Analysis of Casino Online Gambling Data in Relation to Behavioural Risk Markers for High-Risk Gambling and Player Protection

Article *in* International Gambling Studies · September 2011

DOI: 10.1080/14459795.2011.629204

George Tsogas

Simo Dragicevic
Playtech

International Gambling Studies
Vol. 11, No. 3, December 2011, 377–391

Analysis of casino online gambling data in relation to behavioural risk markers for high-risk gambling and player protection

Simo Dragicevic[a]*, George Tsogas[b] and Aleksandar Kudic[c]

[a]*Bet Buddy, UK;* [b]*Cass Business, School, City University London, UK;* [c]*GTECH G2, UK*

The Internet gambling industry has witnessed tremendous growth in recent years. Nonetheless, our understanding of Internet problem gambling behaviour remains in its infancy. In this paper we build on previous research analysing behavioural markers for high-risk Internet gambling using a new casino data set of active real money Internet gamblers. We assess the first month of play following registration using four behavioural markers: trajectory, frequency, intensity and variability. Our findings identify groups of gamblers who show signs of potentially risky behaviours, specifically gambling intensity and frequency. These gamblers also spend time gambling on slots Internet games. These findings provide a basis for using behavioural analysis to educate players about risks associated with gambling. We suggest a framework for how this can be implemented. Further research leading to the identification of risk factors for problem gambling using new methodologies and data sets will increase the clinical understanding of Internet problem gamblers.

Keywords: behavioural analytics; casino games; Internet gambling; problem gambling; screening

Introduction

Internet gambling is one of the fastest-growing sectors of the gambling industry. This growth is primarily being driven by increasing broadband penetration and a growing number of countries beginning to regulate online gambling markets. Global Betting and Gaming Consultants (2011) estimate that the Internet gambling sector will grow to US$41.71 billion by 2013, or 8.9% of global gross gambling yield. As the popularity of Internet gambling grows, concerns have been raised with regards to its effects compared with traditional forms of gambling, inasmuch as it can cause greater levels of problem gambling behaviour due to the very nature and immediacy of the online interaction and gaming experience. The growth in Internet gambling, however, also offers new opportunities to analyse data and player behaviour relative to risk, which can be utilized by academics, policymakers and the gambling industry to understand better gambling behaviours and to protect vulnerable players.

Internet and problem gambling

The assertion that Internet gambling is potentially more harmful than other forms of gambling is controversial. Research into online gambling and levels of problem gambling is an emerging field. One of the first attempts to assess the link between Internet and

*Corresponding author. Email: simo@bet-buddy.com

ISSN 1445-9795 print/ISSN 1479-4276 online

http://dx.doi.org/10.1080/14459795.2011.629204
http://www.tandfonline.com

gambling behaviour was undertaken by Griffiths (2003). He proposed a number of factors that make Internet gambling potentially addictive, including accessibility, anonymity, affordability, convenience, interactivity and disinhibition, arguing that addicts use the Internet to fuel underlying addictions. Griffiths (2003) further argues that the Internet could also potentially be a very dangerous medium, as the structural characteristics of the gambling software itself – including methods for paying and receiving winnings, speed of play, gambling features such as maximum stake allowed, and ambience through the use of stimulating light – might promote further addictive tendencies (see also Orford, 2011).

King, Delfabbro and Griffiths (2009) reviewed different forms of gambling that utilize new digital media and concluded that new gambling technologies could facilitate making gambling more accessible and ubiquitous, especially to the young, thus exposing them to greater levels of risk to problematic involvement in gambling. Jolley, Mizerski and Olaru (2006) assessed the effects of past behaviour (habit) on customer satisfaction in the prediction of actual behaviour (retention) using an Internet gambling experiment. The authors asserted that environments that foster habitual levels of gambling appear to facilitate gambler retention. Wood, Griffiths and Parke (2007) studied 422 UK university student poker participants and found a relatively high level of problem gambling amongst the players. The authors stressed that this was worrying, given the rate at which online Internet gambling is developing globally and given its 24-hour and seven-days-a-week availability. Further analysis of Internet poker players by Parke, Griffiths, Wood and Rigbye (2009) suggests that a new type of problem gambler could be emerging because of the Internet; a problem gambler who is winning a lot of money — for example over $1,000 per month — and also facing other problems related to time spent on the activity. The analysis and interpretation by Griffiths, Wardle, Orford, Sproston, and Erens (2009) of the 2007 UK Gambling Prevalence Survey showed that the problem gambling prevalence rate was significantly higher among Internet gamblers compared with non-Internet gamblers and that the medium of the Internet may be more likely to contribute to problem gambling than gambling in offline environments.

Research utilizing Internet gambling data

Despite arguments suggesting that game characteristics and the Internet can be decisive factors for increasing levels of problem gambling, it remains a limitation of the aforementioned studies that they have not assessed actual player gambling data and behaviour. Peller, LaPlante and Shaffer (2008) analysed 47 papers that embrace the notion that new gambling technology is hazardous to players' health. Their key recommendation was that improved study methods, including analysis of actual player betting data, can increase understanding of how new gambling technology affects player health.

A key development in the study of actual Internet player gambling data has been the research undertaken at The Division of Addictions, an affiliate of the Harvard Medical School, that used Internet gambling data obtained from the Internet gambling operator *bwin* Interactive Entertainment AG. LaPlante, Kleschinsky, LaBrie, Nelson and Shaffer (2009) carried out the first epidemiological study of actual Internet poker gambling behaviour amongst 3,445 participants of *bwin*. According to *bwin* (2010), the conclusions from this analysis refute prior research claiming that online gambling caused greater levels of problem gambling. *bwin* stated that the epidemiological findings contradict the conventional view that most Internet players exhibit excessive gaming behaviour and that evidence shows that the overwhelming majority of *bwin* customers played moderately, e.g. 2.5 sports bets with a stake of four Euros every four days, one casino game every two weeks and one poker session every three days, with an average 13 Euros per session and 1.80 Euros per session, respectively.

