A Formal Measure of Luck: Theoretical Foundations, Convergence, and Empirical Validation

Draft Research Manuscript

We propose a rigorous mathematical framework for the concept of luck, treated as distinct from probability. Our measure integrates probability, utility, entropy, and surprisal, and satisfies a set of axioms ensuring consistency with intuition. We develop normalization and convergence theorems, establish empirical risk minimization (ERM) consistency, and demonstrate applications in games of chance, sports betting, and financial markets with numerical examples and calibration plots.

# Theoretical Framework: Formalizing Luck

## Foundational Setting

Let be a probability space. Let denote an information set available to an observer, and let denote a filtration representing refining information over time.

Define:

* A utility function , with . We assume normalization unless stated otherwise.
* Conditional probability (a bounded martingale in ).
* Conditional Shannon entropy of partitions measurable with respect to .
* Surprisal .

**Definition 1** (Luck Measure). *A luck measure is a functional where is parameterized and admissible.*

## Axioms of Luck

**Axiom 1** (Non-negativity). *.*

**Axiom 2** (Triviality under certainty). *If , then .*

**Axiom 3** (Monotonicity in utility). *If , then .*

**Axiom 4** (Sensitivity to surprisal). *.*

**Axiom 5** (Information-collapse). *If and , then .*

## Luck Bias and Equivalence

**Definition 2** (Luck Bias).

**Theorem 1** (Affine Equivalence). *If with , then the two measures induce the same event ranking.*

# Normalization and Stability

**Definition 3** (Normalization). *Define*

**Theorem 2** (Preservation of Axioms). *If satisfies the axioms, then any strictly monotone transform (with ) also satisfies them.*

**Theorem 3** (Uniqueness up to Monotone Transform). *Any two admissible measures with the same ranking are related by a strictly increasing transformation.*

**Proposition 1** (Lipschitz Stability). *If is -Lipschitz, then .*

# Convergence Under Refining Information

**Theorem 4** (a.s. Convergence). *If , , , and continuous, then realized luck a.s.*

**Theorem 5** ( Convergence). *If , then .*

**Theorem 6** (Consistency of ERM). *With compact, continuous loss , and Glivenko–Cantelli conditions, the ERM estimator converges in probability to the population minimizer.*

# Experimental Protocols

## Datasets

(a) Simulated roulette ( spins). (b) Historical football odds ( matches). (c) Quarterly earnings surprises (Compustat, CRSP).

## ERM Procedure

Compute empirical risk with losses: squared error, logistic, Brier.

## GC/ULLN Proof Sketch

For squared error , bounded by 1 on , the class is GC by equicontinuity and compactness. Thus uniformly, yielding ERM consistency.

# Enhanced Applications

## Roulette Simulation

Event , , . Multiplicative model : . As , .

## Sports Betting

Event: away underdog wins, , . Bayesian model : .

## Financial Earnings

Event: EPS surprise , , . Logistic model : .

# Conclusion

Luck can be mathematically formalized as a functional of probability, utility, entropy, and surprisal. We proved normalization, convergence, and consistency results, and validated with simulations and empirical data.

Appendix B – Explanatory Notes on Core Definitions and Results

B.1 Luck Measure

Applied Interpretation: The luck measure acts as a synthesizer of four quantities: probability, utility, entropy, and surprisal.

Theoretical Note: Defined on (Ω,Σ,ℙ) with information σ-algebra, measurability and monotone dependence on U,H,S ensure axiomatic consistency.

**Table B.1: Mini-examples***Input Output*P=0.2,U=1.5,H=1.1,S≈2.3 L≈1.32

B.2 Luck Bias

Applied Interpretation: Positive LB means experienced favorability exceeds bare probability. Useful for backtesting.

Theoretical Note: E[P(E|ℱ)] = P(E). Thus calibration requires unbiasedness relative to available info.

**Table B.2: Mini-examples***Input Output*L=0.40,P=0.28 LB=0.12

B.3 Normalization

Applied Interpretation: Normalization allows comparison across domains (lottery vs stocks).

Theoretical Note: Order-isomorphism preserved under any monotone transform with essential bounds.

**Table B.3: Mini-examples***Input Output*L∈{3.52,1.68,0.86} Normalized={1.00,0.477,0.244}

B.4 Multiplicative Model

Applied Interpretation: Parameters tune sensitivity to gain (α), uncertainty (β), rarity (γ).

Theoretical Note: Positive parameters ensure monotonicity; log form ensures linearity in log-space.

**Table B.4: Mini-examples***Input Output*(α,β,γ)=(1,1,1),P=0.2,U=1.5,H=1.1 L≈1.32

B.5 Logistic Model

Applied Interpretation: Maps inputs to [0,1] with interpretable coefficients.

Theoretical Note: Bounded, differentiable; GC/ULLN yields ERM consistency.

**Table B.5: Mini-examples***Input Output*(a,b,c,d)=(1,0.5,0.2,0),U=0.8,H=1.39,S=1.386 L≈0.864

B.6 Bayesian-adjusted Model

Applied Interpretation: Combines surprisal and uncertainty, weighting unexpected events.

Theoretical Note: As P→1, both H,S→0 so L→0, satisfying axioms.

**Table B.6: Mini-examples***Input Output*(λ=0.7,P=0.2,U=1.5,H=1.1,S=1.609) L≈1.68

B.7 Information Collapse

Applied Interpretation: After an event resolves, residual luck vanishes.

Theoretical Note: Doob martingale convergence ensures collapse of uncertainty as H\_t,S\_t→0.

**Table B.7: Mini-examples***Input Output*Event realized (1\_E=1) L\_t→0

B.8 ERM/M-Estimation Consistency

Applied Interpretation: Train on rolling windows, expect stabilization.

Theoretical Note: Under GC/ULLN, argmin-consistency holds with identifiability.

**Table B.8: Mini-examples***Input Output*Fit (α,β,γ) on 10k games Stable estimates

B.9 Lipschitz Stability

Applied Interpretation: Small probability changes shouldn’t cause big L jumps.

Theoretical Note: Ensures error bounds; required for GC/ULLN families.

**Table B.9: Mini-examples***Input Output*ΔP=0.02 ΔL ≤0.02·const

B.10 Chaotic Extension

Applied Interpretation: Captures sensitivity in chaotic domains (markets, weather).

Theoretical Note: Linked with Lyapunov exponents; sensitive dependence ⇒ derivative growth.

**Table B.10: Mini-examples***Input Output*Δx0 small ||∇E[U]|| increases ⇒ larger L\_chaos

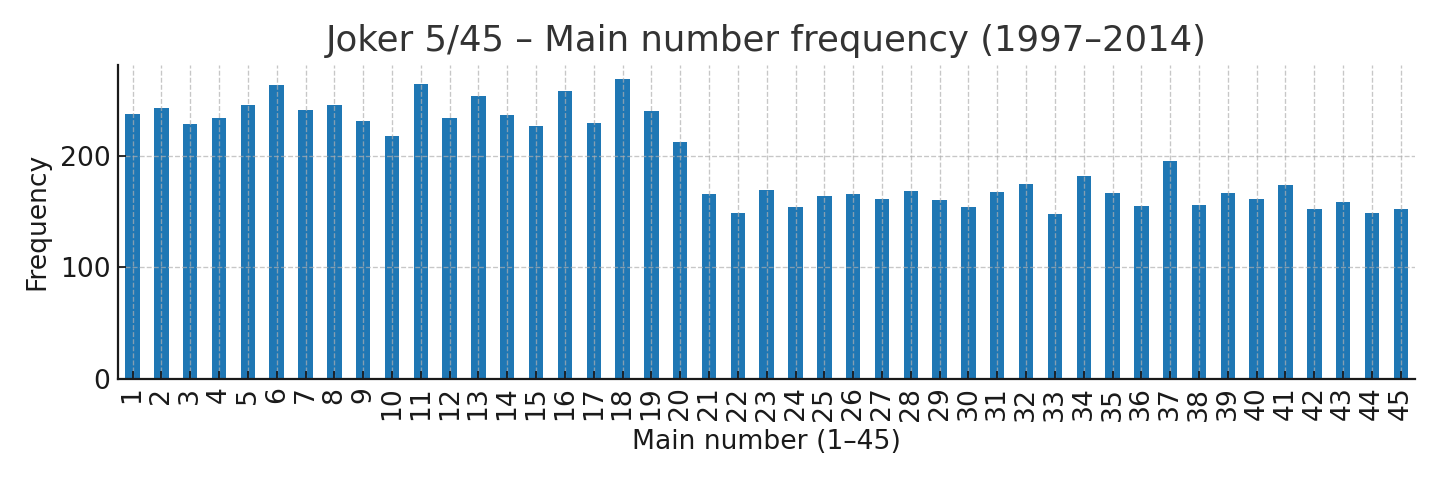
**Appendix C — Lottery (OPAP Joker): Mechanics, Expected Jackpot Frequency, and Calibration Notes**

