in telephone surveys, some people are without landlines and mobile phones. The coverage bias can increase further if only mobile phone numbers are sampled. On the other hand, in countries with high internet penetration rates and high smartphone penetration rates like Australia (Australian Communications and Media Authority, 2022; Deloitte, 2019), this might be less of a problem than in other countries.

2.2 Factors affecting response in web surveys

In their systematic review, Fan and Yan (2010) conceptualized factors affecting response in web surveys into those affecting response rates in survey development, survey delivery (such as contact delivery modes, design of invitations, pre-notifications, reminders, and incentives), survey completion (from socio-demographics, psychographics to participation theories), and survey return. This research investigates survey response maximization strategies, and is thus predominantly focused on survey delivery factors.

In practice, more recent evidence shows that survey delivery factors such as survey structure (i.e., email invitation content, type and format of survey questions), assurance of privacy and confidentiality, interests of participants, and communication method, highly influence response in web surveys (Saleh & Bista, 2017). At the questionnaire design level, factors such as survey length, question difficulty, the content of the first question, and usage of a progress bar are related to completion rates (Liu & Wronski, 2018). More-

veys using mail contact to request survey response via the web mode as the first option (Dillman, 2017). In text-to-web surveys, SMS is used instead to recruit participants to online surveys (e.g., Bucher & Sand, 2021), and as such they are based on web-push principles.

over, Van Mol (2017) reported the effectiveness of extra reminders in an online survey among over-surveyed populations (regardless of the reminder content). On the other hand, Saleh and Bista (2017) reported that email reminders and incentives are effective in only particular socio-demographic groups.

Texting mobile numbers has previously been used as a survey invitation mode, a response maximization technique, and a pre-recruitment method to other survey modes. For example, American Trends Panel panellists are sent either email or SMS invitations if they have previously consented; all initially offline respondents who later received tablets, receive only text message invitations to their device (Keeter, 2019). Under certain conditions, texting is more effective than sending emails. De Bruijne and Wijnant (2014) confirmed that text messaging is more efficient for invitations when considering a response via a smartphone, and equally efficient as email invitations when considering total response in an online panel (but this can lead to a faster response). Moreover, Phillips and Compton (2019) reported that SMS reminders were associated with an increase in response rate in an online survey; in comparison to telephone reminders, they were less efficient, but more cost-effective. On the other hand, SMS reminders can have a positive impact on response in comparison to e-mail reminders (Sala, Respi, & Decataldo, 2018), and Bosnjak, Neubarth, Couper, Bandilla, and Kaczmirek (2008) reported that sending an advance SMS was more effective when compared to email pre-notifications in an opt-in online panel survey. These results were interpreted as SMS being both attention-grabbing and effective for establishing legitimacy.

Offering rewards is one of the most common response maximization strategies in survey delivery. There has been extensive research on the effectiveness of conditional and unconditional monetary/coupon incentives in web surveys, as well as prize draws, but the evidence has been mixed. In the last decade, some authors report that offering unconditional incentives (e.g., Parsons & Manierre, 2014), conditional incentives (e.g., Dykema, Stevenson, Day, Sellers, & Bonham, 2011), and a chance to enter a prize draw/lottery (e.g., Laguilles, Williams, & Saunders, 2011; Morgan, Rapee, & Bayer, 2017) increase response rates in web surveys. Other research did not find an effect of unconditional (e.g., Dykema et al., 2011) or conditional (e.g., Knowles & Stahlmann-Brown, 2021) incentives, or an increase in the conditional incentive amount (Neal, Neal, & Piteo, 2020; Spreen, House, & Gao, 2020). However, Mavletova and Couper (2016) reported that offering conditional differential incentives increased response rates of mobile-completers more than of PC web-completers.

2.3 SMS and text-to-web surveys

Dillman (2018) argues that the 2020s will be the age of smartphones, with most telephone communication being no longer voice conversation but rather texts and emails. This creates problems for RDD telephone surveys, and opportunities for text message surveys and text-to-web surveys. SMS surveys can be considered as a form of mass messaging and are known to have a low response, selection bias, and low data quality (Kongsgard, Syversen, & Krokstad, 2014). However, we have to acknowledge regulatory environments that could make SMS and text-to-web surveys, as well as texting to increase response, quite limited in particular contexts. In the US (Fordyce, Bilgen, & Stern, 2020) and some European countries (Kongsgard et al., 2014), prior consent to text messages is required, even for research purposes. On the other hand, it is not required in countries like Germany (Bucher & Sand, 2021) and Australia.

Texting as an interview mode can be defined as presending survey questions via SMS and also receiving answers from a respondent via SMS. Some of the advantages of this approach to data collection are a quick turnaround, collecting responses close in time to behaviors as the subject of survey research, and the ability for behavioral intervention, i.e., sending both information and reminders (Conrad, Schober, Antoun, Hupp, & Yan, 2017). As well as SMS surveys, data can be collected with text-to-web surveys. Andreadis (2020) demonstrated that it was feasible to conduct a large-scale web survey with SMS as the only contact mode if having mobile numbers of the target group, as well as reported the effectiveness of pre-notifications. Fordyce et al. (2020) compared synchronous text message surveys, i.e., questions and answers are exchanged in text messages, and asynchronous text message surveys, i.e., a text-to-web survey with a URL in the invitation; they reported no significant relationship between completion rates and the type of text message survey. Lastly, Balabanis, Mitchell, and Heinonen-Mavrovouniotis (2007) who examined the use of SMS to recruit respondents to web and telephone surveys, concluded that SMS can be used effectively with mixed-mode methods or as a pre-recruitment method to panels of respondents.

2.4 Aims of this research

The literature explains that telephone surveys based on RDD sampling are generally known for their relatively low response rates (Keeter, Hatley, Kennedy, & Lau, 2017), and the same conclusion can be made for text message surveys (Conrad et al., 2017; Kongsgard et al., 2014). In terms of errors of representation (see the Total Survey Error framework in Groves et al., 2009), nonresponse errors are not the only errors prevalent in internet and smartphone surveys—there is also the issue of undercoverage of people with no internet access in their household or on their mobile devices (Antoun et al., 2019; Couper, 2000). This could lead to a notable representation bias as a result of combining both types of error of representation. It is often challenging to distinguish between nonresponse and coverage errors in mobile surveys, and discussing selection bias is more appropriate (Couper et al., 2017). Thus, in this article I will test this fairly new approach with a focus on nonresponse and, as a result of both nonresponse and undercoverage, socio-demographic representation bias. With an empirical analysis, I will answer the following research questions.

RQ1 What response rates can be expected in an RDD-sampling SMS-invitation web-push survey?

RQ2 What data collection characteristics in survey delivery, such as incentives, text message content, or time of sending SMS invitations, affect response rates in a survey of this type?

RQ3 What level of socio-demographic representation bias is present in a survey of this type?

The evidence presented in this study can be extended to other survey research approaches in similar contexts (e.g., in Australia, with a similar topic) using text messaging. This includes but is not limited to: SMS surveys, SMS recruitment, and SMS pre-notifications and reminders.

3 Methods

3.1 Data

The data from this methodological project were collected and compiled by the Centre for Social Research and Methods at the Australian National University, with the main aim to explore ways to replace existing, expensive survey methods with cheaper, more flexible ones. The RDD-sampling SMS-invitation web-push survey was purposely designed to enable the study of not only response, but also nonresponse/representation bias, and accuracy relative to a number of demographic and non-demographic benchmarks from large-scale government-funded surveys in Australia.

