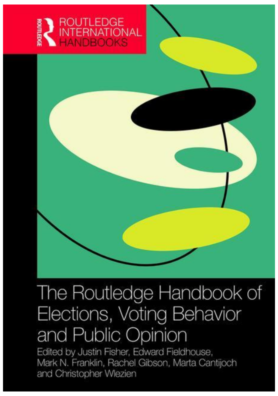

The Routledge Handbook of Elections, Voting Behaviorand Public Opinion

Justin Fisher, Edward Fieldhouse, Mark N. Franklin, Rachel Gibson, Marta Cantijoch, Christopher Wlezien

Horses for courses

Publication details
https://www.routledgehandbooks.com/doi/10.4324/9781315712390.ch37
Edward Fieldhouse, Christopher Prosser
Published online on: 26 Sep 2017

37

HORSES FOR COURSES

Using internet surveys for researching public opinion and voting behavior

Edward Fieldhouse and Christopher Prosser

With the rapid growth of the internet and digital technology in recent decades, public opinion and electoral research has undergone a transformation in how data are collected. As in-person surveys based on random probability samples of the population have become increasingly expensive, they have become increasingly rare. Declining response rates for telephone surveys and unresolved questions about the impact of mobile phones on sampling frames have also raised questions about the continued validity of phone samples (Curtin, Presser, and Singer 2005; Lavrakas et al. 2007). At the same time internet-based surveys, which can be delivered at a fraction of the cost in a short time frame, have become widespread. Online panel surveys have become perhaps the predominant way to collect data about political attitudes and behavior.

However, despite a common assumption that internet surveys would eventually make other forms of data collection redundant, the inexorable march of the online panel and the decline of other forms of survey data collection have been hampered by doubts about the accuracy and representativeness of online data.

Funders of research in political science, and customers of public opinion research more generally, have continued to hedge their bets whilst researchers attempt to resolve the on-going debate about whether online surveys can produce data of the same (or better) quality than more traditional methods. Survey research always involves necessary compromise between the quality of the data, the feasibility of contacting respondents, and the cost of doing so (Kish 1987). Internet surveys have a clear advantage in terms of cost – the cost per respondent and speed of data collection is unparalleled in other survey modes. The question for survey researchers is whether the potential risks of internet survey data outweigh these advantages.

In this chapter, we review some of the burgeoning evidence about the quality of internet-based survey methods.[1] There is strong evidence that online surveys enjoy considerable advantages that make them the mode of choice for a number of research purposes, but at the same time other methods – particularly those based on random probability samples – are better suited for making point estimates of population values. In other words, it is not a case of which mode is better, but a case of “horses for courses,” meaning that, for the time being at least, online and offline modes of data collection will continue to exist in parallel.

Types of internet survey

Traditional survey modes are often linked to particular sampling methods – face-to-face surveys are usually associated with area sampling, and phone surveys are most often conducted by random digit dialing (RDD) (Tourangeau, Conrad, and Couper 2013). This is not the case for internet surveys, which come in various shapes and sizes, and in terms of assessing the advantages and disadvantages of different modes of data collection, the devil is in the detail.

One important characteristic of most internet surveys is the panel element: whereas it might be feasible to carry out one-off internet surveys amongst a small and known target population, because of the lack of a sampling frame for the population, many internet surveys are based on existing larger *access* panels that are regularly drawn on to complete a variety of different surveys. These access panels are distinct from traditional *longitudinal* panel surveys, which repeatedly measure the attitudes and behaviors of the same respondents. Access panels can themselves be used to create longitudinal panels – for example, the British Election Study Internet Panel (Fieldhouse et al. 2015).

Other forms of internet survey do exist, such as river sampling – where respondents are recruited to take a survey via an advertisement or webpage rather than through an access panel. However, the dominant model of internet survey research is conducted using access panels (AAPOR Standards Committee 2010), and we focus our discussion in this chapter on these.[2] As we discuss below, the panel element of internet samples provides a number of advantages when it comes to researching the dynamics of political attitudes and voter behavior.

The most important characteristic of an internet panel is the method of sampling. There are two main types of samples used in internet survey panels – probability samples and nonprobability (opt-in) samples. Many of the criticisms leveled at internet surveys are aimed at the absence of probability sampling rather than the effects of delivery by internet. When comparing different modes of data collection (for example, telephone and internet), it is important to differentiate between the effects of mode from the effects of sampling. Bias in survey data has two main sources: bias in the sample and bias in the responses. Bias in the sample arises because the sample isn't representative of the target population, which may be due to the way the sample was drawn or to variations in the willingness of different groups to respond to the survey. Response bias, on the other hand, arises because those that do complete the survey may give inaccurate or misleading answers, or may be unable (or unwilling) to answer the questions in the way that was intended by the researcher. Both response and non-response bias are, of course, distinct from (and in addition to) sampling error, which refers to the random sampling variability inherent in any sample data.

Although for practical reasons mode and sampling methods may be related to each other, they can be separated. For example, online panels can be selected by random probability sampling, and in-person surveys can be selected on the basis of nonprobability methods such as quotas. However, in practice, due to the substantial costs of probability sampling over the internet and the potential saving from the removal of all interviewer costs, internet surveys are usually based on nonprobability samples.

Sampling and accuracy

Statistical inference is the idea that one can make generalizations about an entire population from a sample that is drawn from it. This is based on the simple premise that everyone in the population has a non-zero and known chance of being included in the sample.