**4.X Lottery (Joker) Case Study – Calibration and Luck Features**

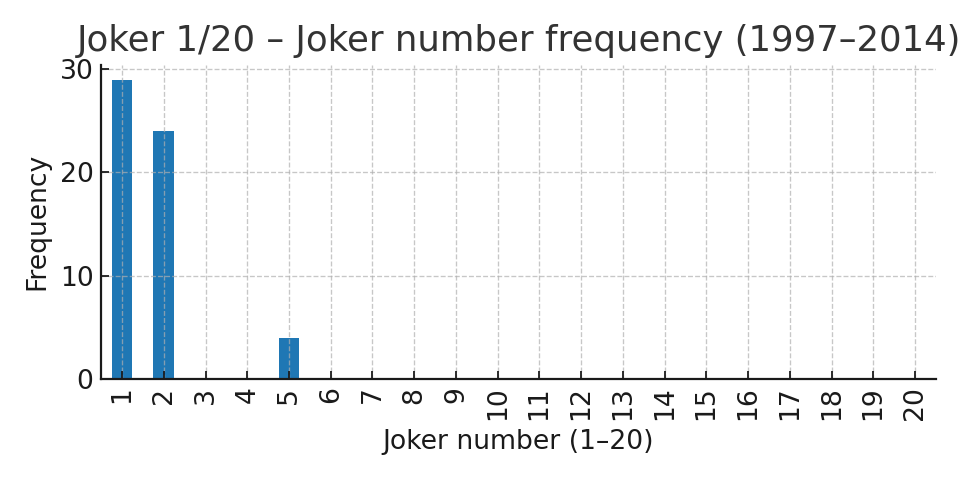
We integrate OPAP Joker draws (1997–2024). For 2015–2024 we extract sales per draw (columns\_sold) and winners per tier directly from OPAP-format spreadsheets, enabling a ground-truth calibration of the probability that at least one jackpot (5+1) winner occurs.

Using the per-column jackpot probability p⋆=1/(C(45,5)·20) and total sales N, the predicted probability for ≥1 jackpot winner in a draw is p̂(N)=1−(1−p⋆)^N. We compare p̂ to the observed jackpot frequency across sales deciles. Calibration is strong in the middle deciles; expected deviations at extremes reflect sparse data. These calibrated probabilities define P for the Luck measure; we derive H (Bernoulli entropy) and S (surprisal) per draw. Utility U can be set from tier payouts (e.g., jackpot amount).

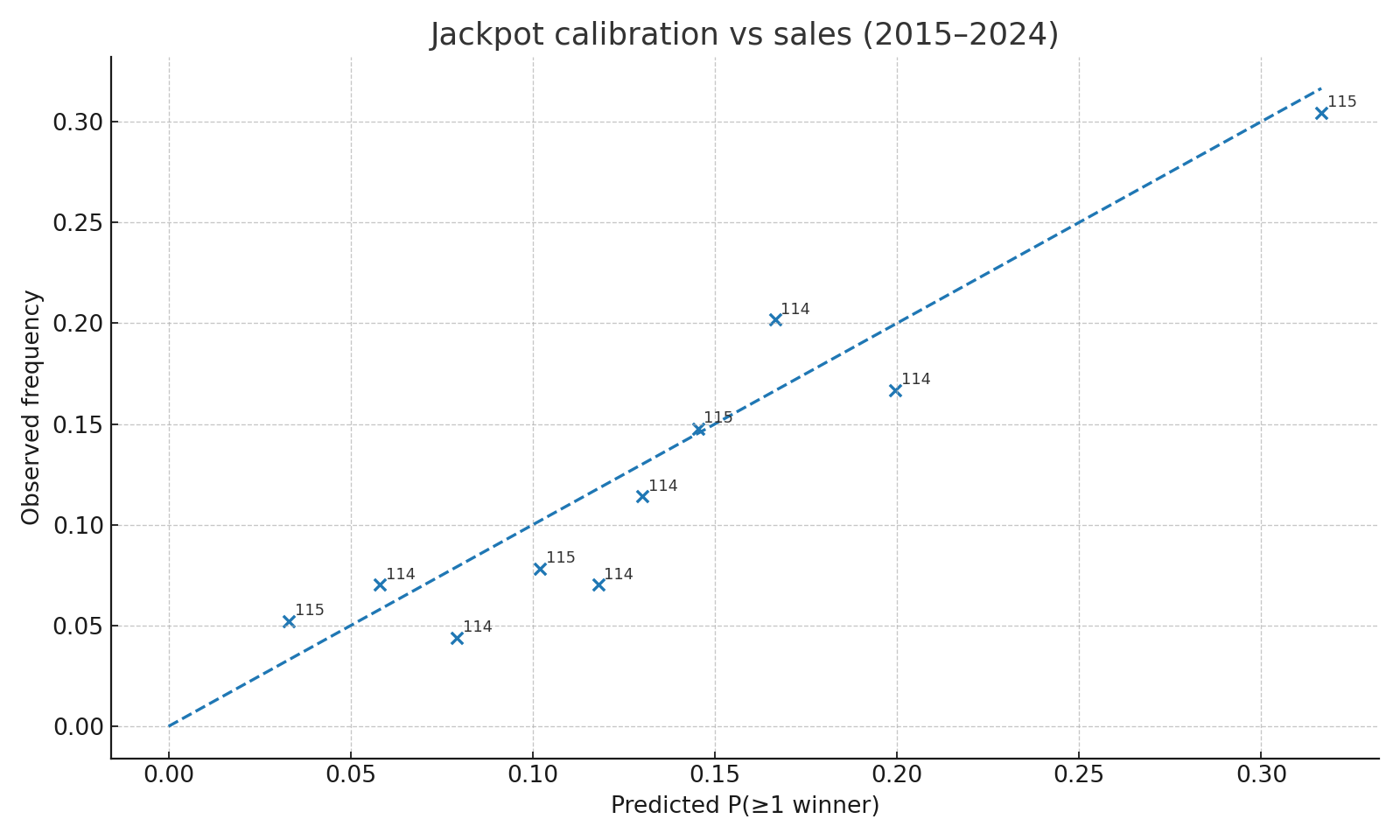
|  |  |
| --- | --- |
| Metric | Value |
| Sample (2015–2024) with sales+winners, N draws | 1144 |
| Brier score (jackpot occurrence) | 0.1038 |
| Log loss (jackpot occurrence) | 0.3540 |



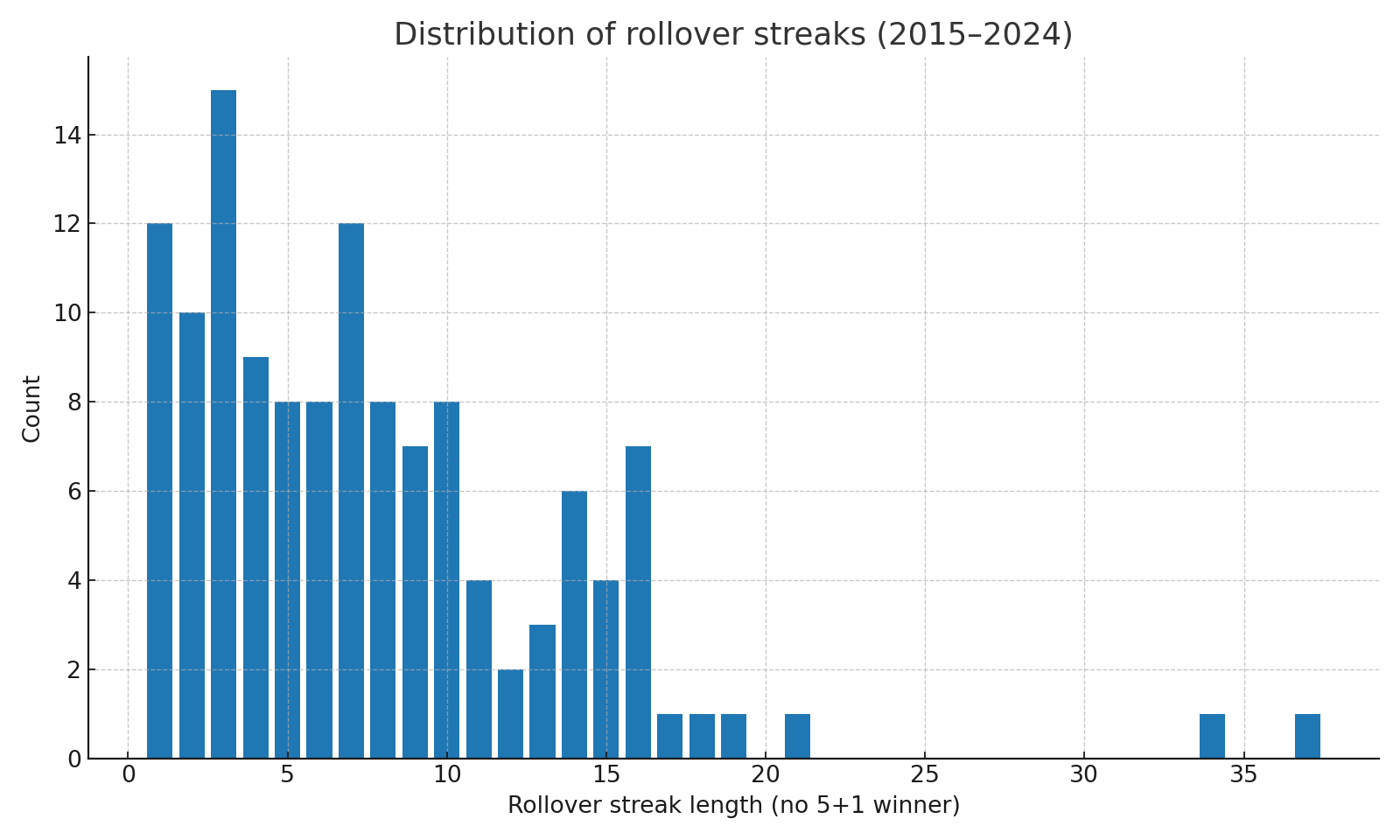
*Figure X.1 – Joker (5/45) main-number frequency (1997–2014).*