To collect survey data in this project, the online "Survey on Wellbeing, Health and Life in general 2020" questionnaire was programmed (see the Online supplementary materials). Besides items measuring health and wellbeing, the questionnaire included items on the use of internet and technology, satisfaction with different dimensions of life, personality traits, primary demographic items like gender, age, and education, and secondary demographics such as country of birth, citizenship, employment, and income. The questionnaire consisted of 35 questions, out of which there was one multiple answer question and no grid or open-ended questions. The median response time was about 8 minutes.

To study factors affecting survey response, the analyzed data file consisted of all randomly generated mobile numbers receiving an SMS invitation ($n = 38,512$) as cases. For each telephone number, the following data collection characteristics were coded prior to data collection: time of day and day of week the SMS was sent, text message content, reminder, incentives offered, type of invitation, appended geodemographics, and stratification information. After the data collection was completed, the survey data file ($n = 631$) was used to identify all mobile numbers belonging to survey participants and derive the response variable survey response (1=unit response, 0=unit nonresponse). Section "Data analysis" describes the dependent variables, independent variables, and statistical modeling in more detail.

3.2 Sampling and sampling frame

Since there is no real connection between probabilistic sampling and web survey collection data from the general population (Callegaro et al., 2015), I combined a sampling approach generally used in telephone surveys and telephone recruitment with online data collection. Thus, I carried out random digit dialing (RDD) generation of Australian mobile numbers. Each Australian mobile number consists of the leading numbers 04 and 8 more digits that can be randomly generated—format 04XX XXX XXX. Hence, there are 100 million possible combinations, and not all of them were used at the time of data collection. There are spare numbers, e.g., 0440 000 000–0444 300 000, numbers allocated to satellite phones, e.g., 0420 100 000–0420 109 999 (Pivotel), and rail corporations, e.g., 0420 000 000—0420 019 999 (Rail Corporation New South Wales) (Australian Communications and Media Authority, n.d.). The remaining numbers, 62 million or 62% of all possible combinations, are allocated to Vodafone, Optus, and Telstra mobile service providers and can be used in general population surveys. However, knowing that there are about 24 million people living in Australia (Australian Bureau of Statistics, 2016) and about 20 million mobile numbers, only about one out of three combinations represent a valid/live mobile number.[2]

To remove invalid mobile numbers and to decrease the cost of texting invitations, i.e., not sending SMS messages to mobile numbers that are not live, I used the service provided by SamplePages (n.d.). They matched my randomly generated mobile numbers[3] to the mobile numbers from the

[2]Estimation based on: 18.5 million Australians aged 18+ (Australian Bureau of Statistics, 2016), 6% of them with two mobile phones, 32% of Australian aged 6–13 and 91% of Australian aged 14–17 own a mobile phone (Roy Morgan, 2015, 2016, 2018), only about one out of three combinations represent a valid/live mobile number.

[3]500,000 mobile numbers were generated using RDD; the number was determined based on the estimated match rate and the required sample size of mobile numbers with a positive match.

general population in their database for validation and to append geo-demographics (with approximately 7% matching rate, 34,734 numbers[4]). They estimated that 90% of all those numbers were live at the time of matching and appending. For "the top-up sample"[5] (Stage 2, please see "Survey experiment" subsection), they also validated 3,778 additional numbers[6] with no matches in their database. 100% of those numbers were active on the day of validation, i.e., 1–3 days before the text messages were sent to the validated numbers. In the end, the total sample consisted of:

- 12,302 numbers with full geo-demographics available (gender, age group, Statistical Area Level 4 (SA4));
- 22,432 numbers with partial geo-demographics available (about 96% of those without age [SA4 only, gender only, or SA4 and gender] and about 4% with age information [age only, age and gender, or age and SA4]);
- 3,778 numbers with no geo-demographics available (added in Stage 2).

For Stage 1, the sample was stratified[7], and for Stage 2, there was just one sample with no stratification carried out. Stratification was primarily used to study its effect on data accuracy, i.e., for the purpose of a separate study.

3.3 Survey experiment

In this study, I collected data with a sophisticated survey experiment, including two associated stages in a responsive survey design.

Stage 1. To study factors affecting and improving response in a survey applying an RDD-sampling SMS-invitation web-push approach to data collection, I divided the sample of 27,000 numbers into 48 experimental groups. There were a number of experimental variables I intended to test the survey (non)response against:

1. survey reminder;
2. day of the week initial SMS was sent;
3. time of day initial SMS was sent;
4. incentives; and
5. SMS invitation text (information on the topic of the survey, information on benefits of participation).

See Table 1 for more information.

While all other conditions were split 50:50, I had to adjust the sizes of the experimental groups based on the types of incentives, which depended on the available budget. With a total budget of $3000, I intended to spend $200 on incentives for the prize draw (target sample size $n = 400$), $1000 on incentives for the $5 conditional incentives group (target sample size $n = 200$), approximately $700 on mobile number validation and appending geo-demographics, and approximately $1100 on text messaging. Assuming a lower response in groups not being offered a (monetary) reward, the no-incentives group of sampled potential respondents was larger (60%) than the lottery (30%), and the $5 incentives (10%) groups. I purposely collected data on all days of the week in an attempt to minimize the effect on the response of particular events on certain days of the week. However, this was done with an intention to aggregate the days into "weekdays" and "weekends", to keep sufficient statistical power for all experimental groups.

Stage 2. The response rate in Stage 1 was about 50% lower than initially expected, and to increase the sample size for a separate benchmarking component of the project, the decision was made to use a top-up sample. For this reason, I decided to use all remaining numbers with appended geo-demographics ($n = 7,734$, with predominantly partial information for stratification), and to validate new numbers with no geo-demographic information ($n = 3,778$, pinged but not matched).

Of 500,000 mobile numbers, SamplePages matched and appended 34,734 numbers, 27,000 of which were sent text invitations in Stage 1 and 7,734 in Stage 2. Of 465,266 mobile numbers with no matches in the SamplePages databases, 12,000 were randomly selected for validation, 3,778 of which were successfully "pinged" for Stage 2 text invitation.

[4]While generation of mobile phone numbers was carried out by the researcher and was random (following RDD principles), validation of mobile numbers provided by SamplePages in Stage 1 excluded Australian mobile numbers not in their database. As such, the selection closely resembled drawing random numbers from the SamplePages database, but with more control over sampling on the researcher's end (and for lower cost). In practice, there is a trade-off between (1) coverage of mobile numbers not in SamplePages database, (2) the ability to carry out stratified sampling of mobile numbers in the Australian context. This potential undercoverage bias is, combined with nonresponse bias, addressed in the Representation bias subsection of the Results.

[5]The term "top-up sample" comes from longitudinal survey methology, and defines a sample that is recruited to refresh an existing longitudinal panel—to either increase sample size, target specific (sub)groups or improve coverage (Watson, 2014).

[6]Mobile number validation also known as "pinging"; 12,000 RDD mobile numbers, 3,778 active, 31.5% live number validation rate, fairly consistent with "20 million mobile owners / 62 million possible combinations" ratio. This type of selection using random generation of mobile numbers with subsequent "pinging" can be considered as a true RDD in the Australian context. The "pinged" sample of mobile numbers was added in Stage 2 to study response rate conditional on the type of validation of mobile numbers (see Table 4 for results).