Further research from The Division of Addictions using *bwin* data was undertaken by Braverman and Shaffer (2010), who studied the betting patterns displayed during the first month of Internet gambling. They analysed betting patterns relative to behavioural markers, in an attempt to predict the development of gambling-related problems by using the k-means clustering methodology to group gamblers with similar betting patterns. Braverman and Shaffer (2010) selected a sample of live action sports bettors from their full research data set that included 48,114 players who opened an account with *bwin* during February 2005. For this study, the final data sample consisted of 530 players. The study was particularly noteworthy as it was the first to analyse actual Internet gambling behaviour during the first month of gambling activity in order to predict problem gambling behaviours. The study revealed that players characterized by high intensity and frequency of gambling and also by high variability of wager (bet) sizes during their first month of gambling were at higher risk than others of reporting gambling-related problems upon closing their accounts.

Braverman and Shaffer (2010) utilized four variables (intensity, frequency, variability and trajectory) as risk factors to analyse problem gambling behaviour in their study. Some of these risk factors arguably denote greater justification for inclusion as signs of problem gambling (i.e. bet variability) than others (i.e. bet frequency). For example, there is considerable research (American Psychiatric Association [APA], 1994; Blaszczynski & Nower, 2002; Cummins, Nadorff, & Kelly, 2009; Ferris & Wynne, 2001; Johansson, Grant, Kim, Odlaug, & Gotestam, 2009) that supports the inclusion of bet variability as a sign of problem gambling, due to either loss chasing, or overconfidence through the illusion of control with early large winnings. Johansson et al. (2009) identified 35 such risk factors; nine of these factors could be supported by two or more studies linking them to pathological gambling, including cognitive distortions. Jacobsen, Knudsen, Krogh, Pallesen, and Molde (2007) asserted that cognitive factors, such as loss chasing, control illusion and big wins, play a role in the development of problem gambling, while Cummins et al. (2009) supported the notion that early wins influence more reckless betting patterns.

Further support for bet variability as a problem gambling risk factor can be found in Blaszczynski and Nower (2002) by analysing their pathways model. They identified three distinct groups of gamblers manifesting impaired control over their gambling behaviour: (a) behaviourally conditioned problem gamblers; (b) emotionally vulnerable problem gamblers; and (c) antisocial, impulsivist problem gamblers. In the first pathway (behaviourally conditioned problem gamblers), Blaszczynski and Nower (2002) illustrate that these problem gamblers fluctuate between regular/heavy and excessive gambling because of conditioning, distorted cognitions surrounding the probability of winning, and/or a series of bad judgments or poor decision-making Whilst many types of betting can have conditioning effects, Blaszczynski and Nower (2002) highlight that members of this subgroup could be involved in chasing losses. The APA (1994) also asserts that problem gamblers chase losses. Ferris and Wynne (2001) specify wager increase as an indicator of problem gambling behaviour. The APA (1994) provided further evidence for the inclusion of intensity, frequency and trajectory as risk factors, asserting that problem gamblers need to increase the amount of their wagers to achieve the desired excitement previously experienced at lower levels of wagering and that problem gamblers report unsuccessful attempts to cut back or control gambling. The Australian Productivity Commission gambling report (Productivity Commission, 2010) highlighted gaming intensity as a problem gambling risk factor and recommended limiting the intensity at which gamblers can gamble on gaming machines given the risks posed by high intensity gambling. In addition, Holtgrave's (2009) analysis of adult survey data in Canada between 2001 and 2005 concluded that the frequency of play was significantly and positively related to problem gambling.

One recommendation from Braverman and Shaffer (2010) was that further studies were needed in order to confirm their findings from other types of gambling (e.g. poker and casino). Consequently, in this paper we employ some newly available Internet casino gambling data to assess whether the behaviours of casino players during the first month of play demonstrated similar betting behaviours to Braverman and Shaffer's (2010) live action sports bettors. Thus, the main purpose of our study is to describe the patterns of gambling behaviour of Internet gamblers relative to a set of problem gambling risk factors, as identified earlier: intensity, frequency, variability and trajectory. We also consider alternative risk factors that could be used in future studies and assess the merits of using such analyses in order to help players make more informed choices about their gambling activities.

Internet gambling and player feedback

The analysis of Internet gambling data relative to risk factors provides opportunities for giving players feedback on their gambling behaviour relative to problem gambling risk factors. Research undertaken by McDonnell-Phillips (2005) to assess the usefulness of responsible gaming tools showed that all players, including problem gamblers, reported that they tried to self-regulate by having some kind of spending limit in mind. The research also reported that most players were in favour of both the option to set their own limits when gambling, as well as receiving feedback on their transactions, such as detailed statements about how much they had spent on a given day or month. Larger global studies (e-Commerce and Online Gaming Regulation and Assurance [e-COGRA], 2007) also indicate that the most popular option for player feedback is receiving regular financial statements, with 75% of 10,000 respondents considering this option to be at least quite useful.