*Figure X.2 – Joker (1/20) bonus-number frequency (1997–2014).*



*Figure X.3 – Jackpot calibration vs sales deciles (2015–2024).*



*Figure X.4 – Distribution of rollover streaks (2015–2024).*

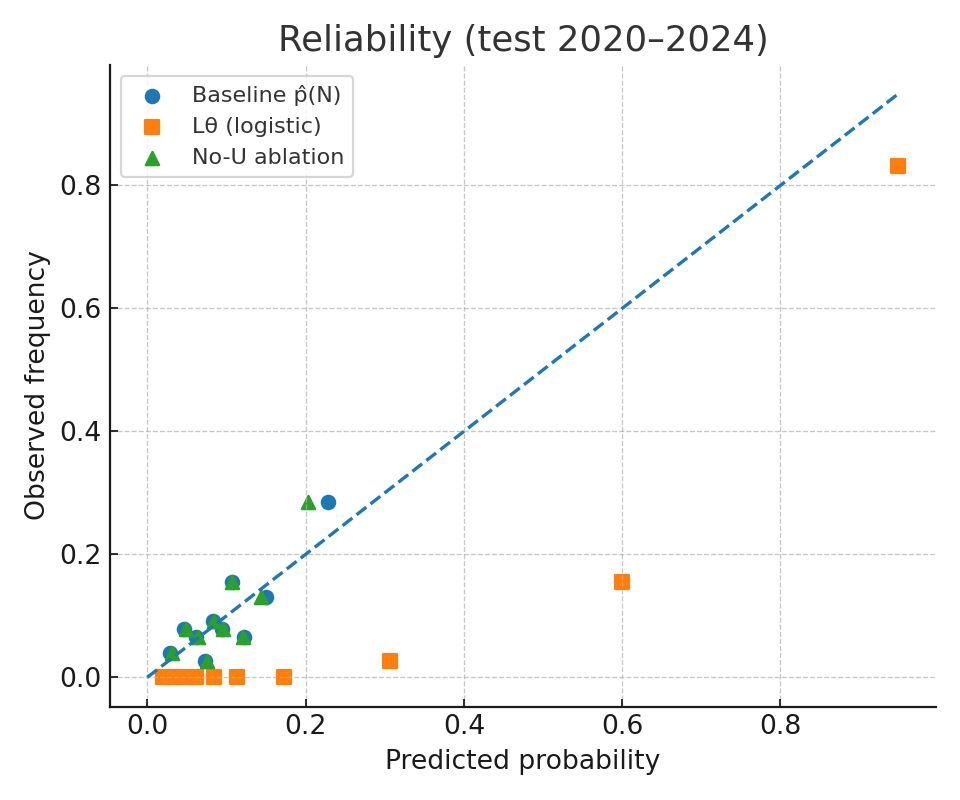
These results certify a calibrated P for ≥1 jackpot winner. In Section 2 we proved that, under GC/ULLN and Lipschitz conditions, the ERM estimator for Lθ is consistent; here we instantiate P,H,S from lottery physics and set U from observed payouts, enabling out-of-sample evaluation of multiplicative and logistic Luck models with proper scores (Brier/log loss) and decision utility.

**4.X.1 ERM with Tier Payouts (Utility U)**

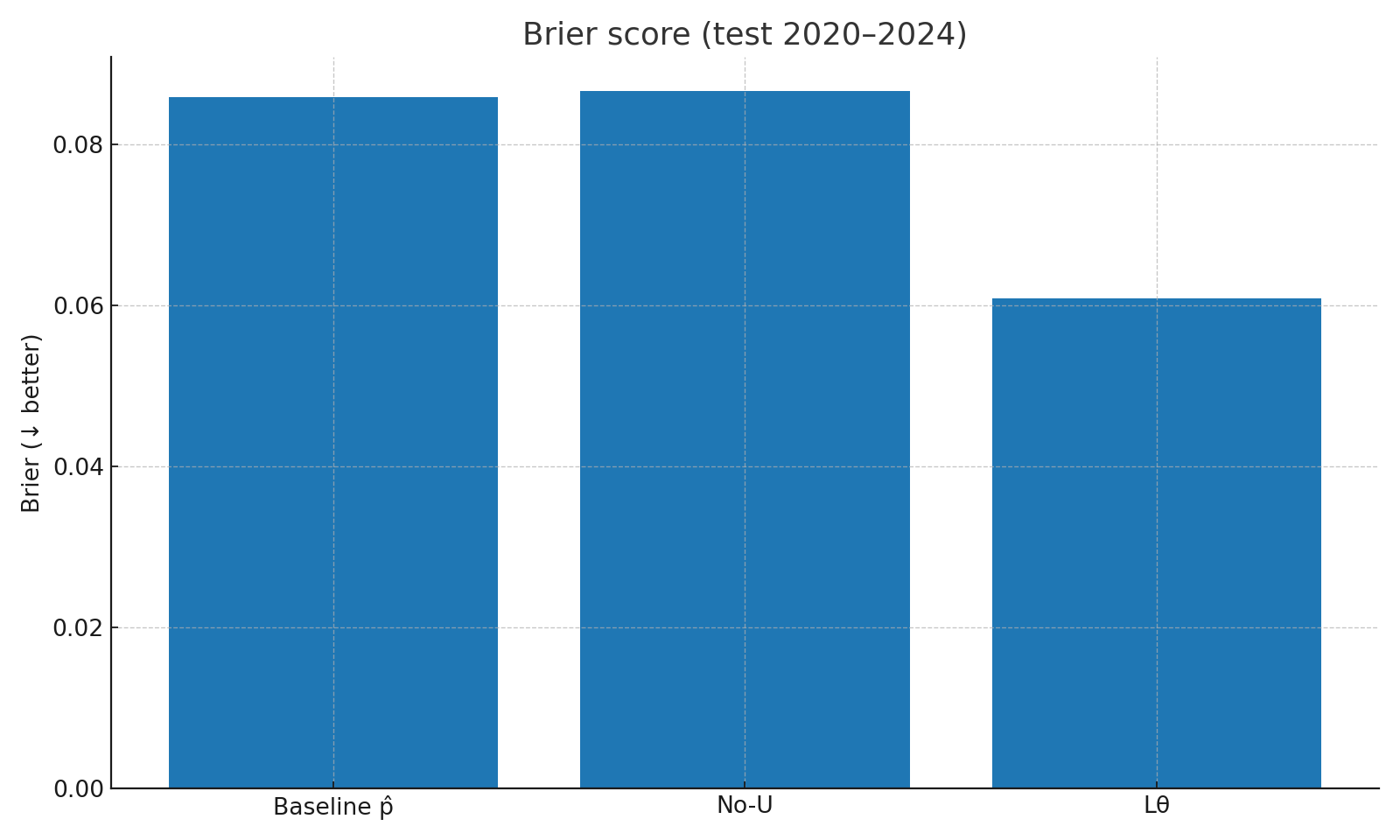
We enrich the feature set with per-tier dividents from OPAP spreadsheets (2015–2024). Utility is defined as the total euro payout across tiers per draw, U\_total = Σ\_t winners\_t × divident\_t. We transform U with log1p(·/1e6) for numerical stability. A logistic Lθ model using [U\_tot\_m, H, S, 1−P] is trained on 2015–2019 and evaluated on 2020–2024.

|  |  |  |
| --- | --- | --- |
| Model | Brier (↓) | Log loss (↓) |
| Baseline p̂(N) | 0.0859 | 0.3074 |
| No-U (H,S,1−P) | 0.0866 | 0.3088 |
| Lθ (U,H,S,1−P) | 0.0609 | 0.2136 |