[7]Based on availability of stratification information: 8,000 numbers by gender*age group*SA4 and 19,000 by state only.

Table 1
Experimental design (Stage 1, $n = 27,000$)

Time of day	Incentive type	SMS invitation text	%[a]
Survey reminder 5 days after first SMS			
Monday to Friday			
Afternoon	No Incentives	Benefits	3.75
Afternoon	No Incentives	Survey topic	3.75
Afternoon	$5 conditional	Benefits	0.63
Afternoon	$5 conditional	Survey topic	0.63
Afternoon	Prize draw	Benefits	1.88
Afternoon	Prize draw	Survey topic	1.88
Evening	No Incentives	Benefits	3.75
Evening	No Incentives	Survey topic	3.75
Evening	$5 conditional	Benefits	0.63
Evening	$5 conditional	Survey topic	0.63
Evening	Prize draw	Benefits	1.88
Evening	Prize draw	Survey topic	1.88
Saturday and Sunday			
Afternoon	No incentives	Benefits	3.75
Afternoon	No Incentives	Survey topic	3.75
Afternoon	$5 conditional	Benefits	0.63
Afternoon	$5 conditional	Survey topic	0.63
Afternoon	Prize draw	Benefits	1.88
Afternoon	Prize draw	Survey topic	1.88
Evening	No incentives	Benefits	3.75
Evening	No incentives	Survey topic	3.75
Evening	$5 conditional	Benefits	0.63
Evening	$5 conditional	Survey topic	0.63
Evening	Prize draw	Benefits	1.88
Evening	Prize draw	Survey topic	1.88
"No survey reminder 5 days after first SMS			
Monday to Friday			
Afternoon	No incentives	Benefits	3.75
Afternoon	No incentives	Survey topic	3.75
Afternoon	$5 conditional	Benefits	0.63
Afternoon	$5 conditional	Survey topic	0.63
Afternoon	Prize draw	Benefits	1.88
Afternoon	Prize draw	Survey topic	1.88
Evening	No incentives	Benefits	3.75
Evening	No incentives	Survey topic	3.75
Evening	$5 conditional	Benefits	0.63
Evening	$5 conditional	Survey topic	0.63
Evening	Prize draw	Benefits	1.88
Evening	Prize draw	Survey topic	1.88
Saturday and Sunday			
Afternoon	No incentives	Benefits	3.75
Afternoon	No incentives	Survey topic	3.75
Afternoon	$5 conditional	Benefits	0.63
Afternoon	$5 conditional	Survey topic	0.63
Afternoon	Prize draw	Benefits	1.88
Afternoon	Prize draw	Survey topic	1.88
Evening	No incentives	Benefits	3.75
Evening	No incentives	Survey topic	3.75
Evening	$5 conditional	Benefits	0.63
Evening	$5 conditional	Survey topic	0.63
Evening	Prize draw	Benefits	1.88
Evening	Prize draw	Survey topic	1.88

[a] Percentage of the combined sample in Stage 1

Table 2
Experimental design (Stage 2, $n = 11,512$)

Time of day	Incentive type	SMS invitation text	Invitation type	%[a]
No reminder, Tuesday to Thursday				
Evening	No Incentives	Benefits	Advance SMS	33
Evening	No Incentives	Benefits	Responsive SMS	33
Evening	No Incentives	Benefits	Standard SMS	33

[a] Percentage of the full sample

In Stage 2, I did not replicate the experimental design from Stage 1. Instead, I analyzed the existing preliminary data from Stage 1 and selected the optimal combination of approaches in terms of response and costs (a responsive approach). If there was no notable difference in rates between experimental groups, I decided to go with the cheapest data collection approach. In the end, I standardized the following conditions: weekday evening invitations (Tuesday-Thursday), text message communicating benefits, no incentives offered, and no survey reminders.

I introduced two alternative approaches to SMS recruitment as follows.

1. Advance SMS/pre-notification. I attempted to increase response by pre-notifying respondents about the upcoming SMS invitation; the literature explains that advance SMS can help establish trust and legitimacy (Bosnjak et al., 2008). At the same time, I only texted a link to the survey to respondents who did not opt-out by responding "STOP".

2. "Responsive" text message survey invitation. I attempted to increase response by not sending the link to the survey in the first text message, but only if the respondents agreed to receive an invitation by responding "YES". This can be considered as a less intrusive approach to survey invitations, similarly to an advance SMS.

The third type of invitation was a "standard" single-invitation SMS including a URL to an online questionnaire. This was the only type of invitation from Stage 1. Please see Table 2 for more information.

3.4 Data collection

Data collection took place between Wednesday November 4th and Saturday November 21st 2020. The initial stage, Stage 1, took place between November 4th and November 17th, and Stage 2 took place between November 17th and November 21st 2020. Due to the experimental design, invitations were sent during most of these periods; the online survey/questionnaires were deactivated two days after the last invitation in a stage was sent. Reminders sent to one half of the Stage 1 sample, excluding those who opted-out of receiving future SMS after the first SMS invitation, were sent between November 9th and November 15th using a unified approach that avoids adding extra variability: SMS was sent 5 days after the first one, in the evening, offering the same type of incentives as in the first SMS, and communicating benefits. For more information about the timeline, please see Figure 1.

Text messages were designed based on the results of a qualitative study by the Social Research Centre, who ran focus groups to test the design of their advance SMS that was later used in recruitment to their online panel (Kellard, 2017). Moreover, the communicated benefits of participation without receiving a reward (i.e., asking to help our research) was based on survey participation theories, such as social exchange or cognitive dissonance theories (for more information, see Keusch, 2015). The text message content is shown in Figure 2.

Stage 2 was conducted in much less time and was built on an optimized survey design heavily based on the results of Stage 1. It was finished in 5 days and included elements of rapid data collection, which could be considered as a key advantage of the proposed RDD-sampling SMS-invitation web-push approach.

3.5 Data analysis and statistical modeling

With the division into experimental survey groups (see Tables 1 and 2), I aimed to create a binary logistic regression model with survey response (0=unit nonresponse, 1=unit response) as the dependent variable and characteristics of data collection as the predictor variables. After combining and coding data from Stages 1 and 2, those predictors were:

- time of day (afternoon, evening);
- day of the week (weekday, weekend);
- SMS invitation text (topic, benefit);
- reminder (reminder sent, reminder not sent);
- type of invitation for maximizing response
  - "standard"single-SMS invitation with no incentives offered (reference group),