In a study focusing on the effectiveness of warning messages to gamblers, Monaghan and Blaszczynski (2010, 2010a) propose that in contrast with signs displaying probabilities or informing players of the risks associated with gambling, signs designed to encourage players to reflect on, appraise, evaluate and self-regulate their actions have greater theoretical and empirical support. Wiebe (2011) also argues for the importance of tailoring specific messages to different types of players as part of an Informed Decision Making (IDM) Framework. Wiebe's IDM Framework classified players into casual, frequent and intensive gamblers, with each group requiring different objectives, content and information delivery mechanisms to assist in the maintenance of healthy gambling behaviour. For example, whilst casual gamblers could benefit from awareness programmes that promote gambling literacy, intensive gamblers could be best served by encouraging them to take advantage of player feedback mechanisms, such as play activity reports and self-assessment tools. Very little empirical research exists that examines the effectiveness of player feedback mechanisms that utilize risk factors to predict gambling behaviour. Griffiths, Woods and Parke (2009) undertook a survey of 2,348 customers of the online gambling site operator Svenska Spel to ascertain the usefulness of such feedback tools. From the respondents, 26% had activated a feedback tool that assessed their game play relative to risk factors associated with problem gambling and 36% found receiving feedback on their gambling profile in this context useful. Whilst the authors state that the accuracy of the tool was extremely high, a limitation of this study was the lack of empirical analysis of the risk factors used by the operator to assess what constituted problem gambling behaviour.

Methodology

Data sets

We obtained de-identified Internet player data from GTECH G2, an Internet gambling software provider for lotteries and commercial Internet gambling operators. We asked GTECH G2 to identify a random and de-identified selection of players considered to be active, real money players at a point in time. There is no standard definition of what an active real money player is within the gambling industry. The definition typically takes into account whether a player has gambled with real or legal currency (i.e. not bonus money) during the last one to three months from a point in time. Other methodologies for segmenting players also include the analysis of a player's behaviour relative to factors within a specified period such as recency (how recently a player gambled within the period), frequency and total monetary spend (analysis of recency, frequency and monetary activity is termed as RFM, from a marketing perspective).

We selected the analytic sample from the full research cohort that included 128,788 people who had an active account with three Internet gambling sites licensed in Malta offering download casino and poker games to regulated markets in Europe. We decided to focus this study on casino players since the poker data only included cash play, not tournament play. We also needed to choose casino sites that used GTECH G2's payments gateway to enable us to map the different payments transaction types. This filtering of the full sample resulted in 26,221 (20.4%) players engaged in casino gambling. From these players our data set included 2,188 active real money players who were active gamblers for the month ending 31 July 2010. Based on this date, we analysed player data over the previous year, from 1 August 2009 to 31 July 2010, and selected players who had opened an account during that period (1,251 players). Finally, from those who had opened an account we selected the players who had gambled with a minimum of two bets during their first month of gambling, consistent with the methodology followed by Braverman and Shaffer (2010). After applying these filters the final research sample to be used in the k-means clustering analysis included 546 players. The casino data set included data resulting from online gambling across five casino game types ('Roulette', 'Slot Game', 'Table Game', 'Video Poker' and 'Other').

Clustering approach

Whilst the full, de-identified data set consisted of all available online player data in respect of demographic, session, game and transaction data, we only used a subset of the data for this analysis. The data attributes required to recreate the Braverman and Shaffer k-means clustering analysis included the player unique identifier, the start and finish times of each session, the number of rounds played during a session and the amount that was bet in real money (or legal currency), excluding any betting with bonus money.

Braverman and Shaffer (2010) put forward four variables or risk factors in their methodology: (i) frequency – the total number of active days (i.e. days on which a participant placed at least one bet during the first month of betting; (ii) intensity – total number of bets divided by frequency; (iii) variability – standard deviation of bets; and (iv) trajectory – the trajectory of first month of bets. We use the same methodology as Braverman and Shaffer (2010) for calculating these variables for our casino data set for frequency, intensity and trajectory. For trajectory, we developed a time trend to the data points in the first month of gambling via regression analysis, with amount bet being the dependent variable, used in conjunction with the frequency risk factor as the predictor variable, i.e. each active betting day was a sequential number. A positive slope value indicates an increasing trend in amount bet during the first month of gambling, a negative slope value indicating a decreasing trend in amount bet. Our method for calculating variability is different, in that we calculated variability of the amount bet per betting day rather than per wager. Whereas there are several methods for determining the number of clusters for a data set when using the k-means clustering method, we replicated Braverman and Shaffer (2010) by using four clusters. We also calculated mean values for variables for the first month's gambling activity. This included calculating the number of days of active gambling, total number of bets placed (n), total bets per day (n), total losses (Euros) and the standard deviation (SD) of the bet amount per bet or wager (Euros). Bet amount or wager was calculated using the total amount bet per session divided by the total number of played rounds per session.

Results

Description of clusters

Table 1 shows the standardized results from the k-means clustering analysis and Table 2 provides mean value of variables that describe gambling behaviours of different clusters for the first month of gambling using our casino data.

Cluster 1 consisted of 80 players (14.7% of the total players) who demonstrated a high frequency of gambling during the first month (mean = 14.89 days) relative to the other players, bet an average of 4.35 bets per day, and exhibited a positive trajectory. The mean SD of bet size for these players was 11.59 Euros. Cluster 2 consisted of 429 players (78.6% of total players) who gambled moderately compared to the players in the other clusters. These players gambled the least frequently in the first month (mean = 4.11 days), placed the least number of bets (mean = 3.72 bets per day), and lost the least amount of money (mean = 571.78 Euros). However, these players did demonstrate a higher mean SD of bet size (20.63) compared with players in cluster 1. Cluster 3 consisted of a small number of players (n = 6 or 1.1% of total players) who showed high variability in the amount bet per day when compared with the other clusters. These players also gambled frequently (mean = 11.33 days) and intensively (mean of total bets = 98.17) and lost the most money during the first month (mean loss = 21,650.32 Euros). These players demonstrated a negative trajectory and recorded a trend of reducing their bet size over the course of the first month more considerably than players in the other clusters. Finally, cluster 4 consisted of 31 players (5.7% of total players) who demonstrated a highly intensive betting pattern when compared with the other players (mean of total bets = 138.45). Whereas they gambled less frequently than the players in clusters 1 and 3, these players placed an average of 24.39 bets per day during their first month of gambling. Players in cluster 4 also displayed the highest mean SD of bet size (55.38 Euros).

