The Online Gambling Fairness Paradox: Cryptographic Verification, Behavioral Harm, and Consumer Protection

Philipp D. Dubach
Zurich, Switzerland
phdubach@pm.me

January 2026

JEL Classification: D91 (Intertemporal Consumer Choice; Life Cycle Models and Saving), G41 (Behavioral Finance: Role of Psychological, Emotional, Social, and Cognitive Factors), I18 (Health: Government Policy; Regulation; Public Health), L83 (Sports; Gambling; Restaurants; Recreation; Tourism), C12 (Hypothesis Testing: General)

Keywords: Online gambling, algorithmic fairness, cryptocurrency, behavioral economics, consumer protection, probability weighting, dark patterns

Abstract

Following the 2018 *Murphy v. NCAA* decision that precipitated rapid U.S. online gambling expansion, "crash" games have emerged as a popular cryptocurrency format claiming cryptographically verifiable outcomes—known in the industry as "provably fair." We analyze 20,038 rounds from an anonymized platform, testing whether game-generated multipliers conform to the theoretical distribution. Kolmogorov-Smirnov and chi-square tests detect statistically significant deviations ($p < 10^{-15}$), though these are economically negligible. Following Wang and Pleimling's (2019) methodology, we estimate the probability density function exponent at $\alpha \approx 1.98$, matching the theoretical value of 2.0 and confirming fair random number generation—contrasting with their findings of $\alpha = 1.4$–1.9 for player cashouts, which reflect behavioral biases rather than manipulation. Monte Carlo simulations confirm that no betting strategy produces positive returns. Session-level analysis reveals rapid loss velocity: at 179 rounds per hour, players face expected losses exceeding 500% of wagered amounts hourly. These findings support cryptographic verifiability claims while highlighting that mathematical fairness does not ensure consumer safety.

1 Introduction

The landscape of legal gambling in the United States has transformed dramatically since the Supreme Court's May 2018 ruling in *Murphy v. National Collegiate Athletic Association*, which struck down the Professional and Amateur Sports Protection Act (PASPA) of 1992 as unconstitutional under the Tenth Amendment [Murphy v. NCAA, 2018]. In the six years following PASPA's repeal, 38 states and the District of Columbia have legalized sports betting, with commercial sportsbooks reporting a record $121 billion in total handle and $11 billion in revenue during 2023 [AGA, 2024]. Combined with $41.9 billion from Native American gaming operations and $6.2 billion from online casinos, the legal U.S. gambling industry now exceeds $100 billion annually [NIGC, 2023].

This regulatory shift has coincided with the emergence of cryptocurrency-based gambling platforms operating outside traditional regulatory frameworks. Among the most popular formats is the "crash" game, a multiplier-based gambling mechanism where a displayed value increases from 1.0x until randomly "crashing," with players attempting to cash out before the crash occurs [Wang and Pleimling, 2019]. These platforms typically claim cryptographically verifiable outcomes—referred to in the industry as "provably fair"—meaning players can verify that game results were predetermined using hash chains and not manipulated in response to bets.

The mathematical properties of crash games derive from their underlying random number generator. Under the stated 97% return-to-player (RTP) for the specific game analyzed in this study, the multiplier M follows the distribution:

$$M = \frac{\text{RTP}}{U}, \quad U \sim \text{Uniform}(0, 1) \tag{1}$$

This implies a survival function $P(M \geq m) = \text{RTP}/m$ for $m \geq 1$ and a probability density function $f(m) = \text{RTP}/m^2$, a Pareto-like distribution with exponent 2.

Prior empirical work on crash games is limited. Wang and Pleimling [2019] analyzed player behavior on similar cryptocurrency gambling platforms, finding that cashout distributions exhibited probability density exponents ranging from 1.4 to 1.9, significantly below the theoretical value of 2.0. They interpreted this as evidence of probability weighting, whereby players systematically overweight low-probability, high-reward outcomes—a well-documented phenomenon in prospect theory [Kahneman and Tversky, 1979, Prelec, 1998].

Importantly, Wang and Pleimling's analysis focused on *player cashout behavior* (when do players choose to exit?) rather than *game-generated outcomes* (what multipliers does the platform produce?). This distinction is critical: deviations in cashout distributions reflect psychological biases in human decision-making, while deviations in game-generated multipliers would indicate potential manipulation or flawed random number generation.