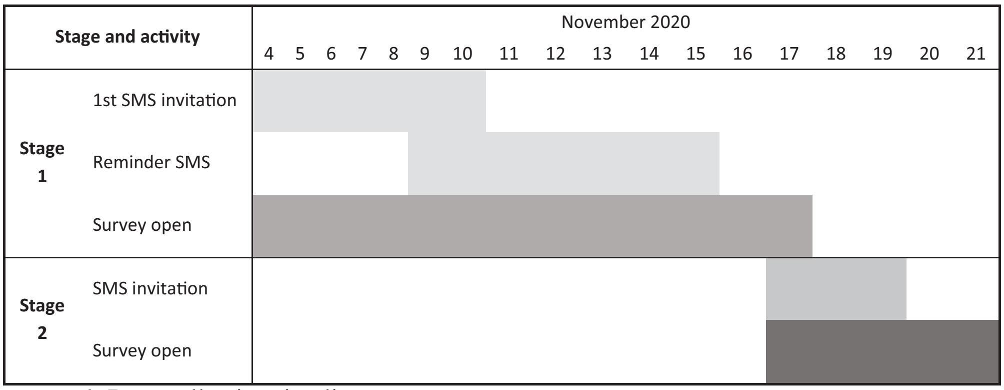

Figure 1. Data collection timeline

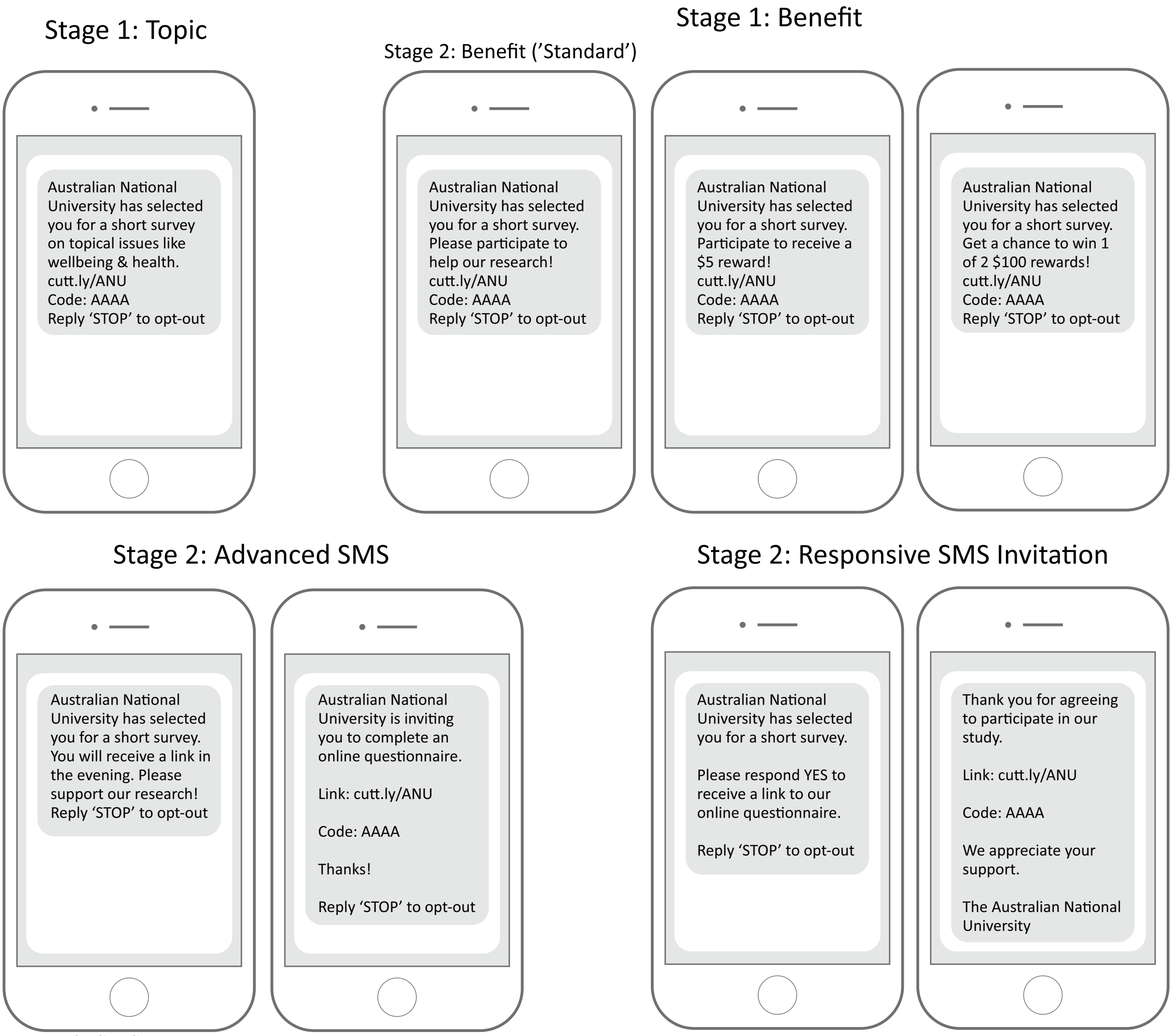

Figure 2. SMS content

– "standard" single-SMS invitation with $5 incentives offered,
– "standard" single-SMS invitation with an offer to enter a $100 prize draw,
– advance SMS with no incentives offered,
– responsive SMS invitation with no incentives offered.

After reviewing the results of the preliminary analysis and receiving feedback from SamplePages, I decided to use different information of appended geo-demographics to a mobile number as a predictor as well. The following groups of mobile numbers were coded:

- geo-demographics including age;
- partial geo-demographics not including age; and
- no geo-demographics appended/available.

The reason for this was the different probabilities of a mobile number being live, and due to eligibility of respondents dependent on their age information (people aged younger than 18 were ineligible).

Taking these probabilities into account, response rate calculation adjusting for unknown eligibility (*e*-value) would be considerably affected. To calculate the rates, I used AAPOR Response Rate Standard definitions, and their proposed calculations of response rates RR2 and RR4 (American Association for Public Opinion Research, 2016, pp. 61–62). In both cases, the RR were calculated by counting partial interviews[8] as survey respondents. To calculate the *e*-value for RR4, I used estimates from SamplePages in combination with the estimates from Roy Morgan Young Australians Survey 2018 (Roy Morgan, 2018) and Roy Morgan Single Source Australia 2016 on phone ownership of minors (Roy Morgan, 2016). More details about the estimates and calculation of response rates are available in the Results section.

Data analysis was carried out in the statistical software Stata 13 (StataCorp, 2013). As well as descriptive analysis, as previously discussed, I conducted binary logistic regression modeling.

4 Results

In this section, I will present the results to answer research questions RQ1-RQ3. Firstly, I will discuss response rates in a survey combining online data collection with RDD sampling and SMS invitation, comparing responses between different fundamental approaches, such as offering incentives and an advance SMS. Secondly, I will dig deeper into what affects response and analyze the data for all experimental groups using binary logistic regression modeling. Finally, I will present nonresponse bias by comparing the distribution of socio-demographic variables between this study, the Australian Census 2016 benchmarks (Australian Bureau of Statistics, 2016), and three probability-based samples from the Online Panels Benchmarking Study (Pennay, Neiger, Lavrakas, & Borg, 2018).

4.1 Response rates

To answer the first research question RQ1, I will present the results on response rates for a few different response maximization approaches from data collection Stages 1 and 2. In this analysis, I will not control for other characteristics of data collection (e.g., with logit regression). For this reason, I calculated AAPOR RR2 and RR4. The difference between these two calculations is the estimate of a portion of the sample with unknown eligibility that are ineligible (represented by the *e*-value). In an RDD-sampling SMS-invitation web-push survey, respondents with unknown eligibility are those who did not respond and who did not break off. Respondents who confirmed that they were at least 18 years of age by starting the survey were considered eligible. Unfortunately, those who clicked on the link but did not start the survey, did not provide enough information to assume their eligibility status.

