SECONDARY DATA ANALYSIS REPORT

CAUSAL INFERENCE METHODS IN GAMBLING RESEARCH

Richard J. E. James
School of Psychology, University of Nottingham, UK

Hyungseo Kim
School of Psychology, University of Nottingham, UK

Lucy Hitcham
School of Psychology, University of Nottingham, UK

Richard J. Tunney
School of Psychology, Aston University, UK

Suggested citation:

James, R.J.E., Kim, H., Hitcham, L., & Tunney, R.J. (2023) Causal inference methods in gambling research. Report prepared for Greo Evidence Insights (Greo), Guelph, Ontario, Canada. https://doi.org/10.33684/2024.004

Funding source:

The funding for this study came from Greo through a grant for secondary data analysis using gambling harm data in Great Britain. Funding for these grants was derived from regulatory settlements in lieu of financial penalties from gambling operators in Great Britain. The Gambling Commission reviews and approves proposals for destinations of regulatory settlements. Under the terms of a settlement, there is to be no publicity or benefit for the operator in connection with the regulatory settlement.

Greo supported the National Strategy to Reduce Gambling Harms in Great Britain as a neutral, arm's-length organisation that managed the solicitation, review, funding, and project management of safer gambling initiatives. Greo was not involved in the ideas presented herein and imposed no constraints on publishing.

Corresponding author: Richard J. E. James, Richard.James4@nottingham.ac.uk

ABSTRACT

The purpose of this project was to review and utilise methods from other disciplines in the social sciences in order to be able make stronger causal claims using cross-sectional gambling data such as gambling prevalence studies. We focused on the question of whether there is a causal relationship between specific gambling products and individual gambling harms, specifically problem gambling. There has been an existing literature that has looked at this issue, but fails to control for selection biases on engagement with specific gambling behaviours. We reviewed and used three approaches: propensity score matching, coarsened exact matching, and sample selection modelling to account for this limitation. Then, we applied them to examine the relationship between 8 types of gambling activity (online betting, offline betting, pools, scratchcards, lotteries, pools, offline bingo, and slot machines/FOBTs) on problem gambling in 9 British gambling prevalence surveys: the Health Survey for England in 2012, 2015, 2016 and 2016; the Scottish Health Survey in 2012, 2015, and 2016, and the British Gambling Prevalence Study in 2007 and 2010. The results showed most gambling activities were, unsurprisingly, associated with an increased risk of addictive behaviour. However, there was a clear gradient of risk. For some gambling activities (i.e., online gambling, slots), the effect sizes were substantially higher than others (i.e., lotteries). The modelling also highlighted covariates that exert strong effects on both the IV and DV of interest at the same time, such as age. The findings of this project thus highlight the importance of controlling for selection mechanisms that influence both engagement and outcomes of interest. There are some activities (e.g., bingo, pools) where these substantially affect relationships between engagement and harms. Nonetheless, there is clear evidence of associations between certain types of gambling product and indicators of harm when these are controlled for.

SECTION 1: BACKGROUND TO THE STUDY

Psychology has become increasing interested in the strength of causal inferences that can be made from observational data (e.g. cross-sectional surveys) (1). Importantly, some commonly used statistical models, such as structural equation modelling (SEM), inherently make causal assumptions that are often ignored in favour of focusing on model fit (2). In other disciplines, such as econometrics, the use of methods that attempt to strengthen the quality of causal inferences being made is more common. Consequently, there is a clear gap between the kind of inferences that we often want to make in gambling studies, and the statistical tools that we use to analyse data. The purpose of this project was to examine the use of these models, and then apply them to existing gambling data in the United Kingdom.

The question we ask in this project is to determine the extent to which different gambling activities are associated with disordered gambling, and thus the most severe individual gambling harms. This is a question that has been asked in a multitude of ways (3-5) and has a clear policy objective as it identifies whether certain gambling products warrant closer attention or regulation. It is one that has had considerable recent interest, particularly on whether certain types of virtual gaming machine (fixed odds betting terminals (FOTB)) are especially associated with gambling harms (6, 7). The existing literature has focused on this using methods such as logistic regression to regress problem gambling (yes or no) onto gambling activities and other covariates (e.g. age, gender, socioeconomic status, wider gambling involvement). The findings from this literature have not been consistent or conclusive (3). There is stronger evidence some forms of gambling (e.g., slots and electronic gaming machines) are more strongly associated with the display of addictive behaviour, even when controlling for involvement, but these are not consistent from study to study. In the UK, LaPlante et al (5) used data from the British Gambling Prevalence Survey 2007 to regress problem gambling status on gambling activities, controlling for involvement (the number of activities engaged in) or not. They found only one activity, fixed odds betting terminal (FOBT) use, was associated with problem gambling. The Gambling Commission subsequently commissioned research to replicate these findings in other datasets (BGPS 2010, HSE/SHeS 2012) (8). They found that findings were inconsistent from dataset to dataset, and there were no activities clearly related to disordered gambling. While this could be interpreted in support for the role of involvement rather than specific activities, another alternative is that stronger analytic methods are required.

The causal inference methods we are interested in for this project principally come from econometrics and political science: propensity score matching, coarsened exact matching, and two-part modelling. All of these acknowledge and try to overcome a critical limitation with cross-sectional data such as gambling prevalence studies, which is that responses on the independent variables of interest (in our case, gambling variables) are subject to selection bias. Factors such as age, gender, education and socioeconomic background affect the extent to which someone is likely to engage in a behaviour such as slot machine play, but also their risk of disordered gambling directly. Thus, when studying the effect of slot machine play on problem gambling in these datasets, it is difficult to ascertain the extent to which the effect is driven by the product being analysed, or the selection biases that affect whether people engage with the aforementioned product. In other words, are these effects driven by the characteristics of the product, or the characteristics of the player?

Matching methods such as propensity score and coarsened exact matching (CEM) both work by trying to match participants on a series of control variables. In doing so, they endeavour to create as closely as is possible the conditions of a randomized controlled trial. The statistical approach in achieving this is slightly different. In propensity score matching, an initial model is generated by regressing the independent variable onto key covariates of interest. This model is then used to calculate propensity scores for each case exposed to the treatment or not. These scores are then used to identify cases that are appropriately matched on these covariates, which are then subject to analysis. In contrast, coarsened exact matching works by collapsing responses across variables to generate exact matches. For instance, where people who engage in a specific activity or not systematically differ by age, age groups are merged (coarsened) to try and match cases as closely as possible. Then, the analysis can either be repeated on the CEM weighted data, or another algorithm applied (e.g., nearest neighbour matching, k-to-k matching) to refine the modelling further and strengthen the causal inference being made.

In contrast, two-part, double hurdle and sample selection models (SSM) take a different approach. Rather than attempting to capture these biases in a single regression equation, these estimate separate equations for engagement and the quantity of interest (in this case, disordered gambling) (9). Often, these approaches are used for outcomes such as income, where there are an excess of zeroes due to non-engagement in the labour market, and then a distribution beyond zero that can be analysed with a generalized linear model. The SSM differs from the other models insofar as it makes slightly different assumptions, specifically that there are variables in the model that explain sample selection, but not the quantity component of the analysis. Double hurdle and two-part models are less restrictive on this, and for this purpose may be preferable.

Causal modelling in gambling research

In the first phase of this project, we conducted a semi-structured literature review of the research that has been conducted using these inference methods to get an idea of how frequently these methods have been employed in the gambling literature so far, the kind of data they have been applied to, and the research questions that they have been tasked with addressing. In this review, the focus was placed on propensity score matching (PSM) (10) and sample selection modelling (SSM) (11) which have been used to varying degrees in the literature.