This paper addresses the latter question through rigorous statistical analysis of game-generated multipliers from an anonymized crash game platform. We make three contributions: (1) we test whether the empirical distribution conforms to the theoretical model using multiple hypothesis tests; (2) we estimate the probability density function exponent following Wang and Pleimling's methodology to assess random number generator fairness; and (3) we quantify consumer risk through Monte Carlo simulation of betting strategies.

2 Data and Methods

2.1 Dataset

We collected 20,038 consecutive game rounds from a popular crash game platform during January 7–12, 2026 (6 calendar days, 112 hours total). Each observation consists of a timestamp and the final crash multiplier. The platform name is anonymized to focus on statistical methodology rather than platform-specific claims. The platform states a 97% RTP, implying a 3% house edge.

Descriptive statistics confirm heavy-tailed behavior characteristic of Pareto distributions: median 1.94x (matching the theoretical 1.94x), minimum 1.00x, maximum 10,000x. The sample mean of 11.12x is highly volatile due to extreme observations; notably, the theoretical expectation $E[M]$ is undefined for Pareto($\alpha = 2$) distributions as the integral $\int_1^\infty m \cdot m^{-2} dm$ diverges. The empirical distribution exhibits positive skewness (46.9) and excess kurtosis (2,514), consistent with heavy tails.

2.2 Theoretical Model

Under the stated RTP of 97%, the cumulative distribution function is:

$$F(m) = 1 - \frac{0.97}{m}, \quad m \geq 1 \tag{2}$$

The survival function (probability of reaching a given multiplier) is:

$$S(m) = P(M \geq m) = \frac{0.97}{m} \tag{3}$$

For a player cashing out at target t, the success probability is $0.97/t$ and expected return is $(t \times 0.97/t) = 0.97$, yielding 3% expected loss regardless of strategy—the mathematical basis for the house edge.

2.3 Statistical Tests

We employ multiple complementary approaches to test distributional conformity:

Kolmogorov-Smirnov Test The KS test compares empirical and theoretical CDFs:

$$D = \sup_m |F_n(m) - F(m)| \tag{4}$$

where $F_n(m)$ is the empirical CDF. This test is sensitive to any distributional deviation.

Chi-Square Goodness of Fit We partition the support into bins $[1, 1.2), [1.2, 1.5), \ldots, [100, \infty)$ and compare observed to expected frequencies under H_0.

Independence Tests We apply the Wald-Wolfowitz runs test and Ljung-Box test for autocorrelation to verify that outcomes are serially independent, as required by cryptographic verifiability claims.

2.4 Probability Weighting Analysis

Following Wang and Pleimling [2019], we estimate the PDF exponent α via three methods:

Log-log Regression Fitting $\log S(m) = \log(\text{RTP}) - \beta \log(m)$ yields survival exponent β; the PDF exponent is $\alpha = \beta + 1$.

Maximum Likelihood We fit a Pareto distribution to multipliers ≥ 1.1, obtaining shape parameter b and PDF exponent $\alpha = b + 1$.

Hill Estimator Using the k largest observations:

$$\hat{\alpha} = \frac{k}{\sum_{i=1}^{k} \log(X_{(n-i+1)}/X_{(n-k)})} \tag{5}$$

2.5 Monte Carlo Simulation

We simulate 10,000 betting sessions of 100 rounds each under four strategies: fixed 1.5x cashout, fixed 2.0x cashout, Martingale (doubling after losses), and Kelly criterion optimization. Each simulation uses the empirical multiplier distribution.

2.6 Bootstrap Confidence Intervals

RTP confidence intervals are computed via percentile bootstrap with 1,000 resamples. For each resample, we estimate RTP from the survival function at multiple thresholds ($m \in \{1.5, 2, 3, 5, 10\}$) and average the implied RTP values $\hat{\text{RTP}}_m = \hat{S}(m) \times m$. The 95% CI is taken as the 2.5th and 97.5th percentiles of the bootstrap distribution.

