

## Theoretical Loss and Gambling Intensity: A Simulation Study

Michael Auer, Andreas Schneeberger, and Mark D. Griffiths

MANY RECENT STUDIES OF INTERNET GAMBLING—particularly those that have analysed behavioral tracking data—have used variables such as “bet size” and “number of games played” as proxy measures for “gambling intensity.”<sup>1</sup> However, neither bet size nor the number of games played takes into account the house advantage of a game. Players are risking less when they play games with low house advantages. A low house advantage, therefore, corresponds to a high payout. Furthermore, data presented from these studies have typically been presented by game type (e.g., data are only presented from online sports bettors or online poker players). However, using a concise simulation analysis of online gamblers playing a variety of games, this short article argues that bet size cannot be reliably used across games and/or game types as a measure of gambling intensity.

Griffiths and Auer<sup>2</sup> outlined the many advantages and disadvantages of using behavioral tracking data in the gambling studies field. The main advantages of behavioral tracking data are that it (a) provides a totally objective record of an individual’s gambling behavior on a particular online gambling Web site; (b) provides a record of events and can be revisited after the event itself has finished; and (c) usually comprises very large sample sizes. These are the main reasons that such data will be used here.

In this article, we describe what we believe is the best and most stable measure for “gambling intensity.” This measure is the “theoretical loss.” In the

long run, outcomes in games of chance are always dependent on the house advantage: games with a big house advantage lead to higher losses for the gambler, while games with a lesser house advantage lead to lower losses. For instance, lottery games typically have relatively high house advantages (e.g., 50%), whereas casino games typically have relatively low house advantages: roulette games with a single “zero (0)” on their wheels, for example, have a house advantage of 2.7%.

The “loss/win” variable—often referred to as the gross gaming revenue (GGR)—is the difference between “total bet” and “total win.” However, as a measure of a player’s gambling intensity, it is not suitable, as it is typically distorted by the occasional winning occurrences by gamblers, particularly in the short-term. In the very long run, GGR is a

---

<sup>1</sup>A. Broda, D.A. LaPlante, S.E. Nelson, R.A. LaBrie, L.B. Bosworth, and H.J. Shaffer, *Virtual harm reduction efforts for Internet gambling: effects of deposit limits on actual Internet sports gambling behaviour*, 5 HARM REDUCTION JOURNAL 27 (2008); R.A. LaBrie, S. Kaplan, D.A. LaPlante, S.E. Nelson, and H.J. Shaffer, *Inside the virtual casino: A prospective longitudinal study of Internet casino gambling*, 18(4) EUR. J. PUB. HEALTH 410–416 (2008); D.A. LaPlante, J.H. Kleschinsky, R.A. LaBrie, S.E. Nelson, and H.J. Shaffer, *Sitting at the virtual poker table: A prospective epidemiological study of actual Internet poker gambling behavior*, 25 COMPUTERS IN HUMAN BEHAV. 711–717 (2009); D.A. LaPlante, A. Schumann, R.A. LaBrie, and H.J. Shaffer, *Population trends in Internet sports gambling*, 24 COMPUTERS IN HUMAN BEHAV. 2399–2414 (2008); S.E. Nelson, D.A. LaPlante, A.J. Peller, A. Schumann, R.A. LaBrie, and H.J. Shaffer, *Real limits in the virtual world: Self-limiting behavior of Internet gamblers*, 24 J. GAMBLING STUD. 463–477 (2008); S. Dragicevic, G. Tsogas, and A. Kudic, *Analysis of casino online gambling data in relation to behavioural risk markers for high-risk gambling and player protection*, 11 INT’L GAMBLING STUD. 377–391 (2011).

<sup>2</sup>M.D. Griffiths and M. Auer, *Online versus offline gambling: Methodological considerations in empirical gambling research*, 7(3) CASINO & GAMING INT’L 45–48 (2011).