In total, 40 relevant papers were identified. Of the 40 papers identified, 33 used a propensity score model to match a control group and condition group on observed baseline co-variates or risk factors (12). For example, the model can be used to generate a matched sample of those not suffering from gambling problems with those that do, based on risk factors (13). For instance, some studies used PSM to control for the effects of other clinical phenomena in small scale samples (14). Of the remaining papers, six used sample selection modelling developed by Heckman (11) to correct for sample selection bias in measuring gambling prevalence. For example, gambling surveys often use “trigger questions” to determine whether to continue with gambling-related questions but this can lead to bias in sampling (15), or conversely lead to measurement error when not asked (16). When corrected, this modelling may change inferences of problem gambling rates and other risk factors or co-morbidities (17) to reveal higher prevalence. The final paper reviewed used a two-part model (TPM) to account for non-consumers when examining the effect of IQ on participation in horse race betting (18). This review did not identify any literature using methods such as coarsened exact matching or directed acyclic graphs, which highlights their potential for future use in research. Structural equation modelling (SEM) literature has not been a focus of this review, but papers can be identified to explore their use in causal inferences. However, there is a large literature that raises caution with the use of structural equation modelling as it is presently employed in social sciences research, and particularly the use of fit indices (19).

The majority of the literature used secondary analysis of existing data sets (e.g., gambling prevalence surveys) when considering these statistical methods. These significantly varied in the topics being examined, but also clustered into themes of inquiry. One such cluster focused on matters of legislation and the effect it has on behaviour. For example, Ludwig et al (20) investigated the impact of regulation amendments in Germany on gambling participation, meanwhile, other papers looked at the legalisation of gambling in the US in relation to fiscal impacts (21), crime rates (22) and alcohol consumption (23). Another cluster focused specifically on gambling products and modality. Gainsbury et al (24) and Papineau et al (25) used PSM to assess the relationship between internet gambling and possible harm as opposed to offline gamblers and found online modalities to be more problematic. Whereas Paton & Vaughan Williams (26) looked at new gambling products (FOBTs) compared to traditional machines and Nichols (21, 22) investigated only casino gambling. Alongside the secondary analysis, there was a subset of international papers that collected original data, mostly with clinical populations of people with problem gambling. These papers used PSM to draw conclusions about the effects and prevalence of co-morbidities with psychiatric illness (27) and other dependencies (14) or the effects of different treatments (28, 29) for problem gambling. Other topics include the impact of religion (30), traumatic brain injury (31) and personality traits (32) on gambling behaviour. This shows the wide range of contexts in which statistical methods have or can be used in gambling literature to strengthen causal inferences or could be applied to other areas of research such as other psychiatric illnesses (17).

As mentioned, some of the literature focused on specific forms of gambling, including online gambling. Specifically, five papers focused on casino gambling, four focused on lottery participation, two on racetrack betting and others included data on multiple types of gambling and games. The use of PSM or SSM allowed researchers to match gamblers and non-gamblers (30), match regions with or without casinos (33), draw conclusions about associated risk factors (34) and make comparisons between the behaviour of different gambling types (35, 36).

Within the wider gambling literature, risk factors associated with gambling have been studied extensively which resulted in there being consistency in what risk factors and covariates are considered within the statistical models. Common risk factors included: age, gender, race, educational attainment, income or socio-economic status (SES), marital status, household location, other dependencies (alcohol, nicotine, drugs) and co-morbidities (anxiety, depression, PTSD). Some papers investigated covariates more specifically to determine their causal link with gambling participation or problems. For example, IQ (18), standard of living (34) or attitudes towards gambling (37) have each been investigated. In this review, many of these risk factors were used in the statistical models to create the matched groups or address selection bias to inform causal inferences.

Of the 40 papers identified in this review, 16 did not include a measure of problematic gambling. This is in part due to the topic of inquiry in some of the papers, but also due to the datasets used within the secondary analyses. Of those papers that include a scale for problem gambling, the Problem Gambling Severity Index (PGSI) (38) was more commonly used with 10 papers reporting it. This was followed by the South Oaks Gambling Screen (SOGS) (39) used by 6 papers, the NORC Diagnostic Screen for Gambling Disorders (NODS) (40, 41) used by 3, with other regional, adapted or bespoke scales being used by others. Following PSM or SSM, some papers found prevalence rates of problem gambling using the scales to increase. For example, Harrison et al (15) found the rate to increase from 1.3% to 7.7% after SSM and Ludwig et al (20) saw an increase from 0.2% to 0.7% after PSM.

Overall, these findings highlight how these methods of modelling causal inference have and can be used in a variety of contexts. While PSM has been most commonly used until present, there is scope to expand the use of SSM, and for researchers to use other methods such as coarsened exact matching or directed acyclic graphs in cross-sectional gambling data.

The present study

As noted in the literature review, these methods have been applied principally to existing data such as gambling prevalence surveys. A number of analyses have attempted to use them to determine if specific kinds of gambling activity are especially harmful, but there is both a clear need for further research here, as well as rigorous application of these methods to a range of gambling activities. For this project, we propose to determine the extent to which specific gambling products are differentially associated with problem gambling. Using data from 9 surveys that measure gambling prevalence, we apply these models to examine the relationship between gambling activities and harm. This analysis will test the extent to which there is a gradient of gambling harms, but also identify robust, meaningful risk factors that feed into subsequent causal modelling of gambling related risk factors.

SECTION 2: METHODOLOGY

Sampling:

Data was taken from nine different British gambling prevalence surveys: The British Gambling Prevalence Surveys in 2007 and 2010, the Health Surveys for England in 2012, 2015, 2016, and 2018, and the Scottish Health Surveys in 2012, 2015, and 2016. The data for each of these surveys was collected by the National Centre for Social Research (NatCen), using a probability sampling approach. The methods for these have been reported in detail elsewhere (42-47), but a brief summary is provided here.

The British Gambling Prevalence Survey series was conducted before and after legislative changes were implemented from the Gambling Act 2005. A computer assisted interview was administered to respondents who agreed to take part. The Health Survey for England is a long running series of interviews about a range of different health topics, conducted annually since 1991. There are some topics that are asked regularly (e.g., biometric measures, alcohol use, smoking, chronic illness), and others that are asked intermittently (e.g., healthy eating). Past year gambling engagement and problem gambling questions were added to the Health Survey for England and Scottish Health Surveys in 2012 following the cessation of funding for the British Gambling Prevalence Survey. The gambling engagement and problem gambling questions in the HSE and SHeS are identical to the BGPS, and have been asked intermittently since. Forthcoming data from the 2021 series will include gambling as well, and is due to be released later in 2023. The primary difference between the BGPS and HSE/SHeS data series is that whereas the gambling content was part of the main interview in the BGPS, in the HSE and SHeS it was included as part of a self-completion questionnaire that had to be mailed back to the survey company. This means that the response rate is lower than for the full survey. Participants in all studies were sampled using a probability sampling frame, that was randomly drawn from a stratified sample of postcodes from the UK Postcode Address File.

