Journal of Gambling Studies (2025) 41:1419–1436
https://doi.org/10.1007/s10899-025-10419-6

ORIGINAL PAPER

Across the Bettor-Verse: an Open Banking Perspective on Gambling in the United Kingdom

Kasra Ghaharian[1] · Jay Peterson[2] · Mana Azizsoltani[1,2] · Richard J. Young[1] · Eric R. Louderback[3,4]

Accepted: 11 July 2025 / Published online: 5 August 2025

Abstract

Open banking provides an opportunity to leverage bank transaction data to support gambling harm prevention by providing an objective and comprehensive view of individuals' gambling activities. In this study, we present a large-scale descriptive analysis of gambling transaction activity from an open banking dataset, representing a sample of 1,030,788 users of a credit scoring service in the UK for 2022. The median gambler ($n = 650{,}502$) transacted with 2 merchants, made 14 deposits, had a typical debit of £10.00, a net loss of £95.00, and deposited 0.53% of income. Consistent with prior gambling literature, our analyses revealed small groups of gamblers (i.e., about 3%) who displayed disproportionately higher gambling activity based on their frequency of debits, total amount debited, and net loss. Males and younger people were more likely to fall into these high-involvement groups, and group members transacted with significantly more gambling merchants (5 to 6 merchants vs. the median of 2). This study establishes a baseline census of gambling activity from open banking data, offering insights to inform researchers and policymakers about opportunities for harm prevention.

Keywords Gambling · Gambling harm · Transaction data · Open banking

Analyzing gambling behavioral data to build models that identify individuals at risk of harm has been a focus of both academic research and industry efforts for harm prevention. The typical approach involves using objective tracking data — most commonly from

✉ Kasra Ghaharian
kasra.ghaharian@unlv.edu

[1] International Gaming Institute, University of Nevada, Las Vegas, NV, USA

[2] William F. Harrah College of Hospitality, University of Nevada, Las Vegas, NV, USA

[3] Chestnut Hill Analytics Insights, Boston, Massachusetts, USA

[4] Harvard Medical School, Boston, MA, USA

online operators, electronic gambling machines, and other digital platforms — to understand behaviors such as how often people gamble and how much money they spend, as well as more nuanced markers of harm like chasing losses or variability in wagering (see reviews in Delfabbro et al., 2023 and Ghaharian et al., 2023b).

These existing efforts to identify gambling-related harm have predominantly relied on information collected from a single gambling operator, which inherently provides a limited view of an individual's overall gambling activity. While this approach has been useful, and has had notable impacts, it is also subject to limitations. Chief among these is the fact that analyzing behavioral data from a single operator fails to capture the full breadth of an individual's gambling involvement. For example, a prior report revealed that high-spending gamblers hold accounts with six operators on average (McNair et al., 2021). In addition to these structural limitations, most gambling research has historically relied on self-reported data, which is subject to recall biases (Heirene et al., 2022) and challenges recruiting representative participants (Pickering & Blaszczynski, 2021).

Accordingly, stakeholders have begun to explore methodologies that capture a broader range of individuals' gambling activities, which could be more effective in identifying those at-risk (Zendle & Newall, 2024a). Efforts to create a "single-customer view", by pooling data across operators, face significant implementation challenges and raise questions about the optimal governance of such a data source (Newall & Swanton, 2024). Using bank transaction data is an alternative option, as a range of transactional activity is available including gambling activities across different merchants (Newall & Swanton, 2024; Swanton et al., 2019). But implementing harm prevention interventions directly through individual banks presents significant challenges. Such an approach would require extending the scope of harm prevention interventions and regulation beyond gambling operators to include financial institutions as well. Additionally, and similar to the challenge of pooling data across gambling operators, each bank would need to adopt and support these interventions, requiring coordinated efforts and buy-in across multiple institutions. However, emerging financial technology (fintech; Pandey et al., 2024) offers a potential solution by allowing aggregation of consumer activities across multiple banks and accounts, such as checking, savings, and credit cards.

Open banking, which facilitates data aggregation across individuals' bank accounts via account holder consent, presents a promising tool for gambling harm prevention. A key strength in this context is its ability to offer a holistic view of an individual's gambling transactions across the operators they gamble with. However, the technology has yet to achieve widespread adoption. In the United Kingdom (UK), open banking is relatively more established, and is being explored to support financial checks in online gambling (Gambling Commission, 2024). In the United States (US), the Consumer Financial Protection Bureau has proposed an "open banking rule" (Gillison, 2024), while in Canada, legislation has been formally approved by the government to develop a framework for "consumer-driven banking" (Duncan, 2024).

The Present Study

As both gambling access and fintech continue to develop, a better understanding of how aggregated financial data can support gambling harm prevention becomes increasingly relevant. Despite this potential, academic research investigating this data source remains limited. Seminal work by Muggleton et al. (2021) first demonstrated associations between gambling spend and a range of negative financial and health outcomes using transaction data from a single UK retail bank. More recent studies have built on this foundation: Zendle and Newall (2024a) linked open banking data to self-reported measures of gambling harm in a small UK sample ($n = 424$), while Marionneau et al. (2024) analyzed financial transactions among over 23,000 Finnish applicants to a debt consolidation service. These studies underscore the value of objective financial data and the need for further research in this area.

Accordingly, in the present paper we present a large-scale descriptive analysis, offering foundational insights into gambling transactional activity in a sample of open banking data from 1,030,788 individuals in the UK for the 12-month period during 2022 (i.e., January 1– December 31). We explore gambling behaviors for the sample as a whole, amongst subgroups according to levels of gambling activity, and across subgroups according to sociodemographic information provided in the data. This study adhered to open science practices; our research questions, hypotheses, methods, and analysis plan were pre-registered (https://osf.io/8tudr). Based on our primary pre-registered research questions, we hypothesized the following:

RQ1: What proportion of our sample engage in gambling activities?

In September 2022, the UK Gambling Commission reported overall participation in any form of gambling during the last four weeks (online or land based) as 44% (Gambling Commission, 2022). The participation rate dropped to 28% when those who only played the National Lottery were excluded. A limitation of our dataset is that there is an inability to observe gambling cash transactions (e.g., cash used in land-based casinos). Notwithstanding, we predicted that the proportion of people within our sample who engage in gambling would be comparable to the participation rates reported by the UK Gambling Commission in their 2022 survey. This expectation was based on the assumption that individuals who gamble tend to do so with some regularity, such that a 12-month observation period (as used in our study) and a 4-week self-reported window (as used in the UK Gambling Commission's survey) may capture a broadly overlapping group of active gamblers.

