Public Health 204 (2022) 63–69

Contents lists available at ScienceDirect

Public Health

journal homepage: www.elsevier.com/locate/puhe

Original Research

How survey mode affects estimates of the prevalence of gambling harm: a multisurvey study

P. Sturgis*, J. Kuha

Department of Methodology, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK

ARTICLE INFO

Article history:
Received 15 July 2021
Received in revised form
16 December 2021
Accepted 16 December 2021

Keywords:
Gambling harm
Survey
Selection bias
Error

ABSTRACT

Recent general population surveys have produced highly variable estimates of the extent of problem gambling in Great Britain, ranging from as low as 0.4% to as high as 2.7% of adults. This level of uncertainty over the true level of problem gambling creates difficulties for policy makers and those planning treatment and support services for individuals and families affected by problem gambling. In this article, we assess the extent to which differences in approaches to sampling and measurement between surveys contribute to variability in estimates of problem gambling. We compare estimates of problem gambling using the Problem Gambling Severity Index across eight different surveys conducted at approximately the same time but which use different sampling and measurement strategies. Our findings show that surveys conducted online produce substantially higher estimates of problem gambling compared with in-person interview surveys. This is because online surveys, whether using probability or non-probability sampling, overrepresent people who are more likely to gamble online and to gamble frequently, relative to the proportions of these groups in the general population.

Introduction

Since 2012, official statistics on the prevalence of gambling and gambling harm in Great Britain have been collected using a combined version of the national health surveys for England and Scotland and a bespoke survey in Wales. These surveys use what are considered “gold standard” methodologies of random sampling and in-person interviewing. They have estimated comparatively low rates of gambling harm in the adult population. The 2016 survey estimated the rate of problem gamblers to be 0.7% and the rate of adults at risk of gambling harm to be 4.2%. Similar rates of 0.4% and 3.9% were estimated in the 2018 Health Survey for England (covering England only).

In 2019, a survey carried out by YouGov found 2.7% of British adults identified as problem gamblers and 13.2% at risk of gambling harm, more than three times higher than the health survey had estimated less than a year previously for England. This survey had a quite different methodological approach using non-probability sampling and online self-completion of questionnaires. Such a large discrepancy in estimates raises questions about what the true level of gambling harm is in the general population, which, in turn, makes it difficult for policy makers and planners to determine the appropriate level of resource allocation for treatment and support services.

In the context of the global trend toward increased online surveying, it is essential to better understand how survey mode affects the accuracy of estimates of problematic gambling. Recent evidence suggests that non-probability online samples substantially overestimate problem gambling compared with probability samples collected in person and by phone because of both selection bias and poor measurement quality.[1,2] However, it is not clear to what extent this is because of online interviewing on the one hand or non-probability sampling on the other.

Our objective in this article is to assess how sample design and survey mode affect estimates of harmful gambling. We do this by comparing estimates of gambling behavior and gambling harm across a set of contemporaneously conducted surveys using a consistent set of questions but different sampling and data collection methodologies. We evaluate how differences in the designs of the surveys are related to variation in estimates of gambling behavior and gambling harm, using this to draw conclusions about the likely prevalence of problem gambling in the adult population of England.

* Corresponding author.
E-mail address: p.sturgis@lse.a.uk (P. Sturgis).

https://doi.org/10.1016/j.puhe.2021.12.014

Survey designs and measures of gambling harm

To assess how survey design features affect estimates of harmful gambling, we consider eight near-contemporaneous surveys, which all included the same measure of gambling harm, the Problem Gambling Severity Index (PGSI). These were the 2016 and 2018 rounds of the Health Survey for England, the 2019 and 2020 GambleAware Treatment and Support surveys carried out by YouGov, and three surveys conducted for the purposes of this study in November and December 2020 by Yonder, NatCen, and Kantar Public. In addition, Ipsos-MORI has provided us with data from a survey they collected for their own purposes in January 2021. The key design features of the surveys are described in detail in Appendix 1 and summarized in Table 1.