Gambling activities:

Gambling activities were reduced from 19 to 9 categories of play. This was determined in part based on a minimum residual exploratory factor analysis using a tetrachoric correlation (the data for this is available on the OSF, see https://osf.io/xqaw5/?view_only=9ae3432b55b24c50bade971bb6af612f), and relevant subject knowledge. In some cases, activities did not load onto a factor, or cross loaded onto a range of different factors. The main reason for this was because these surveys use common sets of gambling activities, which have multiple questions pertaining to similar activities. For instance, the betting questions fractionate into dog racing, horse racing, sports betting, and other betting. Although these potentially capture different structural mechanisms (e.g., lower win rates in dog and horse racing), statistically these pose a problem as the base rate of engagement for many of these activities is low, most below 10% annual engagement and some below 5%. This makes finding a matched sample more difficult. In these cases, decisions were guided based on the overall rate of engagement – for instance there were separate factors of racing-based betting (i.e. dogs and horses) and other betting activities. However, the low endorsement of some of these made modelling difficult, and combining the two made sense (also in light of high rates of covariance).

Covariates:

The following covariates were entered into the analysis. Age (grouped in 5 year age bands), sex (ref = Male), income quintile, qualifications (Degree, FE/HE without degree, A-level or equivalent, GCSE/O-level or equivalent, less than O-level or no qualification), Index of Multiple Deprivation (IMD) quintiles, current smoking status, current drinking status, self-reported health (SRH) (5 point Likert) and whether the respondent has a longstanding limiting illness. IMD and SRH were excluded from the CEM analyses because these in combination with the other covariates led to substantial (90-95%) data loss.

Outcome variable:

The Problem Gambling Severity Index was administered to all past-year gamblers in each of the surveys. The PGSI is a widely used measurement of disordered gambling that has been used in clinical, general population and gambling specific subsamples. We used total PGSI score as the outcome variable in this analysis. Respondents who completed the gambling questions, but reported they did not gamble were imputed with a score of 0.

Data Analysis:

The data were cleaned to remove participants that did not provide complete answers to the variables of interest. In most cases, this was on the socioeconomic variables, or were missing gambling questions entirely from non-engagement with the self-completion questionnaire. The causal inference analyses were conducted in STATA v. 17. The meta-analyses were conducted in *R*.

For each of the gambling activities, six separate analyses were conducted upon the data. To estimate a baseline effect of gambling activities on problem gambling, linear regression without and with covariates of interest were estimated.

Then, a propensity score matching procedure was conducted to control for the effects of the aforementioned covariates on the treatment of interest (i.e., gambling activities). The propensity score matching proceeded as follows: a logistic regression was used to generate propensity scores predicting the gambling activity by the covariates of interest. Respondents were then matched using a caliper matching procedure (starting with a caliper width of .05, and increasing until matches could be found). The effects were then re-estimated on the propensity score matched sample. The procedure for CEM is similar, except categories are coarsened (combined) to exactly match cases on the covariates. Following advice from

Blackwell et al (2010), the nearest matching neighbour algorithm was applied to restrict cases to those with overlapping values on the variables of interest. An additional modification to this procedure is to restrict the sample to a *k*-to-*k* match (i.e., equal sample for treatment and control). This was also applied to the sample.

Finally, a two-part model was estimated. In the first part, a logistic regression procedure is used to test for differences between respondents who scored 0 on the outcome of interest, and those who scored > 0. Then, a linear regression was fit for respondents scoring greater than 0 on the outcome. Because we were interested in discriminating non-gamblers from gamblers that did not engage in a specific activity, PGSI score was transformed by adding 1 to PGSI score for each respondent who reported gambling in the past year i.e., non-gamblers scored 0, gamblers scored 1-28. Thus, this preserved the structural relationship for gambling activities while allowing non-gamblers versus not gambling on a specific gambling activity to be discriminated.

Meta-analysis

To synthesize the findings from the separate analyses, a meta-analysis was conducted on the effect sizes from the 432 analyses (8 types of gambling activity, 6 methods, 9 datasets) reported. To calculate the meta-analysis, the unstandardized regression coefficient was extracted from each of the analyses, alongside the standard deviation of the dependent variable (PGSI score) for each sample analyzed, and the sample sizes for people who gambled on an activity and those who did not. These effects were then transformed into standardized mean difference (Cohen's *d*) using the esc_B function from the 'esc' package in *R*.

A three-level random effects meta-analysis was specified using the rma.mv function in the 'metafor' package, with dataset included as the third level to minimize the effect of systematic biases within a dataset affecting the overall findings. Random effects models are widely acknowledged to be preferable for meta-analysis because it makes less restrictive assumptions of study heterogeneity (48). The third level was added because there are 48 analyses per dataset, and without sufficient control, the results of the meta-analyses might reflect systematic biases between datasets rather than the impact of gambling products. The use of three level models has been supported for this purpose (49). In addition to this analytic structure, the following moderators were entered into the analysis: gambling activity (with lotteries as the reference category), analysis type (with bivariate linear regression as reference), year (centred at 2012), and country (with Britain as the reference versus England and Scotland, which also controls for the effects of gambling prevalence vs. health survey).

SECTION 3: RESULTS

Table 1 reports the descriptive statistics for the analyzed samples. As a first step, linear regressions with and without adjustment for covariates were conducted to estimate the baseline effect of gambling engagement on problem gambling. These showed effects of gambling engagement across all activities. That is to say that all activities were associated with problem gambling to varying degrees. Regression adjustment reduced these effects, but only marginally. A number of demographic factors such as (younger) age and (male) sex exerted strong effects on both engagement with the gambling activity, and problem gambling severity. Full information about these is reported in the OSF at https://osf.io/xqaw5/?view_only=9ae3432b55b24c50bade971bb6af612f.

Propensity Score Matching

The results from the propensity score matching are reported in Table 2. The results show that engagement with almost every single gambling activity was associated with greater problem gambling. However, the size of the effect varied considerably from activity to activity. The effect size for some activities, specifically online gambling and slot machines (including FOBTs) were notably higher. Some activities, specifically offline betting and casino/table games, were slightly higher, whereas football pools and offline bingo were inconsistent from dataset to dataset. Finally, lotteries and scratchcards were associated with lower increases in problem gambling.

Coarsened Exact Matching

The results from the coarsened exact matching are also reported in Table 2. Overall, the results are very similar to the PSM in terms of the effect size. The primary difference between PSM and CEM was that very few of the football pools and bingo effects were significant in the CEM analyses. These analyses were associated with much greater uncertainty, in part due to the level of data loss. Otherwise, the online gambling and slots effects remained significant, and were the activities most strongly linked with problem gambling. Some of the other activities (lotteries, scratchcards, betting) showed inconsistencies from test to test. In part, this is because CEM involves significant data loss. This is particularly notable for the SHeS datasets because they are smaller in the first place. Therefore, it is not clear the extent to which these non-significant effects are driven by a lack of effect, or a lack of statistical power. To this end, the need for meta-analysis is particularly acute as there is evidence of an effect for some gambling activities, but data loss makes it difficult to determine whether this is systematic or noise.

Two Part Modelling

Finally, two-part modelling was applied to the data. The findings from this are also reported in Table 2. In terms of the overall effect size, they are smaller than the bivariate and regression adjusted analyses, but slightly larger than the PSM and CEM analyses. A key difference was that the lottery and scratchcard effects were consistently non-significant. In part this may be because levels of engagement with these are fairly high. For instance, the majority of gamblers play the lottery, so the data are heavily imbalanced on this measure. Overall however, the results were otherwise similar in both direction and magnitude of effect to the other causal inference models.