H1: Our sample will exhibit a participation rate in gambling that is comparable to reported estimates.

RQ2: What are the transactional gambling behaviors of our sample?

We did not make specific hypotheses for our second research question, which was largely exploratory. In this context, we use the term "transactional gambling behaviors" to refer to observable financial activity - such as the number, value, and frequency of gambling-related debits and credits - captured in bank transaction data. These measures reflect patterns of gambling engagement and do not capture psychologically informed constructs such as urges

or motivations. Based on the nature of these indicators, we expected to observe some degree of skewness in their distribution. This expectation was informed by prior studies using similar datasets. For example, Zendle and Newall (2024a) combined Problem Gambling Severity Index (PGSI) scores with open banking data for a sample of people who gambled in the past year ($n=424$). They found that the total number and average size of gambling transactions, as well as net spend, increased with PGSI score. Most of the sample (54.0%) scored 0 on the PGSI, with under 20% scoring above 5. And research by Muggleton et al. (2021), using transaction data from a single UK bank, found a large difference between the mean and median values for the number of transactions and total spent gambling.

RQ3: Is it possible to distinguish subgroups of highly-involved gamblers whose activity is disproportionately high?

Prior studies using gambling behavioral tracking data (LaBrie et al., 2007; Nelson et al., 2021), or transactional information from payments service providers (Ghaharian et al., 2023a) and open banking data (Zendle & Newall, 2024a), have demonstrated that samples of gamblers will typically exhibit subgroups of "highly-involved gamblers". Therefore, we predicted that we would be able to identify subgroups of highly-involved gamblers whose gambling transaction behaviors were disproportionately high. LaBrie et al. (2007) and Nelson et al. (2021) found that 1% of their sample of online sports bettors exhibited discontinuously high gambling behaviors, whereas Nelson et al. (2019) also revealed a top 1% group in their sample of daily fantasy sports players. Although the dataset we analyzed differed from these studies, we believed that these prior findings provided adequate rationale for our following hypotheses:

H2a: We will be able to distinguish subgroups of highly-involved gamblers whose activity (in terms of the number of gambling debits, the total debited amount, and the net balance) is disproportionately high.

H2b: Approximately 1% of our sample will exhibit disproportionately high gambling transaction activity.

We also hypothesized that the highly-involved subgroups would be more likely to conduct transactions across multiple gambling merchants. We base this hypothesis on prior work from Nelson et al. (2021), who found partial support for their hypothesis that heavily involved sports bettors would be more likely to engage in other forms of gambling such as poker and casino games. Additionally, prior work by the Behavioral Insights Team (McNair et al., 2021), using data from a single UK bank found that higher spenders who gamble, on average, hold accounts with six merchants.

H3c: Highly-involved gamblers will be more likely to gamble with multiple merchants.

Methods

The University of Nevada, Las Vegas institutional review board reviewed the authors' plans and confirmed that the study did not meet the definition of research with human subjects.

Data Description

We worked with a financial services provider registered in the UK to gain access to the original dataset — Department of Trust (DoTrust; https://www.dotrust.co.uk/). DoTrust pur chased the dataset from a third-party commercial vendor that offers credit scoring services. The underlying data originated from UK consumers who had voluntarily provided explicit consent to this vendor, via regulated Open Banking protocols, to share their bank transaction history in exchange for accessing the vendor's credit scoring services. Open Banking frameworks in the UK allow consumers to grant time-limited authorization for regulated third-party providers to securely access their account information using standardized Application Programming Interfaces (APIs). This vendor processed and anonymized the consented transaction data for their own commercial purposes, as well as resale.

DoTrust made the data available via an AWS bucket, which consisted of twelve csv files (one file for each month of 2022, each approximately 60GB in size). A data dictionary was also provided that described the contents of each column. We downloaded these files to create the analytic dataset described below.

The analytic dataset used for this research included financial transaction records for the year 2022 from 1,030,788 UK-based users of the credit scoring vendor[1]. Our first step in constructing the analytic dataset was to include only users who exhibited transaction behavior for each month of the year. We performed this preprocessing step due to the nature of open banking, which allows users to opt-out of sharing their data. This meant that, in the originally provided data, some users may have only exhibited transactions for a portion of the year. For example, an individual might opt-in during July and opt-out in October, meaning we would only observe their data in our files for January through to their opt-out date.

Measures

To address our pre-registered objectives, we used self-reported demographic information included in the underlying data for each user (5-year age bands, £5,000 salary bands, and gender [M = Male, F = Female, U = Unknown]) and derived key gambling behavioral measures. Each record in the analytic dataset included detailed information such as the transaction amount, whether it was a debit or credit, and its assigned "purpose". This purpose categorization (e.g., home, mobile, salary, bills, entertainment, gambling) was determined using a proprietary algorithm developed by the credit scoring vendor. While the specific logic of the algorithm constitute confidential vendor information and are therefore not available to the research team, its function is to systematically assign each transaction a purpose category based on merchant data. Given our focus on describing gambling activity, we utilized the vendor-assigned "gambling" category[2] to calculate an aggregated set of gambling transaction variables (for the entire year of 2022) for each user in the analytic dataset (Table 1).

[1] The analytic dataset represents a subset of the original data provided to us. Its construction involved several stages of preprocessing and manipulation, summarized here and detailed in our pre-registration.

[2] Among transactions within the gambling category, there were over 30 unique gambling merchants, including the top 5 operators in the UK market.

Table 1 Variable definitions

Variable	Definition
Total debited amount (£)	Total amount of all gambling debits.
Debits (n)	Count of debit gambling transactions.
Average debit amount (£)	Total debited amount divided by Debits.
Total credited amount (£)	Total amount of all gambling credits.
Credits (n)	Count of credit gambling transactions.
Average credit amount (£)	Total credited amount divided by Credits.
Duration	The difference between the first and last day with a gambling transaction, inclusive.
Transaction days (n)	The number of days with a gambling transaction.
Frequency	The number of days with a gambling transaction divided by Duration.
Merchants (n)	Count of unique gambling merchants debited to by a user during the sample time period.
Average balance (£)	The mean of the values in a column indicating the user's account balance for each transaction.
Overdraw (n)	Count of instances when the account balance transitions from a positive to a negative balance due to a gambling debit.
Overdraft (n)	Count of instances when overdraft (i.e., when account balance is negative) is used to make a gambling debit.
Net balance (£)	Total debited amount minus Total credited amount.
% of income	Total debited amount divided by the midpoint of the user's reported salary band.