---

Michael Auer is Business Unit Manager Predictive Analytics at a2mlab in London. Andreas Schneeberger is CEO of a2mlab. Mark D. Griffiths is Professor of Gambling Studies in the International Gaming Research Unit, Psychology Division, at Nottingham Trent University in Nottingham.

more reliable indicator of gambling intensity, as (statistically) it corresponds to the theoretical loss. This means that theoretical loss is the most reliable and robust indicator of gambling intensity. 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 following formula shows the calculation of the theoretical loss for a bet on a single game ( $g$ ). The house advantage depends on the specific game in question.

$$\text{Theoretical Loss}_g = \text{Total Bet}_g \times \text{House Advantage}_g$$

The more diverse the gambling behavior, the more the bet size will typically deviate from the theoretical loss. For players playing only one game, a high bet size will always be accompanied by a high theoretical loss. However, if the gambling behavior is more diverse, players with high bet sizes will not necessarily have a high theoretical loss. Given the reliance on variables such as bet size and/or number of games played as proxy measures for gambling intensity, this article examines the properties of theoretical loss using a mathematical simulation study of up to 300,000 gamblers, playing as many as 13 different games. The hypothesis was that the robustness of bet size and number of games played as measures of gambling intensity would decline as gaming becomes more hybrid.

## METHOD

The authors programmed a procedure in the statistical package *R* that simulates different numbers of players, playing as many as 13 different forms of gambling games with varying house advantages. A simulation study is most appropriate to show that the bet size and the number of games does not completely explain the theoretical loss, since an empirical study with a sample of gamblers always carries the problem of validity. For this reason, the mathematical approach of a simulation study was chosen. (However, the authors also plan to publish a similar study with real-world online behavioral data.)

### *Game types*

As argued above, casino operators offer a wide variety of games with different house advantages.

The house advantage is a key structural factor that influences game attractiveness to players and helps determine the casino's revenue stream. High house advantages tend to be less attractive to the gambler but yield high profits. Based on our experience in the field of gambling, as well as our knowledge of the different gaming platforms, we chose 13 games to be part in the simulation study. Table 1 displays these 13 games and their house advantages. These house advantages are mean average values, because different operators sometimes modify games slightly. For example, slot machines and video poker, in particular, can vary greatly across operators. Therefore, we provided two house advantages each for slot machines and video poker.

Viewing the different house advantages also reflects the problem of using bet size or number of games as a measure of gambling intensity. Betting 10 euros on keno actually corresponds to betting almost 100 euros on roulette with a single zero. However, to date, all published studies using behavioral tracking data have only used bet size as the measure of gambling intensity. Using the number of games as a measure of gaming intensity would be an even-worse measure, as it also neglects the amount bet.

### *Simulation parameters*

As mentioned above, the first advantage of a simulation study is that gamblers could play up to 13 different games with varying house advantages. A second advantage is that the simulation can be conducted with different numbers of players. We simulated up to 300,000 players, playing the 13 games. For each of the players, the simulation calculates a

TABLE 1. GAME TYPE AND HOUSE ADVANTAGE USED IN THE SIMULATION STUDY

<i>Game Type</i>	<i>House Advantage</i>
Keno	25.0%
Big 6 Wheel	10.0%
Roulette (double zero)	5.6%
Video poker (high house advantage)	4.8%
Sports betting	4.5%
Slots (high house advantage)	3.3%
Roulette (single zero)	2.7%
Slots (low house advantage)	1.8%
Baccarat (banker)	1.7%
Baccarat (player)	1.6%
Blackjack	0.8%
Craps (double odds)	0.6%
Video poker (low house advantage)	0.5%

specific game-preference pattern. On real world platforms, it is known that gamblers also play a variety of different games.<sup>3</sup> Once the game preferences had been determined, up to 5,000 games per player were simulated. Each simulated player had an individual number of games simulated. The “drawing” of games happens according to the game preference pattern. Finally, the bet size per game being played has to be determined. Each player’s range of bet size lies in an individual interval. The minimum bet size for each player is a natural number drawn from the interval (1:100). To compute the maximum bet size, a natural number between 0 and 100 is drawn and added to the minimum bet size.