Meta-analyses

The findings of the meta-analysis show an effect of gambling engagement on problem gambling overall (pooled effect = 0.3205, se = 0.0109, $p < .001$). Moderator analyses for activity and analysis type were significant, but not for year of study nor country of study. The moderator analyses formalize the observations made from the individual analyses. There was a gradient of effects for gambling activities, with lotteries displaying a lower risk than scratchcards, followed by bingo, with pools and offline betting roughly equivalent. This was followed by table games, slot machines/FOBT's, then online gambling. In terms of the analysis moderator, linear regression had the highest effects reported, shortly followed by regression with covariate adjustment. The effect of gambling activities was moderated further by propensity score matching. The most notable effects were for two-part modelling and coarsened exact matching (using weighted and k-to-k matching procedures). The findings of the meta-analysis are reported in Table 3.

Table 1. Descriptive statistics for the analytic sample.

	B2007	B2010	E2012	S2012	E2015	S2015	E2016	S2016	E2018
N	5675	6512	5048	3354	4963	3301	4889	2829	5271
Gambler	67.93%	74.69%	63.99%	67.35%	60.59%	65.43%	56.71%	63.80%	56.21%
PGSI *M*	0.203	0.230	0.102	0.120	0.110	0.105	0.116	0.109	0.105
PGSI *SD*	1.321	1.329	0.814	0.988	0.978	0.927	0.957	1.057	0.943
Activities									
Table	12.51%	13.62%	6.34%	5.90%	5.82%	7.27%	5.58%	5.41%	4.97%
Betting	20.97%	19.88%	13.59%	13.00%	13.56%	15.09%	12.49%	14.32%	11.21%
Online	6.34%	7.65%	6.58%	7.22%	8.28%	10.66%	8.59%	11.03%	9.26%
Lotteries	61.11%	66.14%	56.93%	60.76%	51.02%	57.62%	46.84%	54.83%	44.91%
Scratchcards	20.32%	25.18%	19.10%	17.95%	20.37%	22.21%	19.96%	22.87%	18.33%
Pools	7.49%	7.66%	2.06%	4.23%	1.81%	4.63%	1.83%	3.89%	2.18%
Bingo	2.98%	4.12%	5.71%	6.20%	5.46%	6.54%	4.92%	7.49%	5.29%
Slots	15.86%	14.00%	7.09%	7.39%	7.26%	9.57%	6.58%	8.41%	6.09%
Age									
16-17	2.45%	2.93%	0.85%	1.04%	0.64%	0.67%	0.71%	0.67%	0.78%
18-19	2.59%	2.96%	1.45%	1.04%	1.21%	1.09%	1.18%	1.31%	1.18%
20-24	5.74%	7.68%	5.35%	4.32%	4.78%	5.60%	4.21%	4.88%	4.04%
25-29	7.37%	7.31%	6.93%	6.44%	6.71%	6.27%	7.51%	6.68%	6.07%
30-34	8.44%	8.25%	8.24%	6.50%	8.54%	6.54%	9.05%	7.28%	8.73%
35-39	10.77%	9.77%	9.21%	8.91%	9.57%	7.85%	8.80%	7.74%	9.24%
40-44	10.70%	9.49%	10.04%	10.26%	9.39%	7.97%	8.72%	7.92%	8.35%
45-49	9.34%	8.97%	9.94%	10.52%	9.97%	8.82%	9.26%	8.59%	9.39%
50-54	7.86%	8.85%	9.45%	9.66%	9.37%	9.45%	9.63%	10.18%	9.64%
55-59	9.29%	7.26%	8.64%	9.03%	7.78%	9.82%	8.80%	9.37%	8.94%
60-64	7.24%	7.72%	8.02%	8.80%	8.66%	8.54%	8.84%	9.51%	8.42%
65-69	4.86%	7.08%	8.24%	9.00%	8.64%	10.48%	8.24%	9.51%	8.86%
70-74	5.36%	5.07%	5.51%	5.46%	5.78%	6.82%	5.81%	6.93%	7.63%
75-79	3.91%	3.18%	4.20%	4.03%	4.31%	5.54%	4.77%	5.05%	4.36%
80-84	2.71%	1.93%	2.58%	3.34%	2.90%	2.91%	2.72%	2.90%	2.77%
85-89	1.00%	1.24%	0.99%	1.22%	1.43%	1.24%	1.25%	1.17%	1.20%
90+	0.37%	0.31%	0.36%	0.42%	0.30%	0.39%	0.52%	0.32%	0.42%
Sex (% F)	52.09%	53.95%	54.14%	55.19%	53.56%	54.95%	54.22%	55.99%	54.28%
Qualifications									
Degree	24.32%	31.57%	29.18%	31.54%	31.49%	32.32%	31.89%	35.56%	32.50%
FE/HE wo degree	8.11%	7.46%	12.36%	12.08%	12.51%	12.15%	11.18%	12.34%	12.84%
A-level	12.26%	12.71%	15.79%	14.01%	16.54%	15.18%	16.43%	15.34%	16.58%
O-level.GCSE	29.89%	24.43%	20.46%	18.69%	19.54%	19.21%	20.23%	16.97%	19.18%
Lower/no qual	25.43%	23.82%	22.21%	23.67%	19.91%	21.15%	20.25%	19.79%	18.90%
Smoker (% Y)	22.94%	24.77%	17.89%	22.99%	16.26%	19.30%	17.08%	19.05%	16.49%
Self-reported health	1.92 (0.84)	1.96 (0.89)	1.97 (0.92)	2.01 (0.96)	1.96 (0.91)	2.00 (0.94)	2.01 (0.94)	2.05 (0.96)	2.02 (0.92)
LLI	23.15%	29.15%	39.74%	47.88%	41.48%	49.08%	44.39%	49.73%	44.77%
IMD Quintile									
1 (least deprived)	20.23%	20.36%	24.43%	20.9%	22.57%	18.75%	20.77%	21.92%	19.84%
2	21.96%	21.13%	21.51%	23.76%	21.38%	26.57%	19.90%	23.01%	21.95%
3	20.23%	20.22%	20.48%	22.51%	21.30%	20.30%	21.58%	22.55%	21.00%
4	19.86%	18.66%	19.20%	17.86%	17.53%	18.63%	18.95%	16.97%	20.32%
5 (most deprived)	17.73%	19.63%	14.38%	14.97%	17.23%	15.75%	18.80%	15.55%	16.88%
Drinker (% Y)	74.71%	74.74%	81.32%	80.50%	81.72%	80.37%	82.92%	80.10%	81.43%
NS-SEC									
Managerial			38.59%	35.57%	41.27%	35.81%	38.83%	37.08%	41.81%
Intermediate			24.84%	22.21%	23.92%	23.60%	24.49%	22.20%	23.58%
Routine			36.57%	42.22%	34.82%	40.59%	36.69%	40.72%	34.60%
Income									
1	18.41%	20.22%	16.82%	17.02%	14.89%	17.48%	17.41%	16.33%	17.95%
2	19.56%	18.55%	17.85%	18.78%	17.71%	18.09%	17.89%	18.56%	19.50%
3	16.49%	23.69%	20.42%	20.30%	20.77%	22.08%	20.19%	20.82%	18.73%
4	21.83%	15.89%	23.95%	21.79%	21.64%	22.05%	22.85%	21.14%	22.05%
5	23.70%	21.64%	20.96%	22.09%	24.98%	20.30%	21.66%	23.15%	21.78%

Table 2a. The relationship between gambling activities and problem gambling scores using different kinds of causal inference methods in the Health Survey for England