Data Analysis

To test the hypothesis for RQ1 — the proportion of our sample that engages in gambling — we counted the number of unique users in the analytic dataset who made at least one gambling transaction. Then, we calculated what proportion of our total sample this represented. We also performed this same analysis, but excluding National Lottery transactions (by excluding transactions based on the associated merchant name).

For RQ2, we conducted descriptive analyses to describe the transactional gambling behaviors of our sample. We first created a subset of the analytic dataset to include only the users who made at least one gambling debit during the sample period (i.e., purpose equal to "gambling"). Then, for all variables we computed five-number summary statistics (minimum, lower quartile, median, upper quartile, maximum), as well as (where applicable) the means, standard deviations, counts, and percentages.

To explore associations between variables we conducted bivariate Spearman's correlations. To examine whether gambling transaction variables differed by demographic characteristics, we conducted Kruskal-Wallis (KW) tests. KW tests were used to account for the anticipated skewness in the data.

To test the hypotheses for RQ3, we created percentile plots for debits, total debited amount, and net balance for all gambling transactions. Then, similar to the examination of a scree plot in factor analysis, we identified the points of disproportionality for each variable. Based on this point we created groups of "highly-involved gamblers". Then, using Chi-square tests, we assessed the differences in distribution across gender, salary bands, and age bands between variables for the highly-involved gamblers and the rest of the sample, i.e., the "not highly-involved gamblers". This included a test for the difference in the number of merchants transacted with (H3c).

Results

Gambling Participation - Research Question 1

The number of individuals who made at least one gambling transaction during our study period of 2022 was 659,117. This represented a 63.94% participation rate. Excluding lottery transactions, the participation rate was 49.80% (n = 513,426).

Gambling Activity - Research Question 2

For our analytic subset of gamblers (n = 650,502[3]), 59.96% were male, 37.38% were female, and 2.66% had an undefined gender. 50.04% of the sample were between 18 and 34 years of age (n = 325,549). The most common salary band was £20,000 - £30,000, representing 201,060 users (31.71%). Visual summaries of the distribution of users by gender, age band, and salary band are presented in Figures S1–S3.

In line with our expectations for RQ2, most gambling transaction variables exhibited considerable right (i.e., positive) skew, with much larger means than medians. Descriptive statistics for all variables are provided in Table 2.

The median gambler had a duration of 289 days between their first and last gambling transaction during the 365-day period (mean = 230.64). The median gambler made gambling transactions on 13 days (mean = 30.50), representing 8.91% of their duration (mean = 23.40%). The median gambler transacted with 2 different merchants (mean = 2.41). They made 14 debits (mean = 67.99) with a median amount per debit of £10.00 (mean = £17.22), and 1 credit (mean = 8.74) with a median amount per credit of £30.55 (mean = £92.56). The median net loss was £95.00 (mean = £507.01) and total debited amount was £156.05 (mean = £1,422.63). The median total debited amount as a percentage of income (using the midpoint of the salary band as a proxy for the annual income for each user) was 0.53% (mean = 5.93%).

The median account balance was £298.51 (mean = £1,114.92). Across all gamblers, 11.09% of users overdrew their account (n = 72,119) at least once (made a gambling transaction that transitioned the account balance from positive to negative balance), and 27.30% of users used their overdraft to make a gambling debit (n = 177,559) at least once.

[3] We omitted 8,615 users from the analytic sample, including 7,879 users who made no debit transactions, 122 whose total debited amount was zero, and 614 users whose transaction activity was separated into two time periods; this lapse may have been due to these users opting out and then opting back in again.

Table 2 Descriptive statistics for all gamblers

Variable	Min	Q2	Median	Q3	Max	Mean	SD
Total debited amount (£)	0.01	40.00	156.05	593.00	1,678,795.00	1422.63	6,499.72
Debits (n)	1.00	4.00	14.00	46.00	11,722.00	67.99	202.36
Average debit amount (£)	0.01	7.50	10.00	15.90	111,949.00	17.23	151.39
Total credited amount (£)	0.00	0.00	10.00	175.47	2,109,737.83	915.63	6,648.31
Credits (n)	0.00	0.00	1.00	5.00	2,685.00	8.74	36.45
Average credit amount (£)	0.01	11.00	30.55	77.16	500,015.00	92.56	1,522.23
Duration	1.00	113.00	289.00	346.00	365.00	230.64	133.32
Transaction days (n)	1.00	4.00	13.00	37.00	365.00	30.50	44.47
Frequency	0.01	0.04	0.09	0.25	1.00	0.24	0.32
Merchants (n)	1.00	1.00	2.00	3.00	22.00	2.41	1.83
Average balance (£)	−420,563.72	35.97	298.51	785.44	2,687,929.91	1,114.92	7,972.41
Overdraw (n)	0.00	0.00	0.00	0.00	136.00	0.39	2.22
Overdraft (n)	0.00	0.00	0.00	1.00	4,997.00	9.07	60.86
Net balance (£)	−2,109,147.93	20.00	95.00	339.00	242,402.61	507.01	4,728.83
% of income	0.00	0.14	0.53	2.06	4,117.68	5.93	32.14

Correlation analyses between all gambling transaction variables are provided in Figure S4. All correlations between salary band and each gambling transaction variable, as well as age band and each gambling transaction variable, were statistically significant (see Tables S1 and S2). The results of the KW tests for age, gender, and salary also indicated significant differences across all variables ($p < 0.001$ after Bonferroni correction) (see Tables S3-5). We highlight a few statistics here but present all variable means and medians by age and salary bands in the supplementary material (see Tables S6-15). The median and mean for the total debited amount and net balance increased with salary. The median number of debits increased but tapered down for the highest two salary bands, whereas the mean for this variable decreased as salary increased. The amount debited as a percentage of income decreased as salary increased. Notably, the median user in the lowest salary band (< £10,000) debited 2.35% of their income (mean = 28.40%). The mean number of debits and total amount debited appeared highest for the 30–34, 35–39, and 40–44 age bands. Males exhibited higher mean values than females across all transaction variables. Males, on average, made 76 debits across 34 days for a net loss of £603.34, whereas females, on average, made 57 debits across 25 days for a net loss of £358.18. In terms of the amount debited as a percentage income, the median female user debited 0.43% (mean = 5.05%) and the median male 0.62% (mean = 6.50%). We note that while these results were statistically significant, interpretation should be performed with caution due to small effect sizes.

Highly-involved Gamblers - Research Question 3

Percentile plots for each of the three variables — (1) debits, (2) total debited amount, and (3) net balance for all gambling transactions — are displayed in Fig. 1.

