Published as: Auer, M. & Griffiths, M.D. (2014). An empirical investigation of theoretical loss and gambling intensity. *Journal of Gambling Studies*, 30, 879-887.

Abstract

Many recent studies of internet gambling – particularly those that have analysed behavioural tracking data – have used variables such 'bet size' and 'number of games played' as proxy measures for 'gambling intensity'. In this paper it is argued that the most stable and reliable measure for 'gambling intensity' is the 'theoretical loss' (a product of total bet size and house advantage). In the long run, the theoretical loss corresponds with the Gross Gaming Revenue generated by commercial gaming operators. For shorter periods of time, theoretical loss is the most stable measure of gambling intensity as it is not distorted by gamblers' occasional wins. Even for single bets, the theoretical loss reflects the amount a player is willing to risk. Using behavioural tracking data of 100,000 players who played online casino, lottery and poker games, this paper also demonstrates that bet size does not equate to or explain theoretical loss as it does not take into account the house advantage. This lack of accuracy is shown to be even more pronounced for gamblers who play a variety of games.

Introduction

The issue of how to measure 'gambling intensity' is an important one in the gambling studies field. Over the last few years, this issue has become much more to the fore as researchers in various jurisdictions have been given access to behavioural tracking data. Many of these studies have used proxy measures for gambling intensity including variables such 'bet size' and 'number of games played' (e.g., Broda, LaPlante, Nelson, LaBrie,, Bosworth & Shaffer, 2008; LaBrie, Kaplan, LaPlante, Nelson & Shaffer, 2008). LaPlante, Schumann, LaBrie & Shaffer, 2008; LaPlante, Kleschinsky, LaBrie, Nelson & Shaffer, 2009; Nelson, LaPlante, Peller, Schumann, LaBrie, & Shaffer, 2008; Dragicevic et al., 2011). Another major problem with these studies is that they have tended to present data by single game type (e.g., only data from online poker players or sports bettors are presented). However, as researchers have noted (e.g., Auer, Schneeberger, and Griffiths, 2012; Wardle, Moody, Griffiths, et al, 2011) online gamblers typically gamble on a variety of games.

There are various ways to conceptualize gambling intensity. Such ways could include parameters involving the time spent gambling, the number of gambles made, and/or the amount of money won or lost while gambling. In the studies mentioned above, monetary involvement has tended to be the main proxy measure for gambling intensity. This study proposes a different proxy measure for the money risked while gambling. The present authors define gambling intensity as the amount of money that players are putting at risk when playing. This might be considered easy to do (e.g., by using 'bet size'), but the element of chance is rarely accounted for, especially when a random win occurs. For instance, two gamblers putting the same amount of money at risk might end up with very different wins or losses at the end of similar length gambling sessions because of the chance factor. For this reason, the present authors are using a measure that is completely independent of random events and takes into account the true amount of money that players are prepared to risk. The interesting aspect of this is that most of the time, gamblers themselves are probably not aware of the amount of money they risked at the end of a playing session.

A recent paper using a simulation study by Auer, et al (2012) demonstrated that the most robust and stable measure for 'gambling intensity' is the 'theoretical loss'. Their paper showed that all previous studies using proxy measures for 'gambling intensity' had failed to take into account the house advantage. Outcomes in games of chance over the long-term will always be dependent upon the house advantage of each different type of game. Li (2003) showed that 'at risk' decision-making in the short-term is totally different from decision-making over longer periods of time. Decision making over the long-term can be explained by the expected value whereas short-term decision-making does not seem to be based on any expectation rule. However, studies investigating decision-making in situations where people have to make choices (e.g., Colbert, Murray, & Nieschwietz, 2009; Li, 2003) assume that players have a real choice in which they can truly influence the outcome and (thus) the expected return. However, this is not the case in pure chance games. Whatever the player chooses to do in pure chance situations, the house advantage will determine the expected return in the long-term.