In addition, we examined the mean values of variables that described the gambling behaviours of the different clusters for the data relating to the entire period of gambling

Table 1. Standardized scores of cluster centres on gambling behaviour characteristics (n = 546).

	Cluster 1 (n = 80)	Cluster 2 (n = 429)	Cluster 3 (n = 6)	Cluster 4 (n = 31)
Frequency	1.87636	−0.36761	1.13606	0.02518
Intensity	−0.09532	−0.21948	0.48517	3.18946
Variability	0.02426	−0.12131	7.49805	0.16499
Trajectory	0.10245	−0.00285	−1.47125	0.05978

Source: Bet Buddy

Table 2. Mean values of variables that describe gambling behaviours of different clusters for the first month of gambling (n = 546).

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Period of gambling (days)	14.89	4.11	11.33	6.00
Total bets (n)	67.01	15.2	98.17	138.45
Bets per day (n)	4.35	3.72	7.83	24.39
Total losses (Euros)	2450.23	571.78	21,650.32	2570.45
SD of bet amount per wager (Euros)	11.59	20.63	44.00	55.38

Source: Bet Buddy

and the results confirmed the broad trend for the observed period of gambling. Players in cluster 2 continued gambling with a relatively moderate gambling pattern, compared with players in the other clusters, and they lost the least amount (mean = 2,666.58 Euros). Cluster 3 players continued to gamble with a highly variable bet amount per day and continued to lose more heavily than the other players, losing on average more than five times the amounts players in the other clusters lost. Cluster 4 players continued to gamble with the most intensive betting patterns (mean = 21.26 bets per day compared with an average of 4.75 for the other three clusters). Finally, Table 3 outlines how much time the players gambled on each game type as a percentage of the total time gambled for each cluster for the entire period of gambling.

Table 3. Gambling duration by game type as a percentage of total gambling time for each cluster for the entire period of gambling (n = 546).

	Roulette	Table game	Slot game	Video poker	Other game
Cluster 1	13.5%	34.8%	45.6%	5.2%	1.0%
Cluster 2	13.2%	46.1%	28.6%	11.5%	0.7%
Cluster 3	84.1%	8.6%	7.0%	0.1%	0.2%
Cluster 4	37.8%	9.4%	47.6%	4.4%	0.7%

Source: Bet Buddy.

Discussion

Risk factors

Whilst we observed a group of players who varied the amount bet per day considerably during the first month of gambling (cluster 3) these players' mean SD of bet size (44.00 Euros) was lower than those observed in cluster 4 (55.38). Therefore, although cluster 4 players gambled more often and with a more variable mean SD of bet size, they lost on average over 88% less than players in cluster 3. This posits the question of why the 31 players in cluster 4 – in spite of losing considerably less money – were demonstrating riskier gambling behaviours than the six players in cluster 3, who were the only players to demonstrate a notable negative trajectory of betting during the first month. Can we infer from such analysis that cluster 3 players could have been chasing losses and subsequently decreased their gambling trajectory as their losses increased? There are some similarities in these results when compared with Braverman and Shaffer (2010). Their cluster 1 gamblers (high intensity, high variability), while losing more money over the entire period of gambling compared with cluster 3 (high intensity, low variability), gambled with a marginally lower intensity (13.5 bets per day) compared with cluster 3 (14 bets per day). Interestingly, our players in cluster 1, even though they were betting more frequently and losing more money compared with our relatively moderate players in cluster 2, gambled with a mean SD of bet size of 11.59, which was considerably lower than the more moderate players in cluster 2 (mean SD of bet size = 20.63).

Our results also show that the z-score for cluster 3 with regards to variability (7.48905) was markedly different to those of the other clusters in our analysis and with Braverman and Shaffer (2010), including their cluster 1 (variability = 4.40874) which included 11 gamblers (73% of all gamblers in cluster 1) who closed their accounts due to problem gambling reasons. One reason for this difference could be that our method for calculating variability was based on the standard deviation of bets per day (rather than per bet) as our data set did not allow for calculating the exact amount wagered per bet. This approach is arguably more consistent with most problem gambling screens' approach to assessing loss chasing, including the Diagnostic and Statistical Manual for Mental Disorders-IV (DSM-IV), the Canadian Problem Gambling Index (CPGI) and the South Oaks Gambling Screen (SOGS). These screens assess loss chasing by asking gamblers whether they returned a following day to try to win back losses. Within these screens there is no reference to chasing losses within other time periods, which could be problematic when used in the context of Internet gambling, which allows for players to gamble via multiple sessions within the same day from their homes. Given the growth projections for Internet gambling we believe there is merit in assessing whether these screens should reframe such questions to take into account this industry trend. Griffiths and Whitty (2010) also highlight that in the very specific issue of Internet problem gambling detection, if problem gambling could be identified online without the use of diagnostic gambling screens, then this may also have implications for the development of new problem gambling screening instruments in the future.