One of the main threats to inference from internet panel samples is coverage error (Couper 2000). Coverage error is error that arises due to the people that are missing from a sampling frame (and so have a zero probability of being included in the survey) and differences between people covered by the frame and people who are not (Groves 1989). If the sampling frame covers the entire population, or those who are missing from the frame are missing at random, then no coverage error will occur. However, non-internet users are not missing at random from internet surveys. People with internet access and who use the internet regularly are systematically different from those who do not on socio-demographic, attitudinal, and other characteristics (Ragnedda and Muschert 2013; Zickuhr and Smith 2012).

Internet access and use has increased rapidly since the early days of internet survey research. However, there is still a sizeable population without internet access – according to the World Bank, in 2014, 22 percent of the adult population of OECD countries did not have access to the internet. Increasing internet penetration does not mean the risk of coverage error is decreasing. As the population without internet access becomes smaller, differences between people with and without internet access may become more pronounced (Callegaro, Baker et al. 2014).

An additional challenge for internet surveys is that there is no sampling frame of the general population from which to draw a sample, making it impossible to calculate the probability of selection. In-person surveys can sample from fairly complete registers of addresses, and telephone surveys can use random digit dialing, but there are currently no complete registers of e-mail or IP addresses of internet devices (computers, tablets, smartphones, etc.), and there is no clear link between devices and the population unit of interest (i.e., the individual person or voter). Not only are many devices shared (e.g., home computers) but many individuals have multiple devices and others none at all.

To address the lack of sampling frame, some internet panel probability samples are recruited offline using traditional address based or RDD random sampling methods (web probability panels). Coverage errors due to respondents not having internet access are solved by either providing it for them or supplementing the internet survey with additional survey modes (AAPOR Standards Committee 2010). However, this significantly increases the cost of recruiting and maintaining the panel and the dominant form of the internet sample remains the nonprobability based on opt-in panel.

There are many ways in which respondents might be recruited to a nonprobability panel but exactly how respondents are recruited to different panels is often protected as proprietary information (AAPOR Standards Committee 2010). Known methods for recruitment include advertising on websites, target mailing through partner organizations, and recommendation by other panelists. Panelists are usually offered a small incentive for taking part in surveys, generally entries into a prize draw or points which can be redeemed for cash. Examining the reasons people participate in online surveys, Poynter and Comley (2003) find that most (59 percent) respondents say incentives are an important reason, as well as other reasons such as curiosity (42 percent), enjoyment (40 percent), and having their views heard (28 percent).

Once they have developed an access panel, internet survey operators attempt to emulate a population sampling frame by drawing samples from the larger access panel in numbers or quotas for different groups based on the relative size of those groups in the target population. Regardless of its sophistication, no method can transform a nonprobability sample into a probability sample. The goal of all research is to make inferences about phenomena that are unbiased by "disturbing variables" (Kish 1987) that might lead to incorrect conclusions. Probability sampling diminishes the risks of this problem through randomization. Nonprobability samples do not have the benefit of randomization, and if self-selection into a nonprobability panel is related to the variables of interest then there is the substantial risk that inferences will be wrong (Baker et al. 2013). However, as Kish wrote, "Great advances of the most successful sciences … were, and are, achieved without probability sampling … Probability sampling for randomization is not a dogma, but a strategy ..." (1965: 28–29). Nonprobability samples have a long history of use in public opinion research in the form of quota sampling. History has shown that, although quota samples can come up with the right answer some of the time, they can also go drastically wrong (Mosteller et al. 1949; Jowell et al. 1993). Without a theoretic basis, it is impossible to judge how accurate a nonprobability sample will be.

Some scholars have risen to this challenge and have tried to put the use of nonprobability samples on a firmer theoretical footing (Rivers 2007; Terhanian and Bremer 2012). This work draws from case-control methods and uses matching to a control group (Rosenbaum and Rubin 1983) at the sample selection stage and propensity score weighting after sample selection to deal with potential biases in the sample (see Chapter 36 for a more comprehensive examination of these issues).

The most important element of inference with nonprobability samples is whether the self-selection into the sample is "ignorable" (Rivers 2013). Essentially, the opt-in sample (where probability of inclusion is affected by unobserved effects of self-selection, non-response, and non-coverage) can provide an estimate of a variable *Y* so long as the probability of inclusion is conditionally independent of *Y* (Rubin 1976). In other words, it is ignorable insofar as it is possible to explain any difference between the survey and the population on *Y* with a given set of covariates.

Whether self-selection into nonprobability internet panels is ignorable is an empirical question. Research in Britain suggested that nonprobability internet surveys show great promise (Sanders et al. 2004; Sanders et al. 2007; Twyman 2008). However, the early consensus in American studies of internet surveys was that nonprobability panels performed worse than probability surveys of different modes (Malhotra and Krosnick 2007; Chang and Krosnick 2009; Pasek and Krosnick 2010; AAPOR Standards Committee 2010; Yeager et al. 2011). More recent American studies have been more optimistic (Ansolabehere and Schaffner 2014; Kennedy et al. 2016). In the next section, we examine the performance of nonprobability samples in more detail.

Fit for purpose?