Supporting hypothesis H2a, we observed a subgroup of users who demonstrated disproportionately higher gambling involvement across each of the three variables. We did not find definitive support for hypothesis H2b, as the points of discontinuity for all three variables appeared to occur beyond the top 1% of users.

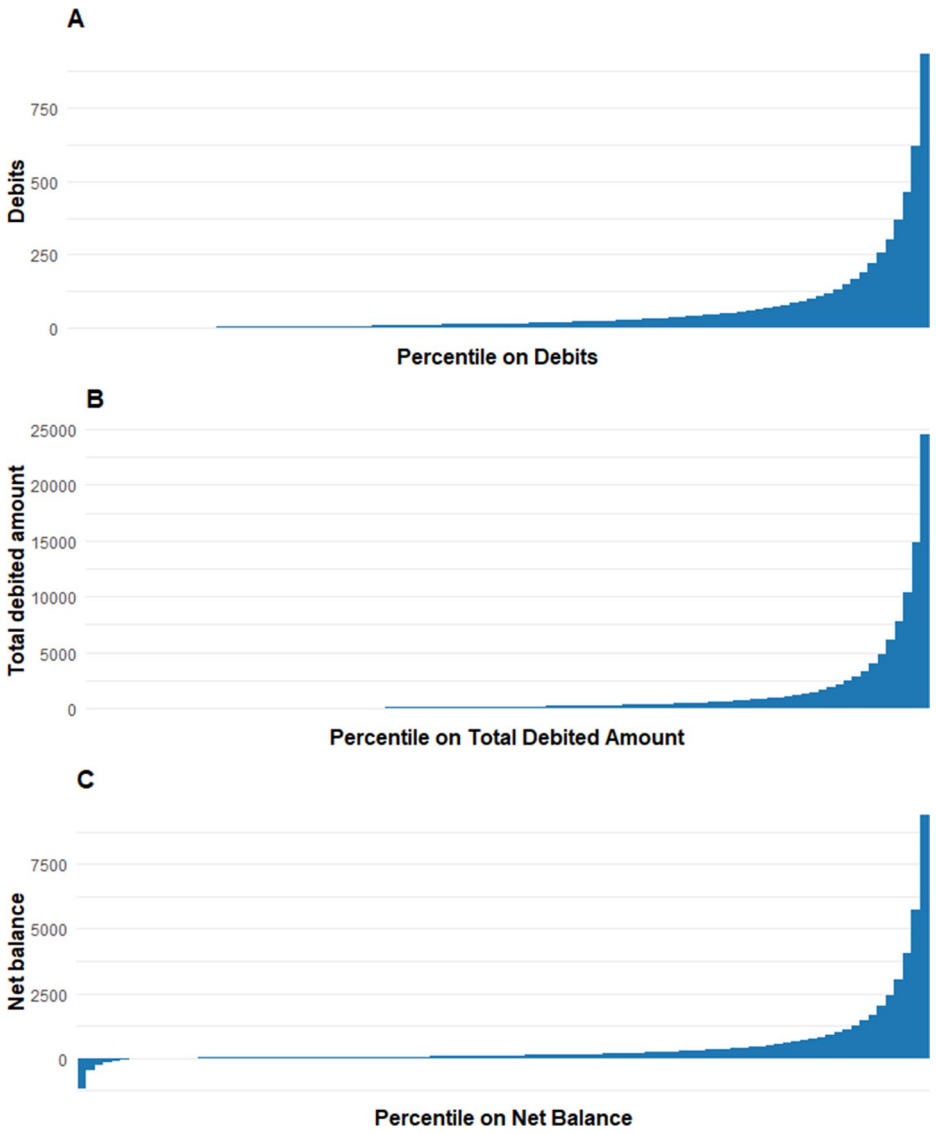

Fig. 1 Percentile plots for debits, total debited amount, and net balance. *Note*. **A** = percentile distribution for Debits (*n*); **B** = percentile distribution for Total debited amount (£); **C** = percentile distribution for Net balance (£).

The exact point where the distinction between percentile groups became more pronounced was not decidedly clear; however, we defined it as the top 3% to create three highly-involved groups: (1) "high-frequency" (HF) users (top 3% according to debits), (2) "high-volume" (HV) users (top 3% according to total debited amount), and (3) "high-losses" (HL) (top 3% according to net balance). Additionally, we identified a fourth group in the percentile plot for net balance: (4) the bottom 1%, representing "big winners". The

Table 3 Highly-involved group(s) overlap

Group	HF	HV	HL	HF + HV	HF + HL	HV + HL	HF + HV + HL
HF (n = 19,499)	6,202 (31.8%)	-	-	3,559 (18.3%)	1,399 (7.2%)	-	8,339 (42.8%)
HV (n = 19,516)	-	2,562 (13.1%)	-	3,559 (18.2%)	-	5,056 (25.9%)	8,339 (42.8%)
HL (n = 19,516)	-	-	6,202 (31.8%)	-	1,399 (7.2%)	5,056 (25.9%)	8,339 (42.8%)

HF high-frequency; *HV* high-volume; *HL* high-losses. Each row represents the total membership of a given highly-involved group. The cells indicate the number of users within that group who also belong to one or more of the other groups, along with the corresponding row-wise percentage

highly-involved groups were not mutually exclusive and Table 4 provides information about the group overlap.

Differences by Sociodemographic Variables

We examined gambling behavior for all three gambling involvement variables — debits, total debited amount, and net balance — to understand how these behaviors differed across salary bands, gender, and age bands. Using Chi-square tests, we found statistically significant differences in demographic representation within the top 3% of users for each variable ($p < 0.001$ for all tests). In Fig. 2, we present the percentages of high-involvement users against users who do not belong to any of these groups within each demographic category, to visually depict the variations across salary bands, gender, and age bands.

In terms of age band, users in the 30–34 and 35–39 bands exhibited notable levels of greater representation across all highly-involved groups. For all three variables, these bands made up ~23% and ~20% for the highly-involved groups vs. ~19% and ~16% respectively for the rest of the sample. Users in the lower salary ranges were overrepresented in the HF group (e.g., 34.88% vs. 31.61% for the £20,000 - £30,000 band). Males were significantly overrepresented among high-involvement users across all three variables. For the "big winners", females were relatively more represented among this group, whereas males were less represented.

Differences in Gambling Transaction Variables

We assessed differences in gambling transaction behavior between users in each highly-involved group and users who did belong to any of these groups (not-HI). Table 4 presents the median and mean for each variable and group.