Opportunities exist for more detailed analysis of online gambling data and behavioural markers (risk factors). Our results for variability do not provide concrete evidence as to whether any players were actually involved in loss chasing, whichever method one uses to measure variability (i.e. by bet size of bet amount per day). The analysis of a chain of losing and of increasingly variable (in terms of bet amount) bets, rather than assessing the overall bet variability for each bet during a day or online gambling session, could provide clearer insights into problem gambling behaviour related specifically to loss chasing. Xuan and Shaffer (2009) provide further considerations in respect to the analysis of loss chasing. They analysed loss chasing in the context of the betting patterns of Internet sports gamblers who closed their accounts due to gambling related problems. The authors stated that regarding the construct of loss chasing, they observed evidence that supports the construct with respect to increasing stake during increasing monetary loss. However, they concluded that whilst problem gamblers experienced increased monetary loss and increased their stake size prior to closing their accounts, they did not chase longer odds. The analysis of bet variability as a risk factor has further considerations. For example, analysing the ratio of a gambler's early wins versus their early losses or total bet amount, rather than the overall bet variability, particularly for games with little or no skill element such as casino, could provide useful feedback for players regarding the risks associated with the illusion of control.

Game structure

The composition of our clusters provides evidence of different gambling patterns when Internet gamblers begin gambling. Cluster 4, as shown in Table 1, comprised a group of players that gambled more intensively than any of the other players, which could be explained by the faster nature of casino games in comparison to other gambling verticals. The most intensive live action sports bettors in Braverman and Shaffer (2010) were in their cluster 3, who bet on average 19 days and placed 14 bets per day, fewer than our intensive casino players in cluster 4, who on average placed over 24 bets per day but only bet on six days during the first month.

Our analysis shows that our most intensive gamblers in cluster 4 spent the majority of their time gambling on slots games (47.6%), which was the highest percentage for slots games across all clusters, followed by roulette (37.8%). Slots games were also very popular with gamblers in cluster 1 (45.6%), who gambled the most frequently in our research cohort. Our cluster 3 gamblers spent the vast majority of their time gambling on roulette; however, this data was highly influenced by one specific player's gambling behaviour and therefore little can be inferred from this. Our most moderate gamblers, cluster 2, gambled across all different game types, with table games (46.1%) being the most popular. We can infer from this analysis that amongst active real money casino Internet gamblers, players who gamble more intensively and frequently spend the majority of their time gambling on slots type games. This is in contrast to more moderate gamblers who, whilst playing across all gaming types, had a preference for table games. Game type, therefore, could be one explanation as to why our intensive gamblers in cluster 4 typically placed more bets than Braverman and Shaffer's (2010) intensive gamblers in clusters 1 and 3, in that the structural characteristics of slots allows for more intensive betting compared with live action sports betting.

Turner (2008) states that continuous games, such as casino and horse betting, as well as instant games, are ongoing and allow players to place multiple bets. Discrete games, such as lottery and traditional sports betting, allow for significant gaps between bets. Therefore, continuous games can condition and reinforce gambling behaviour, and in that way could lead to the development of problem gambling behaviour. Turner (2008) also states that volatile games result in frequent small prizes and infrequent large prizes. Slot machines, lotteries and bingo are volatile, whilst roulette, blackjack, craps and baccarat are not as volatile and typically have prizes that equal the bet. Volatility also depends on how people play the game, e.g. variations in bet size. As problem gamblers frequently increase the size of bets (e.g. doubling up, increasing the bet size after a win) a key consideration is that increased variability increases losses over the long run. In other words, the volatility of the game combined with the player's betting strategy can lead to problem gambling behaviour. This adds additional complication when assessing problem gambling risk profiles of players who play multiple games that are offered in an Internet gambling casino vertical, which typically includes both volatile games (slots) and less volatile games (roulette and table games). Holtgrave's (2009) analysis of adult survey data in Canada between 2001 and 2005 supports Turner's (2008) arguments, asserting that problem gambling was particularly pronounced for volatile games, such as slots. A contrasting perspective on this topic has been argued by Welte, Barnes, Tidwell and Hoffman (2009) who assessed the relationship between the games played by over 2,000 American young adults and the symptoms of problem gambling. Whereas their study showed that the form of gambling that made the largest contribution to problem gambling was casino, in contrast to Turner (2008) and Holtgrave (2009), they, nonetheless, refuted the hypothesis that rapid (or continuous) forms of gambling – such as slot machines – are the most problematic forms of gambling.

In this paper, by examining the types of games played by different types of gamblers, we suggest that a combination of analysis of risk factors, such as variability of bet size and intensity and frequency of betting, in conjunction with a deeper understanding of the gaming vertical, can provide further insight into high-risk gambling strategies and behaviours.

Player education

Internet gambling allows for many effective preventive measures that are arguably more difficult to implement in land-based gambling, such as forms of pre-commitment and self-limitation. We have demonstrated that the very nature of Internet gambling allows for behavioural analysis using player data, which can not only help academics and scientists to understand better gambling issues from a clinical perspective, but also can help the industry to identify at-risk players. Segmentation of customers based on the analysis of customer behaviour data has long been practised in Internet marketing and risk management. However, the use of technology to segment players based on analysing gambling patterns relative to problem gambling risk factors is not a widespread practice. We strongly suggest that such segmentation could provide new opportunities for gambling operators to provide feedback to their players that could potentially help at-risk players to regulate their game play.

Our analysis of active real money casino players highlights that the gambling behaviours of intensive gamblers varies considerably when analysed relative to risk factors. Given this variation, it would be beneficial to consider micro-segmentation of players, rather than using broad segmentation categories, such as causal, frequent and intensive players, as discussed earlier. This would allow for an even greater level of personalization of player feedback.

