Longitudinal Study of Gambling

Consultation and scoping review

NatCen
Social Research that works for society

Authors: Debbie Collins and Heather Wardle
Date: 05/04/2019
Prepared for: Gambling Commission

At **NatCen Social Research** we believe that social research has the power to make life better. By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

NatCen Social Research
35 Northampton Square
London EC1V 0AX
T 020 7250 1866
www.natcen.ac.uk

A Company Limited by Guarantee
Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland (SC038454)
This project was carried out in compliance with ISO20252

Contents

Tables

Figures

1 Background

The Gambling Commission (the Commission) identified understanding changes in gambling behaviour over time as a key theme within their Research Programme 2018-2022. This included suggestions to conduct longitudinal research to address a range of research questions. Subsequently, the Commission appointed NatCen Social Research to scope design options for a potential longitudinal study of gambling in Great Britain to support the Commission's understanding of the feasibility of implementing such a study.

The aims of the scoping project were to:

- Better understand the range and priority of research questions that both the Commission and a wide range of stakeholders think a longitudinal study could address;
- Set out the full range of longitudinal design options, summarising the pros and cons of each;
- Provide a summary of which types of designs are best suited to answering which types of research questions;
- Provide indicative costs for multiple methodologies to help inform the Commission about the scale of resources required;
- Provide some recommendations as to which design options are most likely to answer the priority research questions identified by the Commission.

Figure 1.1 shows the structure of the scoping review project, and of this report.

Figure 1-1 Structure of consultation, scoping review and report

Refine research questions	Identify possible designs	Assessment of design options
• Consultation survey sent to 141 participants to gain views on what key research questions they would like a longitudinal study to address • Valid responses received from 32 participants • Results discussed with the Gambling Commission and key research questions refined	• Range of different methodological approaches identified by NatCen Social Research • A summary of design options sent to 7 experts in longitudinal designs for comments • 6 experts provided comments; no further design options were identified	• Design options assessed against a standardised list of strengths and weaknesses • Design options assessed against research questions • Recommendations made based on these results • Indicative costs for several design options produced

1.1 Research questions

Research questions drive the design of a study. To better understand the needs of the broader research and policy community in Britain, a rapid online user consultation identified common areas of interest. Details of who responded and a summary of responses can be found in Appendix A. The consultation specifically focused on the types of research questions users would like to see a longitudinal study answer and the different subgroups that should be included. Summary findings were provided to the Commission who confirmed that the following are their priority areas of interest:

- Understanding gambling trajectories over time, including how people's gambling behaviour changes over time and what affects these changes. The study would explore pathways into gambling and problem gambling and how individuals move in and out of risky gambling behaviours. It would also look at the trajectories of those in treatment and recovery.
- the relationship between gambling products and gambling behaviour trajectories over time.
- the consequences/harms of gambling and how these may change over time.

These priority areas may change as the Commission's thinking on longitudinal research evolves. However, these were the research questions used in our subsequent assessment of the ability of different research designs to answer these types of questions.

1.2 Generation of design options: some considerations

1.2.1 Defining the target population

A range of different design options are discussed in section 2 and assessed in section 3. An important consideration in the assessment of each of these options is what the target population for the study should be. The online user consultation sought views on the target population for a longitudinal survey of gambling. Views were presented to the Commission, who confirmed that the following are priority groups of interest:

- The general population – including non-gamblers to assess movements in and out of gambling and the gambling products that are played;
- Highly engaged gamblers – to assess movement in and out of harmful behaviour, the relationship between harmful gambling and gambling products and other external factors (e.g. advertising, life events) and the consequences of gambling;
- Problem gamblers and/or those receiving treatment – to assess treatment and recovery pathways (or otherwise) from problems.

Design options are assessed in terms of their coverage of these groups.

1.2.2 Identifying gamblers

There is no list of all people who gamble in Great Britain, however there are lists of people who have signed up to online gambling sites and people who have registered as members of particular gambling venues (e.g. bingo, casinos) and lists of people who have been in contact with gambling treatment services. These lists will not include all gamblers, however. The use of these lists is considered in section 2.1.1.

In the absence of a list of all gamblers, gamblers are identified using self-assessment tools. Currently the Commission collects data on gambling participation and prevalence through questions asked on the Health Survey for England, Scottish Health Survey and the Welsh Problem Gambling Survey. These surveys identify gamblers by asking a set of screening questions, that ask about gambling behaviour retrospectively.

1.2.3 Sample design

Another important consideration is the sample design. There are several factors to consider: the importance of being able to generalise findings to the wider population that the study is meant to represent; the need to identify enough people in specific population groups of interest to be able to say something meaningful about their behaviours and experiences over time; and the kinds of analyses that may be undertaken. This is a key challenge in meeting the research priorities set out by the Commission and something we discuss throughout the remaining sections of this report.

1.2.4 Sample size

In addition to the general population, the Commission are interested in particular sub-groups: young people aged 16-34, engaged gamblers (those who gamble at least once a week) and problem gamblers. Problem gambling is measured using standardised self-assessment tools such as the DSM-IV Multi Response or Problem Gambling Severity Index (PGSI). Table 1.1 shows the prevalence of these subgroups in the general population.

Table 1-1 Prevalence of groups of interest in GB population

Subgroup	% of English & Scottish population 16+
Gamblers	59.9%
Engaged gamblers[1]	14.9%
Problem gamblers	0.6%

Subgroup	% of English & Scottish 16-34
Gamblers	60.0%
Engaged gamblers	12.6%
Problem gamblers	0.9%

Source: Unweighted data from 2016 combined Health Survey for England and Scottish Health Survey. A non-random omnibus survey was used to collected data in Wales and the data were not included in the combined estimates.

The low prevalence of problem gamblers presents a challenge in designing a longitudinal study of gamblers. Different design options are considered: increasing the initial sample size to ensure that it contains enough problem gamblers to track over time; prioritising the follow up on problem gamblers where this information is available; and including a separate cohort of known problem gamblers.

Table 1.2 illustrates the estimated numbers of engaged and problem gamblers that would be identified in a general population sample, depending on the overall size of the achieved sample. The table also illustrates the estimated numbers of young people in the groups of interest. It should be born in mind that these numbers represent the achieved sample size at wave one and assume that likelihood to respond is the same for all subgroups. In practice this is unlikely to be the case as, for example, young people tend to be less likely to take part in surveys than other age groups.

[1] Engaged gamblers are those who gamble (excluding the national lottery) at least once a week.

Table 1-2 Illustration of sample size for general population study at wave 1

	Achieved sample size			
All 16+	**50,000**	**40,000**	**20,000**	**10,000**
Engaged gamblers 16+	**7,450**	**5,960**	**2,980**	**1,490**
Problem gamblers 16+	**300**	**240**	**120**	**60**
16-34 year olds	**11,585**	**9,270**	**4,635**	**2,315**
Engaged gamblers	**1,460**	**1,170**	**585**	**290**
Problem gamblers	**105**	**85**	**40**	**20**

Table 1.3 illustrates the impact attrition has on the achieved sample size over three waves of data collection depending on the size of the achieved sample at wave one. The response rates assumed at waves 2 and 3 are 65% and 75% respectively. As with table 1.2, for the purposes of this illustration it is assumed that likelihood to respond at subsequent waves is constant across subgroups, however this is unlikely to be the case in practice.

Table 1-3 Illustration of achieved sample size over 3 waves

	Achieved sample size		
	W1	**W2**	**W3**
All 16+	**50,000**	**32,500**	**24,375**
Engaged gamblers 16+	7,450	4,840	3,630
Problem gamblers 16+	300	195	145
16-34 year olds	11,585	**7,530**	**5,650**
Engaged gamblers	1,460	950	710
Problem gamblers	105	70	50
All 16+	**40,000**	**26,000**	**19,500**
Engaged gamblers 16+	5,960	3,875	2,905
Problem gamblers 16+	240	155	115
16-34 year olds	9,270	**6,025**	**4,520**
Engaged gamblers	1,170	760	570
Problem gamblers	85	55	40
All 16+	**20,000**	**13,000**	**9,750**
Engaged gamblers 16+	2,980	1,935	1,455
Problem gamblers 16+	120	80	60
16-34 year olds	4,635	**3,010**	**2,260**
Engaged gamblers	585	380	285
Problem gamblers	40	25	20
All 16+	**10,000**	**6,500**	**4,875**
Engaged gamblers 16+	1,490	970	725
Problem gamblers 16+	60	40	30
16-34 year olds	2,315	**1,505**	**1,130**
Engaged gamblers	290	190	140
Problem gamblers	20	15	10

Estimating the number of engaged and problem gamblers who continue to participate in the study is difficult because membership of these groups will be dynamic. For example, someone identified as being a problem gambler at wave 1 may not be when interviewed at wave 2. However, we do not know the extent of churn within these groups – this is something the proposed longitudinal survey could estimate. For this illustration, we have assumed the rate of churn remains constant over time.