With three "types" of mobile numbers based on the availability of appended information, three different *e*-values had to be estimated to calculate AAPOR RR4. Based on the information I received from SamplePages, about 10% of their database contains mobile numbers that are not live, hence the coefficient of 0.9 for the "Complete geo-dem" group with only respondents aged 18+. Without having appended information about mobile owners' age, I estimated the portion of Australians with mobile phones who were not yet 18. I consulted two reports from Roy Morgan (2016: 91% of teenagers aged 14–17 have a mobile phone, 2018: 32% of children aged 6–13 have a mobile phone) and the Australian Census 2016 distribution by age, and estimated that about 9% of all Australian mobile owners were not yet 18 (Australian Bureau of Statistics, 2016). Combined with the portion of inactive numbers, the final *e*-value for "Partial geo-dem group" was 0.82. With pinged (all live) numbers, I only had to adjust the *e*-value for ineligible respondents not yet 18 years of age.

The results from Table 3 show how different actions to increase response in a mobile survey are more or less efficient. Interestingly, there was a very little and statistically not significant difference between respondents who were not offered incentives (1.74% RR2, 2.02% RR4), those who were offered $5 incentives (1.69% RR2, 1.97% RR4), and those who were offered to enter a lottery for a $100 eGift card (1.73% RR2, 2.02% RR4) in Stage 1.

[8] Partial interviews were those respondents who provided enough information for post-stratification weighting, i.e., reached at least question 28 out of 35 (see the questionnaire in the Online supplementary materials), but did not complete the questionnaire.

Table 3
Response rates by different response maximization approaches

	Incentive type			Invitation type		
AAPOR Survey Outcomes	No	$5	Lottery	Standard	Advance	Responsive
All invited	16,184	2,724	8,092	3,838	3,837	3,837
Participated						
Complete	267	44	126	36	80	43
Partial	14	2	14	3	2	0
Known eligibility						
Breakoff	33	1	10	10	12	4
Unknown eligibility (for mobile numbers with different appended information)						
Geo-dem including age	6,630	1,117	3,417	798	510	534
Partial geo-dem not including age	9,240	1,560	4,525	1,738	1,997	2,014
No geo-dem (pinged)	0	0	0	1,253	1,236	1,242
Total	15,870	2,677	7,942	3,789	3,743	3,790
e-value						
Geo-dem including age	0.9	0.9	0.9	0.9	0.9	0.9
Partial geo-dem not including age	0.82	0.82	0.82	0.82	0.82	0.82
No geo-dem (pinged)	0.91	0.91	0.91	0.91	0.91	0.91
Combined	0.85	0.85	0.85	0.87	0.86	0.86
Response rate (in %)						
RR2	1.74	1.69	1.73	1.02	2.14	1.12
RR4	2.02	2.97	2.02	1.18	2.47	1.30

It seems that it is not worth investing money into offering potential respondents a compensation for their participation in an SMS invitation survey, but rather into other approaches like sending an advance SMS (2.14% RR2, 2.47% RR4). When using the approach with an introductory SMS but no reminders, the response rate was more than twice as high than without an introductory SMS (1.02% RR2, 1.18% RR4 in Stage 2), and still notably higher than for the Stage 1 groups with about 45% of respondents receiving a reminder SMS.

Standard invitation group and responsive SMS group (RR2 1.12%, RR4 1.30%) from Stage 2 had a lower response but they also did not receive a reminder SMS in contrast to the groups from Stage 1 and, therefore, response rates cannot be compared directly. It is possible that response rate in a responsive SMS survey with a reminder would be comparable to response rates from the first three approaches. This will be estimated with further analysis.

4.2 Factors affecting response

In the first subsection, I showed how response rates differ between different SMS texting approaches and how unknown eligibility, which I was able to estimate using external data, affects the final response rates as a conditional indicator of data quality. In this subsection, I will show the results of binary logistic regression to provide answers to RQ2. Regression modeling was conducted to showcase how different approaches I tested with the experimental randomized design affect or do not affect (non)response. In this section, only RR2 numbers are presented and discussed. In the unit record file, I cannot assume which numbers were eligible and ineligible, hence no RR4 can be calculated.

The results from Table 4 show that there are a number of ways to maximize response, but there are also unnecessary (and quite expensive) measures that do not successfully convince potential survey respondents to participate. Thus, many of my findings on response maximization methods and techniques were not in line with theoretical expectations. First of all, offering incentives in a single-invitation SMS had no positive or negative effect on response rates in comparison to a single-invitation SMS with no incentives offered. This is in line with findings on response rates (RQ1). Inefficiency of incentives comes as a surprise, as it was anticipated in the survey design phase to give double the response rate for the lottery experimental group, and three times the response rate in the $5 conditional incentives experimental group. This inefficiency resulted in a much lower total response rate in Stage 1, encouraging me to make a decision to use a top-up sample and Stage 2 of data collection. In this stage, I used an opportunity to test two different approaches as an alternative to offering incentives: an advance SMS invitation; and a responsive SMS invitation. While responsive SMS invitations did not improve response in comparison to a standard single-SMS invitation, an advance SMS invitation statistically sig-

Table 4
The caption of the table

Predictor	Coef.	Std. Err.
Type of invitation		
No incentives offered, standard invitation	0	
\$5 conditional incentives, standard invitation	−0.01	0.16
Prize draw for \$100 coupons, standard invitation	0.01	0.1
No incentives, advance SMS invitation	0.67**	0.15
No incentives, responsive SMS invitation	0.02	0.18
Stratification information		
Pinged numbers, no info	0	-
Complete stratification info or incomplete info with age	0.68**	0.18
Incomplete stratification info (with no age info)	0.38*	0.18
Survey SMS reminder		
No SMS reminder sent	0	-
Yes, 5 days after first SMS	0.59**	0.09
Day of the week		
Weekday	0	-
Weekend	−0.05	0.09
Time of the day		
Evening	0	-
Early afternoon	−0.13	0.09
SMS invitation text		
Benefit	0	
Topic	−0.01	0.09
Constant	−4.82	0.19
Pseudo R^2	0.013	

* $p < 0.05$ ** $p < 0.01$

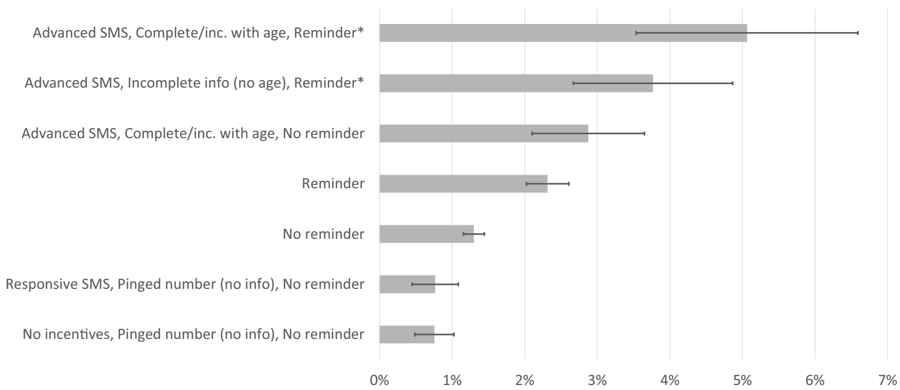

Figure 3. Predictive margins for the best and worst combinations of approaches based on RR2 response rates (binary logistic regression model, see Table 4 for coefficients). Starred groups are not randomly assigned; the predictive margins for these combinations were calculated based on response rates in other similarly structured experimental groups.

nificantly boosted response. The improvement in response was quite similar to an improvement if potential respondents were sent a reminder 5 days after receiving the first SMS invitation.