		F1	F2	F3	F4	F5	F6	F7	F8
E2012	LR	.579	.367	.617	.099	.259	.564	.198	.636
	Adjusted	.531	.347	.574	.086	.228	.480	.188	.589
	PSM	.470	.322	.459	.079	.212	.149 *ns*	.189	.460 *ns*
	CEM	.319	.256	.458	.082	.215	.235 *ns*	.211 *ns*	.454
	CEM *kk*	.460	.265	.508	.094	.217	.601	.177 *ns*	.537
	TPM	.474	.312	.520	-.078 *ns*	.173	.399	.136	.532
E2015	LR	.666	.396	.734	.122	.257	.839	.243	.735
	Adjusted	.600	.362	.680	.110	.212	.750	.229	.675
	PSM	.246	.361	.476	.099	.208	.246 *ns*	.255	.347
	CEM	.373	.143	.537	.116	.166	.226 *ns*	.169 *ns*	.345
	CEM *kk*	.735	.235	.646	.129	.157	.607 *ns*	.325 *ns*	.601
	TPM	.524	.315	.615	-.006 *ns*	.128	.656	.167	.606
E2016	LR	.704	.587	.853	.109	.235	.935	.441	1.243
	Adjusted	.631	.549	.803	.085	.164	.826	.419	1.187
	PSM	.530	.462	.423	.085	.086	.228 *ns*	.672	.875
	CEM	.337	.380	.361	.055 *ns*	.102 *ns*	.490 *ns*	.196 *ns*	.747
	CEM *kk*	.468	.546	.561	.048 *ns*	.101 *ns*	.420 *ns*	.295	1.024
	TPM	.543	.510	.746	-.104 *ns*	.030 *ns*	.705	.352	1.131
E2018	LR	.721	.440	.792	.114	.209	.480	.310	1.050
	Adjusted	.669	.411	.772	.112	.167	.376	.287	1.006
	PSM	.730	.290	.562	.130	.148	.058 *ns*	.202 *ns*	1.027
	CEM	.169 *ns*	.170	.371	.091	.108 *ns*	.130 *ns*	.118 *ns*	.538
	CEM *kk*	.277	.151 *ns*	.493	.100	.137 *ns*	-.036 *ns*	.283 *ns*	.784
	TPM	.586	.353	.733	-.008 *ns*	.054 *ns*	.236 *ns*	.200	.940

Notes: LR = linear regression (without adjustment), Adjusted = linear regression with covariate adjustment, PSM = Propensity Score Matching, CEM = Coarsened Exact Matching, kk = *k*-to-*k* matching, TPM = Two-part model.

F1 = Poker, table and private gambling, F2 = offline betting (at a bookmarker or racetrack, F3 = online gambling, F4 = lotteries, F5 = scratchcards, F6 = football pools, F7 = offline bingo, F8 = slot machines and fixed odds betting terminals.

E2012-2018 – Health Survey for England (year)

Table 2b. The relationship between gambling activities and problem gambling scores using different kinds of causal inference methods in the Scottish Health Survey.

		F1	F2	F3	F4	F5	F6	F7	F8
	LR	.555	.508	.842	.140	.162	.449	.134	1.033
	Adjusted	.479	.454	.802	.124	.125	.346	.147	.991
	PSM	.385	.380	.764	.121	.131	.102	.060 *ns*	.776
S2012	CEM	.295 *ns*	.176	.445	.064 *ns*	.001 *ns*	.284 *ns*	.084 *ns*	.382
	CEM *kk*	.376 *ns*	.258	.722	.047 *ns*	.023 *ns*	.301 *ns*	.171 *ns*	.490
	TPM	.413	.406	.763	.049 *ns*	.060 *ns*	.267	.109 *ns*	.956
	LR	.561	.428	.678	.139	.201	.281	.120	.609
	Adjusted	.504	.375	.634	.125	.161	.148	.108 *ns*	.542
	PSM	.481	.207	.409	.134	.130	.014 *ns*	.089 *ns*	.342
S2015	CEM	.436 *ns*	.336	.491	.119	.266	-.004 *ns*	.054 *ns*	.460
	CEM *kk*	.680	.330	.600	.127	.303	-.091 *ns*	.075 *ns*	.647
	TPM	.450	.330	.595	.137 *ns*	.104 *ns*	.044	.059 *ns*	.483
	LR	.838	.541	.691	.144	.253	.539	.183	.739
	Adjusted	.778	.495	.632	.126	.205	.413	.196	.666
	PSM	.650	.386 *ns*	.458	.115	.095 *ns*	.141 *ns*	.417 *ns*	.496
S2016	CEM	.576	.403	.385	.065 *ns*	.147 *ns*	.085 *ns*	.301 *ns*	.461
	CEM *kk*	.688	.560	.548	.081 *ns*	.152 *ns*	.184 *ns*	.313 *ns*	.579
	TPM	.723	.457	.581	.079 *ns*	.139	.315	.162 ns	.605

Notes: LR = linear regression (without adjustment), Adjusted = linear regression with covariate adjustment, PSM = Propensity Score Matching, CEM = Coarsened Exact Matching, kk = *k*-to-*k* matching, TPM = Two-part model.

F1 = Poker, table and private gambling, F2 = offline betting (at a bookmarker or racetrack, F3 = online gambling, F4 = lotteries, F5 = scratchcards, F6 = football pools, F7 = offline bingo, F8 = slot machines and fixed odds betting terminals

S2012-2016 – Scottish Health Survey (year)

Table 2c. The relationship between gambling activities and problem gambling scores using different kinds of causal inference methods in the British Gambling Prevalence Survey

		F1	F2	F3	F4	F5	F6	F7	F8
	LR	.765	.491	1.187	.233	.476	.732	.822	.821
	Adjusted	.659	.446	1.071	.203	.420	.724	.738	.709
	PSM	.500	.338	.776	.156	.431	1.261	.403	.529
B2007	CEM	.398	.336	.656	.166	.292	.410	.276 *ns*	.479
	CEM *kk*	.622	.362	.936	.185	.364	.558	.547	.578
	TPM	.561	.388	.969	.053 *ns*	.332	.679	.638	.606
	LR	.760	.576	1.138	.202	.388	.304	.1.048	.910
	Adjusted	.670	.521	1.029	.233	.323	.331	.904	.815
	PSM	.475	.415	.813	.249	.275	.324	.493	.714
B2010	CEM	.425	.404	.782	.217	.271	.172 *ns*	.714	.586
	CEM *kk*	.574	.515	.984	.219	.303	.229	.996	.769
	TPM	.576	.448	.931	.090 *ns*	.214	.266	.782	.723

Notes: LR = linear regression (without adjustment), Adjusted = linear regression with covariate adjustment, PSM = Propensity Score Matching, CEM = Coarsened Exact Matching, kk = *k*-to-*k* matching, TPM = Two part model.