Both the mean and median values for all variables were markedly different for the highly-involved groups compared to the not-HI group. The median highly-involved users' duration of activity spanned almost the entire sample period, compared to around three quarters of the period for the median not-HI user. The median user in the highly-involved groups debited at least 50% of their reported salary for the sample period, as compared to 2% for not-HI users. The median net loss for the HF group was £4,023.29, for the HV group £6,454.58, and for the HL group was £7,176.35. In comparison, the median not-HI user's net loss was £85.00.

Table 4 Descriptive statistics by group

	HF (n = 19,499)		HV (n = 19,516)		HL (n = 19,516)		Not-HI (n = 618,663)	
Variable	Mean	Median	Mean	Median	Mean	Median	Mean	Median
Total debited amount (£)	20,122.83	13,093.92	26,172.15	18,444.10	22,139.32	15,093.31	532.71	140.00
Debits (n)	950.96	740.00	789.74	597.50	73.68	464.00	36.72	13.00
Average debit amount (£)	22.48	15.39	81.03	34.17	79.87	32.97	14.84	10.00
Total credited amount (£)	13,804.17	8,171.17	17,812.27	12,141.71	11,957.28	6,448.73	333.97	10.00
Credits (n)	125.63	82.00	111.56	64.00	74.10	32.00	4.59	1.00
Average credit amount (£)	170.03	101.04	360.26	188.40	294.05	168.77	76.04	28.00
Transaction days (n)	185.31	182.00	152.06	148.00	136.97	130.00	24.44	12.00
Frequency	0.53	0.52	0.45	0.44	0.40	0.39	0.23	0.08
Merchants (n)	6.20	6.00	5.74	5.00	5.36	5.00	2.25	2.00
Average balance (£)	580.91	307.51	1,250.55	491.31	1,278.25	453.18	1,117.72	293.46
Overdraw (n)	4.53	1.00	4.55	1.00	4.11	1.00	0.21	0.00
Overdraft (n)	122.41	1.00	96.98	0.00	84.59	0.00	5.02	0.00
Net balance (£)	6,318.65	4,023.29	8,359.87	6,454.58	10,182.04	7,176.35	198.74	85.00
% of income	89.31%	50.75%	105.93%	64.44%	88.08%	49.72%	2.25%	0.48%

All variables were tested for differences in group means. All p-values were <0.001. *HF* high-frequency; *HV* high-volume; *HL* high-losses; *Not-HI* not highly-involved

In support of hypothesis H3c, users in the highly-involved groups transacted with a greater number of gambling merchants. The median HF user transacted with 6 merchants, and both the median HV user and median HL user transacted with 5. In contrast, the median not-HI user transacted with only 2 merchants.

Unplanned Exploratory Analysis

We conducted an exploratory analysis to examine the concentration of deposited amounts among highly-involved users. By combining the three highly-involved groups and removing duplicates, we identified 31,839 unique highly-involved users (4.89% of the 650,502 users). These users accounted for £595,856,139, or 64% of the total deposited amount (£925,426,468).

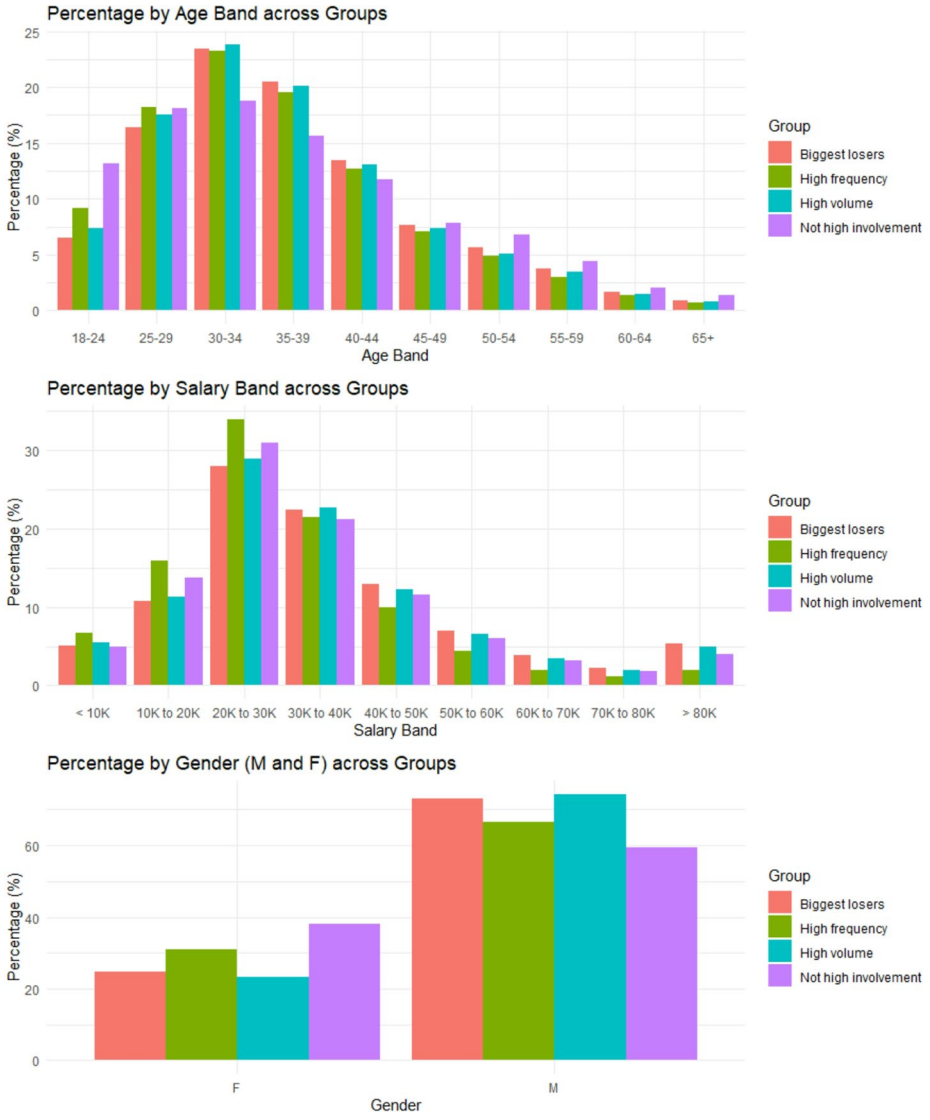

Fig. 2 Percentages for demographic variables by group

Discussion

In this study, we examined the gambling transaction activity across a large sample of open banking data, representing 1-year of transactions for over 1 million users of a credit scoring service. While most users in this sample who engaged in gambling displayed modest levels of involvement, our analysis revealed a considerable skew in transaction activity. Our analysis provides insights into the diverse levels of gambling engagement within this sample and highlights the potential of bank transaction data in understanding varying degrees of gambling behavior.