Moreover, I noticed statistically significant differences in response rates between mobile numbers with different levels of appended demographics. The findings are partially in line with findings in the first subsection, where I calculated *e*-values (estimates of unknown eligibility). The numbers with the highest response rate were those with complete stratification info or incomplete info with age, which comes as no surprise because they do not include mobile owners that are younger than 18. They are followed by the numbers with partial stratification information with no age information and "pinged" numbers with no stratification information appended, although the difference between these two groups is statistically significant at the $p<0.05$ level but not at the $p<0.01$ level. One possible explanation could be that people with mobile numbers with some stratification information available are less concerned with privacy, since they agreed to have their mobile number listed.

On the other hand, I did not notice any statistically significant differences between the time of day I sent SMS invitations, the days I sent SMS invitations, and the content of the text message (communicating topic or benefits, such as offering incentives). In Stage 1 I noticed a slightly higher response rate if texting invitations in the evening, but the difference after combining data is not statistically significant. With days of the week, I noticed that there might be differences between certain days (e.g., Saturday at first seemed to be a bad day for sending the first SMS, but a suitable day for sending a reminder), but the aggregation into weekdays and weekends eliminated these differences. Some of these data collection characteristics should be explored further in future research using larger samples.

In Figure 3, I am extending the analysis, and presenting the predictive margin results for the best and worst combinations of response maximization approaches for SMS invitation web-push data collection. The results show it is as important to have a high-quality list of validated mobile numbers, as it is to select the right approach to SMS invitation. While we did not observe large differences between nonresponse in the results from Table 3, the presented predictive margins showcase how different (hypothetical) strategies can result in quite different response outcomes.

The maximization approach with the highest predicted RR2 is one that I did not test in our study, as I did not send reminders in Stage 2. This combination is an advance SMS using mobile numbers with appended age, and a reminder (RR2 95% confidence interval (CI) $[3.5\%, 6.6\%]$). It is followed by the same approach except using mobile numbers with no age information (RR2 95% CI $[2.7\%, 4.9\%]$), but the difference is not statistically significant. Of all approaches I combined and tested in practice, the best seemed to be "advance SMS using mobile numbers with appended age and no reminder" (RR2 95% CI $[2.1\%, 3.6\%]$).

In terms of using reminders—the difference between sending a reminder SMS five days after the initial text message invitation, and not sending a reminder, is statistically significant. In practice, the difference is about 1 percentage-point, ceteris paribus. This is consistent with the results presented in Table 3.

Out of all the different maximization approaches I tested, the least effective were shown to be using pinged numbers with no reminders and either a responsive or a single-SMS invitation with no incentives offered. In these cases, we could not realistically expect RR2 of more than 1.1% (95% CI upper bound).

4.3 Representation bias

In the following paragraphs, I will address the issue of representation bias, which may or may not be related to low response rates; the issue is identified in the existing literature on the topic (e.g., Conrad et al., 2017). As I discussed previously, extremely low response rates can represent a bigger problem than just low response rates, which are an issue for most survey research nowadays. Also, 90% of the sampled mobile numbers in this study were randomly selected from SamplePages database containing estimated 22% of all mobile numbers from the general population of Australians. Thus, there was a potential for undercoverage bias.

To estimate representation bias, I will compare the distribution of socio-demographics variables, most of which are commonly used in post-stratification weighting like raking, between different probability-based samples. Estimates from these samples will be compared to the benchmarks from the Australian Census 2016 (for the general population, 18+ years of age Australian Bureau of Statistics, 2016) as the highest quality data source for Australia. Representation bias from the RDD-sampling SMS-invitation web-push survey will then be compared to representation bias from the following: a standalone RDD telephone sample; an RDD end of a telephone survey ("piggybacking") sample; and an address-based sample[9]. In the end, I will answer the research question RQ3.

The results in Table 5 show socio-demographic differences between the general adult population (from Australian Population Census 2016) and different sample surveys. These can be interpreted as a combination of nonresponse and coverage bias. In practice, the differences are corrected with post-stratification weighting, which often does but sometimes does not improve the accuracy of non-demographic estimates (Groves et al., 2009).

[9]The samples are from the Online Panels Benchmarking Study (OPBS) 2015; estimates are taken from Pennay et al. (2018, pp. 29–40).

Table 5
Differences in distributions of key primary socio-demographic variables (unweighted)

	Benchmark AU Census 2016	RDD SMS Web-push	Standalone RDD	A-BS sampling recruitment	RDD End of survey recruitment
Sex					
Male	48.8	44	46	39	42
Female	51.2	56	54	61	58
Age (years)					
18–24	11.8	6	7	4	6
25–34	18.5	11	9	10	9
35–44	17.3	13	15	13	15
45–54	17.1	18	15	16	19
55–64	15.1	24	20	22	21
65-74	11.4	23	18	22	21
75+	8.8	6	14	13	9
Education					
Secondary Education or Cert I/II	43.4	25	38	40	35
Cert III/IV, (Advanced) Diploma	29.8	34	29	24	29
Bachelor's degree or higher	26.7	41	33	37	35
Birthplace					
Australia	66.3	76	75	73	75
Other	33.7	24	25	27	25
Region					
New South Wales	32	32	29	34	32
Victoria	25.5	25	24	26	25
Queensland	19.9	18	22	17	20
South Australia	7.3	8	9	8	8
Western Australia	10.5	9	10	9	9
Tasmania	2.2	3	3	2	3
Northern Territory	0.9	< 1	1	1	1
Australian Capital Territory	1.7	4	4	3	1
Response rate (RR2 in %)	-	1.6	14.7	26.2	9.8

We can see that my sample is closer to the benchmarks than other samples for some variables and their categories, and more biased for some other items. Generally speaking, all probability-based samples, i.e., my RDD-sampling SMS-invitation web-push sample and OPBS 2015 samples, are different to the population distributions of the target primary demographics. Firstly, surveys tend to attract more females than males, which was confirmed by my data. Secondly, surveys attract older respondents; in my survey, there was a larger than usual portion of those aged 55–74, but people aged 75+ were less overrepresented than in the other surveys, probably due to a lack of digital literacy. On the other hand, a slightly higher portion of those younger than 35 (18–24 and 25–34 age groups combined) was included in our study, and they were less underrepresented. Thirdly, relatively more educated than less educated respondents participate in surveys, and in my survey, the education-related nonresponse bias was even more severe. The differences could potentially be attributed to people with a university degree being more likely to own a mobile phone (coverage bias); however, due to a high smartphone penetration rate (only about 9% of Australians are without a smartphone), education-related coverage bias can only explain a portion of the total representation bias. Moreover, all surveys in Australia seem to underestimate the portion of foreign-born residents of Australia with a similar magnitude. Finally, my sample was quite accurate in estimating distribution by state, and the difference between the Australian Census 2016 and my estimates was larger than 2 percentage-points for the Australian Capital Territory only.