F1 = Poker, table and private gambling, F2 = offline betting (at a bookmarker or racetrack, F3 = online gambling, F4 = lotteries, F5 = scratchcards, F6 = football pools, F7 = offline bingo, F8 = slot machines and fixed odds betting terminals

B2007/10 = British Gambling Prevalence Survey 2007/10

Table 3. Summary of multilevel meta-analysis

Effect	Estimate	Standard error	*z*	*p*
Intercept	0.2244	0.0280	8.0007	<.001
Gambling activity (REF = lotteries)				
Scratchcards	0.0961	0.0073	13.1909	<.001
Bingo	0.1743	0.0113	15.4348	<.001
Pools	0.2481	0.0123	20.1532	<.001
Betting	0.2455	0.0079	31.0394	<.001
Table/casino	0.3692	0.0096	38.3482	<.001
Online	0.5245	0.0096	54.6943	<.001
Slots	0.5069	0.0090	56.4377	<.001
Analysis (REF = linear regression no adjustment)				
Covariate adjustment	-0.0411	0.0075	-5.4612	<.001
Two part modelling	-0.1751	0.0082	-21.3293	<.001
Propensity Score Matching	-0.1108	0.0075	-14.7881	<.001
Coarsened Exact Matching	-0.1713	0.0086	-19.9884	<.001
Coarsened Exact Matching (*k*-to-*k* matching)	-0.1685	0.0103	-16.2835	<.001
Year	-0.0014	0.0052	-0.2742	0.7840
Country (REF: Britain)				
England	-0.0022	0.0431	-0.0512	0.9591
Scotland	-0.0453	0.0403	-1.1236	0.2612

SECTION 4: DISCUSSION

Results across multiple statistical techniques to make causal inferences on gambling engagement show that there is a gradient of risk associated with certain gambling activities. In other words, some forms of gambling appear to be more addictive, or more attractive to those at risk of gambling problems, than others. While the idea that certain forms of activity are inherently more risky than others is not a new finding, the literature has been plagued with problems about the extent to which this can be attributed to the product itself, or to factors that drive selection of certain gambling behaviours. Critiques of the problem gambling paradigm have often noted that the focus on people with problem gambling often comes with implicit assumptions that individual factors that can be predicted by demographic and socioeconomic indices (e.g., age, sex). Because these frequently predict problem gambling and types of gambling engagement, it is difficult to tease out these factors. The modelling reported here demonstrates comprehensively that there are specific gambling activities that are associated with a greater display of addictive behaviour.

The findings of this exercise concord and contrast with the existing literature on gambling types of problem gambling. The literature on types of gambling and problem gambling has not been particularly consistent except for the most severe cases (i.e., slots/EGMs/FOBTs, online gambling). In contrast, these findings are notable for their consistency from dataset to dataset. One potential reason for this is the use of PGSI score instead of problem gambling classification. Especially in large datasets where the prevalence of problem gambling is relatively small, the use of problem gambling classification is likely to lead to noise as the number of cases is low (< 100), and this has been borne out in the existing literature. One possibility to understand this variance would be to access existing datasets to determine whether these findings are more stable with raw problem gambling scores than PG status. More generally, there is a need to synthesize the existing evidence on this, to determine the extent to which these effects are robust. That slots and online play come out as the largest effects concords with the literature in this area. It is however notable that the association between these forms of activity and problem gambling does not appear to change over time. The period covered encompasses the significant growth of online gambling, and heavy media attention on the dangers posed by FOBTs. Despite these changes, which have affected overall prevalence, the effect size seems to be similar over time.

While most analyses were consistent in terms of the degree of association between the activity and problem gambling, as well as statistical significance, there were two notable exceptions to this. Estimates for football pools and bingo were inconsistent from dataset to dataset, although associated overall in the meta-analysis. There were very strong indications of selection effects in football pools, largely driven by age and gender (i.e. younger men). This meant that while the unconditional effects were consistently strong, when matched these were substantially smaller. This points to the benefits of conducting this analysis. It should be noted that the effect of football pools play is much greater pre-2010. In the BGPS series, engagement with pools sits at around 7-7.5%, whereas in the SHeS it is 3.8-4.5%, and the HSE it is 1.9-2.3%. This drop represents a continued replacement of football pools with sports betting as an activity. In contrast for bingo, associations with PGSI severity are very high in the BGPS data series (similar to slots and online gambling), and lower in the health surveys post 2010, regardless of the type of analysis conducted. For some procedures (i.e., coarsened exact matching), the effects in individual datasets were often non-significant. There is a need to understand this effect further. Over the same period, it is one of three activities, alongside online gambling and scratchcards, that appears to have been more resilient to wider changes in gambling engagement.

These results highlight the importance of controlling for selection biases when making inferences. One particular strength of this study is the ability to control for factors that determine selection of engagement with gambling behaviour. Our findings suggest that a proportion of the variance can be explained by these selection biases, but there remains substantial, significant product-specific effects nonetheless. Typically, analyses that have looked at the effect of gambling products on problem gambling have tried to control for them by including them as covariates in regression analyses. However, the problem here is that for some variables, there are strong effects on both selection (i.e., gambling engagement) and outcome (i.e., problem gambling). To contextualise this with an example, consider the case of sports betting. Is this associated with problem gambling because of the structural features of the activity, which can be modified, or is it because it is most commonly engaged with by younger men with psychological and sociodemographic profiles that make them more vulnerable to addiction? The research on gambling advertising shows that advertising content is geared towards these demographic groups, implying the presence of a selection effect here (50, 51).The methods reported here can go some way to disentangling these competing effects, whereas regression adjustment cannot. Our results suggest that standard covariate adjustment overestimates the effect of gambling activities on problem gambling, and as such recommend that future analyses use stronger approaches to control for these biases. Importantly though, even the more conservative analyses conducted (e.g., CEM) suggest that when these selection effects are controlled for more thoroughly, effects of gambling engagement remain, and are product specific.

This also has implications for our conceptualisation of gambling harms and responsible gambling. Models such as the Conceptual Framework of Harmful Gambling (52) consider that the type of gambling one engages in is an important factor in gambling harm. However, they note a critical limitation of this literature is that it does not distinguish between explanations that focus on the role of the product in creating gambling harm, and those that suggest certain forms of play are especially attractive to those at risk at harm. Our findings have the potential to further our understanding of this relationship because they control for a large set of risk factors that affect predisposition to gambling harms. This approach ought to be expanded into different international settings to determine the boundary conditions of these effects. Furthermore, a distinction in this framework is particularly made between continuous and punctuated forms of gambling. The analyses reported here can be used to identify a gradient of harm within these gambling types.

One limitation of the study is that the items ask only about breadth of engagement with a specific gambling behaviour, and not depth (i.e. frequency of use). Effects of frequency of use, or monthly engagement, are likely to be much higher than past year engagement, and especially salient for activities where engagement effects are lower (i.e. scratchcards). Nonetheless, the propensity score matching shows that gambling activities on a whole are associated with higher PGSI scores. More importantly, there is a gradient of risk, with some activities associated with a smaller increase (e.g. lotteries) than others (e.g. betting, slot machines). Although we did not control for involvement, as some other studies have (5), our analyses control for the mechanisms that determine involvement. Similarly, although the datasets are designed to be representative of the national populations of Britain, England and Scotland respectively, there are some groups these samples systematically do not recruit, such as the homeless, students in halls of residence, and prisoners. There is reason to believe each of these groups reports a heightened risk of problem gambling (53, 54). Further research, perhaps using the British Gambling Prevalence Survey datasets or other international datasets would be beneficial to determine whether the effect of frequency on problem gambling is stronger for some gambling behaviours than others.

These findings cover the most recent publicly available data in the United Kingdom. However, there are a couple of caveats to be noted, particular that since the release of these data, aspects of gambling regulation have changed. Specifically, the maximum odds staked on fixed odds betting terminals were reduced from £100 to £2. Across the range of slot machine categories in the UK, the current highest stake is presently £5. Subsequent iterations of the Health Survey for England and Scottish Health Surveys in 2021 have included gambling questions, and it would be of great interest to examine the effects of these changes on gambling behaviour, and the effects reported here. However, one potential obstacle to this is the potential for wider changes in consumer behaviour in light of the COVID pandemic. Betting offices were shut at several junctures through the pandemic, and this is likely to have changed consumer behaviour, particularly as online and mobile gambling continued, albeit with fewer opportunities to bet.