With respect to RQ1, the proportion of users engaged in gambling in our sample was 63.94%, significantly exceeding the UK Gambling Commission's 4-week self-reported rate of 44%. Excluding National Lottery transactions, our rate was 49.80%, compared to 28%. Some of this discrepancy likely stems from timeframe differences (12-month vs. 4-week[4]) and methodological differences (objective vs. self-report). However, demographic differences also likely contribute. Our sample comprised 59.96% male users, versus approximately 49% male in the general UK population (Office for National Statistics, 2023). Additionally, 50% of our sample were aged 18–34, whereas the median age of the UK population is 40.7 years (Office for National Statistics, 2024). The discrepancy may also reflect the nature of our dataset (comprised of users engaged in digital financial management) among whom online gambling may be more prevalent due to higher digital literacy.

Consistent with our expectations for RQ2 based on prior research, our results indicated a right-skewed distribution in the gambling transaction variables, suggesting at least some portion of users engaged in disproportionately high levels of gambling activity. This finding aligns with Muggleton et al. (2021), who observed similar skewness in gambling behaviors using UK bank transaction data, as well as other research analyzing gambling behavioral data from a variety of sources including daily fantasy sports players (Nelson et al., 2019, online sports bettors (Nelson et al., 2021), and online casino bettors (Edson et al., 2022).

Our descriptive analysis of the transactional gambling behaviors of our sample also revealed differences across demographic groups. We noted higher involvement amongst males and the younger to middle-aged groups, which is consistent with extant research. Humphreys and Perez (2012) found that online gamblers, specifically in the UK, were predominantly younger males, and participation rates amongst this demographic appears to be higher, in general, according to much of the research (Dowling et al., 2015; Gainsbury et al., 2015; McCormack et al., 2013). Additionally, we observed that these demographic groups — as well as users in lower salary bands — debited a greater proportion of their reported income to gambling merchants.

Our analysis contributes to the current discourse surrounding *affordability*. For our sample, the median total amount debited to gambling merchants as a percentage of income was 0.53% (mean = 5.93%), suggesting that most users appear to gamble within their financial means. Gambling within one's means is a hallmark of responsible gambling messaging (Newall et al., 2023), yet the specific thresholds that define this behavior are often unclear or not explicitly stipulated. 6,674 (1.03%) users in our dataset debited amounts exceeding their reported income, indicating they were not gambling within their means. But even gambling a smaller proportion of income could still pose risks for certain individuals. For example, lower-risk gambling guidelines in Canada include the recommendation to "gamble no more than 1% of household income" (Canadian Centre on Substance Use and Addiction, 2024). In our sample, 234,651 users (36.07%) debited more than 1% of their reported income. Similarly, the Netherlands has recently introduced a guideline limiting gambling deposits to 30% of net income (Fletcher, 2024). In our dataset, 26,019 users (4.00%) debited more

[4] We calculated the 4-week participation in our sample post-hoc by: (1) creating 4-week windows across the sample period, (2) counting how many users made a gambling deposit within each window, (3) taking the average (i.e., mean) of the number of users across all 4-week windows, and (4) dividing by the total number of users in the sample. This calculation resulted in an average participation rate of 54.40%, somewhat lower than our "full sample" participation rate. The discrepancy underscores the importance of clearly stating time frames when interpreting gambling prevalence from transactional data. While this 4-week rate is more in line with the UK Gambling Commission's estimate, it is still considerably higher.

than 30% of their reported income, surpassing this threshold. Furthermore, a considerable number of users either used an overdrawn account or went into overdraft to make gambling debit transactions, further underscoring affordability and risk concerns. Survey evidence from the UK Gambling Commission using the PGSI found that overdraft usage for gambling was higher amongst problem gamblers (28%) versus those not at-risk (1%) (Gambling Commission, 2020).

In an effort to prevent gambling-related harms and address affordability concerns, recent regulatory proposals in the UK have introduced thresholds for financial risk checks. These include a £150 net deposit limit over a rolling 30-day period (Zendle & Newall, 2024b). Industry stakeholders have expressed concerns that such thresholds may be overly intrusive. In a post-hoc analysis, we found that 38.90% of users in our sample would exceed this threshold. This proportion is notably higher than the Gambling Commission's own modeling, which estimated that a lower threshold of £125 would affect approximately 20% of accounts (Rhodes, 2023). Future research should examine the impact of these thresholds across additional datasets and explore the characteristics of those who would be flagged under such criteria.

Consistent with prior research, the transaction activity in our sample reveals that a small minority of players engage in gambling at disproportionately high levels compared to the broader population. By analyzing the top 3% of users based on the number of debits, total debited amount, and net balance, we identified three highly-involved sub-groups. In an additional exploratory analysis, we found that unique users across these groups represented just under 5% of the total sample but accounted for 64% of the sample's total deposited amount, which aligns with prior research showing that a small proportion of gamblers are responsible for the majority of gambling expenditure (Deng et al., 2021; Louderback et al., 2024). It is important to note that our sample was drawn from users of a credit scoring service, which may overrepresent individuals with higher levels of financial vulnerability or gambling involvement compared to more representative datasets such as those used in Muggleton et al. (2021). Examining the overlap among these groups provides important practical implications. Notably, a substantial portion (42.8%) of users belong to all three groups, suggesting that extreme gambling behavior is oftentimes characterized across multiple dimensions. That being said, there were a considerable number of users who exhibited extreme gambling behavior across two dimensions or on a single dimension. This finding highlights the need to incorporate multiple behavioral markers when identifying at-risk individuals and tailoring intervention strategies. For example, harm prevention messaging may need to differ for high-volume players versus high-frequency players (e.g., see Saxton et al., 2021).

Limitations and Future Directions

Our analysis was based on a large and recent sample of UK-based individuals; however, our findings may not fully reflect the general population, as the sample consisted solely of users from a single credit scoring service. We relied on the data provider's automated methodology for categorizing transactions, which may have resulted in the omission of some gambling merchants. Future work could look into advancing automated transaction categorization methods to improve the accuracy of identifying gambling-related transactions to ensure that all relevant gambling activity is captured.

We note that while the transaction data potentially provides a more holistic view of users' gambling, some activity remains unobserved, e.g., cash transactions in brick-and-mortar facilities. Additionally, while bank transactions indicate when and where gambling spending occurred, they do not capture platform-specific behavioral patterns such as, for example, session length and game choice. And while much of our analysis was grounded in objective transaction data, the demographic information was self-reported by users, making it subject to the typical limitations associated with self-reported data (e.g., recall bias, inaccurate responses).