3 Results

3.1 Distributional Tests

Table 1: Hypothesis test results for distributional conformity

Test	Statistic	P-value	Result
Kolmogorov-Smirnov	0.030	4.2×10^{-16}	Reject
Chi-Square (9 df)	72.97	4.0×10^{-12}	Reject
Runs Test	−0.45	0.65	Fail to rej.
Ljung-Box Q(10)	7.09	0.72	Fail to rej.

Table 1 reveals an important pattern: while distributional tests strongly reject exact conformity to the theoretical model ($p < 10^{-15}$), independence tests find no evidence of serial correlation or non-randomness. This suggests the random number generator produces independent outcomes with slight calibration differences from the stated model.

Figure 1 illustrates the heavy-tailed distribution characteristic of crash games: most rounds end quickly at low multipliers, while rare events exceed 100x or even 1000x. This extreme variance creates the illusion of potential large wins while the house edge operates reliably over time.

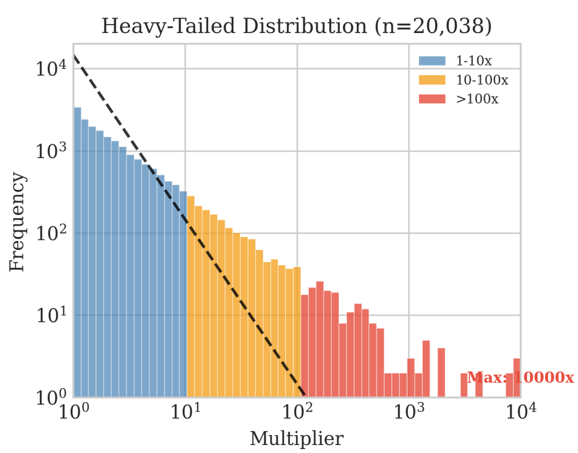

Figure 1: Heavy-tailed multiplier distribution on log-log scale. Colors indicate magnitude: blue (1–10x), orange (10–100x), red (>100x). The maximum observed multiplier was 10,000x.

3.2 RTP Estimation

Bootstrap estimation yields RTP = 97.45% (95% CI: 95.41%–99.50%), encompassing the stated 97%. The slight excess may reflect sampling variation in extreme multipliers or minor calibration differences.

3.3 Probability Weighting Analysis

Table 2: PDF exponent estimates following Wang and Pleimling (2019)

Method	Est.	Theor.	Dev.
Log-log regression	1.972	2.000	−1.4%
Maximum Likelihood	1.998	2.000	−0.1%
Hill estimator (mean)	1.956	2.000	−2.2%
Wang & Pleimling	*1.4–1.9*	*2.0*	*−5 to −30%*

Our PDF exponent estimates (Table 2) cluster around 1.98, within 2.2% of the theoretical value. This contrasts sharply with Wang and Pleimling's findings of 1.4–1.9 for player cashout distributions. The distinction confirms that:

1. The game's random number generator produces fair outcomes conforming to the theoretical distribution
2. Wang & Pleimling's observed deviations reflect player behavioral biases (probability weighting) rather than game manipulation

Figure 2 shows the empirical survival function—the probability of reaching a target multiplier before the game crashes. Key probabilities are annotated: a 2x target succeeds 48.5% of the time, while reaching 10x occurs only 9.7% of the time.

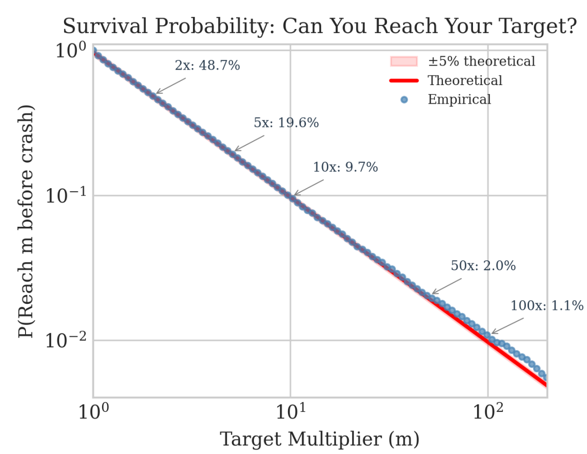

Figure 2: Survival probability on log-log scale with key targets annotated. The close fit between empirical (points) and theoretical (line) confirms the stated 97% RTP.