To sum up, low response rates in the RDD SMS-invitation web-push study and undercoverage of people without internet or smartphones seemed to lead to some demographic representation bias, but this can be reported for other probability surveys as well. The only notable difference between my sample and all other samples was in estimating education distribution, which could potentially be affected by an increasing differential nonresponse five years after the OPBS was carried out.

5 Discussion

The 2020s brings both good news and bad news for survey methodology. The bad news is that response is declining in most survey research, and low response rates in my survey were an example of that. Nevertheless, the good news is that the existing literature (e.g., Groves & Peytcheva, 2008) explains how nonresponse is not necessarily associated with representation bias and does not affect data accuracy in survey research based on probabilistic principles. In this study, being prepared to deal with high nonresponse, my focus was on exploring both the bad news and good news perspectives of errors of representation in a relatively new approach to probabilistic sampling with an SMS invitation directing to an online survey. While nonresponse was even higher than initially predicted in the survey design phase (potential nonresponse bias) and selection was somewhat limited to SamplePages database (potential undercoverage bias), in the end they translated into representation bias comparable to the socio-demographic nonresponse bias in three probability-based surveys with significantly higher response rates. In comparison to a similar text-to-web study with a low response rate conducted by Bucher and Sand (2021), my study experienced similar representation bias for education (overrepresentation of the most educated), but not for age (overrepresentation of the oldest instead of the youngest). Moreover, studying non-demographic bias, namely attitudinal, behavioral, knowledge, and other factual bias, will be an important step for further determining the overall representation bias of the proposed approach.

I firstly showed that the expected response rates in a web survey with text message invitations can be very low. Based on a number of text message responses from recipients of my SMS invitations, it was clear that potential respondents were skeptical that the SMS really came from the Australian National University, and that my survey was a legitimate academic research project. While the first page of the online questionnaire thoroughly explained the practical and ethical aspects of the research, the link-click numbers showed that only about one in ten recipients clicked on the link in the survey invitation. The solution to this problem could be sending a longer SMS and providing more information in these messages, but that could considerably increase data collection costs, especially if there was little increase in link-click rates.

Further, I presented evidence on how response in a study of this kind is difficult to increase substantially by using many approaches known to be effective in other survey modes, e.g., offering conditional incentives. One possible explanation of this could be that in a country with a high standard of living like Australia, $5 incentives in a form of a supermarket gift card do not represent enough value for respondents. At the same time, administering incentives to all respondents in my survey would double the costs per completed survey. It appears that the majority of respondents in this study participated for other reasons, which can be better explained with social exchange or cognitive dissonance theories (see Keusch, 2015, for more information). This is further supported by the fact that only 8 of 44 "$5 conditional incentives" respondents decided to accept the coupon after filling out the survey. The other 36 respondents either did not wish to provide their details to be sent a $5 eGift card, or decided to enter the lottery for a $100 eGift card instead, which was an option given after they completed the questionnaire (and not in the introduction or the invitation SMS). It would be quite interesting to see the potential effect of providing $10 or $15 incentives on response, which are standard gift card amounts used by the only Australian probability-based panel Life in Australia for surveys of similar length (Kaczmirek et al., 2019), although this could significantly increase costs and attract more "professional respondents". On the other hand, a text-to-web survey offering higher incentives might attract more of the less educated (with lower income), which would help reduce some of representation bias.

Due to inefficiency in offering incentives, I later explored a couple of different potential response maximization solutions reported in the literature (e.g., Bosnjak et al., 2008). My findings on data collection characteristics affecting (and not affecting) response were consistent with findings from Andreadis (2020) on the use of pre-notification and reminders, and the day and time of SMS invitations. Sending an advance SMS proved to be the best solution to increase response by supposedly building trust with some respondents, even if being sent from a mobile number not associated with the University and only two hours in advance. The other fairly effective approach was a reminder SMS, which showed a similar increase in response rates. Since sending two or more text messages to each mobile number would exceed my budget, I did not test sending more than one SMS reminder. For experimental groups receiving a reminder, the number of completes in two days after the first SMS (initial invitations) were fairly comparable to the number of completes in two days after the second SMS (reminders). This indicates that sending the second reminder SMS a couple of days later could result in a similar increase of response as the first reminder SMS, which was previously reported by Andreadis (2020). Moreover, I did not have a chance to combine two of the best approaches, texting an advance SMS and sending a reminder. Combining the best survey maximization approaches, as well as texting mobile numbers with appended geo-demographics, should be a subject of future research on this topic.

In this study, I was bounded by the survey budget. Thus, there was a limit to how many mobile numbers I could validate, text, pre-notify or remind, and I could only offer small conditional incentives. With a larger budget, my experimental groups would be of a sufficient size to work with more statistical power in order to identify the best approaches to this kind of data collection (e.g., the best day for texting invitations). However, my study adds value by presenting results of a sophisticated survey experiment, including many experimental groups. In the future, investigating response maximization approaches should be more targeted, while building on the results of this study; e.g., carrying out an individual study on the number of reminders offering the best cost-benefit balance for the data collector.

References

American Association for Public Opinion Research. (2016). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (9th ed.). Washington DC: AAPOR.

Andreadis, I. (2020). Text message (SMS) pre-notifications, invitations and reminders for web surveys. *Survey Methods: Insights from the Field*, *8*. doi:10.13094/SMIF-2020-00019

Antoun, C., Conrad, F. G., Couper, M. P., & West, B. T. (2019). Simultaneous estimation of multiple sources of error in a smartphone-based survey. *Journal of Survey Statistics and Methodology*, *7*(1), 93–117.

Australian Bureau of Statistics. (2016). 2016 census of population and housing household internet access. Census TableBuilder. Retrieved from https://www.abs.gov.au/statistics/microdata-tablebuilder/tablebuilder

Australian Communications and Media Authority. (2022). Mobile-only Australia: Living without a fixed line at home. Retrieved from https://www.acma.gov.au/publications/2020-12/report/mobile-only-australia-living-without-fixed-line-home

Australian Communications and Media Authority. (n.d.). The numbering system. Retrieved from https://www.thenumberingsystem.com.au/#/smartnumbers-login

Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., ... Zahs, D. (2010). Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly*, *74*(4), 711–781. doi:10.1093/poq/nfq048

Balabanis, G., Mitchell, V. W., & Heinonen-Mavrovouniotis, S. (2007). SMS-based surveys: Strategies to improve participation. *International Journal of Advertising*, *26*(3), 369–385.

Bosnjak, M., Neubarth, W., Couper, M. P., Bandilla, W., & Kaczmirek, L. (2008). Prenotification in web-based access panel surveys: The influence of mobile text messaging versus e-mail on response rates and sample composition. *Social Science Computer Review*, *26*(2), 213–223.

Brick, J. M. (2008). Random-digit dialing. In P. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 675–678). Thousand Oaks: Sage.

Bucher, H., & Sand, M. (2021). Exploring the feasibility of recruiting respondents and collecting web data via smartphone: A case study of text-to-web recruitment for a general population survey in Germany. *Journal of Survey Statistics and Methodology*, *10*(4), 886–897.

Callegaro, M., Manfreda, K. L., & Vehovar, V. (2015). *Web survey methodology*. London: Sage.

Conrad, F., Schober, M. F., Antoun, C., Hupp, A., & Yan, H. (2017). Text interviews on mobile devices. In P. P. Biemer, E. D. De Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, ... B. T. West (Eds.), *Total survey error in practice* (pp. 299–318). New York: John Wiley & Sons.