Table 1-3 also illustrates that a very large initial general population sample would be needed to generate enough problem gamblers to study over time.

1.2.5 Impact of sample size on detecting change in gambling behaviour

Table 1-4 illustrates differences that would be detectable (i.e. that are likely to be statistically significant) at different sample sizes, assuming the baseline figures were similar to those from the 2016 survey estimates, shown in Table 1-1. For example, if the wave 1 achieved sample was 10,000 and the prevalence of engaged gambling amongst those 16+ changed by 1.3% or more between waves 1 and 2 (from a baseline of around 15%), we would have 80% power to detect this change. If the achieved sample size were 40,000 a change in the prevalence of engaged gamblers aged 16+ of 0.65% between waves 1 and 2 could be detected with the same statistical power.

Table 1-4 Changes in gambling behaviour wave on wave that would be detectable at different sample sizes (power=0.8; alpha=0.05)

W1 achieved sample size =50,000	**W2 vs W1**
Engaged gamblers 16+	0.55%
Engaged gamblers 16-34	1.1%
Problem gamblers 16+	0.15%
Problem gamblers 16-34	0.15%
W1 achieved sample size =40,000	**W2 vs W1**
Engaged gamblers 16+	0.65%
Engaged gamblers 16-34	1.3%
Problem gamblers 16+	0.15%
Problem gamblers 16-34	0.2%
W1 achieved sample size =20,000	**W2 vs W1**
Engaged gamblers 16+	0.9%
Engaged gamblers 16-34	1.8%
Problem gamblers 16+	0.25%
Problem gamblers 16-34	0.25%
W1 achieved sample size =10,000	**W2 vs W1**
Engaged gamblers 16+	1.3%
Engaged gamblers 16-34	2.6%
Problem gamblers 16+	0.35%
Problem gamblers 16-34	0.4%

1.2.6 Topic coverage and interview length

The topics to be covered in a longitudinal survey of gambling will be determined by the research questions and the availability of existing information of relevance to those topics for sampled individuals. Determining the information needed to answer the research questions fell outside the scope of this project. However, to be able to assess the viability of different design options, information on interview length from three longitudinal studies of gambling is used as an indicator, with details provided below.

For comparative purposes, the average questionnaire length of the British Gambling Prevalence Survey (2010) was around 20 minutes.

- The QUINTE longitudinal study: duration of data collection varied between 30 mins to 1.5 hours per wave
- The Leisure, Lifestyle and Lifecycle project: duration of data collection per wave was between 2-3 hours.
- The Swedish Longitudinal Study of Gambling: average questionnaire duration for telephone interviews were 24 minutes (with a standard deviation of 10 minutes).

Depending on the scale and complexity of the question, and the mode of administration, a general rule of thumb is that you can ask 4-5 survey questions per minute. This can be used as a broad guide when helping to think through what level of content it is possible to administer within different approaches.

1.2.7 Frequency of data collection

An important design consideration for a longitudinal study is the frequency of follow up. Data collection should be frequent enough to pick up changes in behaviour and the factors associated with those changes. Typically, frequency of data collection is mediated by cost, respondent burden and attrition. The frequency of follow up for the four longitudinal studies of gambling reviewed were as follows:

- QUINTE – annual over 5 years
- Leisure, Lifestyle and Lifecycle project – 4 waves of data collection over 5 years
- The Victoria Longitudinal Study – 4 waves of data collection over 5 years
- The Swedish Longitudinal Study of Gambling – general population (epidemiological track) followed up 4 times over 7 years; in-depth track followed up subgroup of general population more frequently (3 times in 4 years)

2 Design options

This section describes each of the design options considered. These options are assessed in section 3. The options presented are the result of a two-stage review process. An initial set of options were generated by the NatCen research team and discussed with experts from NatCen Social Research in longitudinal survey design, gambling research, survey methodology and online panel design. They suggested additional options, refinements to existing options and considered the strengths and limitations of each design. A revised list of design options was then sent to six international experts in longitudinal survey design, some of whom worked on longitudinal studies of gambling. Experts were asked to review them and identify any missing options. No additional options were identified though design considerations and limitations were raised, which are reflected in the assessment of options presented in section 3.

2.1 Use existing data sources

2.1.1 Analysis of industry administrative data

Analysis of industry administrative data was considered. These data include membership data, for example for bingo clubs and casinos, and online gambling account holder data. The latter contains individual behavioural data, collected in real time. The kinds of data available include: date-stamped start and end time of each play session; the number of gambles on each activity; the total staked on each activity; the balance on player account at start/end of each session; the number of deposits and minimum/ maximum made during session; the number of bets placed during the session; and the use of responsible gambling tools during each session. These data are being used as part of a current study of Patterns of Play for GambleAware.

2.1.2 Analysis of existing survey data

Existing surveys that ask questions about gambling include the Health Survey for England, the Scottish Health Survey, the Welsh Problem Gambling Survey, and the 2015 sweep of the Millennium Cohort Study (MCS). Data on gambling participation and problem gambling have also been collected on the Avon Longitudinal Study of Parents and Children (ALSPAC) when young people were aged 17 years and 9 month and when they were 21. A third wave of data collection is planned for those aged 25 but data are not yet available. Apart from the MCS and ALSPAC, these surveys are cross sectional. Only ALSPAC provides longitudinal data but is geographically limited to a specific age cohort of young people living in the South West of England. The MCS has not included gambling questions in its current sweep. Questions asked about gambling in the Health Survey series ask about past year's participation in a range of activities, overall frequency of gambling across all activities and data about problem gambling according to two screening instruments: the DSM-IV-Multi Response and the PGSI. Only two questions about gambling were asked on the MCS. These were past week gambling and use of social media gambling games. Given the absence of existing longitudinal survey data on gambling in GB this option is not considered further.

2.2 Build on existing sources

2.2.1 Survey from industry administrative data

In addition to using administrative (i.e. online gambling industry account) data for analysis purposes, these data could also be used to draw a sample of gamblers, who could be invited to take part in a longitudinal survey. The administrative data would provide a means of identifying different types of gambler, for example, those involved in different gambling activities, at different levels of frequency and with different playing behaviours. The inclusion of offline and online data sources would broaden the coverage of gamblers that could be included in a survey. A similar approach was taken in 2014 and 2016 with the survey of customers who held a loyalty card for a bookmaker's and the one-year follow-up survey.

2.2.2 Add questions to existing longitudinal surveys

Adding questions to existing longitudinal surveys was considered. Three longitudinal studies were identified as potential vehicles for running new questions on gambling due to their population coverage.

The UK Household Longitudinal Study (Understanding Society) follows c.70,000 people living in private households in Great Britain and Northern Ireland. Interviews are carried out annually with adults aged 16 and over, with a short self-completion questionnaire completed by young people aged 10-15 years. The questionnaire includes questions on topics such as financial behaviour and attitudes, employment, health, psychological traits, social support and wealth and debts[2]. The study is now in its ninth wave and has to date not included any questions on gambling.

The Longitudinal Study of Young People in England (known as Next Steps) follows a cohort of Year 9 students (aged 14). Cohort members were first interviewed in 2004 and at the last round of fieldwork in 2015 around 16,000 were followed up. No questions on gambling have been included in previous rounds of Next Steps. At the time of writing it is unknown when the next sweep of Next Steps will take place.

The Millennium Cohort Study (MCS) follows young people born in Great Britain and Northern Ireland in 2000-01. Currently the study follows around 19,000 young people, with sweep 17 of data collection coming to an end now. As mentioned in section 2.1.2, MCS included a couple of questions on gambling in sweep 16 (2015).

2.2.3 Use existing survey(s) as a sampling frame

An existing survey or number of surveys could be used as a sampling frame for the first wave of a longitudinal study of gambling. Such an approach was used by the Longitudinal Study of Ageing (ELSA), which invited Health Survey for England (HSE) respondents who agreed to recontact and were aged 50 or over to take part in the new study. Three (recent) years of HSE were selected as the initial sampling frame to ensure the sample size was large enough. The sample has been refreshed at various points since the study started using HSE.