The Health Survey for England uses probability sampling with in-person interviewing, and the NatCen, Kantar, and Ipsos-MORI surveys are online probability surveys, which draw random samples from established panels of respondents who have been pre-recruited to complete surveys on a regular basis for monetary incentives.[3] The panels are established via a "recruitment survey," which also uses probability sampling, although the mode of contact differs between postal (Ipsos-MORI and Kantar) and face-to-face interview (NatCen). Of particular note is the markedly lower response rates achieved for the probability panels (4%–15%) compared with the health surveys (~55%). The YouGov and Yonder surveys use a similar approach, but the established panels of respondents are not drawn randomly. Instead, these panels comprise people who have signed up to take surveys in return for monetary incentives through a range of online and offline recruitment strategies.[4]

The key variable of comparison is the PGSI.[5] It is based on answers to nine questions about gambling, each with four response alternatives: 0 = never, 1 = sometimes, 2 = most of the time, and 3 = almost always. The total PGSI score is the sum of the individual items. The total score is recoded into four categories, indicating "non-gambler," "low-risk," "moderate-risk," and "problem gambling" for scores of 0, 1–2, 3–7, and ≥8, respectively. We focus here primarily on the proportion with a score of ≥1 on the PGSI, which we refer to hereafter as PGSI+1.

Before the PGSI, respondents were asked a set of questions asking whether they had participated in a range of gambling activities during the previous 12 months. Those who reported no gambling were not administered the PGSI and are given a score of zero. Respondents were also asked how frequently they gamble. The small number of respondents who did not provide responses to the PGSI are excluded from analyses. The question wordings and response alternatives are provided in the Appendix.

Results

Fig. 1 shows that estimates for the two health surveys, at 3.9% and 4.1%, are substantially lower than all of the online surveys[a], which range from a low of 7.4% for Ipsos-MORI to a high of 16% for Yonder.[b] The 95% confidence intervals for the health surveys do not overlap with any of the online surveys, so sampling variability can be ruled out as a potential cause of the differences.

[a] For simplicity, we refer to the surveys that used online self-completion as "the online surveys," although the Kantar and NatCen surveys used both online and telephone interviews.

[b] For the random probability surveys, confidence intervals are calculated using Taylor series linearization to account for complex design features. For the non-probability samples, the same approach is used to account for the calibration weights. Although this is technically not correct due to the non-random selection of population elements, it serves as a reasonable approximation.

The Ipsos-MORI estimate is the lowest of the online surveys, but it is not directly comparable because it uses a 4-week reference period for previous gambling behavior, whereas all other surveys refer to 12 months. This likely reduces the PGSI+1 by 1–2 percentage points for this survey.[6] Note also that the health surveys use a target population of adults aged ≥16 years, whereas the online surveys, apart from Kantar (which also uses 16+), use 18+. It is not possible to derive equivalent bands because the health surveys and the Kantar survey do not contain a continuous age in years variable. Given the small size of the 16–17 years age group and the low incidence of PGSI+1 in the general population, this difference will have little or no effect on the point estimate for the general population.

True change over time

Although we cannot rule the possibility of true change in gambling behavior, it does not seem likely to be a major contributory factor for two reasons. First, 12 months is an implausibly short interval to accommodate such a substantial increase in gambling harm. Second, independent surveys conducted during the first lockdown in 2020 found a *decline* in the frequency of gambling.[7,8] It therefore seems highly unlikely that the increase in harmful gambling observed between the 2018 HSE and the online surveys conducted in 2019/20 could be because of a real increase in gambling harm in the population.

Coverage error

There are differences in the covered populations between the health surveys and the YouGov survey, which might have caused some of the difference in estimates. For example, the Postcode Address File (PAF), which is the sampling frame for the health surveys, excludes people who live in institutional addresses, such as halls of residence, hospitals, prisons, and military barracks. The YouGov and Yonder surveys, on the other hand, can include members of these groups but exclude the offline population completely. However, the Kantar, NatCen, and Ipsos-MORI surveys also draw their samples from PAF and therefore have the same coverage properties as the health survey. This means that coverage error can also be ruled out as a potential cause of the differences in estimates.

Measurement error

It is possible that some of the variability in estimates of gambling harm derives from differences in the measurement properties of the survey instruments. For example, answers to the gambling questions might have been differentially affected by the content of questions that preceded them, so-called "order effects."[9] The gambling questions in the health surveys were preceded by questions focusing on mental health and well-being, whereas for all but the Ipsos-MORI survey (which first asked questions about politics and vaccination), the online surveys asked the gambling questions first. Although this pattern is consistent with the possibility that preceding the gambling items with questions about mental health and well-being reduces the frequency of self-reported gambling harm, there is no obvious theoretical reason why this should be so. Without experimental evidence to support such a hypothesis, we conclude that the case for order effects of any notable magnitude is weak.