It is important to distinguish between two purposes that surveys might have – estimating population values of certain variables and of estimating the relationship between sets of variables. Accurate point estimates *can* be estimated from nonprobability samples (Wang et al. 2015; Kennedy et al. 2016) but that does not mean that they *will* be estimated accurately (see Chapter 36 in this volume). Recent research (Kennedy et al. 2016) has emphasized that certain types of nonprobability panels, particularly those that employ sophisticated matching and propensity score weighting methods, perform much better than other types of panel. The key question is whether those methods can adequately adjust for the effects of selection. The consensus is that if the research goals are point estimates of particular variables in the population the risks of incorrect inferences are substantially lower using probability sampling methods (Baker et al. 2013; Callegaro, Villar et al. 2014).

A perennial problem for surveys of all types is that survey respondents tend to be politically and civically more engaged than the general population, and this is likely to be especially true for opt-in panels (Groves, Presser, and Dipko 2004; Tourangeau, Groves, and Redline 2010; Kennedy et al. 2016). The population distribution of political engagement is essentially unknowable and so if political engagement is a non-ignorable conditioning variable this might present a serious problem. An example of such a problem is demonstrated by Mellon and Prosser (forthcoming) in their examination of the 2015 British polling failure. Mellon and Prosser find that the key explanation for the polling miss was the under sampling of non-voters. The demographics of turnout and party support are correlated – for example, younger people are less likely to vote in general, but more likely to vote Labour when they do vote. When a pool of respondents that does not contain a representative number of non-voters is quota sampled or weighted to look like the population, voters that demographically resemble non-voters (and the parties they support) will be overrepresented in the sample.

Whilst the evidence on the accuracy of point estimates from internet surveys based on nonprobability panels is decidedly mixed, many have argued that researchers are often more interested in understanding the relationships between variables, rather than their prevalence in the population. In this respect, the debate is similar to that over the use of convenience samples such as college students in experimental psychology (Mook 1983; Sears 1986). Again the key concern is "ignorability" – there is no basis for assuming the homogeneity of relationships between nonprobability samples and the population. In psychology, for instance, there is some evidence that college student samples lead to effect sizes of different magnitude, direction, and significance compared to non-student samples (Peterson 2001). Similarly, a recent comparison of several nonprobability internet samples found that the samples with the least accurate point estimates were also least accurate in terms of relationships between variables (Kennedy et al. 2016).

Comparing probability and nonprobability samples, some studies have found substantial variation in the strength and significance of different coefficients (Malhotra and Krosnick 2007; Pasek and Krosnick 2010) whilst others have found more similar results (Berrens et al. 2003; Sanders et al. 2007; Stephenson and Crête 2011; Ansolabehere and Schaffner 2014; Simmons and Bobo 2015; Bytzek and Bieber 2016; Pasek 2016).

A common finding with internet surveys is higher levels of concurrent and predictive validity on the internet compared to telephone surveys (Chang and Krosnick 2009; Simmons and Bobo 2015; Pasek 2016). For example, Chang and Krosnick (2009) find that vote choice measured in internet surveys was more highly correlated with widely accepted predictors of vote choice, such as government approval and party identification. Whether or not this increased explanatory power is a function of the mode or an artefact of sample composition (Simmons and Bobo 2015) is not clear.

Malhotra and Krosnick (2007) suggest nonprobability samples can be more troublesome for some outcome variables than others. For example, models of turnout are especially problematic as commonly over 90 percent of internet panels (claim to) vote, leaving a very small (and likely unrepresentative) proportion of the sample to explain the differences between voters and non-voters. How important differences in coefficients are is a matter of judgment. Many of the studies which have argued in favor of nonprobability samples find differences in coefficient strength but ultimately argue that you "would make the same policy inference" (Berrens et al. 2003: 20) or reach the same conclusions about different models of vote choice (Sanders et al. 2007).

Even if nonprobability panels can, and often do, result in the same inferences as probability panels, this is not the same thing as saying they *will*. Even studies supporting the use of nonprobability panels contain contradictory results. For example, researchers would reach the opposite conclusions about the effect of gender on turnout depending on whether they used the probability or nonprobability samples reported in Sanders et al. (2007). Although nonprobability samples may be perfectly adequate for many, even most, research purposes, findings from nonprobability samples that contradict theoretical expectations and long-standing trends should be treated with caution.

The use of nonprobability internet samples for research is a rapidly developing area. Increasingly nonprobability samples can, and do, perform just as well as probability samples – and in some cases actually outperform them (Kennedy et al. 2016). The crux of the debate is therefore not whether you can make inferences from nonprobability samples, but about the assumptions that one must make, and whether such assumptions are any less relevant to probability samples with non-response.

Benefits of internet surveys

Despite some of the reservations about the representativeness of (nonprobability) internet samples, delivery of a survey instrument by internet has a number of potential advantages compared to interviewer-based methods of collection. While in-person surveys offer the benefits of personal rapport between interviewer and respondent – encouraging participation, clarifying questions, interpreting responses, and helping reduce item non-response – there are also associated disadvantages. Collection of data by trained interviewers is expensive, particularly with the increasing costs of human labor and travel. Even telephone and mail costs tend to be expensive compared to online completion of surveys (Callegaro, Baker, et al. 2014). But quite apart from costs, online surveys can be fielded very quickly and turnaround times can be kept short as they do not rely on the availability of interviewers. Moreover, respondents can complete the survey at their own convenience, taking breaks when they like, making the experience less demanding. The use of the internet means that complex routing, randomization, and attractive presentation can be used to smooth the survey experience. It also facilitates use of complex visual and audio tools, and the straightforward delivery of survey experiments (Skitka and Sargis 2006). Given the current popularity of experimental methods for making causal inferences, the convenience of being able to implement fairly complex survey experiments (for example, including multimedia) on substantial samples with a panel is an attractive feature of internet surveys (Mutz 2011).