Under this option, those who agreed to recontact and, if applicable, to their details being passed to a third party would be invited to take part in the new study[3]. The size of sample, geographical and population coverage of existing surveys should be assessed

[2] See https://www.understandingsociety.ac.uk/sites/default/files/downloads/general/long-term-content-plan.pdf for more detail on questionnaire content

[3] This would apply in cases where the existing survey is conducted by a different contractor to the one undertaking the new longitudinal survey.

to identify potential options. For illustrative purposes, Table 2.1 shows the achieved sample sizes of three large scale, random probability surveys.

Table 2-1 Coverage and achieved sample sizes of three existing surveys

Survey	Coverage	Achieved sample size
Health Survey for England	England, adults 16+	c.8,000[a]
Family Resources Survey	GB & NI, adults 16+	c.30,000[a]
Labour Force Survey	GB & NI, adults 16+	c.38,000[b]

a Interviews with adults per annum

b Interviews with households per quarter

2.2.4 Use an existing online panel survey

An online or web access panel is a particular type of web survey in which people are specifically recruited to join a panel and are invited to complete regular online surveys. Online panels can be very large, boasting 100,000s of panellists in some cases (e.g. YouGov). Panels are typically omnibus surveys: space is sold by the question or minute and funders of questions have little or no control over other content in the rest of the questionnaire or the other kinds of surveys panellists are invited to participate in through the panel. There are several online panels that cover the general population in GB as well as panels that focus on specific groups such as black and minority ethnic groups (Ethnos panel) and young people (e.g. Youthsight). Many online commercial panels in GB use non-random probability sampling methods. An exception is the NatCen Panel, which is a mixed mode, web-telephone panel. The NatCen Panel recruits participants 18+ who have previously taken part in the British Social Attitudes survey, a random probability survey of adults in Great Britain[4]. The panel currently includes c8,000 individuals.

2.2.5 Sample from treatment services' records

There are currently three organisations funded by GambleAware to deliver treatment services to problem gamblers in Britain (with a fourth opening in Leeds in Summer 2019). These are GamCare, Gordon Moody and the National Problem Gambling Clinic, based within the Addictions Service at Central North West London NHS Trust. Between these three organisations, face to face treatment is delivered to c.8,000 people per year. As well as the National Problem Gambling Clinic, GamCare has a partner network of currently 15 treatment organisations across Great Britain providing counselling. Both GamCare and Gordon Moody offer/ are piloting online services. In addition, GamCare operates the National Problem Gambling Helpline, which received c.29,000 calls in 2017/18. Local support groups are run by Gamblers Anonymous UK and other smaller charities. It may be possible to recruit a cohort of people accessing these resources.

Gamblers may receive treatment through other addiction clinics or through generic counselling and support services. Using records to identify this other group would be much more difficult as it is unlikely they will be easily identifiable as gamblers.

2.3 Create new data source

[4] Jessop, C. (2017) Developing the NatCen Panel: August 2015-July2017. http://www.natcen.ac.uk/media/1484228/Developing-the-NatCen-Panel-V2.pdf

2.3.1 Create a new longitudinal survey representative of the general population

This option would involve the design of a new longitudinal survey that covers the general population of GB aged 16+. The sampling design could use random or non-random probability methods. Data collection could involve a single mode or a mixed mode design. If a random probability sampling design were chosen, the sampling frame would be the Postcode Address File (PAF) small user file, which is a list of mail delivery points that receive fewer than twenty pieces of mail a day (used as a proxy for addresses containing private households). Irrespective of the sampling method, adults would be screened to identify those in age groups of interest, engaged and problem gamblers. As engaged and problem gamblers are relatively rare, the sample size would need to be quite large.

2.3.2 Build bespoke opt-in panel survey

Rather than use an existing commercial panel, the Commission could contract the design, build and promotion of its own opt-in panel. An opt-in panel is one where individuals seek out the panel and sign up to take part. This type of panel offers rewards, typically financial, to encourage participation.

3 Assessment of design options

The design options considered in section 2 are assessed in this section in terms of methodological considerations, their compatibility with the proposed research questions and in terms of other factors.

3.1 Methodological strengths and weaknesses

3.1.1 Analysis of industry administrative data

The main strengths and limitations of analysing existing industry administrative data are summarised in Figure 3.1. This option would be very cost effective, as there would be no data collection costs.

Figure 3-1 Overview of strengths and limitations of industry administrative data

+	−
More cost-effective than surveys because of size, scope and fact that data have already been collected for operational purposes	Data would only contain gamblers with an account. Would not be able to generalise to the wider population, which may limit impact for policy and practice
Data on patterns of play continuously updated	Data may contain very limited contextual/ background information about individuals
Includes individuals who may not take part in surveys	Metadata (e.g. variable labels, value labels) may be missing/poor quality
	Success in being able to identify and link multiple accounts to an individual is unknown

As discussed in section 2.1.1 online gambling account data would provide detailed data on individuals' patterns of play. However, industry administrative data do not cover the whole population of interest and as such there would be significant coverage error. For example, the most practical industry administrative data to use would be account data generated from online gamblers' accounts. This remains the only data currently available which tracks all gambling activity on a particular company's online games for an individual within a single account. Equivalent data for land-based sectors does not exist. Whilst online gambling is the largest growth sector in Britain, data from the 2016 English and Scottish health surveys shows that only 17% of gamblers gamble online, the rest being conducted in land-based venues where tracked account data is not available. Furthermore, 94% of online gamblers also gambled at land-based venues in the past year. In addition, many online gamblers have more than one account and can frequently switch between providers. This kind of movement is difficult to track. Research which used this data alone would only represents a minority of gamblers and would not represent their full gambling repertoires.

Moreover, this data is likely to contain little contextual information, which limits the research questions that could be answered, refer to section 3.3. Without contextual data it would not be possible to identify particular groups of interest: i.e. problem gamblers and those in treatment/recovery. The quality of metadata may also pose issues for secondary analysis and lessons should be learned from the current GambleAware Patterns of Play project.

The identification and linking of online accounts belonging to the same individual will rely on each account containing common unique identifies, such as the person's name, address and date of birth. It was beyond the scope of this study to assess the quality of online gambling account administrative data or the feasibility of linking multiple accounts held by an individual. However, the GambleAware Patterns of Play project currently taking place will provide an indication of the willingness of gambling operators to make their data available and the quality of these data.

3.1.2 Survey from industry administrative data

Using industry administrative data as a sampling frame for a survey has several advantages, summarised in Figure 3.2. By linking individual's survey responses to their industry administrative data there is the potential to enhance the survey data with more detailed online play data. Individuals would need to consent to this and provide details of their online accounts.

Figure 3-2 Overview of strengths and limitations of surveying gamblers from industry administrative data

+	−
Cost effective means of identifying engaged gamblers	Sample would only contain gamblers with an account and as such sample coverage would be poor
With consent, could link survey and administrative data, providing detailed information on patterns of play	Would not be able to generalise to the wider population, which may limit impact for policy and practice
Information needed to answer research questions not captured in admin data can be collected in the survey	Quality issues: some variables may be of lesser importance to administrators (e.g. address information may not be updated); contact information may be missing or erroneous
	Willingness of industry to allow use of administrative data for a longitudinal survey is unknown, as is whether industry may place restrictions on the survey design, e.g. frequency of data collection, length of interview

The underlying assumption in data linkage is that an individual appears in both data sets: in the survey and the administrative record and that by using a unique identifier or

set of identifiers, an individual's record can be identified, and the data of interest appended to the survey data set.[5]

Linkage is achieved using unique identifiers, such as name and address. The data linkage process suffers from two stages of losses: not all survey respondents consent to data linkage; and not all survey responses of those who have given consent can be linked to their individual administrative record. There is also a risk that industry may not be willing to allow online gambling account data to be used to draw a sample for a longitudinal survey or may want to limit the burden such a study might place on its players, for example by limiting the length of questionnaire and or frequency of data collection.

As discussed in section 3.1.1, the biggest limitation with this design is population coverage. Sampling from industry administrative data would only include gamblers, and particular types of gambler (mainly online) at that.

3.1.3 Add questions to existing longitudinal surveys

Figure 3.3. summarises the strengths and limitations of adding question to existing longitudinal studies.