#### Data analysis

The data analysis was performed with the statistical package *R*. *R* is a language and environment for statistical computing and graphics. It is a GNU project that is similar to the *S* language and environment that was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies).

## RESULTS

The explanatory quality of bet size and number of games with respect to the theoretical loss can most easily be determined by performing a correlation analysis. Table 2 shows the amount of variance of the theoretical loss explained by bet size and number of games played. Table 2 represents the results for different numbers of players, wagering on up to 5,000 single games and playing up to 13 different game types. The amount bet also varies

TABLE 2. EXPLAINED PROPORTION OF THEORETICAL LOSS BY BET SIZE AND NUMBER OF GAMES PLAYED FOR DIFFERENT NUMBERS OF PLAYERS

n	R <sup>2</sup>	
	Bet Size	Number of Games Played
500	59%	38%
1,000	61%	32%
5,000	58%	34%
10,000	56%	32%
50,000	56%	34%
100,000	56%	32%
200,000	56%	32%
300,000	56%	32%

TABLE 3. EXPLAINED PROPORTION OF THEORETICAL LOSS BY BET SIZE AND NUMBER OF GAMES PLAYED FOR DIFFERENT NUMBERS OF GAMES PLAYED

# Games	R <sup>2</sup>	
	Bet Size	Number of Games Played
[1:50]	52%	29%
[1:100]	55%	31%
[1:1,000]	56%	32%
[1:2,500]	56%	32%
[1:5,000]	56%	32%

from 1 to 200 monetary units. It may clearly be seen that the  $R^2$  converges to a specific value for both parameters. The simulation showed that bet size explains 56% of the variance of the theoretical loss, while the number of games played explains 32% of the variance of the theoretical loss. This means that when using bet size alone, 44% of the gambling behavior remains unexplained. When using the number of games played alone, 68% of the variance is left unexplained. The error when using bet size or the number of games played is considerable; it is clearly not justified to use only one of these two indicators as a measure of gaming intensity. All conclusions that have been drawn from previous studies should therefore be treated with considerable caution, as a large amount of the behavior has not been accounted for.

Table 2 clearly shows that the explained variance converges towards the two values 56% (bet size) and 32% (number of games played). Another flexible parameter is the number of single games being played per individual. In Table 2, this number was drawn from the uniform distribution (1:5,000). In Table 3, the number of players simulated was fixed at 300,000, but the number of games played was varied. The first two values of  $R^2$  (52% and 29%) are the results from restricting the maximum number of single games played per individual to 50. The higher the number of single games, the more valid the result. For up to 1,000 single games, the two values of  $R^2$  remain stable (see Table 3).

<sup>3</sup>H. Wardle, A. Moody, M.D. Griffiths, J. Orford, and R. Volberg, *Defining the online gambler and patterns of behaviour integration: Evidence from the British Gambling Prevalence Survey 2010*, 11 INT’L GAMBLING STUD. 339–356 (2011).

TABLE 4. CONTINGENCY TABLE OF THEORETICAL LOSS AND BET SIZE

	<i>Bet Size</i>										<i>Total</i>
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	
<i>Theoretical Loss</i>											
<i>1</i>	23,687	5,371	742	143	39	15	1	—	—	—	29,998
<i>2</i>	5,583	13,384	7,295	2,403	838	303	132	46	16	—	30,000
<i>3</i>	668	7,195	8,941	6,987	3,622	1,528	639	262	121	37	30,000
<i>4</i>	53	3,085	6,178	6,999	6,267	4,026	1,981	903	364	144	30,000
<i>5</i>	7	787	4,299	5,275	5,896	5,721	4,167	2,382	1,070	396	30,000
<i>6</i>	—	144	1,885	4,419	4,777	5,430	5,440	4,445	2,555	905	30,000
<i>7</i>	—	31	533	2,744	4,466	4,665	5,230	5,433	4,565	2,333	30,000
<i>8</i>	—	3	105	843	3,022	4,778	5,152	5,594	6,017	4,486	30,000
<i>9</i>	—	—	20	170	943	3,022	5,467	6,348	6,706	7,324	30,000
<i>10</i>	—	—	2	17	130	512	1,791	4,587	8,586	14,375	30,000
<i>Total</i>	29,998	30,000	30,000	30,000	30,000	30,000	30,000	30,000	30,000	30,000	299,998