There are also differences between surveys in the questions and response alternatives. Respondents who report no gambling in the previous 12 months on these questions are assigned a score of zero on the PGSI, so differences in these questions could affect the

Table 1
Summary information on the sample designs of the eight surveys.

Survey	Sample design	Mode	Sample size	Fieldwork	Age range	Response rate	Question order
HSE 2016	Probability sample, Postcode Address File (PAF) as the first-stage sampling frame, all adults in a household are interviewed, £10 unconditional incentive	Paper self-completion in face-to-face (f-t-f) interview	6691	Annual continuous	16+	55%	After mental health questions at the end of f-t-f interview
HSE 2018	As for 2016 HSE	Paper self-completion in f-t-f interview	6927	Annual continuous	16+	54%	After mental health questions at the end of f-t-f interview
Kantar	Probability, PAF, up to two adults, £5 conditional incentive	Online + phone	1795	November 24, 2020, to December 13, 2020	16+	5%	First in questionnaire
Ipsos	Probability, PAF, up to two adults, £10 conditional incentive	Online		January 21, 2021, to January 27, 2021	18+	4%	After politics, vaccination, views of local area
NatCen	Probability, PAF, one adult, £10 conditional incentive	Online + phone	2049	November 19, 2020, to December 20, 2020	18+	14%	
YouGov 2019	Quota sample (with age, gender, ethnicity, social grade, and region as quota variables), incentive = points toward money	Online	10499	September 24, 2019, to October 13, 2019	18+	N/A	First in questionnaire
YouGov 2020	Quota (age, gender, ethnicity, social grade, region), incentive = point toward money	Online	16401	November 19, 2020, to December 11, 2020	18+	N/A	First in questionnaire
Yonder	Quota (age, gender, region, social grade), incentive = points toward money	Online	6944	November 18, 2020 to November 29, 2020	18+	N/A	Not known

Sample sizes for England only.

estimates of gambling harm. That being said, the two sets of questions cover a large range of gambling activities, and both include an "any other type of gambling" question, so it is not clear why they would produce strongly different rates of gambling prevalence. Our assessment is, therefore, that these differences in question content and format are unlikely to be a notable contributory factor.

The health surveys include a skip instruction at the bottom of the page of questions on gambling activities. The instructions advise respondents who answered "no" to all these questions to skip further forward in the questionnaire. It is possible this led some respondents to answer "no" to all the questions to proceed more quickly to the end of the questionnaire. However, these instructions are at the bottom of the page and are not especially prominent. As there had been no similar filter questions in the self-completion questionnaire up to that point, there was no opportunity for respondents to learn that skipping questions in this way could help them to progress faster. We therefore consider it unlikely that this had a material impact on the estimates of gambling prevalence in the health surveys.

It is well known that people are less willing to admit to socially undesirable attitudes and behaviors in the presence of another person.[10] For this reason, we might expect online surveys to be more accurate because no interviewer is present. To minimize the risk of this kind of bias, the health surveys use a paper self-completion questionnaire for the gambling questions. Nonetheless, it is still possible that the presence of an interviewer or other household members might lead to underreporting of gambling in the self-completion questionnaire.

We can obtain some insight on this by comparing PGSI+1 between respondents who completed the 2016 HSE questionnaire alone or in the presence of another household member. This shows a small difference for the 2016 HSE; of the 63% who completed the gambling questions in the presence of another household member,