Figure 3-3 Overview of strengths and limitations of adding questions to existing longitudinal studies

+	−
Cost efficient means of obtaining a random probability, representative sample	Limited room for new questions
Rich in contextual data (i.e., full range of other questions asked)	Substantial gaps in questionnaire coverage
Gambling questions can be embedded into a study not specifically focused on gambling could reduce non-response bias	Loss of control – cannot guarantee questions remain or are repeated. Unlikely to get permission for qualitative follow-up/bio markers etc
	Frequency of data collection may not be suitable

Including questions on an existing longitudinal study has several benefits. It is a cost-effective way of collecting data, as the Commission would not need to set up its own longitudinal study. The Commission would only pay for the gambling questions yet would have access to a wealth of contextual data on individuals and their families for analysis purposes. Using an existing longitudinal study, with a known and trusted brand, might also reduce unit non-response. An existing longitudinal study will include questions on a range of topics, many of which participants will be familiar with. By not being specifically concerned with gambling the risk of unit non-response bias could be

[5] The data set provided to analysts does not contain the unique identifiers used for matching: this protects the anonymity of the survey respondents.

reduced. The sample size of Understanding Society would also provide a cost-effective way of identifying a sizeable sample of problem gamblers and those in treatment/ recovery. However, this would only be possible if the questions needed to identify these groups can be included in the Understanding Society questionnaire.

However, there are issues with using existing longitudinal studies. Currently questions on gambling are not included and so it is likely a significant number of new questions would need to be added to address the proposed research questions. For example, the recently developed Gambling Participation Instrument (which has similarities to the participation questions asked in the British Gambling Prevalence Survey 2010) contains over 30 items on gambling participation alone, without measurement of gambling problems or gambling harms. It is unlikely that the existing longitudinal surveys would have space to accommodate all the questions required as they typically have limited room for new questions. Moreover, even if questions could be accommodated in one wave there is no guarantee that there would be space for gambling questions in future waves of data collection because of the way in which existing studies are funded and their scientific priorities. This lack of control over questionnaire content presents a significant risk for the Commission.

One of the aims of a new longitudinal gambling study is to look at trajectories. The frequency of data collection is important. A further limitation of using existing longitudinal studies is that the frequency of data collection might not be appropriate for the Commission's needs.

3.1.4 Use existing survey(s) as a sampling frame

This option would provide a cost-efficient means of obtaining a random probability sample of the general population and, potentially, of key groups of interest, depending on the survey(s) used as the sampling source. Another benefit of this approach is that the characteristics of non-responders will be known, as the source survey will contain data on all those invited to take part in the gambling study. This information could be used to weight the wave one longitudinal data, to correct for non-response bias (Figure 3.4).

Figure 3-4 Overview of strength and limitations of using existing survey(s) as a sampling frame

+	−
Cost efficient means of identifying target population & subgroups if survey already includes questions on gambling behaviour	Survey sample size and % giving consent to recontact (by a third party) could limit size of wave 1 sample and of subgroups of interest, i.e. problem gamblers, those in treatment and recovery
(Potentially) reduced response burden as some information will have been collected in the earlier survey interview	Survey owners may not be willing to grant permission.
Provides information on characteristics of survey non-responders that can be used for weighting	Depending on the size of the existing survey & its geographic coverage multiple surveys/ survey years may need combining to create a sufficiently large sample. Combining different survey samples introduces complexity
	Selling a study focused on gambling to the general public may be difficult. Significant non-response bias could affect the accuracy of findings

A limitation with this option is that the source survey would need to be large if it were to yield sufficient numbers in key groups of interest, particularly problem gamblers and those in treatment and recovery. Survey owners generally only allow participants to be followed up once. The need for a large sample may deter survey owners from allowing it to be used as the sampling frame for a longitudinal study of gambling, particularly if gambling is not a priority topic for them. This may rule out using large surveys such as the Labour Force Survey and Family Resources Survey. If the English and Scottish health surveys were willing, several survey years could be combined to create a large enough sample, as was done for ELSA (refer to section 2.2.3). Moreover, as these surveys include questions on gambling behaviour already, this information could be used to oversample particular groups of interest, e.g. problem gamblers, but such a strategy would impact negatively on statistical power. A further consideration is that there is no Welsh Health Survey and gambling questions could not be included in the equivalent National Survey for Wales in 2016. Instead, data was collected via a non-probability omnibus study. Combining data from different surveys, with different sampling designs introduces additional complexity that can affect statistical power and the kinds of analyses that can be undertaken.

3.1.5 Use an existing online panel survey

Figure 3.5 summarises the strengths and limitation of using an existing (online) panel survey.

Figure 3-5 Overview of strengths and limitations of using an existing online panel survey

+	−
Cost efficient means of data collection	Limits on sample size may curtail some analyses and affect the precision and confidence in estimates
If use a random probability panel, a cost-efficient way of achieving good population coverage	Using an online panel may limit questionnaire length (to 30 mins or under)
	Non-probability panels miss certain groups of the population and favour others

Commercial web panels can be large and as such offer a cost-efficient means of data collection. However, panels typically limit the size of the sample available for any one study and the number of follow ups that can be undertaken. The sample coverage will be poor for this type of panel because it relies on people opting in and only uses online data collection. There are also restrictions on questionnaire length that could constrain any longitudinal study of gambling. An indication of the potential length of any one wave's questionnaire is provided in section 3.1.3. It is unlikely that a commercial online panel would be able to accommodate all the questions required to answer all three research questions. It might be feasible to break the questionnaire into chunks, which are run over several waves, but this would introduce differential non-response, as not all respondents will take part in each wave. This would have a detrimental impact on the utility of the data obtained. Whilst commercial panels include socio-demographic data, the information may have been collected a while ago. For the proposed longitudinal study of gambling details on, for example, individual's income and economic activity would need to be checked and updated at each wave. As such there is likely to be only a very modest saving on questionnaire space by using a panel.

There are also quality issues with some web access panels, particularly those using non-random probability samples. A review of web panels by the American Association for Public Opinion Research (AAPOR) concluded that the quality of online panels can vary greatly due to differences in sampling and refreshment and as a consequence estimates produced for the same variables across panels can vary considerably.[6] (AAPOR, 2010).

The NatCen panel is a high quality, random probability panel that includes those people missed by purely online panels. However, its size would limit the numbers of people in key groups of interest (e.g. problem gamblers) identified and the size of changes in behaviours observed over time, refer to sections 1.2.4 and 1.2.5.

[6] Baker, R., Blumberg, S.J., Brick, J.M. (2010) Research Synthesis: AAPOR Report on Online Panels. Public Opinion Quarterly Advanced Access. https://pprg.stanford.edu/wp-content/uploads/2010-AAPOR-Report-on-Online-Panels.pdf

3.1.6 Sample from treatment records

Gambling treatment records could provide a cost-efficient means of identifying gamblers in treatment and recovery. However, as with industry administrative data, using these records would limit sample coverage to a specific subset of the population of interest (i.e. those receiving treatment for gambling problems). The BGPS 2010 found that less than 0.5% of the population used a gambling help group/service/advisor or counsellor for gambling problems. Among those who felt that they had ever had a problem with gambling, the estimate was 9%. In short, most problem gamblers tend not to seek formalised help for their problems. Focusing on this group might also pose challenges for long term retention in a longitudinal study, as they may have more chaotic lives (see Figure 3.6).

Figure 3-6 Overview of strengths and limitations of sampling from clinic records

+	−
Cost efficient way to identify gamblers who are in treatment	Sample coverage issues as only available sampling frames are lists for gambling treatment services. Those using other services or receiving help from family/friends would not be covered
Could use records to oversample particular groups	Retention may be especially difficult, as this vulnerable group may have more chaotic lives
Scope to include other follow up elements, such as qualitative, bio-samples but only with ethical approval	Willingness of service providers to allow the Commission to sample from its patient lists is unknown

Treatment lists will not include subgroups of interest to the proposed study: people who have not received treatment; those not being treated through Gamble Aware treatment services; and individuals who are being treated but where gambling is not listed as the condition. Using treatment records as a sampling frame would also limit the extent to which all three research questions could be answered, see 3.3.