Although these analyses have notable strengths, there are some limitations with regard to the samples that were analyzed. The data were matched after being listwise deleted. This means that there is some systematic missingness in the data. In particular, people who did not reporting income are missing from these analyses. This might bias the findings as the relationship between these activities and problem gambling might be exaggerated or diminished. One especially desirably sensitivity analysis to conduct on the data would be to use multiple imputation to control for this systematic missingness. Another potential limitation is the analysis was done on the unweighted data. The process of generating propensity score weights is similar to the estimation procedure used for the generation of sampling weights, and the use of sampling weights in inferential analysis has been brought into question (55). Again, further sensitivity analyses here may be beneficial. The PSM and CEM analyses do not make the distinction between people who gamble but did not engage in a specific behaviour, and non-gamblers. Although the findings of the TPM suggest this effect is small, this might be an area again where sensitivity analyses would be beneficial. However, this comes with some additional downsides. In particular, for some types of gambling (e.g. lotteries), and CEM in general, this additional data loss may reduce the sample size to the point where the analyses are underpowered. Risk of harm was assessed using PGSI severity. While the PGSI performs well at identifying individuals meeting the criteria for gambling disorder, the performance of the intermediate interpretative categories has been a cause for concern (56) as this is the group where a wider burden of population wide harms may occur (57). The use of raw PGSI score mitigates against this to some extent, as it means the interpretive categories that are less empirically robust are not used. However, it does mean the analysis assumes the PGSI captures a meaningful latent continuum of gambling problems. There are issues with this, especially given the levels of skew in PGSI scores,

SECTION 5: IMPLICATIONS FOR KNOWLEDGE TRANSLATION AND EXCHANGE

Our findings show there are consistent, stable product related impacts on gambling harms across the British population. There are some gambling activities that are specifically harmful, and a gradient of risk across all of the activities questioned. These could be used by policymakers to develop an evidence-based calculus of risk when it comes to regulating gambling activities. As there have been significant shifts in the gambling market over the time period studied, this can highlight certain activities for further investigation. This is a topic of interest at the moment as the Government is consulting on creating parity between the online and offline gaming sectors, in part by reducing restrictions on land-based operators. The results of this analysis suggest that for some types of gambling, especially slot machines that are partially the subject of this consultation, that such moves are premature.

The findings also have the potential for knowledge exchange with people who have experienced gambling harms. Specifically, we show when selection and demographic biases are controlled for that certain activities are associated with a greater risk of harm. Some of our previous research has highlighted how in discussing their gambling journeys, people with lived experience appear to converge on topics discussing specific gambling products as salient in their experiences of harm (58). However, this ought to be supplemented with additional work with people with lived experience of gambling harm to understand how they conceptualize the role of gambling products, either through further analysis of existing data (e.g. forum posts, social media) or novel, co-produced research with stakeholders in building a shared understanding of these impacts.

One possible next step for this work with knowledge exchange implications involves the profiling and segmentation of gamblers at risk of harm. Because engagement with most forms of gambling is uncommon at the population level, and these findings show clear evidence of product-based effects, it would be possible to scale these up to identify patterns of engagement specifically associated with harm, rather than simply general involvement. The development of robust, replicable subgroups could be used to inform stakeholders, from policymakers, third sector bodies or groups interested in promoting safer gambling. For instance, if people engaging in specific combinations of gambling activity are shown to be at much greater risk of harm, these can be flagged earlier for enhanced monitoring, or for tailored intervention.

This project creates the potential for further policy-based analysis to examine the impact of major changes of gambling legislation and regulation on the general population. These models could be applied, with careful selection of outcomes of interest, to test population wide impacts, a key ambition of public health approaches to gambling. As noted in the discussion, this could be applied to the impacts of the 2007 Gambling Act (which has not been done yet), or to changes on FOBT licensing with forthcoming data. It could also be used prospectively to test whether future changes following the Gambling White Paper have mitigated gambling harms. On top of the findings of this work, this project also signposts to policymakers and bodies that commission gambling research on methodological approaches for evaluating the effectiveness of an intervention over and above existing methods.

SECTION 6: REFERENCES

1. Rohrer JM. Thinking clearly about correlations and causation: Graphical causal models for observational data. Advances in methods and practices in psychological science. 2018;1(1):27-42.
2. Pearl J. The causal foundations of structural equation modeling. Handbook of structural equation modeling. 2012:68-91.
3. Griffiths MD, Auer M. The irrelevancy of game-type in the acquisition, development and maintenance of problem and pathological gambling. Frontiers in psychology. 2013;3.
4. Binde P, Romild U. Risk of problem gambling among occupational groups: A population and registry study. Nordic Studies on Alcohol and Drugs. 2020;37(3):262-78.
5. LaPlante DA, Nelson SE, LaBrie RA, Shaffer HJ. Disordered gambling, type of gambling and gambling involvement in the British Gambling Prevalence Survey 2007. The European Journal of Public Health. 2011;21(4):532-7.
6. Dowling N, Smith D, Thomas T. Electronic gaming machines: are they the 'crack-cocaine' of gambling? Addiction. 2005;100(1):33-45.
7. Livingstone C, Woolley R. Risky business: A few provocations on the regulation of electronic gaming machines. International gambling studies. 2007;7(3):361-76.
8. The Gambling Commission. Types of gambling and gambling involvement 2018 [Available from: https://www.gamblingcommission.gov.uk/statistics-and-research/publication/types-of-gambling-and-gambling-involvement.
9. Farewell VT, Long DL, Tom BDM, Yiu S, Su L. Two-Part and Related Regression Models for Longitudinal Data. Annu Rev Stat Appl. 2017;4:283-315.
10. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70(1):41-55.
11. Heckman JJ. Sample selection bias as a specification error. Econometrica: Journal of the econometric society. 1979:153-61.
12. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate behavioral research. 2011;46(3):399-424.
13. Browne M, Rawat V, Newall P, Begg S, Rockloff M, Hing N. A framework for indirect elicitation of the public health impact of gambling problems. BMC Public Health. 2020;20:1-14.
14. Kräplin A, Behrendt S, Scherbaum S, Dshemuchadse M, Bühringer G, Goschke T. Increased impulsivity in pathological gambling: Considering nicotine dependence. Journal of clinical and experimental neuropsychology. 2015;37(4):367-78.
15. Harrison GW, Lau MI, Ross D. The risk of gambling problems in the general population: A reconsideration. Journal of Gambling Studies. 2020;36(4):1133-59.
16. Xiao L, Newall PWS, James RJE. To screen, or not to screen: An experimental comparison of two methods for correlating video game loot box expenditure and problem gambling severity. OSF Preprints. 2023.
17. Harrison GW, Jessen LJ, Lau M, Ross D. Gambling problems in the general Danish population: Survey evidence. 2016.
18. Suhonen N, Saastamoinen J, Forrest D, Kainulainen T. Does IQ predict engagement with skill-based gambling? Large-scale evidence from horserace betting. Journal of Behavioral Decision Making. 2022.