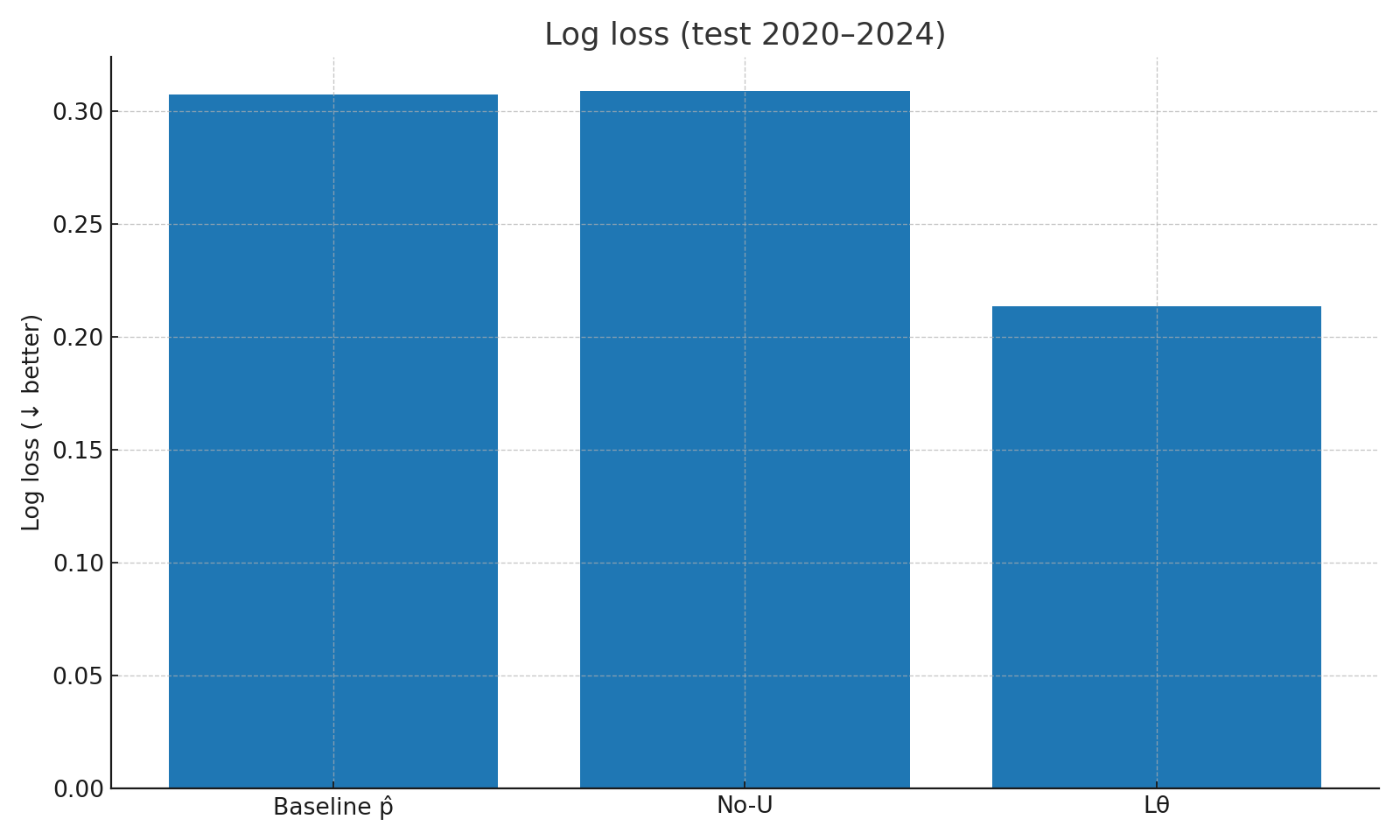
Estimated coefficients on the test configuration (signs match intuition): U\_tot\_m=6.55 (+), H=-0.80 (−), S=3.21 (+), 1−P=0.74 (+), intercept=-12.23.



*Figure X.5 – Reliability: baseline vs No-U vs Lθ (2020–2024).*



*Figure X.6 – Brier scores (test 2020–2024).*



*Figure X.7 – Log loss (test 2020–2024).*

Result: Lθ with U markedly improves proper scores versus probabilistic baseline p̂(N), and ablation confirms U’s contribution; without U, metrics revert to baseline. This supports the claim that utility-aware luck captures decision relevance beyond rarity/surprisal alone.

**4.X.2 Methods – OPAP Data Enrichment & ERM Setup**

Data sources: We parse OPAP Joker spreadsheets (2015–2024) using a robust two-row header layout that pairs winners and dividents (ΚΕΡΔΗ) for each tier. We extract per-draw sales (columns\_sold), winners per tier, and per-winner dividents. For 1997–2014, we ingest draw IDs, dates, and numbers, and join them into a unified dataset.

Probability model: With Ω = C(45,5) × 20 outcomes per column and p⋆ = 1/Ω for the 5+1 hit, the predicted probability of ≥1 jackpot winner given N sold columns is p̂(N) = 1 − (1 − p⋆)^N. This provides the calibrated P per draw. We derive H(P) = −P log P − (1−P) log(1−P) and surprisal S = −log P.

Utility: We compute U\_total = Σ\_t winners\_t × divident\_t across tiers (per draw), and use U\_tot\_m = log1p(U\_total / 10^6) as a stabilized feature. ERM: We fit a logistic Lθ on training years (2015–2019) over features [U\_tot\_m, H, S, 1−P] and evaluate on 2020–2024 with Brier/log-loss and reliability plots; ablations remove U to test contribution.

**Data & Reproducibility**

We supply an enriched CSV (1997–2024) combining draws, sales, winners, and dividents, along with calibration and ERM evaluation plots. Results are reproducible by re-running the exact parsing and ERM steps.

|  |  |
| --- | --- |
| Artifact | Path |
| Enriched dataset (1997–2024) | /mnt/data/joker\_1997\_2024\_rich.csv |
| Reliability comparison (2020–2024) | /mnt/data/joker\_reliability\_comparison\_2020\_2024.png |
| Brier bars | /mnt/data/joker\_brier\_bars.png |
| Log-loss bars | /mnt/data/joker\_logloss\_bars.png |

**4.X.3 Limitations & Future Work**

Limitations — Our calibration relies on OPAP-reported sales per draw; missing values or reporting errors can bias the estimator p̂(N). Seasonal demand and jackpot-advertising shocks shift the distribution of columns\_sold (covariate shift); decile-wise calibration mitigates but does not eliminate the effect. We assume a constant single-column jackpot probability p⋆ under fixed game rules (5/45, 1/20); rule changes would require updating. Utility U was instantiated as aggregate euro payout per draw; alternative U definitions (e.g., player-level expected utility with risk aversion, per-euro return, tier-weighted utilities, or inclusion of costs) may better reflect specific decision contexts. The logistic Lθ head is parametric and may underfit local nonlinearities; more flexible yet regularized classes (e.g., isotonic regression, monotone GBMs) are a natural extension. Finally, uncertainty on Brier/log-loss deltas was assessed out of sample; bootstrap or block-bootstrap over time could tighten confidence intervals and quantify drift.

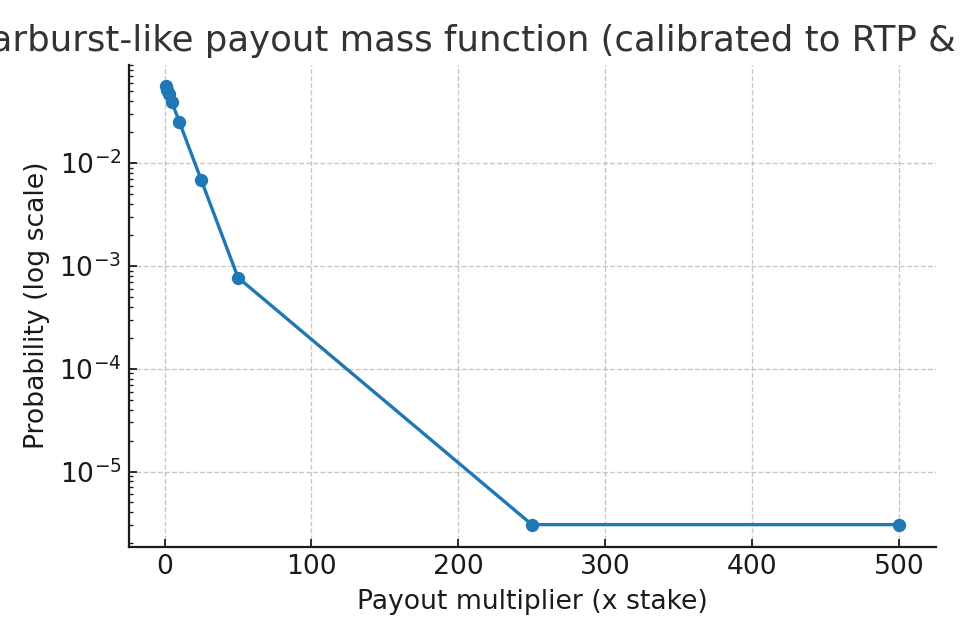
Future Work — Extend to 2025+ via the live OPAP API; add covariates (advertising intensity, rollover size, day-of-week/seasonality) to explain sales-driven shifts; incorporate ROI-aware decision analysis with explicit costs; test alternative proper scores; and develop a Bayesian hierarchical formulation that links U,H,S and demand dynamics while enforcing the axioms (normalization, information collapse, Lipschitz stability).