3.1.7 New, representative general population longitudinal survey

The strengths and limitations of a new, representative general population longitudinal survey are summarised in Figure 3.7. This option affords the Commission control over all aspects of the survey design and the flexibility to design a study to meet all its data needs. However, this control and flexibility would come at a cost. An important design issue for this option, and for the options where a sample is recruited from participants to an existing survey or the creation of a new opt-in panel, is how the new survey would be marketed to participants, to minimise non-response bias. Some gambling studies have marketed themselves as studies of health and leisure to address this issue.

Figure 3-7 Overview of strengths and limitations of a new, representative general population longitudinal survey

+	−
Control over survey design (population coverage, sample size, questionnaire length & content, frequency & mode of data collection)	May be prohibitively expensive
Obtain random probability, representative sample of population in Great Britain	Large sample size needed to pick up enough problem gamblers and those in treatment/recovery
Scope to include other follow up elements, such as qualitative, bio-samples	Recruitment and retention into a new bespoke longitudinal study of gambling may be difficult. Significant non-response bias could affect the accuracy of findings
	Frequency of data collection will involve trading off cost, logistical issues depending on sample size, attrition with research questions and ability to detect change

One way to reduce costs would be to prioritise certain subgroups for follow up, such as those at risk of harm from gambling. In this scenario, high priority groups would be given a greater chance of selection for follow up than other groups: such an approach was used on the Canadian QUINTE longitudinal study. This would reduce the size of the overall sample followed up at subsequent waves, and so reduce costs. However, this approach would introduce complexity at the analysis stage, as the differential selection weights used at subsequent waves would need to be considered.

3.1.8 Build bespoke opt-in panel

Rather than using an existing online panel, the Commission could build its own. This would afford the Commission control over the study design, particularly questionnaire content and frequency of data collection. It would also be cheaper than a new representative survey (see section 3.1.7) or following up respondents to existing surveys (see section 3.1.4) as recruitment methods would be less intensive e.g. advertising for panel members online and through venues rather than directly approaching all sample members. However, the use of such recruitment methods means sample coverage of the general population, and of subgroups of interest would be an issue due to the opt-in sampling methods used. These strengths and weaknesses are summarised in

Figure 3-8.

Figure 3-8 Overview of the strengths and limitations of building a bespoke opt-in panel

+	−
Cost efficient as can recruit from a range of sources, including gambling venues, local communities etc.	Not representative
Can focus recruitment on particular groups of interest	Sample coverage issues, with hard to reach or marginalised groups not included
More control over study design (questionnaire content, sample sizes, qualitative follow-up etc).	Recruitment and retention into a new bespoke longitudinal study of gambling may be difficult
Scope to include other follow up elements, such as qualitative, bio-samples	If an online panel, limit on questionnaire length, particularly at the first wave

3.1.9 Mixed method/multi-cohort approaches

The eight design options assessed so far all have pros and cons. A range of different sampling and design methods are likely to be needed to answer all the research questions of interest (see section 3.3). Other longitudinal studies of gambling have recognised this and developed a mixed method, and in some cases, mixed sample frame (multi-cohort) approach. For example, the Victorian Longitudinal Study of Gambling used a population-based sampling frame as the basis of its design but also included in-depth qualitative work to supplement the insights from the survey findings. The Swedish Longitudinal Study of Gambling included three different sample types: a main track; an in-depth track where a sub-group of people were followed-up more regularly and a follow-up group recruited from a previous study. The most recent national study of Canadian gambling includes a general population random probability prevalence survey, a survey of people seeking treatment and a longitudinal study with people recruited from an online panel to look at engaged gamblers over time. Results from the online panel will be benchmarked against the national random probability survey.

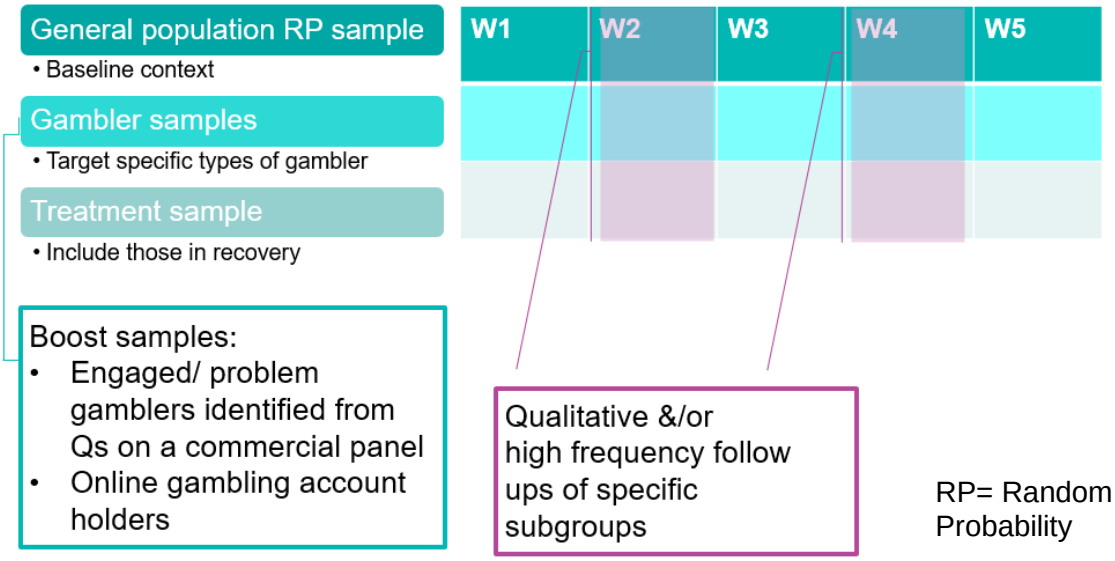
Notably, all three of these examples include a national, random probability survey as their core vehicle for answering the research questions and then employ other methods to look at sub-groups of interest. The main benefit of this approach is that the core, nationally representative data, can be used to benchmark and contextualise data generated using other sampling frames (in the Canadian example, they are planning on using the random probability data to adjust their weighting strategy for their online panel). Each strand is then used to answer specific research questions. The main limitation of this approach is that greater resources are required to fund several different studies and to synthesis results within and between strands.

In Britain, qualitative longitudinal methods have been used for a study of gamblers in Glasgow and examined gambling career trajectories over a five-year period. This study has been repeated in Denmark, with both studies demonstrating that stasis in gambling behaviour is not the norm and situating changes in gambling within people's broader lives and contexts. Other qualitative methods, such as life history methods, have been successfully used in other fields (such as substance misuse) to explore changes in behaviours from a life-course perspective. Any consideration of a longitudinal study of

gambling should consider which methods are most likely to produce meaningful insight for each research question. This includes considering the use of qualitative approaches.

A mixed methods approach could also include the use diaries/Apps to collect behavioural data in more detail and or at greater frequency from subgroups of interest. These subgroups might include, for example individuals suffering harm from gambling or at risk of suffering harm, or individuals where a change in the kinds of gambling products played has been observed. Figure 3-9 illustrates how the different methods/ cohorts might be combined.

Figure 3-9 Illustration of a mixed method, multi-cohort longitudinal gambling study design



This type of design has several strength and limitations, which are summarised in

Figure 3-10.

Figure 3-10 Overview of strengths and limitations of a mixed method, multi-cohort design

+	−
Random probability general population survey can be used to benchmark and contextualise data using other sampling frames	More costly, as multiple studies are combined
Other sampling sources can be used to look at subgroups of interest, e.g. problem gamblers, those in treatment/recovery	Increased complexity in terms of design, data collection, management and analysis
Qualitative methods can be used to provide deeper insights/ exploration of research questions	Greater respondent burden, which would require careful management
High-frequency data collection methods could be used to follow up groups of individuals of interest to look at changes in more detail	Recruitment and retention into a new bespoke longitudinal study of gambling may be difficult

3.2 Summary of strengths and weaknesses of research designs

Nine options have been identified and the strengths and weakness of each considered in section 3.1. These strengths and weaknesses relate to how well or otherwise each design:

- covers the target population i.e. the general population, highly engaged gamblers, problem gamblers and those in treatment and recovery
- generates data that allows the Commission's proposed research questions to be answered. This point relates to the availability of existing information/ the ability to collect information needed to answer the research questions (e.g. through a survey questionnaire) and the ability to detect changes in gambling behaviour (e.g. through frequent and/or in-depth data collection)
- other considerations, such as ethics, logistics and cost

summarises the extent to which each design option covers the target population (general population, engaged gamblers, problem gamblers and those in treatment/recovery), provides information on the topics of interest, and collects data frequently enough to detect changes in behaviour. A traffic light visual representation is used to indicate how well each design option fairs against each methodological criterion. It is acknowledged that within each option there will be many further design considerations and choices that could strengthen its design. However, for simplicity, a high-level comparison is made of the designs, with the strongest variant of a design being compared.