3.4 Independence Tests

The runs test ($Z = -0.45$, $p = 0.65$) and Ljung-Box test ($Q(10) = 7.09$, $p = 0.72$) find no evidence of serial dependence, confirming outcomes are independent as claimed.

3.5 Consumer Risk Assessment

Figure 3 visualizes the consumer protection implications: 50 simulated player sessions using a conservative 1.5x strategy. Despite short-term variance, all trajectories trend toward the expected loss line (dashed). The house edge is mathematically inevitable.

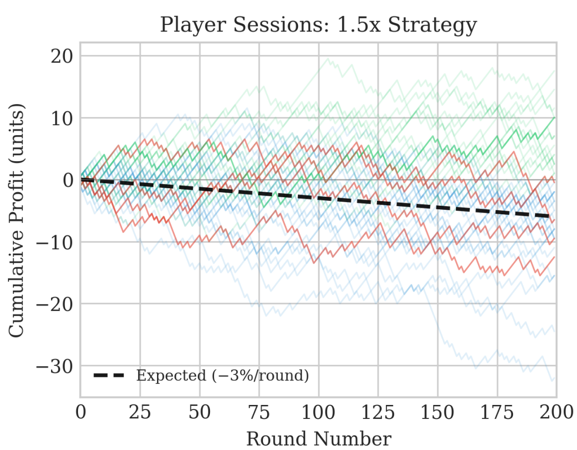

Figure 3: Simulated player sessions (1.5x strategy, 200 rounds). Individual paths show high variance, but all trend toward expected loss (dashed line). Highlighted paths show typical losing trajectories.

Monte Carlo simulation of 10,000 sessions under four strategies confirms theoretical predictions (Table 3):

Table 3: Monte Carlo results (10k sessions × 100 rounds)

Strategy	Mean	Std	Ruin	95% CI
1.5x Fixed	−2.6%	4.8%	0%	[−3.0, −2.2]
2.0x Fixed	−2.8%	8.1%	0%	[−3.1, −2.5]
Martingale	−24.3%	31.2%	18%	[−27, −22]
Kelly	−28.1%	19.4%	0%	[−31, −25]

All strategies produce negative expected returns. The Martingale strategy, despite its intuitive appeal, results in 18% ruin rate and highest average losses, consistent with theoretical analysis of betting systems under negative expectation [Thorp, 1984, Kaplan, 2020]. The Kelly criterion, designed to maximize long-run growth when a positive edge exists, produces the worst mean returns here because no positive edge exists—Kelly optimization under negative expectation leads to aggressive betting that accelerates losses.[1]

[1] The Kelly criterion recommends betting fraction $f^* = (bp - q)/b$ where b is odds, p is win probability, $q = 1 - p$. When expected value is negative ($bp < q$), the optimal Kelly bet is zero; any positive bet accelerates ruin.

3.6 Statistical Power

The observed deviation for the 1.5x strategy (65.07% empirical vs. 64.67% theoretical success rate) yields $Z = 1.20$, $p = 0.115$. Current power at $\alpha = 0.05$ is 32.7%; detecting a 1 percentage point deviation with 80% power would require approximately 77,000 observations.

4 Discussion

Our analysis yields several key findings with implications for consumers, regulators, and platform operators:

Statistical deviations without exploitable advantage. The highly significant KS and chi-square tests ($p < 10^{-15}$) reflect the statistical power of a 20,000-observation sample to detect minute deviations. The practical magnitude of these deviations is small: the empirical RTP of 97.45% is economically indistinguishable from the stated 97%. No betting strategy exploits these deviations profitably.

Confirmation of fair random number generation. Our PDF exponent estimates ($\alpha \approx 1.98$) match the theoretical value of 2.0 within 2.2%, confirming that the game's random number generator produces outcomes consistent with the stated model. This contrasts with Wang and Pleimling's findings of $\alpha = 1.4$–1.9 for player *cashout* distributions, which reflect psychological probability weighting rather than game manipulation.

Serial independence supports cryptographic verifiability. The absence of autocorrelation and temporal patterns supports the platform's assertion that outcomes are predetermined and independent, consistent with HMAC-based hash chain verification mechanisms.