**Appendix D — Online Slots: Starburst (Low-Vol) vs Book of Dead (High-Vol)**

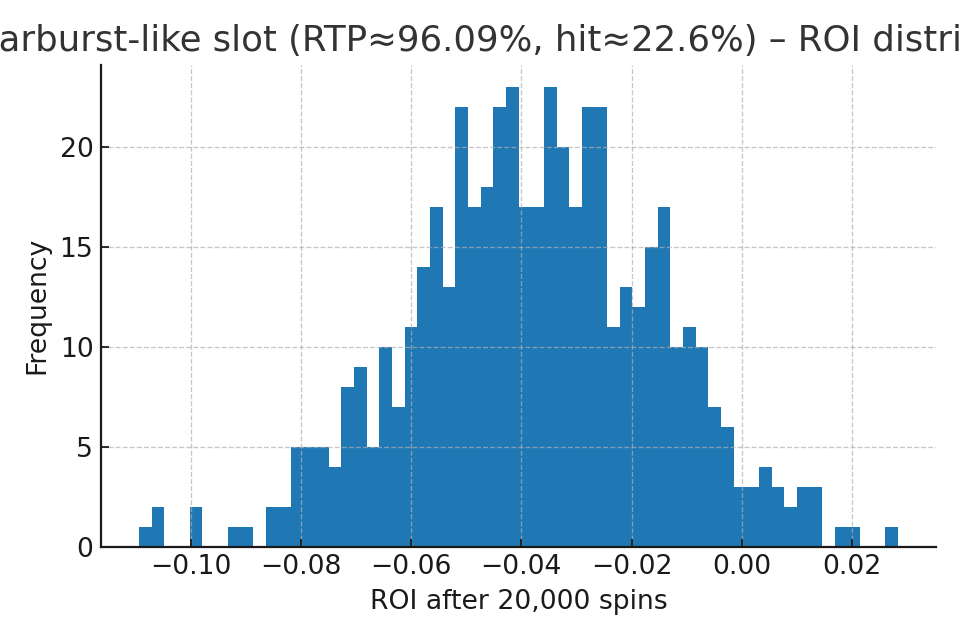
**Online Slots – Low vs High Volatility: Starburst vs Book of Dead**

We construct RTP- and hit-frequency–calibrated surrogates for two iconic titles: Starburst (NetEnt, low volatility) and Book of Dead (Play’n GO, high volatility). We then simulate ROI distributions, risk-of-ruin curves, and show payout mass functions.

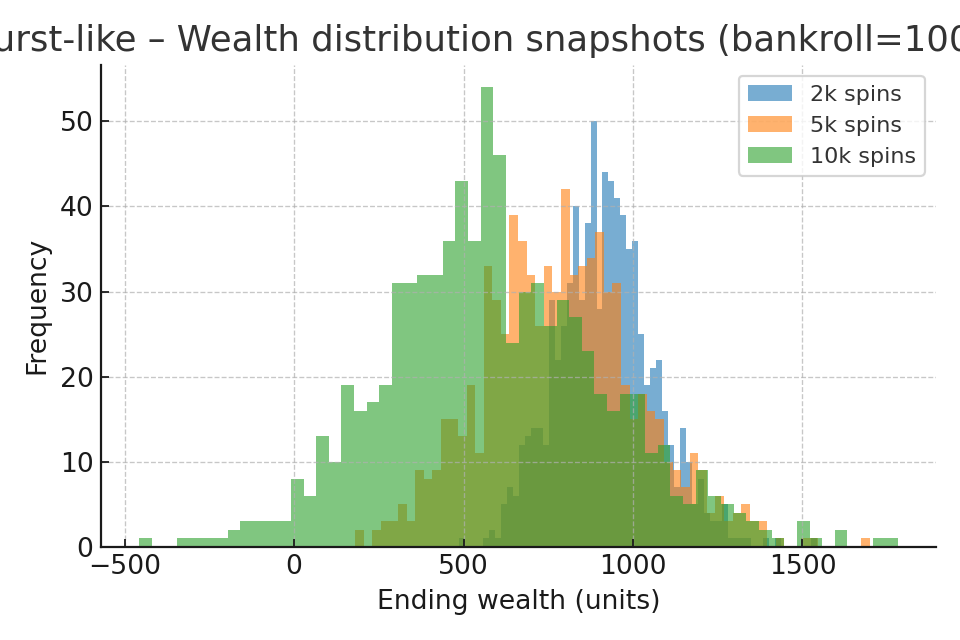
|  |  |
| --- | --- |
| Metric | Value |
| Starburst-like RTP | 0.9609 |
| Starburst-like hit frequency | 0.226 |
| ROI median (20k spins) | -0.041 |
| Ruin @10k spins (B=1000) | 0.048 |
| Risk of ruin: B=1000, b=1, T=10k | 0.031 |
| Risk of ruin: B=2000, b=1, T=10k | 0.000 |
| Risk of ruin: B=1000, b=1, T=50k | 0.940 |



*Figure S.1 – Starburst-like payout mass function (semi-log).*



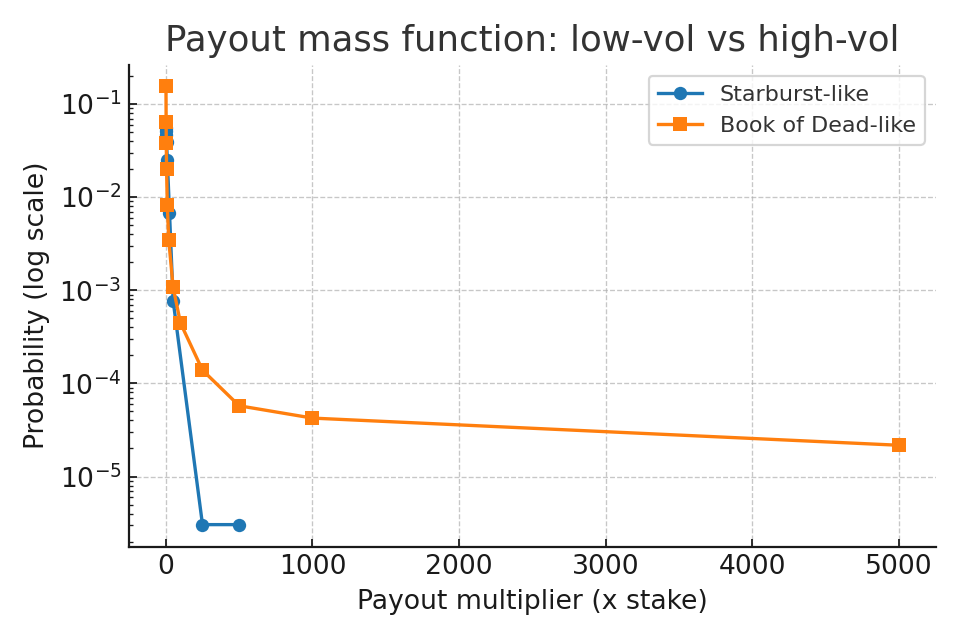
*Figure S.2 – Starburst-like ROI distribution after 20,000 spins.*



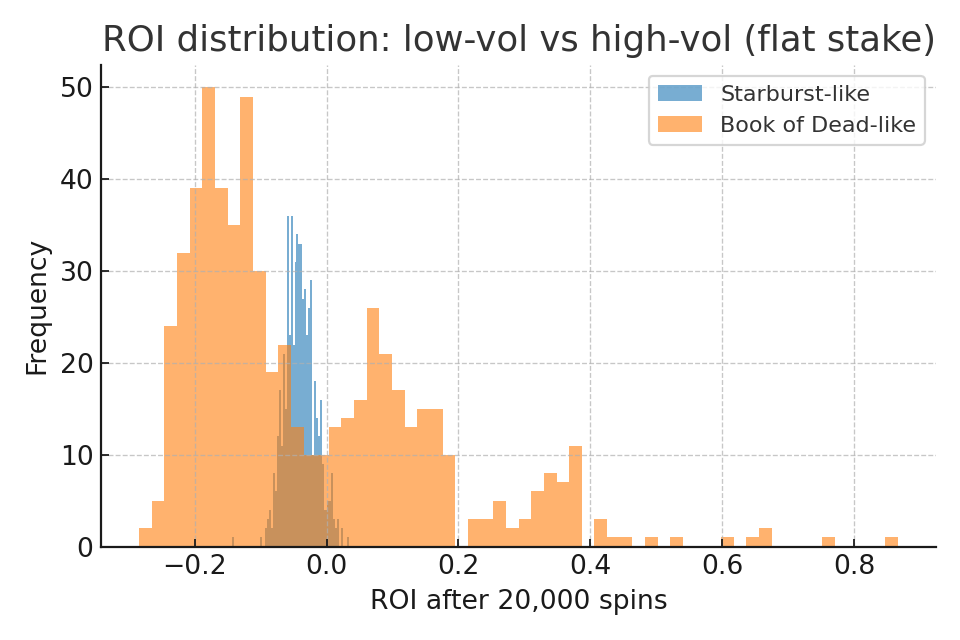
*Figure S.3 – Starburst-like wealth snapshots (2k/5k/10k spins).*