To enhance generalizability, future work could replicate these findings using a broader sample, i.e., beyond users of a single credit scoring service. Incorporating other sources of data could also be worthwhile, for example, by integrating this banking transaction data with data from validated gambling screens (e.g., Zendle & Newall, 2024a) or with wagering data from an online gambling operator. The latter may be particularly useful, allowing for calibration of established online gambling "markers of harm" with objective measures of financial risk.

There is significant potential for further analysis using the existing dataset. Whereas our study was cross-sectional, an analysis of users' behavior over-time would provide insights into longitudinal patterns of gambling involvement and financial risks. Additionally, examining the relationship between gambling and other financial transactions — such as discretionary spending, savings, or the use of financial products — could help inform our understanding of the interconnectedness of gambling behavior, financial harms, and broader financial habits.

Finally, translating these findings into actionable policy recommendations represents a critical next step. For instance, further exploration of thresholds for gambling within one's means could potentially inform the development of evidence-based affordability checks and more effective responsible gambling interventions. Moreover, new insights about sociodemographic differences in gambling behaviors based on open banking data can help to inform the development of more effective risk detection algorithms and targeted harm prevention messaging. As this line of research progresses, careful attention must also be paid to the ethical handling of sensitive financial data, particularly in the context of policy development and implementation.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10899-025-10419-6.

Acknowledgements The authors would like to thank Charles Cohen for providing the data used in this study, as well as for being available to answer questions and provide valuable insights about the dataset.

Funding Statement This study did not receive any funding. The data analyzed in this study were purchased by a financial services provider in the UK and provided to the authors under a data-sharing agreement.

Data Availability The data used in this study are not publicly available. Access to the data was granted under specific agreements that limit its sharing or redistribution. For further inquiries, please contact the corresponding author.

Declarations

Conflicts of interest During the past 5 years, International Gaming Institute (IGI) at University of Nevada, Las Vegas (UNLV), has received research and program funding from ESPN, MGM Resorts International,

Draft Kings, American Gaming Association, Wynn Resorts Ltd, Las Vegas Sands Corporation, Entain Foundation, Aristocrat Gaming, San Manuel Band of Mission Indians, Axes.ai, Sports Betting Alliance, Playtech, Sightline Payments, Global Payments, the State of Nevada Knowledge Fund, and State of Nevada Department of Health and Human Services. IGI runs the triennial research-focused International Conference on Gambling and Risk Taking, whose sponsors include industry, academic, and legal/regulatory stakeholders in gambling. A full list of sponsors for the most recent conference can be found at https://www.unlv.edu/igi/conference/18th/sponsors. IGI maintains a strict research policy (https://www.unlv.edu/igi/research-policy), as well as partnership and transparency framework (https://www.unlv.edu/igi/policies/partnership) to ensure appropriate firewalls exist between funding entities—no matter the entity's classification—and IGI's research and programs.
During the past 5 years, Kasra Ghaharian has received funding for research and/or consulting services from the Nevada Department of Health and Human Services, the Nevada Governor's Office of Economic Development, the Massachusetts Gaming Commission, AXES.ai, Playtech, IGT, Differential Labs, Focal Research Consultants, Yaspa, GP Consulting, and the Massachusetts Gaming Commission. From 2020 to 2022 Ghaharian worked as a Research Assistant for UNLV IGI and conducted research that was supported by funding from the IGI Payments Research Collaborative and Sports Betting Alliance. Ghaharian has received honoraria/travel reimbursement from the Responsible Gambling Council, the Illinois Council on Problem Gambling, and Kindred Group. Ghaharian also holds an equity interest in the Department of Trust and is a member of their advisory network. None of these entities played roles in the design, analysis, or interpretation of this study, and impose no constraints on publishing.
Jay Peterson has no conflicts of interest to disclose.
During the past 5 years Mana Azizsoltani has received funding for research AXES.ai and the State of Nevada Department of Health and Human Services.
During the past 5 years Richard Young has received research funding from the State of Nevada Department of Health and Human Services.
During the past 5 years, Eric R. Louderback has provided paid consulting services on player safety programs for Premier Lotteries Ireland, and has received travel reimbursement and speaker honoraria fees from the International Center for Responsible Gaming (ICRG). He has also received travel reimbursement and speaker honoraria fees from the Responsible Gaming Association of New Mexico.

References

Canadian Centre on Substance Use and Addiction (2024). The Lower-Risk Gambling Guidelines. https://gamblingguidelines.ca/lower-risk-gambling-guidelines/what-are-the-guidelines/

Delfabbro, P., Parke, J., & Catania, M. (2023). Behavioural tracking and profiling studies involving objective data derived from online operators: A review of the evidence. *Journal of Gambling Studies*. https://doi.org/10.1007/s10899-023-10247-6

Deng, X., Lesch, T., & Clark, L. (2021). Pareto distributions in online casino gambling: Sensitivity to timeframe and associations with self exclusion. *Addictive Behaviors, 120*, 106968.

Dowling, N. A., Lorains, F. K., & Jackson, A. C. (2015). Are the profiles of past-year internet gamblers generalizable to regular internet gamblers? *Computers in Human Behavior*, *43*, 118–128. https://doi.org/10.1016/j.chb.2014.10.019

Duncan, E. (2024). Canadian open banking legislation receives Royal assent. Open Banking Expo. https://www.openbankingexpo.com/news/canadian-open-banking-legislation-receives-royal-assent/

Edson, T. C., Tom, M. A., Louderback, E. R., Nelson, S. E., & LaPlante, D. A. (2022). Returning to the virtual casino: A contemporary study of actual online casino gambling. *International Gambling Studies,22*(1), 114–141. https://doi.org/10.1080/14459795.2021.1985581

Fletcher, R. (2024). Online deposit limits come into force in Netherlands. iGB. https://igamingbusiness.com/legal-compliance/dutch-online-gambling-deposit-rules/

Gainsbury, S. M., Russell, A., Hing, N., Wood, R., Lubman, D., & Blaszczynski, A. (2015). How the internet is changing gambling: Findings from an Australian prevalence survey. *Journal of Gambling Studies*, *31*(1), 1–15. https://doi.org/10.1007/s10899-013-9404-7

Gambling Commission. (2020). Consultation on gambling with credit cards. Gambling Commission. https://consult.gamblingcommission.gov.uk/author/consultation-on-gambling-with-credit-cards/