A further advantage of internet surveys is the inherent benefits of a panel design. Whilst in-person survey and telephone surveys can be designed as a panel, the maintenance of the panel element adds to the higher cost compared to internet surveys. When studying political behavior, there are always dangers to causal inference brought about by endogeneity. It is widely accepted that both these problems can be ameliorated, if not eliminated, by a panel design using a range of specialized statistical methods for panel data analysis (Finkel 1995). Internet surveys offer a fast and cost-effective way of delivering repeated measure data for large numbers of citizens and offer a uniquely valuable resource for researchers concerned with understanding within-person change. Just as the representativeness of a sample may be less critical for the relationship between variables than it is for making point estimates, the risks associated with analyzing within-person change are lower still as many unobserved characteristics are constant. For example, if we are interested in factors that affect voters switching from party A to party B, we do not require the sample to be perfectly representative with respect to party support, or even other variables correlated with party support; we only require that factors related to switching are not correlated with the probability of being included in the sample. Because these risks are relatively low compared to the risk of obtaining misleading point estimates, we suggest that the question of which is the better choice (internet versus non-internet) depends on the purpose for which it is needed, as well as the type of sample.

Mode effects

Internet surveys have a great deal in common with other visual and self-administered survey modes (Tourangeau, Conrad, and Couper 2013). A particular concern with all surveys, but particularly self-administered modes is *satisficing* (Krosnick 1991). Satisficing is the use of cognitive shortcuts to reduce required effort to answer survey questions and varies with motivation, ability, and the difficulty of questions. Face-to-face surveys reduce satisficing because interacting with an interviewer in person increases motivation (Holbrook, Green, and Krosnick 2003), whilst internet surveys may be more cognitively demanding because they require respondents to know how to use a computer and answer questions without help from an interviewer (Heerwegh and Loosveldt 2008). Others, however, have argued that visually administered surveys actually reduce the cognitive burden of a survey because respondents do not have to hold information in their working memory (Tourangeau, Conrad, and Couper 2013).

In common with other visually administered surveys, internet surveys are prone to one particular form of satisficing – primacy bias – the tendency for respondents to pick early response options in lists (in contrast to orally administered surveys, which are more likely to suffer recency bias because respondents have to hold information in working memory [Chang and Krosnick 2009]). Primacy bias was originally observed in self-administered paper surveys (Krosnick and Alwin 1987). Evidence from eye-tracking data shows that some respondents in computer-administrated surveys pay more attention to items at the top of a list (Galesic et al. 2008). There is also the possibility of computer-specific primacy effects from certain design options like scrolling answer lists where not all answers are visible at the same time (Couper et al. 2004). However, set against this, internet delivery offers a relatively simple way to randomize the order of response options compared to interviewer-based methods.

Another common concern with internet surveys is *speeding:* the tendency for some respondents to go through a survey as quickly as possible without paying adequate attention to the questions. Tourangeau, Conrad, and Couper (2013) argue that speeding is a problem in any self-administered survey but is associated with internet surveys because it is *detectable*. A recent evaluation of speeding (Greszki, Meyer, and Schoen 2014) found that, although speeding occurs, it does not do so at the high levels some have assumed and is not so prevalent or systematic that it affects estimates.

One form of satisficing that might be of particular concern for some research questions is the tendency for some internet respondents to cheat on knowledge questions. A consistent finding from many studies is that respondents completing surveys online tend to score higher on knowledge questions (Ansolabehere and Schaffner 2014; Fricker et al. 2005; Strabac and Aalberg 2011). These studies have generally concluded that people with internet access are more informed than their offline counterparts. Recent evidence suggests, however, that some respondents cheat on knowledge questions by researching the answers while they take the survey (Clifford and Jerit 2014; Burnett 2016). Despite the dangers of various forms of satisficing, much of the research has demonstrated that overall satisficing is lower in internet surveys than in telephone surveys (Chang and Krosnick 2009).

A widely cited advantage of internet surveys is an improvement in response accuracy for sensitive items. The presence of an interviewer can induce social desirability bias by reducing the sense of privacy (Holbrook, Green, and Krosnick 2003; Tourangeau and Yan 2007; Kreuter, Presser, and Tourangeau 2008; Heerwegh 2009). Mainly this means that interviewees are inclined to hide opinions and behaviors that they perceive to be socially undesirable, or in some sense transgress accepted social norms (Krysan 1998). For example, this might include unwillingness to admit to socially conservative values, such as against immigration or support for extremist parties, and to overestimate socially desirable attributes such as voter turnout. In contrast in internet surveys, respondents can express their socially undesirable opinions in private. For example, Tourangeau, Groves, and Redline (2010) found that self-reported illicit drug use showed consistently higher rates of reporting with self-completion compared to interviewer administration. Chang and Krosnick (2009) find lower levels of socially desirable responding in internet surveys, as measured by the answers provided by white respondents to questions about government programs to support African Americans. Whilst there have been exceptions to this finding (see Ansolabehere and Schaffner 2014), the general consensus is that self-completion internet surveys are likely to have fewer problems of social desirability bias than interviewer-based modes.