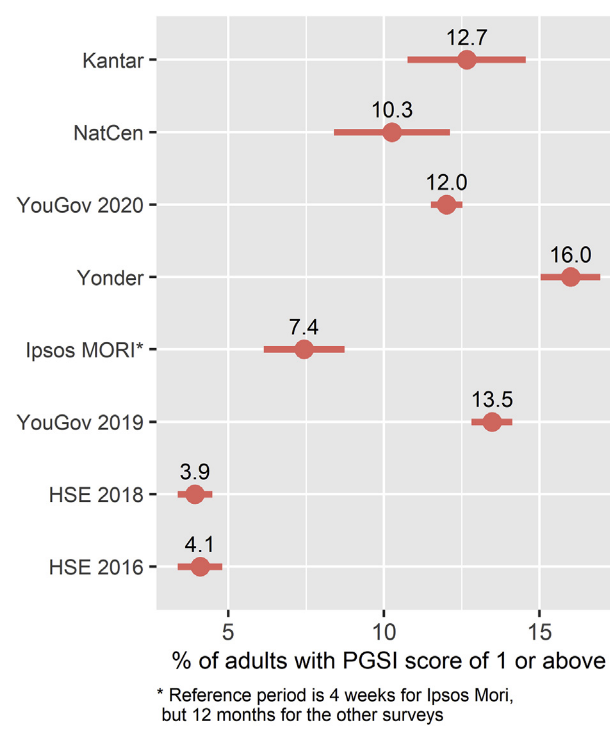

Fig. 1. Estimates of the percentage of adults with PGSI+1 across surveys.

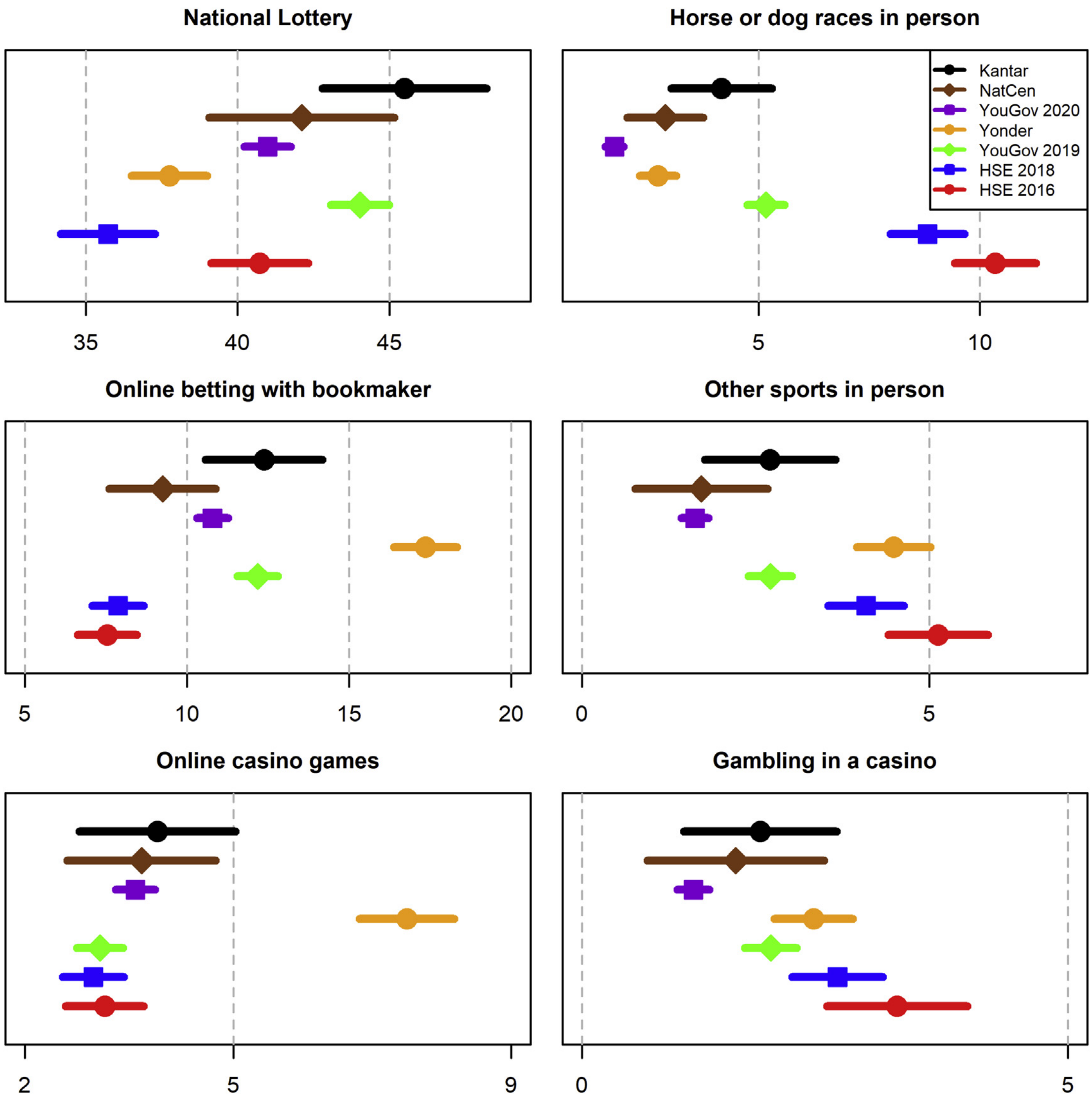

Fig. 2. Estimates of the percentage of adults who have taken part in different gambling activities over the previous 12 months.