As Auer et al (2012) point out, games with a high house advantage lead to higher player losses and games with a low house advantage lead to lower player losses. Theoretical loss is the same measure that the gaming industry describes as Gross Gaming Revenue (GGR). This is the difference between 'Total Bet' and 'Total Win'. The 'theoretical loss' of any given game is represented by the product of the bet size and the house advantage. Over very long periods of time, the theoretical loss corresponds to the GGR with increasing accuracy. The more diverse the gambling behaviour, the more that bet size deviates from the theoretical loss.

In this paper, gambling intensity is defined by the authors as the amount risked by a player. By incorporating the theoretical loss, the amount risked can be measured at a very detailed level. For instance, French roulette has a house advantage of 2.7% and keno has a house advantage of 10% (Auer, Griffiths and Schneeberger, 2012). This means that a player who repeatedly bets $100 on roulette will end up with a loss of $2.7, and a player who repeatedly bets $100 on keno will end up with a loss of $10. Therefore, the product of bet size and theoretical loss represents the amount of money that player will lose in the long run. Previous studies that have used bet size 9as a proxy measure for gambling intensity) would assign the same gambling of $10 intensity to the two players in the aforementioned example (and which obviously is not the case). The bet size is the one risk parameter that players are most likely to be aware of during gambling. However, it is deceptive as it does not take into account the expected return/loss that is controlled by the gaming operator via their house advantage.

This paper is direct a follow-up to the recent simulation study by Auer et al. (2012). In the previous study, Auer et al (2012) showed that 'bet size' and 'number of games' were not appropriate measures of gaming intensity. Their simulation study of 300,000 online gamblers showed that bet size explained 56% of the variance of the theoretical loss and the number of games played explained 32% of the variance of theoretical loss. This means that when using bet size alone, 44% of the gambling behaviour remains unexplained. When using the number of games played alone, 68% of the variance is left unexplained. As this study was a simulation, the present study replicated the study by Auer et al (2012) using real online gambler behavioural tracking data. There are many advantages and disadvantages with using data collected via behavioural tracking (Auer, et al, 2012; Griffths & Whitty, 2010). However, the main advantages outlined by Auer et al (2012) are that behavioural tracking data (a) provide a totally objective record of an individual's gambling behaviour on a particular online gambling website, (b) provide a record of events and can be revisited after the event itself has finished, and (c) usually comprise very large sample sizes.

Given the reliance on variables such as bet size and/or the number of games played as proxy measures for 'gambling intensity', this paper examines the properties of theoretical loss using actual data from the behavioural tracking of gamblers at a real online gambling site. It was hypothesised that the bet size would not explain all of the theoretical loss. It was also hypothesised that the more diverse the individual's gaming behaviour, the less important bet size would become in explaining theoretical loss. This study provides an analysis of real online gambler data (as opposed to simulated data) to highlight the differences between bet size and theoretical loss in relation to actual gamblers who play different types of online games.

Method

Participants: The anonymous sample comprised 100,000 online gamblers who played casino, lottery or poker games during one month (February 2012). All games played by these gamblers were recorded and subsequently analysed.

Procedure: The authors were given access to a large anonymized data set by a commercial gaming operator (*win2day Entwicklungs- und Betriebsgesellschaft m.b.H*), the online casino and lottery portal of *Österreichische Lotterien GmbH* and *Casinos Austria AG*. *win2day* has been online since 2003. win2day offers a wide range of lottery and casino games (as well as poker) to Austrian citizens. During the registration process, there is a mandatory requirement for all players to set time and cash-in limits. Furthermore, the weekly cash-in limit cannot exceed 800 Euros at any time during and after registration. *win2day* offers a wide range of lottery and casino games (as well as poker) to Austrian citizens.