Panel conditioning and trained respondents

Internet panelists are often members of multiple panels (Stenbjerre and Laugesen 2005; Vonk, van Ossenbruggen, and Willems 2006) and a small number of panel members account for a large proportion of survey responses (Craig et al. 2013). The effect of repeated participation in survey research can be divided into two types – the effect on respondents' attitudes and behaviors, and the way in which they respond to the survey itself. Sturgis, Allum, and Brunton-Smith (2009) argue that repeatedly administering attitude questions causes respondents to reflect and deliberate on the issues raised by the questions they are asked. The results of this are stronger and more internally consistent attitudes in the later waves of a panel survey. There is also some evidence that being asked about electoral participation might increase the likelihood of actually voting (Greenwald et al. 1987; Granberg and Holmberg 1992).

Most panel conditioning research, however, has examined the effect of survey participation on the survey response process itself. Trained respondents (i.e., those who have taken many surveys) tend to answer surveys more quickly than fresh respondents (Toepoel, Das, and van Soest 2008). This could be in part because trained respondents are used to the question-answering process and learn how to interpret questions but it could also be due to satisficing and trained respondents may not read questions properly and thus make more mistakes. Toepoel, Das, and van Soest (2008) suggest the latter is the case and trained respondents are less likely to notice reverse-worded questions. Other research has found similar results, with a strong correlation between "professional" (those that take a large number of surveys) and inattentive respondents (Vonk, van Ossenbruggen, and Willems 2006). Whether this occurs due to training effects or the type of respondents in nonprobability samples is unclear. Chang and Krosnick (2009) find evidence of practice effects for probability internet panel respondents but not for nonprobability panel respondents. Hillygus, Jackson, and Young (2014) suggest that satisficing behavior of professional respondents is not due to panel conditioning but to differences in motivation – professional respondents take surveys because they want compensation rather than because they are interested in the survey topic.

There is also evidence from offline surveys that repeatedly administering questions about sensitive items can lead to more socially desirable reporting (Sharpe and Gilbert 1998; Halpern-Manners and Warren 2012). Sharpe and Gilbert (1998) find that repeated administration of the Beck Depression Inventory leads to more socially desirable responses and Halpern-Manners and Warren (2012) find a similar effect for labor market status. How these effects interact with sensitive item reporting is not yet clear.

Bias from panel conditioning may be exacerbated by differential attrition in panels and varying levels of response to invitations from panel members to specific surveys. For example, panel members who choose to complete political surveys tend to be more interested in politics. This is akin to non-response bias in probability surveys, although response rates in internet surveys are hard to define because of lack of a sampling frame, and there is no agreed way to define response metrics (see Callegaro and DiSogra 2008). Cavallaro (2013) suggests that the problems with attrition – the non-random dropping out of some types of respondents – pose a more serious problem to internet panels than the problem of panel conditioning.

Conclusions

Given the problems we have outlined in this chapter, the reader would be forgiven for thinking that our advice would be to avoid using nonprobability internet panels. This is not our intention and we follow other authors (Farrell and Petersen 2010; Callegaro, Villar et al. 2014) in saying that internet research should not be stigmatized. It is important, however, for users to be aware of the potential pitfalls of particular survey modes and threats to inference that may emerge from internet survey research. A recent report into nonprobability internet panels by the American Association for Public Opinion Research (AAPOR) identified a continuum of the expected accuracy of estimates from nonprobability samples, with the highest risk of false inferences arising from the use of uncontrolled convenience samples but treating the sample as if the respondents were a random sample of the population. The accuracy of estimates is likely to be much higher when using surveys that select respondents and adjust the data for non-ignorable conditioning variables. As the study notes, the challenge for researchers "arises in placing surveys between the two extremes. This is largely uncharted territory for social, opinion, and market research surveys" (Baker et al. 2013: 100).

All survey research is imperfect (Weisberg 2005). The potential problems with internet samples – particularly nonprobability samples – should not be taken to mean that other survey modes are without risks, nor that the problems with internet panels are insurmountable obstacles to inference.

The question for researchers must be what is the aim of their research? An earlier AAPOR report into internet surveys recommended that "researchers should avoid nonprobability online panels when one of the research objectives is to accurately estimate population values" (AAPOR Standards Committee 2010: 758), a conclusion echoed by many others (see, for example, Callegaro, Villar et al. 2014; Simmons and Bobo 2015; Bytzek and Bieber 2016, but see Stoker and McCall in this volume). Although some internet panels have shown great promise and this advice may change in the future, for the time being at least we see no reason to dissent from this view.

For other research purposes, internet panels offer greater potential. Seemingly intractable debates over causal ordering in electoral research mean that cross-sectional surveys are increasingly inadequate to answer research questions (Hillygus 2011). No research is without risk and users of internet surveys should approach internet research with their eyes open, but the potential for low-cost longitudinal panels and embedded survey experiments provides an unparalleled opportunity for researchers to answer important questions about electoral behavior.

Notes

1 For a more detailed look at the potential problems with nonprobability sampling and its similarities with non-response in probability samples, see Chapter 36 in this volume.

2 A further type of internet data collection that is seeing an increasing number of users is what might be termed internet convenience samples – running surveys and experiments on crowdsourced labor platforms such as Amazon Mechanical Turk. We do not discuss this type of data here but many of the same concerns and considerations about data quality in internet panel surveys might equally apply to internet convenience samples (see, for example, Clifford, Jewell, and Waggoner 2015).

References

AAPOR Standards Committee (2010) "AAPOR Report on Online Panels," *Public Opinion Quarterly*, vol. 74, no. 4, Winter, 711–781.

Ansolabehere, S. and Schaffner, B. F. (2014) "Does Survey Mode Still Matter? Findings from a 2010 Multi-Mode Comparison," *Political Analysis*, vol. 22, no. 3, June, 285–303.