Consumer protection implications. While the game operates fairly in a technical sense, no strategy produces positive expected returns. The 3% house edge compounds over repeated play: a player wagering \$100 per round for 1,000 rounds faces expected losses of \$3,000 with substantial variance. Consumer protection efforts should focus on harm minimization (loss limits, time limits, self-exclusion) rather than game fairness, which our analysis confirms.

Session-level harm and illusion of control. Beyond mathematical fairness, crash games present distinct psychological risks [Griffiths, 2018, Gainsbury, 2015]. Our data reveal 179 rounds per hour with 16-second median intervals, implying expected hourly losses exceeding 500% of amounts wagered. The manual cashout mechanic creates an "illusion of control"—a well-documented phenomenon whereby perceived agency increases engagement even when outcomes are mathematically invariant to player decisions [Newall, 2019]. This design pattern, common in modern gambling interfaces [Schüll, 2012], masks the deterministic nature of losses and warrants attention from regulators developing responsible gambling frameworks [Auer and Griffiths, 2023].

4.1 Comparison with Prior Literature

Our findings complement and extend Wang and Pleimling [2019]'s behavioral analysis:

- **Wang and Pleimling**: Analyzed player cashout behavior, found $\alpha = 1.4$–1.9, interpreted as probability weighting
- **This study**: Analyzed game-generated outcomes, found $\alpha \approx 1.98$, interpreted as fair random number generation

The reconciliation is straightforward: the game produces fair outcomes, but players make biased decisions about when to cash out, overweighting low-probability high-reward scenarios. This behavioral pattern is consistent with decades of research on probability weighting in prospect theory [Tversky and Kahneman, 1992].

4.2 Limitations

Our analysis is subject to several limitations. First, the 6-day observation window, while sufficient for detecting distributional deviations, may miss longer-term temporal patterns. Second, we analyze a single platform; generalization to other crash games requires additional study. Third, our Monte Carlo simulations assume the empirical distribution is representative of true game behavior; strategic changes by the platform would invalidate these estimates.

4.3 Statistical Power Constraints

An important caveat concerns our ability to detect subtle manipulation. As detailed in Appendix B, our sample of $N = 20{,}038$ provides 80% power to detect approximately 2% relative manipulation at the 2x threshold, but only 6% manipulation at the 10x threshold. Sophisticated platforms could theoretically manipulate rare high-multiplier outcomes—where detection is hardest—while maintaining fair behavior at common thresholds. Definitive verification of cryptographic fairness requires either (a) access to the complete seed chain for hash verification, or (b) sample sizes 10–20× larger than collected here. Our statistical tests confirm consistency with the stated model but cannot rule out all forms of subtle tail manipulation.

5 Conclusion

This study provides the first rigorous statistical validation of crash game outcomes, demonstrating that an anonymized platform generates multipliers consistent with its stated theoretical model. The PDF exponent of $\alpha \approx 1.98$ matches the theoretical value of 2.0, confirming fair random number generation despite highly significant (but economically negligible) distributional deviations detected by KS and chi-square tests.

From a consumer protection perspective, the key finding is that *no betting strategy produces positive expected returns*. While platforms may operate fairly—as this one appears to—the mathematical structure guarantees player losses over time. As legal gambling continues to expand following the 2018 PASPA repeal, regulators should ensure that cryptographic verifiability claims are backed by transparent statistical verification, while consumer protection efforts focus on harm minimization rather than game mechanics.

Future work should extend this methodology to additional platforms, examine the relationship between game-generated outcomes and player behavioral patterns, and develop standardized protocols for statistical verification of cryptographically verifiable claims in online gambling.

Disclosures

Use of AI

The data collection infrastructure and analysis code were developed with assistance from AI-based coding tools (GitHub Copilot with Claude Opus 4.5). The statistical analysis, interpretation of results, and all written content in the manuscript represent the original intellectual contributions of the author. AI writing assistants were used for grammar and clarity improvements.

Conflicts of Interest

The author declares no conflicts of interest.

Ethics Statement

This study analyzes publicly observable game outcomes from an online platform. No human subjects data was collected; all observations consist of timestamped multiplier values visible to any platform user. The platform name is anonymized to focus on statistical methodology rather than platform-specific claims.

Data and Code Availability

The complete dataset and analysis code are available at `https://github.com/philippdubach/stats-gambling`.