The game types were categorized into eight distinct groups: (i) Lottery – Draw/Instant, (ii) Casino – Card, (iii) Casino – Slot, (iv) Casino – Videopoker, (v) Casino – Table, (vi) Casino Other, (vii) Bingo and (viii) Poker. For each of the game types and each player, the 'bet size' and the 'theoretical loss' were computed for the recorded time period (February 2012). In terms of house advantage these game types are very different. In general, lottery games have a relatively high house advantages (typically 50%) whereas slot machines have house advantages in the range of 1 to 5% depending on the gaming platform and the specific game. Poker on the other hand does not have a house advantage as such. In poker, the gaming involvement can be measured via the rake. The rake is a fixed percentage of the stake (bet size) that goes to the casino. The overall theoretical loss is thus comprised of the theoretical loss across all game types plus the poker rake.

Data analysis: Data analysis: The data analysis was performed with the statistical package *'R'*. *R* is a language and environment for statistical computing and graphics.

Results

The correlation between the 'bet size' and the overall 'theoretical loss' across the eight game types for the 100,000 players was 0.85 (d.f.=100,000, $p<0.0001$). Though this correlation is significant, the bet size alone explains only 72% of the variance of the theoretical loss. In order to be able to make further inferences on the difference between the theoretical loss and the bet size, a measure of difference was computed. Theoretical loss and bet size cannot be compared directly as the theoretical loss is always a percentage of the bet size. For that reason it does not make sense to compute the difference between these two measures. If the bet size was a legitimate measure of the theoretical loss, players with high bet sizes should also have high theoretical losses. This means that the ranking of players should be the same for the theoretical loss and the bet size. The higher the difference in the ranking, the less the bet size accounts for the theoretical loss. Consequently the difference in ranks can be used as a proxy indicator of the difference between these two measures. Furthermore the sign of the ranking difference is not important. It does not matter whether the rank for the theoretical loss is higher than the one for the bet size or vice versa. Consequently the absolute value of the difference was computed.

Table 1 shows the distribution of the ranked theoretical loss and the ranked bet size. This shows that the two distributions are equal. The maximum is higher than the number of observations (N) because of ties. Ties occur if two players have the same value and two different ranks are assigned. Also the minimum is not '0' but 1,242 and 1,275, respectively. This corresponds to the number of gamblers who have either no gaming behaviour or very little but equally high gaming involvement. All of these gamblers get assigned the same ranks.

INSERT TABLE 1 ABOUT HERE

The third measure in Table 1 represents the difference between the two rank variables. If all players were ranked equally, the differences would be zero. But obviously this is not the case. A difference of '1' means that the players are either ranked one step higher according to the bet size or one step lower. The mean difference is 13,519. The 90th percentile shows that 10% of the players are more than 29,791 ranks apart. This is quite a high difference particularly as the maximum difference is 118,730 (the difference between the maximum rank 119,971 and the minimum rank is 1,242). A total of 5% of the players are more than 64,479 ranks apart.

The computed ranks were then used to check if the diversity of play correlated with the difference between the theoretical loss and the bet size. A high correlation would mean that players engaging in a variety of different games are not being correctly measured via the bet size. In order to analyse this, the game type specific involvement was measured. The percentage of the theoretical loss per game type was computed for each game type across the 100,000 gamblers in the sample. Table 2 displays the correlation between the difference in ranks and each game type specific involvement measure.

INSERT TABLE 2 ABOUT HERE

Table 2 shows there is a correlation between the playing intensity of the different game types and the prediction error of the total bet. These underlying seven relative measures sum up to '1' for each player as they measure the relative gaming involvement. This means that the higher the involvement in one measure the lower the involvement in the other measures. The biggest correlations between the relative gaming game type specific involvement and the difference in ranks occurred in lottery and poker games. The higher the involvement in lottery games the smaller the difference between the total bet ranking and the theoretical loss ranking. The opposite was found regarding poker involvement (i.e., the higher the involvement in poker games, the higher the difference between the total bet ranking and theoretical loss ranking.