Baker, R., Brick, J. M., Bates, N. A., Battaglia, M., Couper, M. P., Dever, J. A., Gile, K. J., and Tourangeau, R. (2013) "Summary Report of the AAPOR Task Force on Non-Probability Sampling," *Journal of Survey Statistics and Methodology*, vol. 1, no. 2, November, 90–143.

Berrens, R. P., Bohara, A. K., Jenkins-Smith, H., Silva, C., and Weimer, D. L. (2003) "The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples," *Political Analysis*, vol. 11, no. 1, Winter, 1–22.

Burnett, C. M. (2016) "Exploring the Difference in Participants' Factual Knowledge between Online and in-Person Survey Modes," *Research and Politics*, vol. 3, no. 2, June, 1–7.

Bytzek, E. and Bieber, I. E. (2016) "Does Survey Mode Matter for Studying Electoral Behavior? Evidence from the 2009 German Longitudinal Election Study," *Electoral Studies*, vol. 43, September, 41–51.

Callegaro, M. and DiSogra, C. (2008) "Computing Response Metrics for Online Panels," *Public Opinion Quarterly*, vol. 72, no. 5, December, 1008–1032.

Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. S., Krosnick, J. A., and Lavrakas, P. J. (2014) "Online Panel Research: History, Concepts, Applications and a Look at the Future," in Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. J., Krosnick, J. A., and Lavrakas, P. J. (eds.) *Online Panel Research: A Data Quality Perspective*, London: Wiley: 1–22.

Callegaro, M., Villar, A., Yeager, D. S., and Krosnick, J. A. (2014) "A Critical Review of Studies Investigating the Quality of Data Obtained with Online Panels Based on Probability and Nonprobability Samples," in Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. J., Krosnick, J. A., and Lavrakas, P. J. (eds.) *Online Panel Research: A Data Quality Perspective*, London: Wiley: 23–53.

Cavallaro, K. (2013) "By the Numbers: Theory of Adaptation or Survival of the Fittest?," *Quirks Marketing Research Review*, vol. 27, January, 24–27.

Chang, L. and Krosnick, J. A. (2009) "National Surveys Via Rdd Telephone Interviewing Versus the Internet Comparing Sample Representativeness and Response Quality," *Public Opinion Quarterly*, vol. 73, no. 4, Winter, 641–678.

Clifford, S. and Jerit, J. (2014) "Is There a Cost to Convenience? An Experimental Comparison of Data Quality in Laboratory and Online Studies," *Journal of Experimental Political Science*, vol. 1, no. 2, January, 120–131.

Clifford, S., Jewell, R. M., and Waggoner, P. D. (2015) "Are Samples Drawn from Mechanical Turk Valid for Research on Political Ideology?," *Research and Politics*, vol. 2, no. 4, December, 1–9.

Couper, M. P. (2000) "Web Surveys: A Review of Issues and Approaches," *Public Opinion Quarterly*, vol. 64, no. 4, Winter, 464–494.

Couper, M. P., Tourangeau, R., Conrad, F. G., and Crawford, S. D. (2004) "What They See Is What We Get Response Options for Web Surveys," *Social Science Computer Review*, vol. 22, no. 1, February, 111–127.

Craig, B. M., Hays, R. D., Pickard, A. S., Cella, D., Revicki, D. A., and Reeve, B. B. (2013) "Comparison of US Panel Vendors for Online Surveys," *Journal of Medical Internet Research*, vol. 15, no. 11, November, e260.

Curtin, R., Presser, S., and Singer, E. (2005) "Changes in Telephone Survey Nonresponse over the Past Quarter Century," *Public Opinion Quarterly*, vol. 69, no. 1, Spring, 87–98.

Farrell, D. and Petersen, J. C. (2010) "The Growth of Internet Research Methods and the Reluctant Sociologist," *Sociological Inquiry*, vol. 80, no. 1, February, 114–125.

Fieldhouse, E., Green, J., Evans, G., Schmitt, H., van der Eijk, C., Mellon, J., and Prosser, C. (2015) "British Election Study, 2015: Internet Panel Survey," Available online: www.britishelectionstudy.com/data-objects/panel-study-data/ (doi: 10.15127/1.293723) [accessed June 16, 2017].

Finkel, S. E. (1995) *Causal Analysis with Panel Data*, 1st Edition, Thousand Oaks: Sage Publications.

Fricker, S., Galesic, M., Tourangeau, R., and Yan, T. (2005) "An Experimental Comparison of Web and Telephone Surveys," *Public Opinion Quarterly*, vol. 69, no. 3, Autumn, 370–392.

Galesic, M., Tourangeau, R., Couper, M. P., and Conrad, F. G. (2008) "Eye-Tracking Data New Insights on Response Order Effects and Other Cognitive Shortcuts in Survey Responding," *Public Opinion Quarterly*, vol. 72, no. 5, December, 892–913.

Granberg, D. and Holmberg, S. (1992) "The Hawthorne Effect in Election Studies: The Impact of Survey Participation on Voting," *British Journal of Political Science*, vol. 22, no. 2, April, 240–247.

Greenwald, A. G., Carnot, C. G., Beach, R., and Young, B. (1987) "Increasing Voting Behavior by Asking People If They Expect to Vote," *Journal of Applied Psychology*, vol. 72, no. 2, May, 315–318.