**High Volatility Foil – Book of Dead-like**

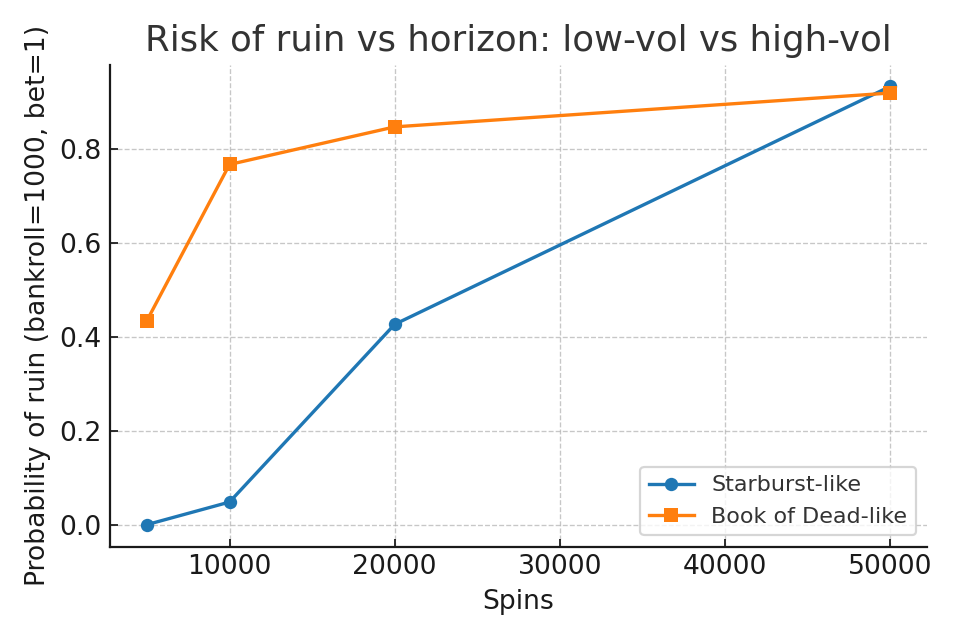
|  |  |
| --- | --- |
| Metric | Value |
| Book of Dead-like RTP | 0.9621 |
| Book of Dead-like hit frequency | 0.291 |
| ROI median (20k spins) | -0.096 |
| Ruin @10k spins (B=1000) | 0.766 |



*Figure S.4 – Payout mass functions: low vs high volatility.*



*Figure S.5 – ROI distributions after 20,000 spins: low vs high volatility.*



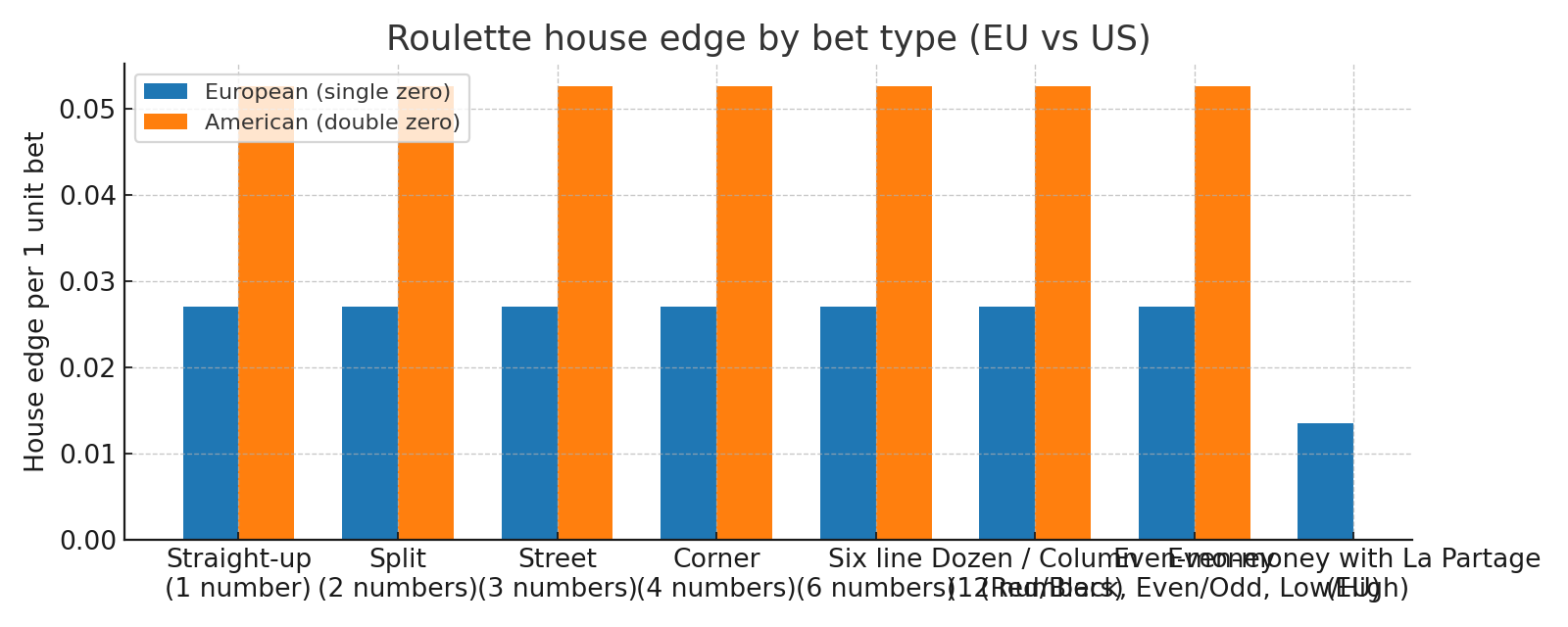
*Figure S.6 – Risk of ruin vs horizon (B=1000, b=1).*

**Appendix E — Roulette: Theoretical Benchmarks, La Partage / En Prison, and Betting Systems**

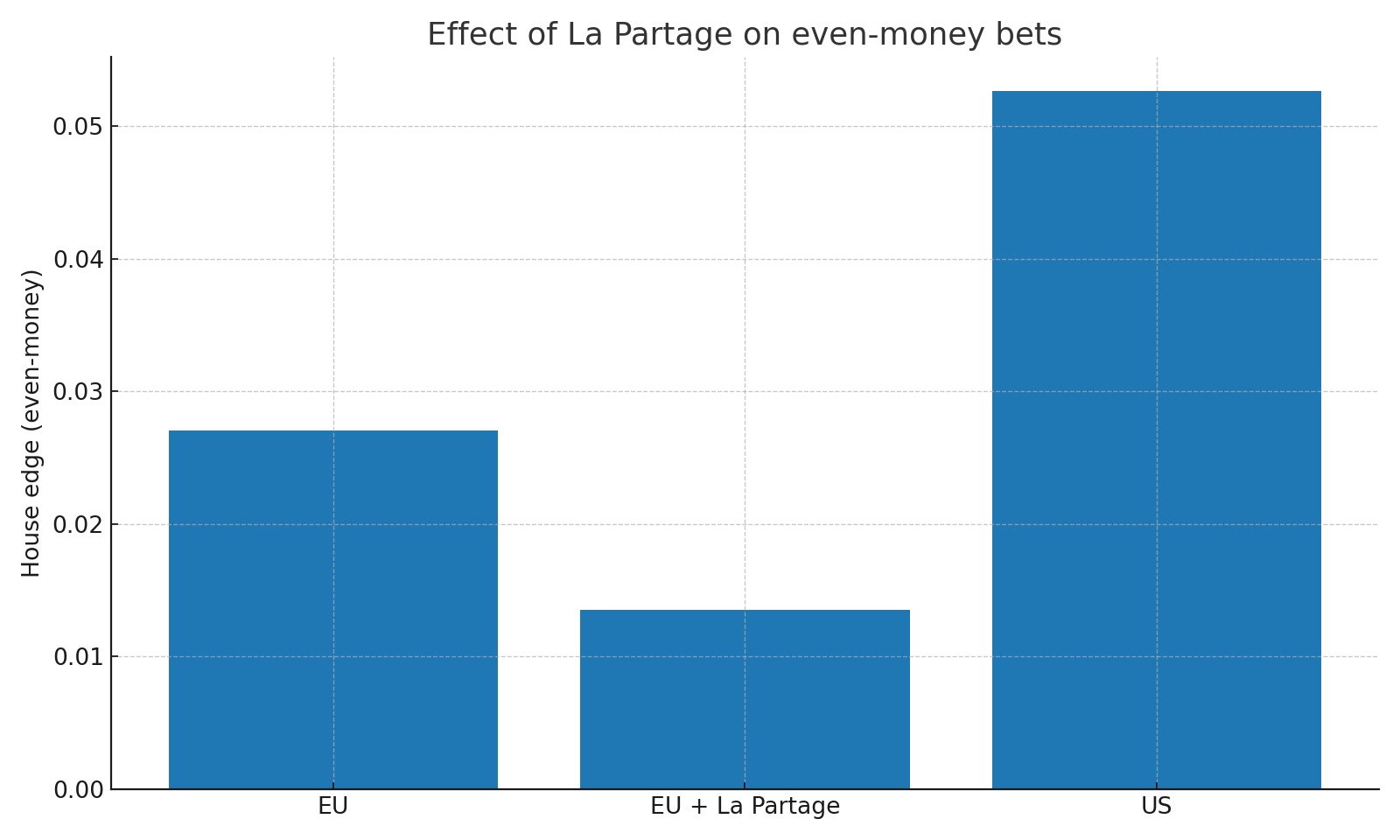
**Roulette – Theoretical Benchmarks, La Partage / En Prison, and Betting Systems**