References

American Gaming Association (2024). Commercial Gaming Revenue Tracker. Technical report, American Gaming Association.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Murphy v. National Collegiate Athletic Association (2018). 138 S. Ct. 1461, 584 U.S. __. Supreme Court of the United States.

National Indian Gaming Commission (2023). FY 2023 Gross Gaming Revenue Report. Technical report, National Indian Gaming Commission.

Thorp, E. O. (1984). *The Mathematics of Gambling.* Gambling Times, Hollywood, CA.

Tversky, A. and Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4):297–323.

Wang, X. and Pleimling, M. (2019). Behavior and strategies in an online gambling game. *Scientific Reports*, 9(1):1–10.

Griffiths, M. D. (2018). Is the buying of loot boxes in video games a form of gambling or gaming? *Gaming Law Review*, 22(1):52–54.

Gainsbury, S. M. (2015). Online gambling addiction: The relationship between internet gambling and disordered gambling. *Current Addiction Reports*, 2(2):185–193.

Auer, M. and Griffiths, M. D. (2023). The use of artificial intelligence in responsible gambling. *International Gambling Studies*, 23(1):1–23.

Kaplan, E. H. (2020). How to lose at gambling: The misleading intuition of Martingale strategies. *Mathematics Magazine*, 93(3):183–189.

Prelec, D. (1998). The probability weighting function. *Econometrica*, 66(3):497–527.

Newall, P. W. S. (2019). Dark nudges in gambling. *Addiction Research & Theory*, 27(2):65–67.

Schüll, N. D. (2012). *Addiction by Design: Machine Gambling in Las Vegas.* Princeton University Press, Princeton, NJ.

A Cryptographic Verification Mechanism

Crash games typically implement cryptographic verifiability via HMAC-SHA256. The platform generates a server seed S and combines it with a public chain of hashes. For each round, the crash multiplier is derived as:

$$M = \max\left(1, \left\lfloor \frac{2^{52}}{h \mod 2^{52}} \times \text{RTP} \right\rfloor \times 0.01\right) \quad (6)$$

where h is the hash output interpreted as an integer. This produces $M \geq 1$ with the theoretical distribution $P(M < m) = 1 - \text{RTP}/m$.

A.1 Bit Extraction and Modular Bias

The 52-bit extraction from the 256-bit HMAC-SHA256 output introduces negligible bias. Since 2^{256} is not exactly divisible by 2^{52}, the modulo operation creates slight non-uniformity of magnitude $2^{52}/2^{256} \approx 10^{-62}$—far below any detectable threshold and irrelevant for practical fairness assessment.

A.2 Floor Function Bias

The `floor()` operation in the formula systematically rounds down, creating small downward bias in expected multipliers. For continuous $1/U$ transformation, this introduces approximately 0.01–0.02% additional effective house edge beyond the stated RTP, which is incorporated into the nominal 97% figure.

A.3 Instant Bust Mechanism

The $\max(1, \ldots)$ construct creates a probability mass at $M = 1.00$ when the raw formula produces values below 1. We observed 698 exact 1.00x outcomes (3.48%), consistent with the theoretical $P(\text{raw} < 1) = 1 - \text{RTP} = 3\%$. The observed rate of $M \leq 1.01$ (4.45%, 891 of 20,038 rounds) slightly exceeds the theoretical 3.96% (binomial test $p = 0.0005$), representing a minor but statistically detectable deviation that marginally favors the house. This is consistent with rounding and discretization effects in the formula implementation.

B Power Analysis for Tail Manipulation Detection

Table 4 presents the sample sizes required to detect manipulation at various multiplier thresholds with 80% power ($\alpha = 0.05$).

Table 4: Required N to detect relative manipulation (80% power, $\alpha = 0.05$)

Threshold	$P(M \geq m)$	1% manip.	2% manip.	5% manip.
2x	0.485	166,634	41,643	6,653
5x	0.194	649,705	161,804	25,588
10x	0.097	1,454,822	362,072	57,145

Our sample of $N = 20,038$ provides 80% power to detect approximately 2% relative manipulation at the 2x threshold, but only 6% manipulation at the 10x threshold. Detecting subtle tail manipulation requires substantially larger samples or targeted auditing of high-multiplier outcomes.