Another way of demonstrating the difference between theoretical loss, bet size, and number of games played is via a contingency table (*see* Table 4). The contingency table was produced from the simulation comprising 300,000 players playing a maximum of 5,000 single games each. Players were grouped into 10 groups according to theoretical loss and bet size. There was a correlation between the two measures, but it was clearly far from perfect. Of the 30,000 most gaming intense players with respect to theoretical loss, 14,375 of them are also among the most gaming intense players with respect to bet size (therefore, 15,625 are not). The difference is even greater for moderate gamblers. For example, of the 30,000 gamblers in Group 6 of the theoretical loss group (*see* Table 4), only 5,430 are also in Group 6 of the bet size group. The error here amounts to 82%  $((30,000 - 5,430)/30,000)$ .

Table 5 shows that the concordance between theoretical loss and number of games is even less. Of the 30,000 most gaming intense players with respect to theoretical loss, only 8,649 are also among the most gaming intense players with respect to number of games played (therefore, 21,351 are not). This amounts to an error of 71%  $((30,000 - 8,649)/30,000)$ .

## DISCUSSION

As far as the authors are aware, no paper published in the gambling studies literature has used theoretical loss as a measure of gambling intensity, and when gambling intensity has been examined, proxy measures, such as bet size and/or number of games played have been incorporated. The hypothesis that the robustness of bet size and number of games played as measures of gambling intensity

TABLE 5. CONTINGENCY TABLE OF THEORETICAL LOSS AND THE NUMBER OF GAMES PLAYED

	<i>Number of Games</i>										<i>Total</i>
	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	
<i>Theoretical Loss</i>											
<i>1</i>	20,203	5,220	1,949	974	574	351	247	174	142	100	29,934
<i>2</i>	6,783	9,199	5,169	2,954	1,980	1,287	913	735	523	457	30,000
<i>3</i>	2,108	6,859	6,275	4,425	3,121	2,291	1,785	1,253	1,062	821	30,000
<i>4</i>	646	4,333	5,509	4,986	3,998	3,085	2,452	2,086	1,590	1,315	30,000
<i>5</i>	175	2,511	4,220	4,772	4,472	3,851	3,163	2,699	2,233	1,904	30,000
<i>6</i>	34	1,218	2,988	4,143	4,285	4,314	3,976	3,421	3,003	2,618	30,000
<i>7</i>	4	536	2,173	3,336	3,844	4,318	4,245	4,126	3,853	3,565	30,000
<i>8</i>	—	142	1,226	2,503	3,421	4,193	4,574	4,684	4,711	4,546	30,000
<i>9</i>	—	35	425	1,567	2,930	3,827	4,527	5,068	5,659	5,962	30,000
<i>10</i>	—	2	68	385	1,268	2,552	4,115	5,678	7,283	8,649	30,000
<i>Total</i>	29,953	30,055	30,002	30,045	29,893	30,069	29,997	29,924	30,059	29,937	299,934

would decline as gaming becomes more hybrid was confirmed.

This article demonstrates that previous measures used are inadequate for explaining players' gambling intensity and that theoretical loss appears to be the most reliable and robust indicator of gaming intensity. The results presented using a concise simulation study clearly demonstrate that neither bet size nor (particularly) number of games played are robust measures of gambling intensity. This suggests that future studies should not consider such variables as proxy measures for gambling

intensity, and that theoretical loss is a much more robust measure.

In order to make statements about overall gambling behavior, a complete view of gambling behavior has to be generated (including, for example, sports betting, casino games, and poker). Results from previous studies outlined in the introduction were only derived singularly for sports betting, casino, or poker playing behavior. The analysis in the study presented here clearly demonstrates that researchers in the gambling studies field need to examine individual gambling behavior across all game types and not just one type.