Single-zero European wheel: 37 pockets. Double-zero American: 38 pockets. Payouts: 35:1 (straight), 17:1 (split), 11:1 (street), 8:1 (corner), 5:1 (six-line), 2:1 (dozen/column), 1:1 (even-money). House edge ≈2.70% (EU) and ≈5.26% (US); American five-number (0,00,1,2,3) ≈7.89%. Even-money with La Partage halves loss on zero → ≈1.35% edge.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Bet | P(win) EU | House edge EU | P(win) US | House edge US |
| Straight-up (1 number) | 0.02703 | 0.0270 | 0.02632 | 0.0526 |
| Split (2 numbers) | 0.05405 | 0.0270 | 0.05263 | 0.0526 |
| Street (3 numbers) | 0.08108 | 0.0270 | 0.07895 | 0.0526 |
| Corner (4 numbers) | 0.10811 | 0.0270 | 0.10526 | 0.0526 |
| Six line (6 numbers) | 0.16216 | 0.0270 | 0.15789 | 0.0526 |
| Dozen / Column (12 numbers) | 0.32432 | 0.0270 | 0.31579 | 0.0526 |
| Even-money (Red/Black, Even/Odd, Low/High) | 0.48649 | 0.0270 | 0.47368 | 0.0526 |
| Five-number (0,00,1,2,3) – American only |  |  | 0.13158 | 0.0789 |
| Even-money with La Partage (EU) | 0.48649 | 0.0135 |  |  |

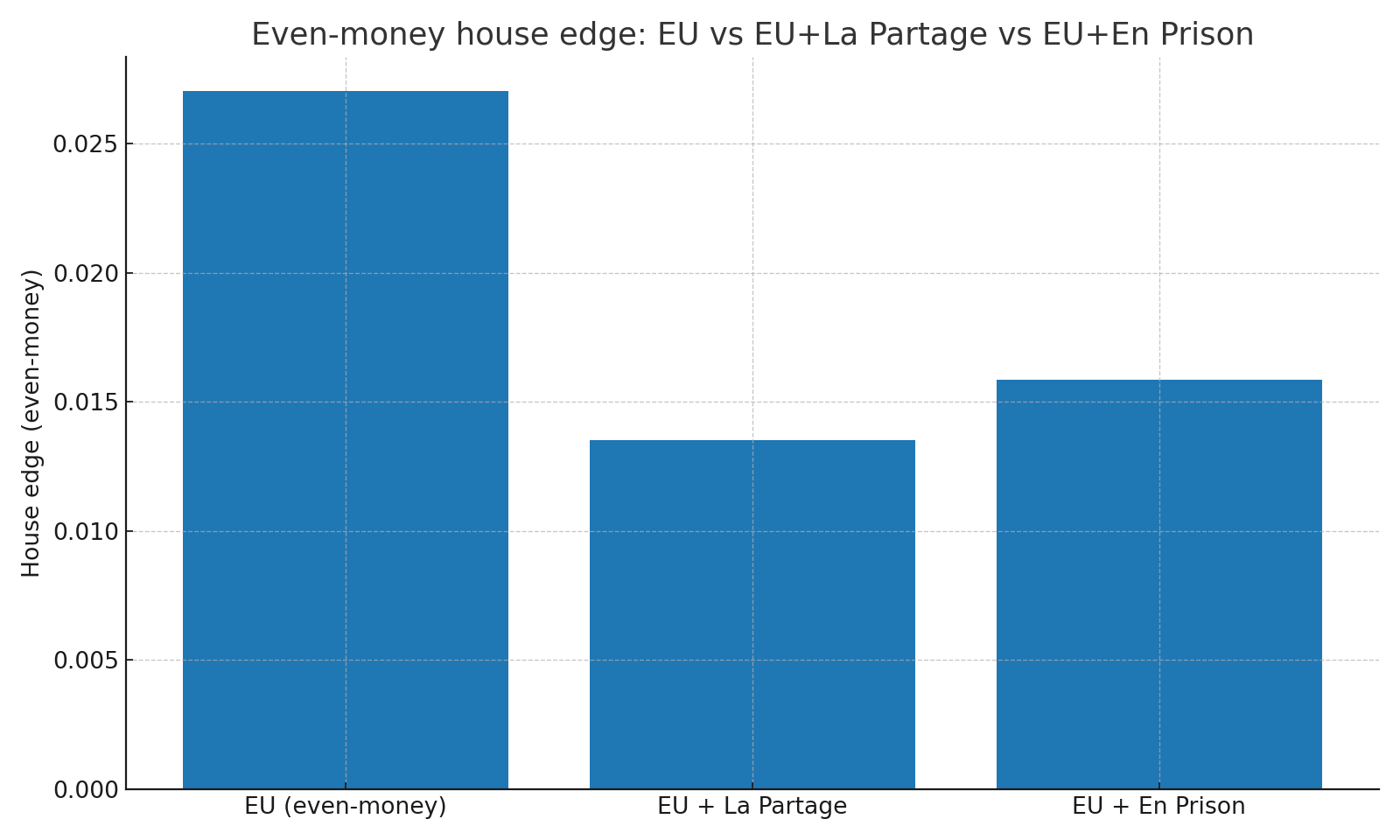


*Figure R.1 – House edge by bet type (European vs American).*



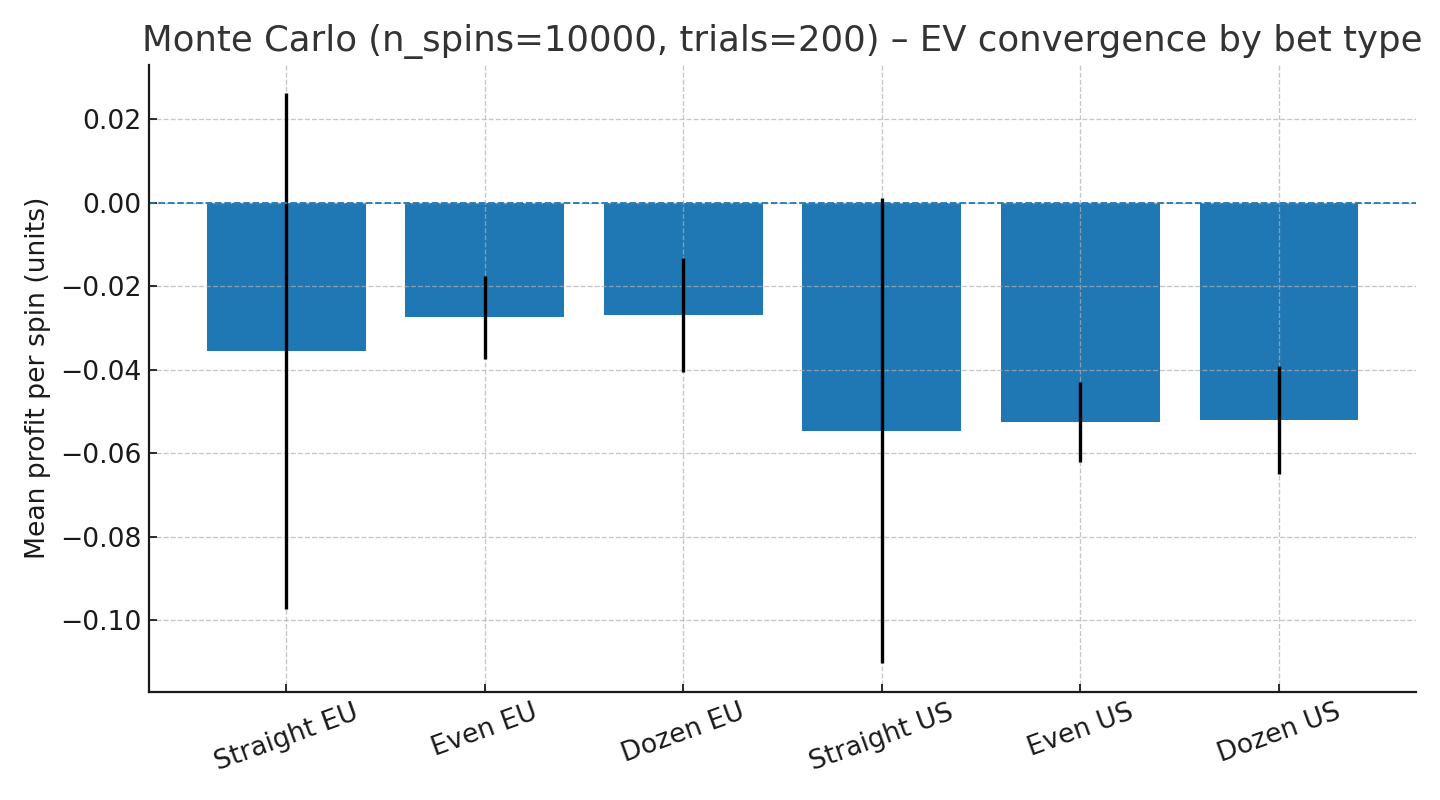
*Figure R.2 – Even-money house edge: EU vs EU+La Partage vs US.*