Building on Goetz's (2010) conceptual model of behaviour change we can describe how behavioural analysis and feedback mechanisms can help players to regulate their gambling behaviour: (i) player Internet gambling data is made available for behavioural analysis; (ii) analysis of the data in undertaken in the context of problem gambling risk factors; (iii) the resulting analysis is fed back to the player and, where appropriate, a series of choices from the operator's existing player protection tools and features are made available to the player (e.g. to set a betting monetary limit or to take a self-diagnostic test); and (iv) from the choices available the player can execute any series of actions to help regulate their game play.

McKinsey and Company (2011) draws attention to the further opportunities that data analysis can offer, beyond traditional applications in marketing and risk management. It states that applying advanced data analysis in new contexts (e.g. healthcare) can identify individuals who would benefit from proactive intervention or lifestyle changes. The conceptual model discussed above builds on this concept and when integrated alongside an Internet operator's existing player protection features (such as setting deposit and loss limits) can function as an ongoing and iterative feedback loop, with new data and the results of player actions being assessed regularly to ensure feedback is up-to-date and relevant. Our analysis also demonstrates that gambling behaviours for the entire period of gambling were consistent with that of the first month of gambling. Early feedback and intervention, therefore, could provide opportunities to influence more moderate and sustainable gambling behaviours over longer periods of time.

Clinical analysis

Notwithstanding the opportunities available to provide gamblers with feedback on their game play relative to problem gambling risk factors, an analysis of four clusters suggests that it is difficult to assign any of the clusters to specific clinical groups with a high degree of certainty with respect to problem gambling. Using Blaszczynski and Nower's (2002) pathways model we can argue that the growth in Internet gambling is an important ecological factor that is increasing gambling availability. Also, some of the behaviours exhibited in clusters 1, 3 and 4 could provide evidence to suggest behaviours consistent with pathway 1 (behaviourally conditioned problem gamblers, notably heavy or excessive gambling and loss chasing). It is not possible to link any of our clusters with pathway 2 (emotionally vulnerable problem gamblers), whilst the low trajectories exhibited across all four clusters could suggest evidence against classifying any of our clusters with pathway 3 (antisocial, impulsive problem gamblers).

The challenge for using these results in a clinical capacity with a large degree of confidence is that we analysed players' exhibited Internet gambling behaviour relative to four problem gambling risk factors. We could not assess, for example, whether any of the gamblers in our study suffered from poor coping or problem solving skills, or negative family experiences, or suffered from behavioural problems such as substance abuse, which are important elements in pathways 2 and 3. In order to enable that, we would have had to augment our exhibited gambling data with gamblers' declared behaviour, such as data obtained from interviews using a gambling screen. That would have enabled a specific assessment of any negative consequences our players may have suffered as a result of their gambling behaviours.

Analysis of alternative data sources from the gambling operator, such as Internet chat-room data and call-centre data, could provide additional evidence suggesting classification into pathway 2 (e.g. evidence of anxiety) and pathway 3 (e.g. irritability, antisocial behaviours). Research from Haefeli, Lischer and Schwarz (2011) supports this, stating that along with the analysis of Internet gambling data, communication-based indicators can be used to predict the early signs of problem gambling behaviour. Their analysis combined interviews with customer services staff to identify problem gambling risk indicators, together with an analysis as to whether these factors could accurately predict whether a gambler closed his or her account due to problem gambling related issues, by assessing the communications of 150 random Internet gambling self-excluders from Internet gambling operators. Their findings state that several predictors correlated with gamblers who self-excluded, including frequency of customer contact and threatening tonality in customer contact.

Limitations

Whilst our data sample was specifically selected to identify active real money players, the methodology of how the vendor calculated who was an active real money player was not available. Further analysis of less active players relative to problem gambling risk factors would provide additional insight into the early behavioural patterns of gamblers, which could provide a more holistic overview of the gambling habits of the broader population of Internet casino players. Although our analysis of player data provides new insights into Internet gambling behaviour, it is, nonetheless, possible that some of the 546 players in our sample played on multiple gambling websites and via other, land-based venues. Given this possibility, the analysis of player behaviour from a single Internet gambling operator risks not providing a truly holistic view of a player's problem gambling risk profile. Unlike Braverman and Shaffer (2010), our data sets did not include a full audit trail to the point when the players closed their accounts, and for what reasons. Even with a full audit trail, challenges exist when trying to identify problem gambling behaviour through the analysis of Internet player data. For example, not all problem gamblers will necessarily self-exclude or close their account. Even though k-means clustering is a useful statistical technique for dividing data into meaningful groups it has some limitations. For example, k-means clustering has difficulty in grouping data with large outliers, such as skewed and non-normally distributed populations like our casino data set (which could explain the results of cluster 3). It also produces more meaningful and natural clusters with the application of greater numbers of variables and sub-clusters. For example, Experian, the credit rating agency, has used the k-means clustering methodology to cluster populations into 45 types and 13 groups using 350 measures (Cameron, Cornish, & Wilson, 2005).