3.9% had a PGSI+1 compared with 4.7% for the 37% who completed the questions alone. This difference, however, is not statistically significant (Chi-square = 0.92, df = 1, $P = 0.354$), which leads us to conclude that socially desirable responding in the health surveys is unlikely to be a significant contributory factor to the lower estimates of gambling harm.

Non-response error/selection bias

In probability sampling, non-response bias results from the failure to contact sampled individuals or from their refusal to take part in the survey once contacted. If the propensity to respond to the survey is correlated with the population parameter of interest, estimates will be biased.[11] In general, the magnitude of non-response bias is unknown, and we can only say that the *risk* of it increases as the response rate declines.

In non-probability sampling, there is no directly equivalent number to the response rate because recruitment typically continues until the sampling quotas are filled, and it is therefore more appropriate to refer to the more general concept of selection bias. If, after weighting adjustments, the kinds of people who agree to complete the survey are different from people in the target population on the characteristic(s) of interest, estimates will be biased.[12] A number of existing studies have found that, on average, non-probability surveys tend to be more biased than probability samples because of unrepresentative samples.[13]

Fig. 2 presents point estimates and 95% confidence intervals for a selection of gambling activities. At the top of the chart, we see that all surveys give similar estimates of the proportion who purchased a National Lottery ticket, ranging from 36% in the 2018 HSE to 46%

Table 2
Frequency of spending money on gambling (estimated percentage of adults in England).

Frequency	Kantar	NatCen	YouGov 2020	Yonder	HSE 2018	HSE 2016
More than once a week	15.1	14.6	18.8	25.7	10.2	12.7
Once a week	22.9	27.9	26.4	27.8	23.7	27.3
Less than once a week	8.6	10.9	8.6	9.3	10.8	10.1
Once a month	18.9	19.0	17.7	16.3	13.3	12.0
Every 2–3 months	14.9	11.2	12.7	10.8	14.2	13.6
Once-twice a year	19.5	16.4	15.8	10.0	27.9	24.2
Total	*100*	*100*	*100*	*100*	*100*	*100*

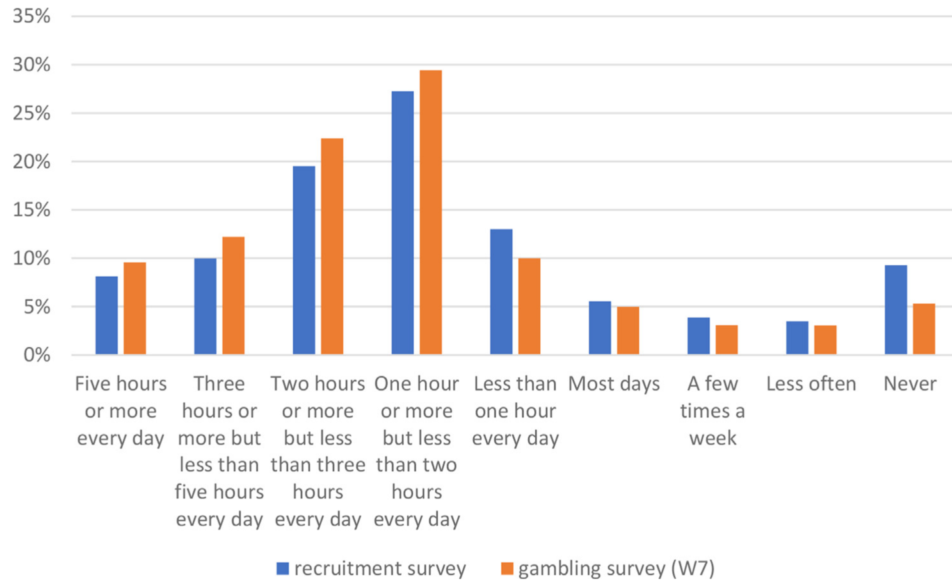

Fig. 3. Frequency of Internet use, Kantar Public Voice survey: percentages of adults estimated from the recruitment survey and wave 7.

in the Kantar survey. For in-person betting on horse or dog races, however, the estimates are notably and significantly higher for the health surveys (9%–10%) than for the online surveys (1%–5%).