Greszki, R., Meyer, M., and Schoen, H. (2014) "The Impact of Speeding on Data Quality in Nonprobability and Freshly Recruited Probability-Based Online Panels," in Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. J., Krosnick, J. A., and Lavrakas, P. J. (eds.) *Online Panel Research: A Data Quality Perspective*, London: Wiley: 238–262.

Groves, R. M. (1989) *Survey Errors and Survey Costs*, 2nd Edition, New York: Wiley.

Groves, R. M., Presser, S., and Dipko, S. (2004) "The Role of Topic Interest in Survey Participation Decisions," *Public Opinion Quarterly*, vol. 68, no. 1, Spring, 2–31.

Halpern-Manners, A. and Warren, J. R. (2012) "Panel Conditioning in Longitudinal Studies: Evidence from Labor Force Items in the Current Population Survey," *Demography*, vol. 49, no. 4, November, 1499–1519.

Heerwegh, D. (2009) "Mode Differences Between Face-to-Face and Web Surveys: An Experimental Investigation of Data Quality and Social Desirability Effects," *International Journal of Public Opinion Research*, vol. 21, no. 1, Spring, 111–121.

Heerwegh, D. and Loosveldt, G. (2008) "Face-to-Face versus Web Surveying in a High-Internet-Coverage Population Differences in Response Quality," *Public Opinion Quarterly*, vol. 72, no. 5, December, 836–846.

Hillygus, D. S. (2011) "The Evolution of Election Polling in the United States," *Public Opinion Quarterly*, vol. 75, no. 5, December, 962–981.

Hillygus, D. S., Jackson, N., and Young, M. (2014) "Professional Respondents in Nonprobability Online Panels," in Callegaro, M., Baker, R., Bethlehem, J., Göritz, A. J., Krosnick, J. A., and Lavrakas, P. J. (eds.) *Online Panel Research: A Data Quality Perspective*, London: Wiley: 219–237.

Holbrook, A. L., Green, M. C., and Krosnick, J. A. (2003) "Telephone Versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Response Bias," *Public Opinion Quarterly*, vol. 67, no. 1, Spring, 79–125.

Jowell, R., Hedges, B., Lynn, P., Farrant, G., and Heath, A. (1993) "The 1992 British Election: The Failure of the Polls," *Public Opinion Quarterly*, vol. 57, no. 2, Summer, 238–263.

Kennedy, C., Mercer, A., Keeter, S., Hatley, N., McGeeney, K., and Gimenez, A. (2016) "Evaluating Online Nonprobability Surveys," Pew Research Center, May 2, 2016, available online: www.pewresearch.org/files/2016/04/Nonprobability-report-May-2016-FINAL.pdf [accessed July 18, 2016].

Kish, L. (1965) *Survey Sampling*, 2nd Edition, New York: Wiley-Interscience.

Kish, L. (1987) *Statistical Design for Research*, New York: Wiley.

Kreuter, F., Presser, S., and Tourangeau, R. (2008) "Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity," *Public Opinion Quarterly*, vol. 72, no. 5, December, 847–865.

Krosnick, J. A. (1991) "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys," *Applied Cognitive Psychology*, vol. 5, no. 3, June, 213–236.

Krosnick, J. A. and Alwin, D. F. (1987) "An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement," *Public Opinion Quarterly*, vol. 51, no. 2, Summer, 201–219.

Krysan, M. (1998) "Privacy and the Expression of White Racial Attitudes: A Comparison Across Three Contexts," *Public Opinion Quarterly*, vol. 62, no. 4, Winter, 506–544.

Lavrakas, P. J., Shuttles, C. D., Steeh, C., and Fienberg, H. (2007) "The State of Surveying Cell Phone Numbers in the United States 2007 and Beyond," *Public Opinion Quarterly*, vol. 71, no. 5, December, 840–854.

Malhotra, N. and Krosnick, J. A. (2007) "The Effect of Survey Mode and Sampling on Inferences about Political Attitudes and Behavior: Comparing the 2000 and 2004 ANES to Internet Surveys with Non-probability Samples," *Political Analysis*, vol. 15, no. 3, Summer, 286–323.

Mellon, J. and Prosser, C. (forthcoming) "Missing Non-Voters and Misweighted Samples: Explaining the 2015 Great British Polling Miss," *Public Opinion Quarterly*.

Mook, D. G. (1983) "In Defense of External Invalidity," *American Psychologist*, vol. 38, no. 4, 379–387.

Mosteller, F., Hyman, H., McCarthy, P., Marks, E., and Truman, D. (1949) *The Pre-Election Polls of 1948: The Report to the Committee on Analysis of Pre-Election Polls and Forecasts*, New York: Social Science Research Council.

Mutz, D. C. (2011) *Population-Based Survey Experiments*, Princeton: Princeton University Press.

Pasek, J. (2016) "When Will Nonprobability Surveys Mirror Probability Surveys? Considering Types of Inference and Weighting Strategies as Criteria for Correspondence," *International Journal of Public Opinion Research*, vol. 28, no. 2, Summer, 269–291.

Pasek, J. and Krosnick, J. A. (2010) *Measuring Intent to Participate and Participation in the 2010 Census and Their Correlates and Trends: Comparisons of RDD Telephone and Non-Probability Sample Internet Survey Data*, Washington, DC: Statistical Research Division of the US Census Bureau.

Peterson, R. A. (2001) "On the Use of College Students in Social Science Research: Insights from a Second-Order Meta-Analysis," *Journal of Consumer Research*, vol. 28, no. 3, December, 450–461.