In terms of selection of the strongest variant, this applies to design option 3 (adding questions to an existing longitudinal survey). Understanding Society is assumed to be the existing survey used, as it covers all Great Britain and has a large sample size. It also applies to design option 5 (adding question to an existing panel), where use of the NatCen panel is assumed, as it uses a random probability mixed mode panel that will reduce sample coverage bias compared with non-random online only panel designs.

Table 3-1 Summary of methodological strengths and weaknesses of design options

Design options / Methodological criteria	1. Analyse industry admin data	2. Sample from industry admin records	3. Add Qs to existing longitudinal survey	4. Existing survey as sampling frame	5. Existing random probability panel	6. Treatment records	7. New, representative longitudinal survey	8. Bespoke opt-in panel	9. Mixed method/multi cohort
Coverage of general population	Poor	Poor	Good	Good	Good	Poor	Good	Poor	Good
Coverage of engaged gamblers	Poor	Poor/Medium	Good	Medium	Medium	Poor/Medium	Medium	Poor	Good
Coverage of problem gamblers	?	Poor/Medium	Good	Medium	Poor/Medium	Poor/Medium	Medium	Poor	Good
Coverage of those in treatment/ recovery	?	Poor/Medium	Poor/Medium	Poor/Medium	Poor	Good	Poor/Medium	Poor	Good
Coverage of topics/items to answer research Qs	Poor	Medium	Poor/Medium	Good	Medium	Poor	Good	Good	Good
Frequency of data collection to detect change in behaviour	Good	Medium	Poor/Medium	Medium	Good	Medium	Medium	Good	Good

Poor | Medium | Good | ? Unknown

The coverage of problem gamblers and those in treatment/recovery in industry administrative data (online gambling account holder data, refer to section 3.1.1) is unknown as the administrative data do not contain information that would allow the Commission to identify these groups. It would be possible to identify engaged gamblers from the data, looking at the frequency with which individuals play online games. However, no information on their play with other forms of gambling product would be available, and as such coverage would be poor.

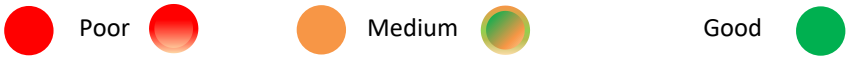
Of the nine designs considered, the mixed method/multi cohort design (described in section 3.1.9) has the greatest potential to identify and understand changes in behaviour through targeted use of high-frequency follow ups, as it offers the best population and topic coverage.

3.3 Compatibility with the research questions

Table 3-2 shows the compatibility of each design option with each proposed research question. The mixed method, multi-cohort design provides the best compatibility with the proposed research questions because of the methodological advantages of this design, discussed in section 3.2.

Table 3-2 Compatibility of design options with proposed research questions

Design option	1. Gambling trajectories, pathways into/ out of gambling, problem gambling, treatment & recovery	2. Relationship between gambling products & behaviours over time	3. Consequences /harms & how these change over time
1. Analysis of industry administrative data			
2. Survey from industry administrative data			
3. Add questions to existing longitudinal surveys			
4. Use existing surveys as a sampling frame			
5. Use an existing random probability panel			
6. Sample from treatment records			
7. New representative longitudinal survey			
8. Build bespoke opt in panel			
9. Mixed methods approach			



3.3.1 Gambling trajectories

To be able to look at gambling trajectories and pathways into and out of gambling a longitudinal study would need to include a sample of the general population and to obtain information on their gambling behaviour and circumstances over time. Analysis of industry administrative data (option 1), as discussed in section 3.1.1, would not provide information on the general population and measures of gambling behaviour. As such this design would not provide data on pathways into and out of gambling or on gambling trajectories.

The use of industry administrative data as a sampling frame for a survey (option 2), would provide the opportunity to collect information about gambling behaviour, but only from those with an online gambling account, as discussed in section 3.1.1. This design would not provide information on pathways into gambling as the administrative data sampling source excludes non-gamblers. It might be possible to identify some pathways into and out of problem gambling, treatment and recovery using option 2 but

because the administrative data only cover online gamblers, who are a small proportion of all gamblers, its likely that other pathways into and out of problem gambling will be missed.

Pathways into and out of treatment could be explored by sampling individuals from treatment records (option 6). However, adopting such a design would limit the study to only those individuals that seek treatment and support through specific centres (see section 2.2.5). Those using alternative services, perhaps for comorbid conditions such as alcohol or drug dependency and those who make use of self-help or the support of family and friends would be excluded. This may be acceptable, depending on what the Commission want to use the data for. Moreover, adopting this option solely would not provide information on pathways into gambling and provide limited data on pathways into problem gambling, as those who are not in treatment/recovery would not be included.

The remaining design options (3, 4, 5, 7, 8 and 9) would all be able to provide data on the general population. However, the success of options 3 and 5 (adding questions to an existing longitudinal study or random probability panel) in providing data on trajectories would depend on the number of questions that could be included to measure individuals' gambling behaviours and the size of the sample. As illustrated in sections 1.2.4 and 1.2.5, the size of the achieved sample has an impact on both the numbers of problem gamblers identified and the confidence with which changes in behaviour can be detected. The extent to which pathways into and out of problem gambling, treatment and recovery can be explored, therefore, will be limited by the sample size.

There is a risk with option 8 (building a bespoke opt-in panel) that the opt-in recruitment process results in certain kinds of individuals being omitted from the study, whose gambling trajectories or pathways may be different to those who opt into the panel study. This systematic bias would limit the ability of this design to fully answer the research question on trajectories and pathways.

Options 4 and 7 (sampling off the back of an existing survey or setting up a new longitudinal study that is representative of the general population) reduce the risk of systematic bias in the composition of the sample associated with a building a bespoke opt in panel (option 8). Moreover, options 4 and 7 afford greater control over questionnaire content than options 3 and 5, so minimising the risk that insufficient data are collected to be able to identify trajectories and pathways into and out of gambling. However, sample size for options 4 and 7 may limit the ability to look in detail at pathways into and out of problem gambling, treatment and recovery, as might the frequency of data collection.

A mixed methods approach (option 9) would provide the potential to explore trajectories and pathways not only among the general population but also for particular subgroups, such as those receiving treatment identified by selecting a separate cohort from treatment records. Pathways into and out of problem gambling could be explored through selecting an additional sample from industry online gambling data, for example and or by identifying and tracking problem gamblers through an online panel. Qualitative methods could explore in more detail pathways and trajectories that would not be possible using survey methods. For these reasons option 9 offers the greatest potential for understanding gambling trajectories and pathways.

3.3.2 Relationship between gambling products and behaviours over time

Design options 1-9 offer varying potential to generate data that could be used to explore the relationship between gambling products and behaviours over time. Design option 1 (analysis of industry administrative data) offers detailed near real-time data on patterns of play for online account holders (see section 2.1.1). However, to answer the research question data are need on individuals' gambling behaviour in the round, including their use of land-based products as well as on-line products. Option 1 does not include data on land-based products. Design option 2 would allow for collection of information on use of land-based products but only for those with online gambling accounts. Those gamblers who do not have online gambling accounts would be excluded. Design option 6 would only provide data for those in treatment/recovery, who may no longer play or have very different playing behaviours and use of gambling products to the wider population.

Options 3-5, 7 and 8 would involve asking questions about use of gambling products. Constraints on questionnaire length for options 3 and 5 (discussed in sections 3.1.3 and 3.1.5) may limit the data available for analysis (i.e. the level of detailed information collected about play behaviour for different products). Asking survey questions about play involves respondents having to recall their gambling behaviour over a specified period (the reference period). Recall can be affected by the length of the reference period and the frequency with which the respondent engages in the behaviour. Make the reference period too long, and respondents may forget episodes or not recall the detail required by the question. Make the reference period too short, and less frequent behaviours may not be captured and the pattern of play or changes in the pattern may not be visible. One strategy would be to collect data more frequently, using a shorter reference period. However, this may not be practical, particularly for options 3 and 5, where the Commission would not have control over the design of the data collection vehicle.