The same pattern is evident for gambling at "other sports event in person," for which the health survey estimates are generally higher compared with the online surveys. Some of this difference likely reflects the cessation of in-person events in March 2020, although the 2019 YouGov survey also shows a lower estimate than the health surveys for in-person gambling activities, so change in gambling behavior due to lockdown restrictions does not completely account for the difference.

The opposite pattern is evident for online betting at bookmakers, for which the health surveys have lower estimates than the online surveys and for online casino games, where the health surveys are among the lowest estimates. The health surveys, then, also detect different types of gambling activities, with in-person gambling more common and online gambling less common compared with the online surveys.

Table 2 reveals a marked difference in the reported frequency of gambling, with the online surveys showing a range of 15%–26% gambling more than once a week, compared with 10% for the 2018 HSE. The higher rate of gambling in the online surveys is also evident at the opposite end of the scale, with 10%–20% reporting gambling only once or twice a year compared with 28% in the 2018 HSE.

Existing studies have found that online and higher frequency gambling are associated with an increased risk of gambling harm.[3,14] This is also the case here, where in all the surveys, the

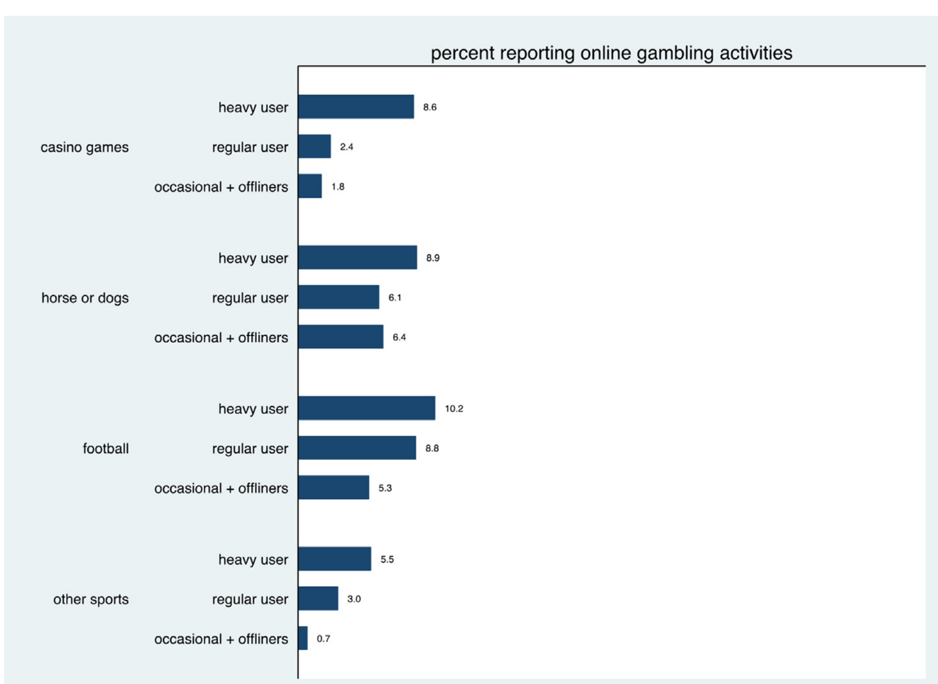

Fig. 4. Percentages of adults who engage in different online gambling activities, by frequency of Internet use, estimated from the Kantar Public Voice survey.

estimated proportions of people with PGSI+1 broadly increase with higher frequency of gambling (these figures are shown in Table A2 of Appendix 2). However, it is also the case that at all levels of frequency, this proportion is higher in the online surveys than in the face-to-face health surveys. It therefore seems likely that differences in sample composition in both frequency of gambling and type of gambling activity are responsible for the higher rates of problem gambling in the online surveys. The online surveys contain more people more likely to gamble online and to gamble frequently, and these characteristics are associated with an elevated risk of harmful gambling.