Couper, M. P. (2000). Web surveys: A review of issues and approaches. *The Public Opinion Quarterly*, *64*(4), 464–494.

Couper, M. P., Antoun, C., & Mavletova, A. (2017). Mobile web surveys. In P. P. Biemer, E. D. De Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, ... B. T. West (Eds.), *Total survey error in practice* (pp. 133–154). New York: John Wiley & Sons.

Datareportal. (2020). Digital 2020: Australian. Retrieved from https://datareportal.com/reports/digital-2020-australia

De Bruijne, M., & Wijnant, A. (2014). Improving response rates and questionnaire design for mobile web surveys. *Public Opinion Quarterly, 78*(4), 951–962.

Deloitte. (2019). Mobile Consumer Survey 2019. Retrieved from https://www2.deloitte.com/au/en/pages/technology-media-and-telecommunications/articles/mobile-consumer-survey.html

Dillman, D. A. (2017). *The promise and challenge of pushing respondents to the web in mixed-mode surveys*. Statistics Canada.

Dillman, D. A. (2018). *How web-push surveys are changing survey methodology*. Presentation at the NatCen/ESS/City University methods seminar, City University of London, London, October 2.

Dykema, J., Stevenson, J., Day, B., Sellers, S. L., & Bonham, V. L. (2011). Effects of incentives and prenotification on response rates and costs in a national web survey of physicians. *Evaluation & the health professions*, *34*(4), 434–447.

Elevelt, A., Lugtig, P., & Toepoel, V. (2019). Doing a time use survey on smartphones only: What factors predict nonresponse at different stages of the survey process? *Survey Research Methods*, *13*(2), 195–213.

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review. *Computers in human behavior*, *26*(2), 132–139.

Fordyce, E., Bilgen, I., & Stern, M. J. (2020). *The use of synchronous and asynchronous text messaging during survey recruitment and screening*. Presentation on the 75th annual conference of the American Association for Public Opinion Research, Online, May 11–14.

Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology*. New York: John Wiley & Sons.

Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, *72*(2), 167–189.

Hsia, J., Zhao, G., & Town, M. (2020). Estimating undercoverage bias of internet users. *Preventing Chronic Disease*, *17*, 200026. doi:10.5888/pcd17.200026

Kaczmirek, L., Phillips, B., Pennay, D. W., Lavrakas, P. J., & Neiger, D. (2019). *Building a probability-based online panel: Life in Australia*. CSRM and SRC Methods Paper, 2–2019.

Keeter, S. (2019). *Growing and improving Pew Research Center's American trends panel*. Pew Research Center.

Keeter, S., Hatley, N., Kennedy, C., & Lau, A. (2017). *What low response rates mean for telephone surveys*. Pew Research Center.

Kellard, K. K. (2017). *What should (and shouldn't) we tell them? Qualitative pre-testing of different approaches for recruitment of a new probability panel*. Presentation on the 7th biennial conference of the European Survey Research Association, Lisbon, Portugal, July 17–21.

Keusch, F. (2015). Why do people participate in web surveys? Applying survey participation theory to internet survey data collection. *Management Review Quarterly*, *65*(3), 183–216.

Knowles, S., & Stahlmann-Brown, P. (2021). Cash is not king in incentivizing online surveys. *Applied Economics Letters*, *28*(2), 105–108.

Kongsgard, H. W., Syversen, T., & Krokstad, S. (2014). SMS phone surveys and mass-messaging: Promises and pitfalls. *Epidemiology: Open Access*, *4*. doi:10.4172/2161-1165.1000177

Laguilles, J. S., Williams, E. A., & Saunders, D. B. (2011). Can lottery incentives boost web survey response rates? Findings from four experiments. *Research in Higher Education*, *52*(5), 537–553.

Liu, M., & Wronski, L. (2018). Examining completion rates in web surveys via over 25,000 real-world surveys. *Social Science Computer Review*, *36*(1), 116–124.

Mavletova, A., & Couper, M. P. (2016). Device use in web surveys: The effect of differential incentives. *International Journal of Market Research*, *58*(4), 523–544.

Morgan, A. J., Rapee, R. M., & Bayer, J. K. (2017). Increasing response rates to follow-up questionnaires in health intervention research: Randomized controlled trial of a gift card prize incentive. *Clinical Trials*, *14*(4), 381–386.

Neal, Z., Neal, J. W., & Piteo, A. (2020). Call me maybe: Using incentives and follow-ups to increase principals' survey response rates. *Journal of Research on Educational Effectiveness*, *13*(4), 784–793.

Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, *33*(201), 101–116.

Parsons, N. L., & Manierre, M. J. (2014). Investigating the relationship among prepaid token incentives, response rates, and nonresponse bias in a web survey. *Field Methods*, *26*(2), 191–204.

Pennay, D. W., Neiger, D., Lavrakas, P. J., & Borg, K. (2018). *The online panels benchmarking study: A total survey error comparison of findings from probability-based surveys and nonprobability online panel surveys in Australia*. CSRM and SRC Methods Paper, 2–2018.

Phillips, B., & Compton, S. (2019). *Text messages and reminder calls in student and alumni web surveys*. CSRM and SRC Methods Paper, 4–2019.

Roy Morgan. (2015). Over a million Australians now use two mobile phones. Retrieved from http://www.roymorgan.com/findings/6024-a-million-australians-use-two-mobile-phones-november-2014-201501160100

Roy Morgan. (2016). 9 in 10 Aussie teens now have a mobile (and most are already on to their second or subsequent handset). Retrieved from http://www.roymorgan.com/findings/6929-australian-teenagers-and-their-mobile-phones-june-2016-201608220922

Roy Morgan. (2018). Kids and mobiles, how Australian children are using mobile phones. Retrieved from https://app.powerbi.com/view?r=eyJrIjoiNzA4Njc4YWMtZDk5oc00M2EwLWI3MDktMjkzMGRmOGNhOThkIiwidCI6IjBkYWM3ZjM5LWQyMGMtNGU3ms04YWYzLTcxZWU3ZTI2OGEyYiJ9&pageName=reportsectionf48b9bb17d51b38a13e1

Sala, E., Respi, C., & Decataldo, A. (2018). Non response in web surveys. The role of SMS reminders. *Rassegna italiana di Sociologia*, *59*(1), 5–24.

Saleh, A., & Bista, K. (2017). Examining factors impacting online survey response rates in educational research: Perceptions of graduate students. *Journal of MultiDisciplinary Evaluation*, *13*(29), 63–74.

SamplePages. (n.d.). What we do. Retrieved from https://samplepages.com.au/#service

Spreen, T. L., House, L. A., & Gao, Z. (2020). The impact of varying financial incentives on data quality in web panel surveys. *Journal of Survey Statistics and Methodology*, *8*(5), 832–850.

StataCorp. (2013). *Stata statistical software: Release 13*. College Station: StataCorp LP.

Van Mol, C. (2017). Improving web survey efficiency: The impact of an extra reminder and reminder content on web survey response. *International Journal of Social Research Methodology*, *20*(4), 317–327.

Watson, N. (2014). Evaluation of weighting methods to integrate a top-up sample with an ongoing longitudinal sample. *Survey Research Methods*, *8*(3), 195–208.