As discussed in section 3.1.9, option 9 offers the potential for high frequency bursts of data collection among groups of interest, and the potential to use multiple methods, such as diaries and qualitative methods, as well as asking survey questions. Linking individual's survey responses to their online gambling account data would provide additional data on patterns of play for this group. Using a mixed method approach would allow for a more in-depth exploration of the relationship between gambling products and behaviours over time.

3.3.3 Consequences and harms of gambling over time

In looking at consequences and harms of gambling over time, the Commission will need to decide how these should be defined. For example, is the Commission interested only in the consequences and harms of gambling on the respondent over time, or also on their family? These definitions will need to be operationalised by those undertaking data collection/analysis. Information on the use of responsible gambling tools during each session of online play is included in online gambling industry administrative data (options 1), but without additional information on individuals' circumstances and how these change over time, it is not enough to answer the research question.

As discussed in sections 3.3.1 and 3.1.6 options 2 and 6 would only provide data for subgroups – those sampled from industry administrative data for online gambling account holders and from gambling treatment/recovery services. However, without information about the wider population, specifically non-gamblers, it will not possible to assess the consequences and harms of gambling, as there would not be a

counterfactual. Options 3-5 and 7-9 would all include non-gamblers, though limitations in sample coverage for option 8 (opt in panel) and questionnaire content for options 3 and 5 (include questions on existing longitudinal survey or random probability panel) may limit the extent to which this research question could be addressed. The Commission would have greater control over questionnaire content with Options 4 and 6 (use existing survey as a sampling frame or set up a new, representative longitudinal survey), and as such there is potential to collect a wider range / more detail about consequences and harms of gambling. Option 9 (mixed methods) offers the greatest potential to explore this research question however, with the use of a multi-cohort design to explore consequences and harms for particular groups, such as problem gamblers, with different trajectories and pathways. The use of qualitative methods is likely to provide greater depth of understanding of harms and consequences of gambling, and the opportunity to include multiple perspectives (the respondent, different family members/friends etc) over time, if of interest to the Commission.

3.4 Costs

Indicative costs for three design options, agreed with the Commission were generated by NatCen Social Research to assist the Commission in assessing the extent of financial support required for a longitudinal study of gambling, and to indicate the magnitude of the likely difference in cost between the different options. The three options costed were:

- The first wave of a new, representative general population longitudinal survey (option 8).
- A follow up to an existing survey (option 4), involving 3 waves of fieldwork
- Using an existing online panel (option 5), collecting 5 waves of data

Indicative costs for these three options are shown in Table 3-3, and include VAT. In generating these costs, NatCen used a limited set of assumptions and its own costing tools and procedures. The costing assumptions were as follows:

- First wave of a new, random probability general population longitudinal survey option with an achieved sample size of 50,000, and an average face-to-face interview length of one hour.
- A follow up to an existing survey (the Family Resources Survey is assumed), involving a web-telephone survey. At wave one the responding sample is assumed to be 10,690 and costs include 2 further rounds of web-telephone data collection with a similar length interview.
- Use of an existing online (non-random probability) panel, with an assumed questionnaire length of 30 minutes. The achieved sample size at wave 1 is assumed to be 40,000. Costs include four further rounds of data collection with a similar length questionnaire.

Ball park costs for the use of a non-random online panel were obtained from a commercial online panel provider, and additional research costs were estimated by NatCen.

Table 3-3 Comparison of indicative costs for different design options

New, random probability general population 60 minute interview 50,000 respondents F2F interview at W1 1 wave	Follow up of existing survey (FRS) 30 minute interview 10,690 respondents W1 Mixed mode Web/CATI 3 waves	Existing online (non-random) panel 30 minute interview 40,000 respondents at W1 5 waves
~£6-8 million	~£1.5-1.8 million	~£0.5-1 million

Costs include estimates for:

- Researcher time to develop the sampling strategy, data collection instruments, manage fieldwork, check data, undertake descriptive analysis and produce a summary report;
- Fieldwork, including incentives
- Scripting
- Data management
- Printing and postage (where appropriate)
- Survey management

The main drivers of survey costs are number of interviews, interview length, mode of data collection and incentive strategy. A large sample, relatively long interview, carried out face to face would be very expensive, whereas using an existing online (non-random) panel is relatively inexpensive, but has several methodological limitations, see section 0. Following up participants to an existing survey, who have given their consent to recontact (by a third party) using cheaper data collections methods (e.g. web and telephone) can offer a cost-effective way of undertaking a random probability study. Response rates to a mixed mode (web-telephone) longitudinal study at wave 1 are likely to be lower than if the first wave were conducted face-to-face. It is also worth noting that evidence from CLOSER longitudinal studies on the effect of introducing mixed-mode designs on single mode surveys is mixed. 'Some have found that using mixed modes reduces attrition, others that it increases attrition compared to unimode designs. Most have found no support for the expectation that mixed modes reduce nonresponse bias, by bringing in respondents who would not participate in the original unimode design.'[7]

The use of incentives has been shown to boost response rates and reduce attrition rates over time[8]. It is recommended that the Commission ask tenders to propose an incentive strategy that seeks to provide value for money. Some experimentation with the value and nature of incentives in early waves of the longitudinal study maybe desirable, particularly to identify the most cost-effective strategy for retaining key groups of interest.

If qualitative fieldwork were to be undertaken, then an indicative cost per face-to-face depth interview of £1200-£1500 should be assumed. This cost includes design, data collection, travel expenses, incentives, data management, analysis and reporting.

[7] Jäckle, A., Gaia, A. and Benzeval, M. (2017). Mixing modes and measurement methods in longitudinal studies. London, UK: UCL Institute of Education: p11 http://openaccess.city.ac.uk/18597/1/Mixing-modes-and-measurement-methods-in-longitudinal-studies.pdf

[8] See for example Gaia, A (2017) The effect of respondent incentives on panel attrition in a sequential mixed-mode design. Understanding Society Working Paper Series 2017-03 https://www.understandingsociety.ac.uk/research/publications/524253

4 Recommended approach

Our recommendations to the Gambling Commission are three-fold:

1) That to look at pathways into and out of gambling a general population sample is needed;
2) To address the other research questions of interest to the Commission, additional studies drawing on samples of people from other, non-representative sources (like treatment providers or from industry account records), are also needed;
3) Qualitative data collection, as well as quantitative data collection, should be considered to give greater insight into the relationships and patterns observed.

We explain our rational for these recommendations in more detail in the rest of this chapter.

To look at pathways into gambling robustly, giving high-quality data that policy makers can trust, a sample of the general population is needed. Whilst other options are available, for example, recruiting from an existing online panel or developing a bespoke online panel, these have significant limitations in our ability to draw generalisable conclusions about pathways into and out of gambling. Other research has shown that online panels routinely over-estimate gambling behaviours and gambling harms, meaning there is likely to be considerable bias when using this as a sampling frame. This issue is considered sufficiently problematic that the new National Study of Gambling in Canada is only using an online panel for longitudinal data collection in the context of also collecting general population data using a random probability sample against which to benchmark and consider the online panel results. Notably, all six of the experts consulted on design options recommended using a general population sample, if it was possible.

A general population study could also provide data on the relationship between gambling products and gambling behaviour trajectories over time, and the consequences/harms of gambling and how these may change over time. A random probability sample of the general population would allow statistical tests to be applied correctly and would minimise coverage error. We recommend this requirement be included in any tender specification.

However, such a design would be expensive, especially if this was used as the sole vehicle for data collection as sample sizes would need to be prohibitively large to include enough problem gamblers to track over time. These issues would be even more acute if analysis of problem gamblers in treatment was also a requirement. Rather than have one cohort, a multi-cohort design could be used. Such a design could involve three separate samples:

- general population sample – sampled either from an existing survey or from PAF
- gambler sample – sampled from loyalty cards/online account holders
- treatment sample – sampled from clinics

Each sample would address different elements of the research questions and each sample would enable the Commission to look at different types of gambler and problem gambler. The studies combined would enable the Commission to tailor specific research content to the most appropriate population group and the inclusion of the general population sample would contextualise findings from the other studies. In addition, each sample type could have its own design and schedule of data collection reflecting what is most appropriate for the questions it is trying to answer. For example, for very engaged gamblers, current evidence suggests that behaviour is fluid and it is not clear that collecting data from these people once a year (or some other regular interval) will capture dynamic change in behaviour that may occur over the short term.