We can also examine differences between the surveys in other characteristics of the respondents, although this is limited to a small number of variables, which are consistently available for them. Table A3 in Appendix 2 shows the estimated distributions of four demographic characteristics. For gender and age, these are similar by construction because these variables are typically incorporated in the survey weights. Estimated distributions of ethnic group (as White vs non-White) are also very similar. Larger differences are observed only for educational qualifications, where the online surveys estimate more people with degree-level qualifications and fewer with no qualifications than do the health surveys. Higher education is in turn associated with more online betting (results not shown here), which could account for some of the differences discussed previously.

How might these differences in sample composition have come about? First, non-probability online panels have been shown to produce substantially biased estimates of behaviors relating to the Internet and technology use.[15,16] Two possibilities are germane to the question of why an online bias might be evident for the probability panel surveys. First, although the offline population and infrequent Internet users *can* join these panels, they may still be underrepresented. Fig. 3 shows the amount of time people spend on the Internet is higher at wave 7 than at the recruitment interview survey of the Kantar Public Voice panel. Because the comparison here is on a variable measured at the recruitment survey, this change over time is driven by less frequent Internet users dropping out of the panel rather than an increase in Internet use by panel members.

Finally, Fig. 4 shows the rates of four types of online gambling in the Kantar survey for heavy, regular, and occasional users combined with offliners. Heavy Internet users are considerably more likely to report all four online gambling activities. This lends additional support to the contention that the online surveys select for people who are more likely to be online and frequent gamblers and who, in turn, are more likely to report gambling harm.

Discussion

Until 2018, official statistics on gambling in Great Britain were delivered using probability sampling and in-person interviewing, an approach that produced comparatively low estimates of gambling harm. However, a survey carried out by YouGov in 2019 estimated a total of more than 6 million adults falling in the “at risk” category. Such wide variability in estimates raises questions about what the true level of harmful gambling is in the general population and what the most appropriate approaches are for estimating gambling harm. The question of how survey mode affects the accuracy of estimates of gambling behavior is particularly pressing as the COVID-19 pandemic has accelerated the shift from interviewer administered to online interviewing.[17,18]

Our objective in this article has been to provide insight on the likely rate of gambling harm in England by identifying the sources of error that are driving disparities in estimates. To do this, we have made comparisons between eight surveys containing a consistent set of gambling questions but varying approaches to sample design and data collection. For six of the surveys, data collection was done via online self-completion with two using a mixed-mode (online and telephone) design, although for the mixed-mode surveys, the vast majority of interviews (90%) were carried out online. Three of the online surveys used probability sampling, and three used non-probability (quota) sampling.

These comparisons have enabled us to identify selection bias as the primary source of the differences in estimates of gambling harm. Comparisons across a range of estimates revealed a systematic pattern: the online surveys contained gamblers who were more likely to gamble online and to gamble frequently. Other potential causes of the differences, including true change in harmful gambling, sampling variability, coverage error, and differential measurement error, seem unlikely to exert a notable influence.

These differences in sample composition are likely to be driving the discrepancies in rates of problem gambling between surveys, with online surveys—whether based on probability or non-probability samples—tending to overestimate gambling harm relative to interviewer-administered in-person surveys. A similar pattern of online surveys overstating the true level of problem gambling has also been observed in a recent systematic review[1] (but see also Russell et al[19]). When samples contain disproportionate quantities of online and frequent gamblers (compared with the general population), surveys will tend also to overestimate gambling harm because online and frequent gambling are independently associated with a higher probability of gambling harm.

Author statements

Ethical approval

None declared.

Funding

The research reported here was funded by GambleAware.

Competing interests

None declared.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.puhe.2021.12.014.