Figure 1 shows the average relative game type involvement for different sizes of the ranking difference. This figure highlights the information that was used to compute the correlations in Table 2. The figure shows that the correlation between the two measures is non-linear for some game types. For players that were equally ranked according to total bet and theoretical loss, the lottery gaming involvement was low (20% lottery involvement on the left side of the graph in Figure 1). However, this is also the case for players who are completely differently ranked (less than 10% lottery involvement) on the right side of the graph in Figure 2. This highly non-linear pattern produced an overall negative correlation of -0.37 (see Table 2). For this reason, the correlation that measures linear relationships has to be interpreted with caution, although Figure 1 clearly shows that there is a distinct pattern.

INSERT FIGURE 1 ABOUT HERE

Involvement in online slot games also showed a negative correlation with the ranking difference (see Table 2). Poker players had a peak at high ranking differences (right side of the graph in Figure 1). The group on the right hand side of the graph in Figure 1 showed the highest difference between the ranks of bet size and theoretical loss. This group of players showed the highest average involvement in poker, followed by Casino Slot games and Other Casino games. This group did not show any involvement in Lottery games.

Table 3 shows the actual numbers that were used to plot the graph in Figure 2. The higher the ranking difference the less valid the bet size as a measure of the theoretical loss. For instance, the data relating to poker clearly shows that high differences occurred with relatively high poker involvement. The last three groups showed 18%, 33% and 42% poker involvement.

INSERT TABLE 3 ABOUT HERE

The first three groups (i.e., low difference between bet size and theoretical loss) showed significant involvement in lottery games and casino slot games. Groups 4 to 17 almost exclusively played lottery games. High differences are associated with multiple game involvement and significant poker involvement.

Discussion

This is the first empirical study to ever examine theoretical loss using data from real gamblers. The study generally confirms the findings of a simulation study carried out by Auer et al (2012). More specifically, this study showed that bet size alone explained only 72% of the variance of the theoretical loss (i.e., 28% of the variance was unaccounted for by bet size). Auer, et al's (2012) simulation study showed that 46% of the variance of the theoretical loss was unaccounted for by the bet size. The error found in the empirical analysis is lower, but this is because the house advantages are not as different as assumed in Auer et al's simulation study.

The results of this study also showed that there is a correlation between game type specific involvement, and the difference between the total bet ranking and the theoretical loss ranking. But this correlation cannot be explained by one number as it is highly non-linear. Conclusions (such as the higher the involvement in lottery games the bigger the difference between the total bet ranking and the theoretical loss ranking) cannot be drawn because of the nature of the relationship.. The one exception is poker involvement (see Table 3). Poker involvement is only slightly increased in the first three ranking difference groups but very high in groups 18 to 20.

This means that players who among other games such as casino and lottery games engage up to 40% in poker games should never be analysed using the bet size. The mix between poker and other game types therefore appears to be especially poor in the predictive power of the bet size. The right hand side of Figure 1 shows the highest difference between the ranks of bet size and theoretical loss. This group of players showed the highest average involvement in poker, followed by the game types casino slot games and other casino games. This group did not show any involvement in lottery games. The occurrence of significant correlations shows that the difference between total bet ranking and theoretical loss ranking is highly associated with game type specific involvement. This again leads to the conclusion that the total bet is not an appropriate measure of the theoretical loss as it does not account for the different house advantages across different game types (or by the rake in poker games).

The data presented in Tables 2 and 3 (and Figure 1), clearly show that the diversity of play correlates with a deterioration of the bet size as a predictor of theoretical loss. This is especially interesting if we look at the inferences that have been made in earlier behavioural tracking studies. Earlier studies on behavioural tracking identified highly involved players to play a variety of games (Nelson et al. 2008) and they made inferences based on the bet size. The data presented here show that these inferences might not hold true or at the very least should be re-analysed using the most stable and robust measure of gaming intensity (i.e., theoretical loss).