Implications for further research

We believe that future research into problem gambling behaviour that will utilize Internet gambling data is both urgently required as well as socially significant and necessary. Such research would benefit from exploring in greater detail a number of areas. Even though we analysed the data in terms of five game types, our casino data set included 78 different games within the game types. Further analysis of each cluster in the context of the 78 games could provide additional insight into which games the players in each cluster participated in and whether these influenced behaviour. Analysis of Internet gambling data for additional gaming verticals needs to be undertaken in order to assess gambling behaviours relative to risk in popular games, such as poker, lottery and instant games. In addition, the use of four risk factors to assess problem gambling behaviour could be extended, by including additional risk factors related to spend, losses, time, type of games and the odds that players bet against. Assessing trends in how quickly players spend cash deposits, trends in player losses, changes in player funding sources, gambling behaviour relative to the time of the day, and how much time is spent gambling can all be assessed relatively easily, given the Internet gambling data available. That data is typically available from the operator's player account management systems and also includes demographics, session, game and transaction data. Gainsbury (2011) describes how Internet gambling behaviour can be tracked and analysed to identify behavioural patterns that may be related to risky play or signify potential problems, arguing that this approach offers considerable benefits for researchers, operators and regulators.

Although exhibited Internet gambling data is likely to continue to be the basis for future research in this area, we believe that there is merit in assessing whether other data sets, including data sets that are not owned by the gambling operator, can be used in conjunction with exhibited data sets in order to try to understand better what drives problem or risky gambling behaviours. Such data could be structured (i.e. data that resides in fixed fields such as relational databases and spreadsheets) – for example, external credit risk, sporting event and climate data – or unstructured (i.e. data that does not reside in fixed fields) – for instance, relating to call centres, online chat and email data (which is another form of exhibited behaviour data). There will, however, be challenges to analysing new data sets. Acquiring external data could incur cost and involve data privacy issues. Data mining techniques required for the analysis of some of these new data sources, particularly unstructured data, could consign additional technical challenges for researchers.

Thus, we recommend that future research using Internet exhibited gambling data should be augmented with actual interviews of gamblers, where feasible, in order to obtain a better understanding as to whether gamblers have suffered negative consequences as a result of their gambling behaviour. This data, along with unstructured exhibited data and external data, would provide additional evidence to enable more effective clinical analysis of Internet gambling behaviour.

Further, the analysis of exhibited data behaviour in relation to problem gambling risk factors should be undertaken using different statistical methodologies. There are many alternative techniques that could be applied to these types of data sets that are arguably more appropriate than k-means clustering. For example, an alternative methodology could entail the use of regression analysis, by using the four risk factors as variables in order to assess whether they can predict which players self-exclude themselves from gambling. In this scenario, a logistic regression would be required given the categorical nature of the outcome, i.e. self-excluded versus not self-excluded.

Finally, using the results of such analysis in the context of providing Internet gamblers with feedback on their game play relative to risk is an emerging field with little empirical evidence. Consequently, further research in collaboration with Internet gambling operators is required in order to assess whether the provision of tailored messaging can positively influence gambling behaviour by helping players to make more informed decisions about their game play.

Acknowledgements

We would like to thank GTECH G2 for supplying the de-identified player data that was used for this study.

Notes on contributors

Simo Dragicevic is co-founder of Bet Buddy, a firm that develops behavioural analytics software to predict gambling behaviour. Simo has over 14 years of experience developing risk and data solutions in the investment banking and gambling industries. Prior to Bet Buddy Simo was a director at Barclays Bank and prior to that was lead manager for Accenture's UK capital markets CRM practice. Simo has a BSc (1995) from Loughborough University and an Executive MBA (2010) from Cass Business School, City University London.

Dr George Tsogas is a Senior Lecture in Management at Cass Business School. Dr Tsogas's specialisms and research focus includes corporate social responsibility, industrial relations and labour theory, and gambling. Dr Tsogas has a PhD (1998) from Birkbeck College, University of London, an MA Labour Studies (1991) from Rutgers, The State University of New Jersey, and a BSc Production and Management Engineering (1989) from The Technical University of Crete.

Aleksandar Kudic is a product manager at GTECH G2, a provider of gaming products and services. He is an engineer with an MBA degree from INSEAD, France.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual for mental disorders* (4th ed.). Washington, DC: Author.

Blaszczynski, A., & Nower, L. (2002). A pathways model of problem and pathological gambling. *Addiction*, *97*, 487–499.

Braverman, J., & Shaffer, H.J. (2010). How do gamblers start gambling: Identifying behavioural markers for high-risk internet gambling. *European Journal of Public Health*. Advance online publication. doi: 10.1093/eurpub/ckp232.

bwin Interactive Entertainment AG. (2010). *Incentives for a paradigm shift: New ways for the scientific analysis of online gaming: Implications for addiction research and responsible gaming*. Vienna: Author. Retrieved from https://media.itsfogo.com/media/sccwupload/Files/Responsibility/2009_WhitePaper_EN_001.pdf

Cameron, F., Cornish, C., & Wilson, N. (2005). A new methodology for segmenting consumers for financial services. *Journal of Financial Services Marketing*, *10*, 260–271.

Cummins, L.F., Nadorff, M.R., & Kelly, A.E. (2009). Winning and positive affect can lead to reckless gambling. *Psychology of Addictive Behaviors*, *23*, 287–294.

Dowling, N., Smith, D., & Trang, T. (2009). Electronic gaming machines: Are they the 'crack cocaine' of gambling? *Addiction*, *100*, 33–45.

e-Commerce and Online Gaming Regulation and Assurance (2007). *Global online gambler survey*. London: Author. Retrieved from http://www.ecogra.org/PressRoom/PressReleases/PressReleasesView.aspx?ID=fcc3eaa2-48cc-4240-b8a4-a48132656ee4

Ferris, J., & Wynne, H. (2001, 19 February). *The Canadian Problem Gambling Index: Final report*. Ottawa, ON: Canadian Centre on Substance Abuse, Retrieved from http://www.ccsa.ca/2003%20and%20earlier%20CCSA%20Documents/ccsa-008805-2001.pdf

Gainsbury, S. (2011). Player account-based gambling: Potentials for behaviour-based research methodologies. *International Gambling Studies*, *11*, 153–171.