References

1. Pickering Dylan, Blaszczynski Alex. Paid online convenience samples in gambling studies: questionable data quality. *Int Gambl Stud* 2021;**21**(3):516–36. https://doi.org/10.1080/14459795.2021.1884735.
2. Pew Research Center. *Assessing the risks to online polls from bogus respondents*. 2020. https://www.pewresearch.org/methods/2020/02/18/assessing-the-risks-to-online-polls-from-bogus-respondents/.
3. Blom AG, Bosnjak M, Cornilleau A, et al. A comparison of four probability-based online and mixed-mode panels in europe. *Soc Sci Comput Rev* 2016;**34**(1):8–25. https://doi.org/10.1177/0894439315574825.
4. Callegaro M, Baker R, Bethlehem J, Göritz AS, Krosnick JA, Lavrakas PJ. Online panel research: history, concepts, applications and a look at the future. In: Callegaro M, Baker R, Bethlehem J, Goritz AS, Krosnick JA, Lavrakas PJ, editors. *Online panel research: a data quality perspective*. UK: John Wiley & Sons; 2014. p. 1–8.
5. Orford J, Wardle H, Griffiths M, Sproston K, Erens B. PGSI and DSM-IV in the 2007 British Gambling Prevalence Survey: reliability, item response, factor structure and inter-scale agreement. *Int Gambl Stud* 2010;**10**:31–44.
6. Sturgis P, Kuha J. *Methodological factors affecting estimates of the prevalence of gambling harm in the United Kingdom: a multi-survey study*. GambleAware; 2021. available at: https://www.begambleaware.org/sites/default/files/2021-05/Methodology_Report_%28FINAL_14.05.21%29.pdf.

7. Gunstone B, Gosschalk K, Joyner O, Diaconu A, Sheikh M. *The impact of the COVID-19 lockdown on gambling behaviour, harms and demand for treatment and support*. GambleAware; 2020. available at: https://www.begambleaware.org/sites/default/files/2020-12/yougov-covid-19-report.pdf.
8. Wardle H, Donnachie C, Critchlow N, Brown A, Bunn C, Dobbie F, et al. The impact of the initial Covid-19 lockdown upon regular sports bettors in Britain: findings from a cross-sectional online study. *Addict Behav* 2021;Volume 118.
9. Schuman H, Presser S, Ludwig J. Context effects on survey responses to questions about abortion. *Publ Opin Q* 1981;**45**(2):216–23.
10. Tourangeau R, Yan T. Sensitive questions in surveys. Sep *Psychol Bull* 2007;**133**(5):859–83. https://doi.org/10.1037/0033-2909.133.5.859. PMID: 17723033.
11. Groves R. Nonresponse rates and nonresponse bias in household surveys. *Publ Opin Q* 2006;**70**(5):646–75. https://doi.org/10.1093/poq/nfl033.
12. Sturgis P, Kuha J, Jennings W, Baker N, Callegaro M, Fisher S, et al. An assessment of the causes of the errors in the 2015 UK General Election Polls. *J Roy Stat Soc* 2017;**181**(3):757–81.
13. Cornesse C, Blom A, Dutwin D, Krosnick JD, De Leeuw E, Legleye S, et al. A review of conceptual approaches and empirical evidence on probability and nonprobability sample survey research. *J Surv Stat Methodol* 2020;**8**(1):4–36.
14. Howe PDL, Vargas-Saénz A, Hulbert CA, Boldero JM. Predictors of gambling and problem gambling in Victoria, Australia. *PLoS One* 2019;**14**(1):e0209277. https://doi.org/10.1371/journal. pone.0209277.
15. Keeter S, McGeeney K, Mercer A, Hatley N, Patten E, Perrin A. *Coverage error in internet surveys: who web-only surveys miss and how that affects results*. Pew Research Center; 2015. Available at: http://www.pewresearch.org/files/2015/09/2015-09-22_coverage-error-in-internet-surveys.pdf.
16. Herzing JME, Blom AG. The influence of a person's digital affinity on unit nonresponse and attrition in an online panel. *Soc Sci Comput Rev* 2019;**37**(3):404–24.
17. Olson Kristen, Smyth Jolene D, Horwitz Rachel, Keeter Scott, Lesser Virginia, Marken Stephanie, et al. Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR task force report. *J Surv Stat Methodol* 2020. https://doi.org/10.1093/jssam/smz062.
18. The UK Gambling Commission is proposing to move its gambling prevalence survey from in person and telephone to online self-completion from 2022: https://www.gamblingcommission.gov.uk/manual/participation-and-prevalence-research/proposal-5-explore-more-future-proof-methods.
19. Russell Alex MT, Browne Matthew, Hing Nerilee, Rockloff Matthew, Newall Philip. Are any samples representative or unbiased? reply to Pickering and Blaszczynski. *Int Gambl Stud* 2021. https://doi.org/10.1080/14459795.2021.1973535.