Poynter, R. and Comley, P. (2003) "Beyond Online Panels," Technovate: CRM, Internet Research and New Media, 2003 Cannes.

Ragnedda, M. and Muschert, G. W. (eds.) (2013) *The Digital Divide: The Internet and Social Inequality in International Perspective*, Abingdon: Routledge.

Rivers, D. (2007) "Sampling for Web Surveys," Paper presented at the 2007 Joint Statistical Meetings of the American Statistical Association, Salt Lake City, August 1.

Rivers, D. (2013) "Comment," *Journal of Survey Statistics and Methodology*, vol. 1, no. 2, November, 111–117.

Rosenbaum, P. R. and Rubin, D. B. (1983) "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, vol. 70, no. 1, April, 41–55.

Rubin, D. B. (1976) "Inference and Missing Data," *Biometrika*, vol. 63, no. 3, December, 581–592.

Sanders, D., Clarke, H. D., Stewart, M. C., and Whiteley, P. (2007) "Does Mode Matter for Modeling Political Choice? Evidence From the 2005 British Election Study," *Political Analysis*, vol. 15, no. 3, Autumn, 257–285.

Sanders, D., Clarke, H. D., Stewart, M. C., Whiteley, P., and Twyman, J. (2004) "The 2001 British Election Study Internet Poll," *Journal of Political Marketing*, vol. 3, no. 4, 29–55.

Sears, D. O. (1986) "College Sophomores in the Laboratory: Influences of a Narrow Data Base on Social Psychology's View of Human Nature," *Journal of Personality and Social Psychology*, vol. 51, no. 3, September, 515–530.

Sharpe, J. P. and Gilbert, D. G. (1998) "Effects of Repeated Administration of the Beck Depression Inventory and Other Measures of Negative Mood States," *Personality and Individual Differences*, vol. 24, no. 4, April, 457–463.

Simmons, A. D. and Bobo, L. D. (2015) "Can Non-Full-Probability Internet Surveys Yield Useful Data? A Comparison with Full-Probability Face-to-Face Surveys in the Domain of Race and Social Inequality Attitudes," *Sociological Methodology*, vol. 45, no. 1, August, 357–387.

Skitka, L. J. and Sargis, E. G. (2006) "The Internet as Psychological Laboratory," *Annual Review of Psychology*, vol. 57, 529–555.

Stenbjerre, M. and Laugesen, J. N. (2005) "Conducting Representative Online Research: A Summary of Five Years of Learnings," Paper presented at the 2005 ESOMAR Worldwide Panel Research Conference, Budapest, April 17–19.

Stephenson, L. B. and Crête, J. (2011) "Studying Political Behavior: A Comparison of Internet and Telephone Surveys," *International Journal of Public Opinion Research*, vol. 23, no. 1, Spring, 24–55.

Strabac, Z. and Aalberg, T. (2011) "Measuring Political Knowledge in Telephone and Web Surveys: A Cross-National Comparison," *Social Science Computer Review*, vol. 29, no. 2, May, 175–192.

Sturgis, P., Allum, N., and Brunton-Smith, I. (2009) "Attitudes Over Time: The Psychology of Panel Conditioning," in Lynn, P. (ed.) *Methodology of Longitudinal Surveys*, London: John Wiley & Sons: 113–126.

Terhanian, G. and Bremer, J. (2012) "A Smarter Way to Select Respondents for Surveys?," *International Journal of Market Research*, vol. 54, no. 6, 751–780.

Toepoel, V., Das, M., and van Soest, A. (2008) "Effects of Design in Web Surveys Comparing Trained and Fresh Respondents," *Public Opinion Quarterly*, vol. 72, no. 5, December, 985–1007.

Tourangeau, R. and Yan, T. (2007) "Sensitive Questions in Surveys," *Psychological Bulletin*, vol. 133, no. 5, September, 859–883.

Tourangeau, R., Conrad, F., and Couper, M. (2013) *The Science of Web Surveys*, New York: Oxford University Press.

Tourangeau, R., Groves, R. M., and Redline, C. D. (2010) "Sensitive Topics and Reluctant Respondents Demonstrating a Link between Nonresponse Bias and Measurement Error," *Public Opinion Quarterly*, vol. 74, no. 3, Autumn, 413–432.

Twyman, J. (2008) "Getting It Right: YouGov and Online Survey Research in Britain," *Journal of Elections, Public Opinion and Parties*, vol. 18, no. 4, October, 343–354.

Vonk, T., van Ossenbruggen, R., and Willems, P. (2006) "The Effects of Panel Recruitment and Management on Research Results," ESOMAR World Research Conference, Panel Research 2006, Barcelona.

Wang, W., Rothschild, D., Goel, S., and Gelman, A. (2015) "Forecasting Elections with Non-Representative Polls," *International Journal of Forecasting*, vol. 31, no. 3, September, 980–891.

Weisberg, H. F. (2005) *The Total Survey Error Approach*, Chicago: University of Chicago Press.

Yeager, D. S., Krosnick, J. A., Chang, L-C., Javitz, H. S., Levendusky, M. S., Simpser, A., and Wang, R. (2011) "Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples," *Public Opinion Quarterly*, vol. 75, no. 4, Winter, 709–747.

Zickuhr, K. and Smith, A. (2012) "Digital Differences," Pew Research Center, April 13, 2012, available online: www.pewinternet.org/~/media//Files/Reports/2012/PIP_Digital_differences_041312.pdf [accessed July 19, 2016].