Global Betting and Gaming Consultants. (2011). *Global gambling report no. 5* GBGC Global Forecasts. Isle of Man, British Isles: Author.

Goetz, T. (2010, October). *It's time to redesign medical data* [Video file]. Retrieved from http://www.ted.com/talks/lang/eng/thomas_goetz_it_s_time_to_redesign_medical_data.html

Griffiths, M. (2003). Internet gambling: Issues, concerns, and recommendations. *CyberPsychology and Behavior*, *6*, 557–568.

Griffiths, M., Wardle, H., Orford, J., Sproston, K., & Erens, B. (2009). Internet gambling, health, smoking and alcohol use: Findings from the 2007 British Gambling Prevalence Survey. *International Journal of Mental Health Addiction*, *9*, 1–11.

Griffiths, M.D., & Whitty, M.W. (2010). Online behavioural tracking in Internet gambling research: ethical and methodological issues. *International Journal of Internet Research Ethics*, *3*, 104–117.

Griffiths, M., Woods, T.A., & Parke, J. (2009). Social responsibility tools in online gambling: a survey of attitudes and behavior among internet gamblers. *CyberPsychology and Behavior*, *12*, 413–421.

Haefeli, J., Lischer, S., & Schwarz, J. (2011). Early detection items and responsible gambling features for online gambling. *International Gambling Studies*. Advance online publication. doi: 10.1080/14459795.2011.604643

Holtgrave, T. (2009). Gambling, gambling activities, and problem gambling. *Psychology of Addictive Behaviors*, *23*, 295–302.

Jacobsen, L.H., Knudsen, A.K., Krogh, E., Pallesen, S., & Molde, H. (2007). An overview of cognitive mechanisms in pathological gambling. *Nordic Psychology*, *59*, 347–361.

Johansson, A., Grant, J.E., Kim, S.W., Odlaug, B.L., & Gotestam, G.K. (2009). Risk factors for problematic gambling: A critical literature review. *Journal of Gambling Studies*, *25*, 67–92.

Jolley, B., Mizerski, R., & Olaru, D. (2006). How habit and satisfaction affects player retention for online gambling. *Journal of Business Research*, *59*, 770–777.

King, D., Delfabbro, P., & Griffiths, M. (2009). The convergence of gambling and digital media: Implications for gambling in young people. *Journal of Gambling Studies*, *26*, 175–187.

LaPlante, D.A., Kleschinsky, J.H., LaBrie, R.A., Nelson, S.E., & Shaffer, H.J. (2009). Sitting at the virtual poker table: A prospective epidemiological study of actual internet poker gambling behaviour. *Computers in Human Behavior*, *25*, 711–717.

McDonnell-Phillips Pty Ltd. (2005). *Analysis of gambler precommitment behaviour*. Report prepared for Gambling Research Australia. Melbourne: Victoria Department of Justice.

McKinsey & Company (2011). *Big data: The next frontier for innovation, competition and productivity*. Retrieved from http://www.mckinsey.com/mgi/publications/big_data/pdfs/MGI_big_data_full_report.pdf

Monaghan, S., & Blaszcynski, A. (2010). Electronic gaming machine warning messages: Information versus self-evaluation. *Journal of Psychology*, *144*, 83–96.

Monaghan, S., & Blaszczynski, A. (2010a). Impact of mode of display and message content of responsible gambling signs for electronic gaming machines on regular gamblers. *Journal of Gambling Studies*, *26*, 67–88.

Orford, J. (2011). *An unsafe bet – The dangerous rise of gambling and the debate we should be having*. Chichester: Wiley and Sons.

Parke, J., Griffiths, M.D., Wood, R.T.A., & Rigbye, J. (2009). Online poker gambling in university students: Further findings from an online survey. *International Journal for Mental Health Addiction*, *5*, 39–52.

Peller, A.J., LaPlante, D.A., & Shaffer, H.J. (2008). Parameters for safer gambling behavior: Examining the empirical research. *Journal of Gambling Studies*, *24*, 519–534.
Productivity Commission (2010). *Gambling* (Report No. 50). Canberra: Author. Retrieved from http://www.pc.gov.au/__data/assets/pdf_file/0004/95683/01-preliminaries.pdf
Turner, E.N. (2008). Games, gambling and gambling problems. In M. Zangeneh, A. Blaszczynski, & N.E. Turner (Eds.), *The pursuit of winning: Problem gambling theory, research and treatment* (pp. 33–64). Berlin: Springer.
Welte, J.W., Barnes, G.M., Tidwell, M.O., & Hoffman, J.O. (2009). The association of form of gambling with problem gambling among American youth. *Psychology of Addictive Behaviors*, *23*, 105–112.
Wiebe, J. (2011). *Informed Decision Making Framework*. Toronto, ON: RGC Centre for the Advancement of Best Practices. Retrieved from http://www.responsiblegambling.org/articles/Jamie_Wiebe_2011.pdf
Wood, R.T.A., Griffiths, M.D., & Parke, J. (2007). Acquisition, development, and maintenance of online poker playing in a student sample. *CyberPsychology and Behavior*, *10*, 354–361.
Xuan, Z., & Shaffer, H. (2009). How do gamblers end gambling: Longitudinal analysis of internet gambling behaviors prior to account closure due to gambling related problems. *Journal of Gambling Studies*, *25*, 239–252.