While behavioural tracking has many advantages (e.g., it provides (i) a totally objective record of an individual's gambling behaviour on a particular online gambling website, (ii) a record of events and can be revisited after the event itself has finished, and (iii) very large sample sizes), there are a number of limitations to the data. The main limitations are that behavioural tracking data (i) collects data from only one gambling site and tells us nothing about the person's Internet gambling in general because Internet gamblers typically gamble on more than one site, (ii) always comes from unrepresentative samples becasue the players that use one particular internet gambling site, and (iii) does not account for the fact that more than one person can use a particular account. However, none of these specific limitations have much effect on the issue explored in this paper.

This study broadly confirmed the findings from Auer et al's (2012) simulation study. The results of this study suggest that future studies and particularly those that utilize behavioural tracking approaches should measure their participants' gambling intensity by incorporating the game-specific theoretical loss instead of using proxy measures such the bet size and/or the amount of money staked. Another implication is that previously published research could be re-analysed using the more robust measure of gambling intensity presented here (i.e., theoretical loss) rather than the proxy measures that were used in the original published studies. This study demonstrates that bet size does not reliably indicate the amount of money that players are willing to risk as it does not take into account the house advantage of each individual game that gamblers engage in. The house advantage represents the percentage held back by the gaming operator and is essential for the amount lost in the long-term and will eventually be equal to the total losses that a player accumulates.

Given the findings presented here and in previous simulation studies (i.e., (Auer, Griffiths and Schneeberger, 2012), future research that sheds light on the influence of game type specific involvement would appear to be useful. In order to further generalize the results, further empirical research utilizing data from other online gaming platforms as well as land-based casino premises should be performed. Research is especially lacking behavioural analysis on Video Lottery Terminal gambling behaviour though a few jurisdictions are already using mandatory or voluntary personalized card systems which are tracking individual gambling behaviour (e.g., Norway that have implemented mandatory player cards).

Conflict of interest: The authors declare that they have no conflict of interest.

References

Auer, M., Schneeberger, A., & Griffiths, M.D. (2012). Theoretical loss and gambling intensity: A simulation study. *Gaming Law Review and Economics, 16*, 269-273.

Broda, A., LaPlante, D. A., Nelson, S. E., LaBrie, R. A., Bosworth, L. B. & Shaffer, H. J. (2008). Virtual harm reduction efforts for Internet gambling: effects of deposit limits on actual Internet sports gambling behaviour. *Harm Reduction Journal*, *5,* 27.

Colbert, G., Murray, D., Nieschwietz, R. (2009). The use of expected value in pricing judgements. *Journal of Risk Research, 12, 199-208.*

Dragicevic, S., Tsogas, S.B., & Kudic, A. (2011). Analysis of casino online gambling data in relation to behavioural risk markers for high-risk gambling and player protection. *International Gambling Studies, 11,* 377–391.

Griffiths, M.D. & Auer, M. (2011). Online versus offline gambling: Methodological considerations in empirical gambling research. *Casino and Gaming International,* 7(3), 45-48.

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.

LaBrie, R.A., Kaplan, S., LaPlante, D.A., Nelson, S.E., & Shaffer, H.J. (2008). Inside the virtual casino: A prospective longitudinal study of Internet casino gambling. *European Journal of Public Health, 18,* 410-416

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 behavior. *Computers in Human Behavior*, 25, 711-717.

LaPlante, D. A., Schumann, A., LaBrie, R. A., & Shaffer, H. J. (2008). Population trends in Internet sports gambling. *Computers in Human Behavior, 24,* 2399–2414.

Li, S. (2003). The role of Expected Value illustrated in decision-making under risk: Single-play vs multiple-play. *Journal of Risk Research, 6, 113-124.*

Nelson, S. E., LaPlante, D. A., Peller, A. J., Schumann, A., LaBrie, R. A., & Shaffer, H. J. (2008). Real limits in the virtual world: Self-limiting behavior of Internet gamblers. *Journal of Gambling Studies, 24,* 463-477.

Wardle, H., Moody, A., Griffiths, M.D., Orford, J. & and Volberg, R. (2011). Defining the online gambler and patterns of behaviour integration: Evidence from the British Gambling Prevalence Survey 2010. *International Gambling Studies, 11,* 339-356.