Gambling Commission. (2022). Statistics on participation and problem gambling for the year to Sept 2022: Official statistics. Gambling Commission. https://www.gamblingcommission.gov.uk/statistics-and-research/publication/statistics-on-participation-and-problem-gambling-for-the-year-to-sept-2022

Gambling Commission. (2024). Open banking data modelling of gambling spend thresholds. Gambling Commission. https://www.gamblingcommission.gov.uk/manual/open-banking-data-modelling-of-gambling-spend-thresholds/data-and-model-open-banking-data-modelling-of-gambling-spend-thresholds

Ghaharian, K., Abarbanel, B., Kraus, S. W., Singh, A., & Bernhard, B. (2023a). Players gonna pay: Characterizing gamblers and gambling-related harm with payments transaction data. *Computers in Human Behavior*, *143*, 107717. https://doi.org/10.1016/j.chb.2023.107717

Ghaharian, K., Abarbanel, B., Phung, D., Puranik, P., Kraus, S., Feldman, A., & Bernhard, B. (2023b). Applications of data science for responsible gambling: A scoping review. *International Gambling Studies*, *23*(2), 289–312 https://doi.org/10.1080/14459795.2022.2135753

Gillison, D. (2024). Open banking rules for consumer data unveiled by US watchdog. Reuters. https://www.reuters.com/business/finance/open-banking-rules-consumer-data-unveiled-by-us-watchdog-2024-10-22/.

Heirene, R. M., Wang, A., & Gainsbury, S. M. (2022). Accuracy of self-reported gambling frequency and outcomes: Comparisons with account data. *Psychology of Addictive Behaviors*, *36*(4), 333.

Humphreys, B. R., & Perez, L. (2012). Participation in internet gambling markets: An international comparison of online gamblers' profiles. *Journal of Internet Commerce,11*(1), 24–40. https://doi.org/10.1080/15332861.2012.650987

LaBrie, R. A., LaPlante, D. A., Nelson, S. E., Schumann, A., & Shaffer, H. J. (2007). Assessing the playing field: A prospective longitudinal study of internet sports gambling behavior. *Journal of Gambling Studies*, *23*(3), 347–362. https://doi.org/10.1007/s10899-007-9067-3

Louderback, E. R., Tom, M. A., Edson, T. C., & LaPlante, D. A. (2024). The Stability of Gambling Expenditure Distributions Over Time and Associations with the Use of Gambling Self-regulatory Tools. *International Journal of Mental Health and Addiction*, 1-23.

Marionneau, V. K., Lahtinen, A. E., & Nikkinen, J. T. (2024). Gambling among indebted individuals: An analysis of bank transaction data. *European Journal of Public Health,34*(2), 342–346.

McCormack, A., Shorter, G. W., & Griffiths, M. D. (2013). Characteristics and predictors of problem gambling on the internet. *International Journal of Mental Health and Addiction*, *11*(6), 634–657. https://doi.org/10.1007/s11469-013-9439-0

McNair, S., Chonaire, A. N., Barry, R., Leak-Smith, L., Larkin, C., Loke, T., & Lawrence, J. (2021). Dealing new data: What bank transactions can tell us about gambling behaviour. The Behavioural Insights Team. https://www.bi.team/blogs/dealing-new-data-what-bank-transactions-can-tell-us-about-gambling-behaviour/

Muggleton, N., Parpart, P., Newall, P., Leake, D., Gathergood, J., & Stewart, N. (2021). The association between gambling and financial, social and health outcomes in big financial data. *Nature Human Behaviour*, *5*(3), 319–326.

Nelson, S. E., Edson, T. C., Louderback, E. R., Tom, M. A., Grossman, A., & LaPlante, D. A. (2021). Changes to the playing field: A contemporary study of actual European online sports betting. *Journal of Behavioral Addictions,10*(3), 396–411. https://doi.org/10.1556/2006.2021.00029

Nelson, S. E., Edson, T. C., Singh, P., Tom, M., Martin, R. J., LaPlante, D. A., Gray, H. M., & Shaffer, H. J. (2019). Patterns of daily fantasy sport play: Tackling the issues. *Journal of Gambling Studies*, *35*(1), 181–204. https://doi.org/10.1007/s10899-018-09817-w

Newall, P. W. S., Rockloff, M., Hing, N., Thorne, H., Russell, A. M. T., Browne, M., & Armstrong, T. (2023). Designing improved safer gambling messages for race and sports betting: What can be learned from other gambling formats and the broader public health literature? *Journal of Gambling Studies,39*(2), 913–928. https://doi.org/10.1007/s10899-023-10203-4

Newall, P., & Swanton, T. B. (2024). Beyond 'single customer view': Player tracking's potential role in Understanding and reducing gambling-related harm. *Addiction*. https://doi.org/10.1111/add.16438

Office for National Statistics (2023, March 31). Male and female populations. Retrieved July 10, 2025, from https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/male-and-female-populations/latest/

Office for National Statistics (2024, March 26). Population estimates for the UK, England, Wales, Scotland and Northern Ireland: mid–2022. Retrieved July 10, 2025, from https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2022

Pandey, D. K., Hassan, M. K., Kumari, V., Zaied, Y. B., & Rai, V. K. (2024). Mapping the landscape of fintech in banking and finance: A bibliometric review. *Research in International Business and Finance*, *67*, 102116.

Pickering, D., & Blaszczynski, A. (2021). Paid online convenience samples in gambling studies: Questionable data quality. *International Gambling Studies*, *21*(3), 516–536.

Rhodes A. (2023). Your questions answered on the financial risk checks consultation. Gambling Commission.https://www.gamblingcommission.gov.uk/blog/post/your-questions-answered-on-the-financial-risk-checks-consultation. Accessed 25 April 2024

Saxton, J., Rodda, S. N., Booth, N., Merkouris, S. S., & Dowling, N. A. (2021). The efficacy of personalized normative feedback interventions across addictions: A systematic review and meta-analysis. *PLoS One,16*(4), Article e0248262.

Swanton, T. B., Gainsbury, S. M., & Blaszczynski, A. (2019). The role of financial institutions in gambling. *International Gambling Studies,19*(3), 377–398. https://doi.org/10.1080/14459795.2019.1575450

Zendle, D., & Newall, P. (2024a). The relationship between gambling behaviour and gambling-related harm: A data fusion approach using open banking data. *Addiction,* (10), Article Article add.16571. https://doi.org/10.1111/add.16571

Zendle, D., & Newall, P. (2024b). Better data access can lead to better collaborative conclusions: Results of a discussion with Heirene. *Addiction*. https://doi.org/10.1111/add.16657