It may be likely that, for a sub-sample of engaged gamblers, the Commission may want to monitor changing behaviours more regularly over a shorter time scale, especially if the Commission are looking to capture and understand the relationship between changing products and the experience of harms. This could be done both quantitatively, using for example, diary methods and more dynamic methods of data capture or in-depth using either prospective or retrospective qualitative methods (or both). The linking of individual's online gambling account data to their survey responses would provide detailed data with which to look at the relationship between online gambling products and gambling behaviour over time.

The frequency of data collection is important. Typically, longitudinal gambling studies follow up survey participants every year or so. Other longitudinal studies, such as Understanding Society use a similar interval. More frequent data collection at scale is likely to be impractical, costly and potentially undesirable, as it could exacerbate drop out. However, when looking at pathways into and out of problem gambling, treatment and recovery it could be argued that high frequency follow-ups, perhaps every few days or weeks would be desirable to prospectively collect data on changes in behaviour. We recommend tenderers be asked to provide proposals for high frequency follow ups of sub groups, providing details on: the criteria they might use to select individuals for such a follow up; the kinds of data that would be collected and how; the length and frequency of follow ups; and what steps they would take to limit the impact of high frequency data collection on individual's on-going participation in the study.

In addition to survey data collection, targeted qualitative follow ups should be used to provide more in-depth exploration of issues throughout the life of the panel. We recommend tenders be asked to provide proposals for targeted qualitative follow ups and to consider which qualitative methods would be most appropriate.

Appendix A. Findings from online consultation

A.1 Background

The Commission commissioned NatCen Social Research and Dr Heather Wardle to advise on the different methods and approaches that can help to better understand changes in gambling behaviour over time, including the determinants of change. The first phase of this study involved an online consultation with government, academic and voluntary sector that sought views on the research questions such a longitudinal study should address and the target population of such a study. Stakeholders were identified as individuals and organisations who were known to Dr Wardle and or the Commission as having an active interest in gambling behaviour and research in this area. 141 stakeholders were invited by email to click on a link to a short (7 question) online questionnaire. The consultation was open between 21-28th February 2019.

A.2 Who responded

A total of 37 responses were received, however 5 of these were found to be void (random text was entered in response to each question). It is possible that these voids were people wanting to see t the consultation questions before going on to complete the questionnaire. Voids were received from 1 academic, 1 policy-maker/civil servant and 3 others. In addition, 1 partial response was received from an academic. This has been included, and the rest of this paper includes response from 32 individuals. The type of responder is shown in Table 1. Note that 4 responders described themselves as belong to more than one group: three were academics who also wore other hats, and one was a non-academic researcher who was also involved in a gambling charity.

Table A.1 Type of responder

Type of responder (multi-coded)	Only	Also	Also	Also	Row Total
1. Academic researcher	13	1	1	1	16
2. Non-academic researcher	2	1	1		4
3. Regulator	0				0
4. Policy maker/civil servant	3				3
5. Faith group member	0				0
6. Member of gambling-related charity/advocacy group	5		1		6
7. Member of non-gambling related charity/advocacy group	2				2
8. Member of the public	0		1		1
9. Other	4		1	1	6
Total Responders	29	1	1	1	32

A.3 Research questions study should focus on

Responses to this question were varied but broadly related to the following five themes

a) Trajectories of gambling over time, including trajectories of problem gambling and recovery

Suggestions included:

- Looking in-depth at people's gambling behaviours and how these change over time: changes in what games they are playing, what modes of access they are using, frequency, expenditure etc
- Establishing the incidence of problem gambling (new cases)
- Looking at how stable (or not) problem gambling rates are (churn rates)
- Looking at early onset of gambling and first gambling experiences
- Motivational changes
- Looking at recovery pathways among problem gamblers

b) Understanding the risk and protective factors for problem gambling

Suggestions included:

- Examining the actions of industry, especially advertising and marketing
- The role of peers
- Family contexts
- Personality traits
- Genetics
- Regulatory changes and environment
- Life events
- Experience of stress

c) Understanding the consequences of gambling

Suggestions included:

- Financial harms
- Impact on wellbeing
- Impact on mental and physical health
- Neglect of other roles and responsibilities
- Neglect of family

d) Understanding the relationship between gambling and other comorbid conditions (especially for problem gamblers)

Specific suggestions included:

- Exploring the relationships between gambling, anxiety, depression, substance abuse and misuse and the trajectories of each as they relate to gambling behaviour change.

e) Understanding the relationship between product type and gambling behaviour change
   - Suggestions included focus on different types of games played and also mode of access.

F) Other

Other suggestions included: generating data to help create safer gambling guidelines; exploring the impact on affected others; understanding the benefits of gambling as well as the costs; how people feel about their gambling and using the study to look at impact of regulatory change or prevention activities.

A.4 Rationale for proposed research questions

Whilst there were a number of viewpoints given (see below) there were three common themes in respondent's rationale for their proposed research questions.

The first theme was simply the lack of knowledge that we currently have in Great Britain about the 'natural history' of gambling, the stability or otherwise of gambling behaviour and our limited understanding of gambling and gambling-related behaviour.

The second theme, related to this, highlighted how it is essential to better understand trajectories of behaviour to inform policy, regulatory and preventative activity. As noted by a couple of participants, this understanding could have implications for resource allocation - for example, if problem gambling is found to be relatively stable state it might suggest that more resource is placed in treatment whereas if there is a high degree of churn, then more resource should be placed on prevention. Using the insight generated from a longitudinal study to help shape prevention activities and to better understand 'what works' was noted by some.

A third commonly mentioned aspect was the need to understand how new technologies and certain gambling products are related to changes in gambling behaviour as this has implications for a) regulatory policy and b) thinking through implications of new societal/technological development on gambling.

Several other aspects were also mentioned:

- Better understanding of the pathways to addiction, especially with comorbid conditions, as this has implications for treatment.
- The need to better understand recovery pathways among problem gamblers, as recovery is a new concept and not yet well understood in gambling.
- To understand the effectiveness of treatment and to help inform what treatment and support is needed.
- International context – it was suggested that any study link up with other international longitudinal studies of gambling (via an advisory group) to ensure consistency/sharing knowledge etc.
- The importance of knowing the impact of prevention campaigns on public health.
- To better understand the role of life stressors, including minor stressors, on gambling behaviour.

A.5 Priority research questions

Respondents clearly found it difficult to identify a single research priority. However, the most common priorities mentioned were (all mentioned by five or more respondents):

- Understanding how problem gambling develops and changes (including exploring the causes of problem gambling)
- Understanding factors in recovery and/or relapse from treatment
- Which types of gambling products are most harmful?
- What are the consequences/harms of gambling?

The following were mentioned by three or fewer respondents:

- Why do people start gambling?
- Do the harms outweigh the benefits?
- What are the costs of gambling-related harms?
- What are the causes of gambling-related harms (as oppose to the causes of problem gambling)?
- What is the relationship between gambling and mental health?
- How to prevent harm – what is most effective?
- What is the relationship between risk and protective factors and gambling behaviour (i.e., life events, experience of ACES etc)?
- How effective is treatment?
- What is the impact of prevention on harms?

A.6 What population should the study cover and why

Views differed on whether a longitudinal study should cover the whole population or focus solely on specific groups, reflecting stakeholders' views on what research question a longitudinal study should focus on. Among those who thought the study should focus solely on specific subgroups, the groups mentioned were:

- Gamblers at different points in their gambling 'career', such as those in treatment, in lapse or relapse
- Children and young people

These groups were not necessarily mutually exclusive.

Those who suggested the study cover the whole population ideally wanted the sample to be large enough to allow subgroup analysis of key groups of interest. Subgroups mentioned were:

- Young people
- Other groups identified as at risk under the 2005 Gambling Act (e.g. older people, migrants)
- Problem gamblers
- Individuals in recovery
- People with comorbidities/health conditions related to addiction, substance misuse, mental health problems and disabilities
- Groups under-represented in other studies, such as the families and friends of people with problem gambling behaviours
- Ethnic minorities

There was an acknowledgement among some stakeholders that the sample sizes needed would need to be large and that this might prove difficult to achieve.

A.7 Priority population groups

Respondents clearly found it difficult to identify a single population group. The priorities groups most commonly mentioned were problem gamblers and young people. However, there were differences of opinion, as discussed in section A.6, about whether a longitudinal study should focus on one or both groups or include them as part of a wider general population study. The decision depends on what the purpose of the study should be.