19. Steiger JH. Understanding the limitations of global fit assessment in structural equation modeling. Personality and Individual Differences. 2007;42(5):893-8.
20. Ludwig M, Kraus L, Müller S, Braun B, Bühringer G. Has gambling changed after major amendments of gambling regulations in Germany? A propensity score analysis. Journal of Behavioral Addictions. 2012;1(4):151-61.
21. Nichols MW, Tosun MS, Yang J. The fiscal impact of legalized casino gambling. Public Finance Review. 2015;43(6):739-61.
22. Nichols MW, Tosun MS. The impact of legalized casino gambling on crime. Regional Science and Urban Economics. 2017;66:1-15.
23. Sabbagh JJ. Drunk gamblers: the impact of legalized gambling on alcohol consumption. 2011.
24. Gainsbury SM, Liu Y, Russell AMT, Teichert T. Is all Internet gambling equally problematic? Considering the relationship between mode of access and gambling problems. Computers in Human Behavior. 2016;55, Part B:717-28.
25. Papineau E, Lacroix G, Sevigny S, Biron J-F, Corneau-Tremblay N, Lemétayer F. Assessing the differential impacts of online, mixed, and offline gambling. International Gambling Studies. 2018;18(1):69-91.
26. Paton D, Williams LV. Do new gambling products displace old? Evidence from a postcode analysis. Regional Studies. 2013;47(6):963-73.
27. Vita A, Bastiani L, Turrina C, Benedetti E, Bergamini A, Molinaro S. At-risk gambling in patients with severe psychiatric illness and in community subjects matched for age and sex. Psychiatry Research. 2021;304:114142.
28. Soberay AD. Comparing short-term outcomes of three problem gambling treatments: A multi-group propensity score analysis: University of Denver; 2015.
29. Smith D, Harvey P, Humeniuk R, Battersby M, Pols R. Effects of affective and anxiety disorders on outcome in problem gamblers attending routine cognitive–behavioural treatment in South Australia. Journal of Gambling Studies. 2015;31:1069-83.
30. Hoffmann JP. Religion and problem gambling in the US. Review of Religious Research. 2000:488-509.
31. Bhatti JA, Thiruchelvam D, Redelmeier DA. Traumatic brain injury as an independent risk factor for problem gambling: a matched case-control study. Social Psychiatry and Psychiatric Epidemiology. 2019;54(4):517-23.
32. Cerasa A, Lofaro D, Cavedini P, Martino I, Bruni A, Sarica A, et al. Personality biomarkers of pathological gambling: A machine learning study. Journal of neuroscience methods. 2018;294:7-14.
33. Wenz M. Matching estimation, casino gambling and the quality of life. The Annals of Regional Science. 2008;42:235-49.
34. Bhat VN. Anxiety about Standard of Living and Expenditures on Lottery Tickets.
35. Hou J, Kim K, Kim SS, Ma X. Disrupting unwanted habits in online gambling through information technology. Journal of Management Information Systems. 2019;36(4):1213-47.
36. Brosowski T, Olason DT, Turowski T, Hayer T. The gambling consumption mediation model (GCMM): A multiple mediation approach to estimate the association of particular game types with problem gambling. Journal of Gambling Studies. 2021;37(1):107-40.

37. Ji Q, Quan X, Yin H, Yuan Q. Gambling preferences and stock price crash risk: Evidence from China. Journal of Banking & Finance. 2021;128:106158.
38. Ferris J, Wynne H. The Canadian problem gambling index. Ottawa, ON: Canadian Centre on Substance Abuse. 2001.
39. Lesieur H, Blume S. The South Oaks Gambling Screen (SOGS): a new instrument for the identification of pathological gamblers. The American Journal of Psychiatry. 1987;144(9):1184-8.
40. Gerstein D, Hoffmann J, Larison C, Engelman L, Murphy S, Palmer A, et al. Gambling impact and behavior study. Report to the National Gambling Impact Study Commission National Opinion Research Center at the University of Chicago, Chicago. 1999:b43.
41. Toce-Gerstein M, Gerstein DR, Volberg RA. A hierarchy of gambling disorders in the community. Addiction. 2003;98(12):1661-72.
42. Wardle H, Moody A, Spence S, Orford J, Volberg R, Jotangia D, et al. British Gambling Prevalence Survey 2010. London, UK: The Stationery Office; 2011.
43. Wardle H, Sproston K, Orford J, Erens B, Griffiths M, Constantine R, et al. British Gambling Prevalence Survey 2007. London, UK: The Stationery Office; 2007. Report No.: 0117037923.
44. Craig R, Mindell J. Health Survey for England, Volume 2: Methods and Documentation. Leeds, UK: The Health and Social Care Information Centre; 2013.
45. NatCen Social Research, University College London. Health Survey for England 2015: Methods. London: Health and Social Care Information Centre; 2016.
46. NatCen Social Research, University College London. Health Survey for England 2016: Methods. London: Health and Social Care Information Centre; 2017.
47. NatCen Social Research, University College London. Health Survey for England 2018: Methods. London: Health and Social Care Information Centre; 2019.
48. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Research synthesis methods. 2010;1(2):97-111.
49. Cheung MW. A Guide to Conducting a Meta-Analysis with Non-Independent Effect Sizes. Neuropsychol Rev. 2019;29(4):387-96.
50. Thomas SL, Bestman A, Pitt H, Cassidy R, McCarthy S, Nyemcsok C, et al. Young people's awareness of the timing and placement of gambling advertising on traditional and social media platforms: a study of 11–16-year-olds in Australia. Harm Reduction Journal. 2018;15(51):1-13.
51. O'Loughlin I, Blaszczynski A. Comparative effects of differing media presented advertisements on male youth gambling attitudes and intentions. International Journal of Mental Health and Addiction. 2018;16:313-27.
52. Hilbrecht M, Baxter D, Abbott M, Binde P, Clark L, Hodgins DC, et al. The Conceptual Framework of Harmful Gambling: A revised framework for understanding gambling harm. Journal of Behavioral Addictions. 2020;9(2):190-205.
53. Bowden-Jones H. Commentary on Petry et al. (2013): Actus Reus – why it matters to pathological gambling treatment. Addiction. 2013;108(3):582-3.
54. Sharman S, Aitken M, Clark L. Reply to Griffiths: The Relationship Between Gambling and Homelessness. Journal of Gambling Studies. 2014:1-3.
55. Johnson DR. Using weights in the analysis of survey data. Presentation prepared for the Population Research Institute, Pennsylvania State University, November. 2008.

56. Currie SR, Hodgins DC, Casey DM. Validity of the problem gambling severity index interpretive categories. Journal of Gambling Studies. 2013;29(2):311-27.
57. Browne M, Volberg R, Rockloff M, Salonen AH. The prevention paradox applies to some but not all gambling harms: Results from a Finnish population-representative survey. Journal of Behavioral Addictions. 2020;9(2):371-82.
58. Bradley A, James RJE. Defining the key issues discussed by problematic gamblers on web-based forums: a data-driven approach. International Gambling Studies. 2020:1-15.

SECTION 7: APPENDICES

The full analytic code and output are available at the Open Science Framework at https://osf.io/xqaw5/?view_only=9ae3432b55b24c50bade971bb6af612f. The data is available from the UK Data Archive at the following sources:

British Gambling Prevalence Survey 2007:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5836

British Gambling Prevalence Survey 2010:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=6843

Scottish Health Survey 2012:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7417

Scottish Health Survey 2015:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8100

Scottish Health Survey 2016:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8290

Health Survey for England 2012:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7480

Health Survey for England 2015:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8280

Health Survey for England 2016:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8334

Health Survey for England 2018:
https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=8649