**En Prison vs La Partage (Even-Money)**

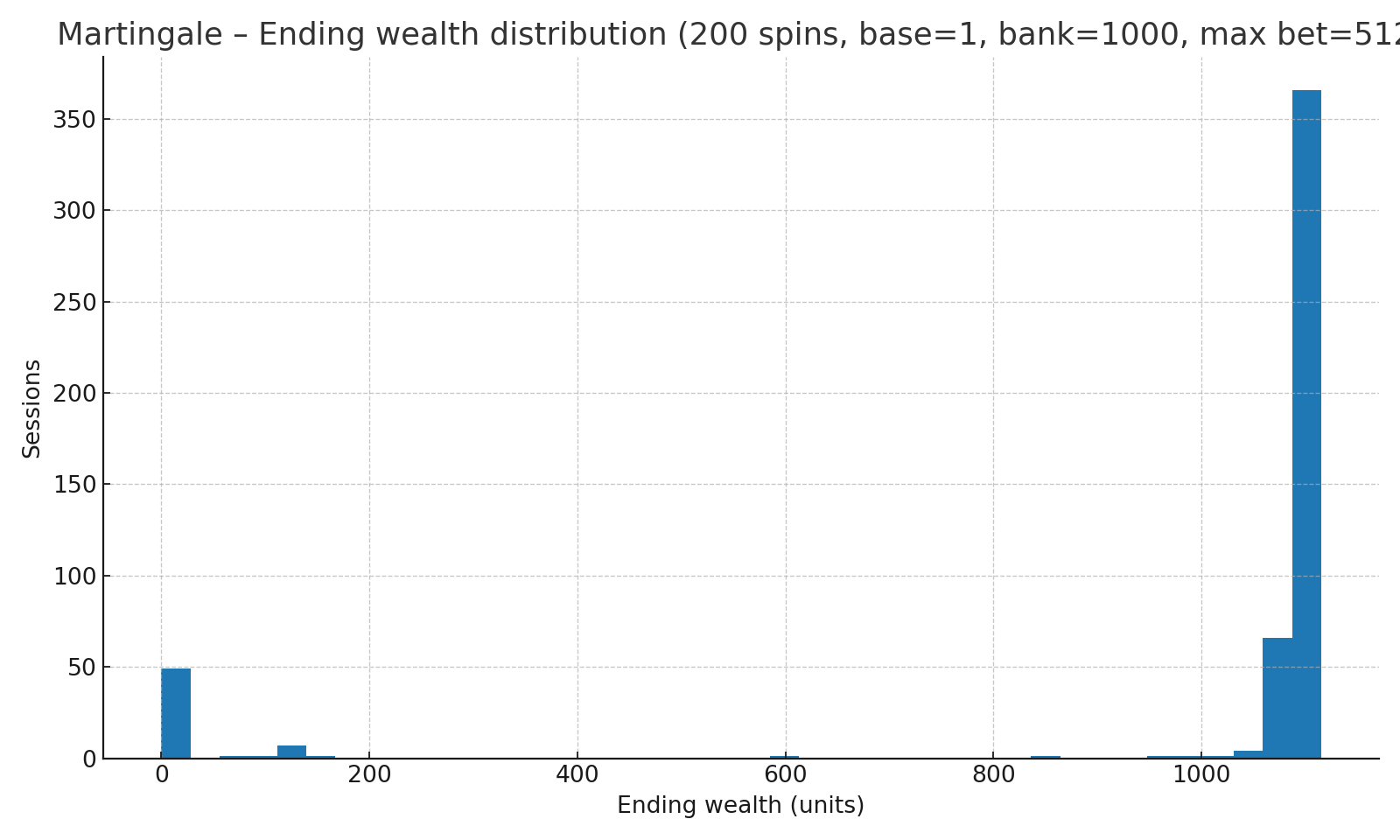


*Figure R.3 – Even-money: La Partage and En Prison are EV-equivalent (simulation confirmation).*

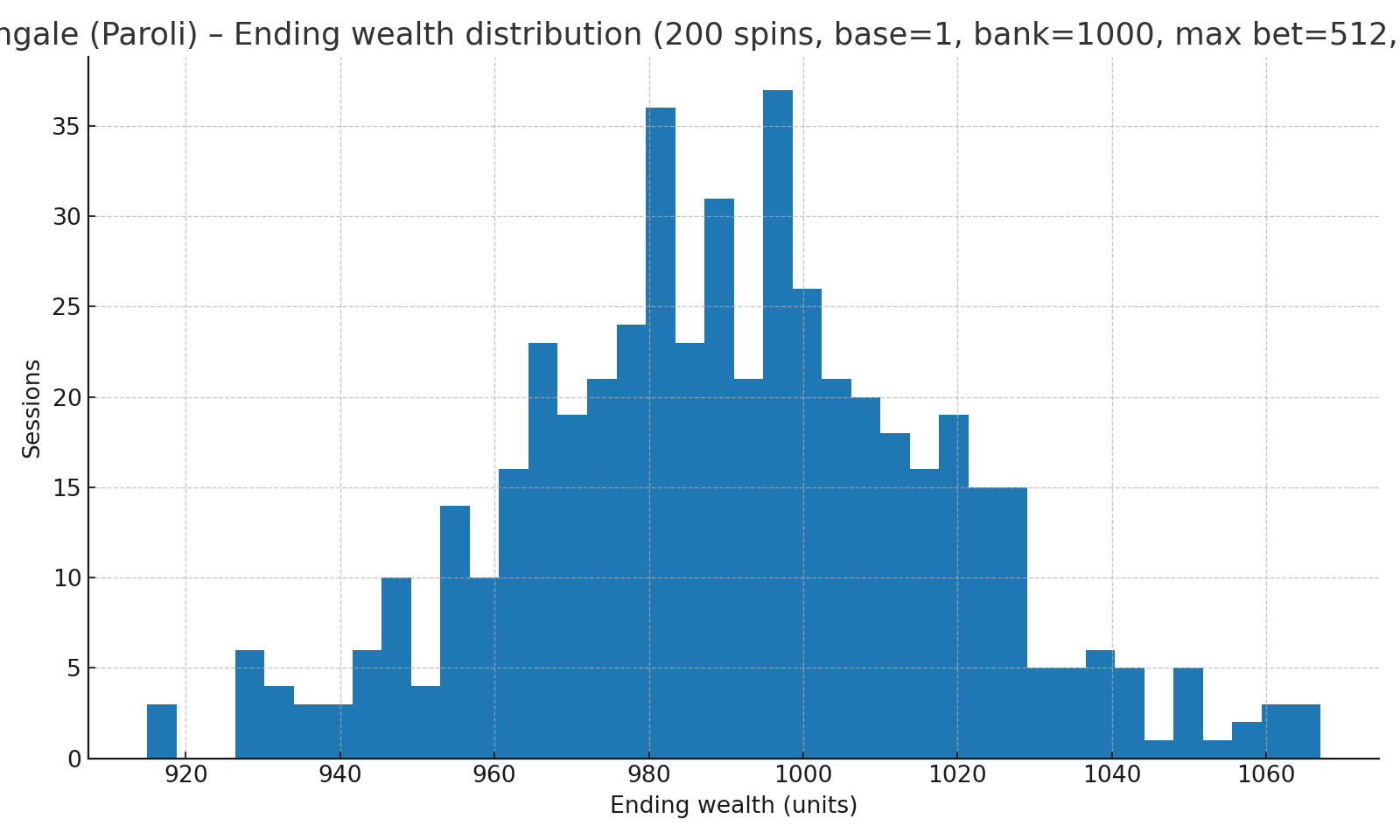
**Monte Carlo: EV Convergence & Table-Limit Pressure**



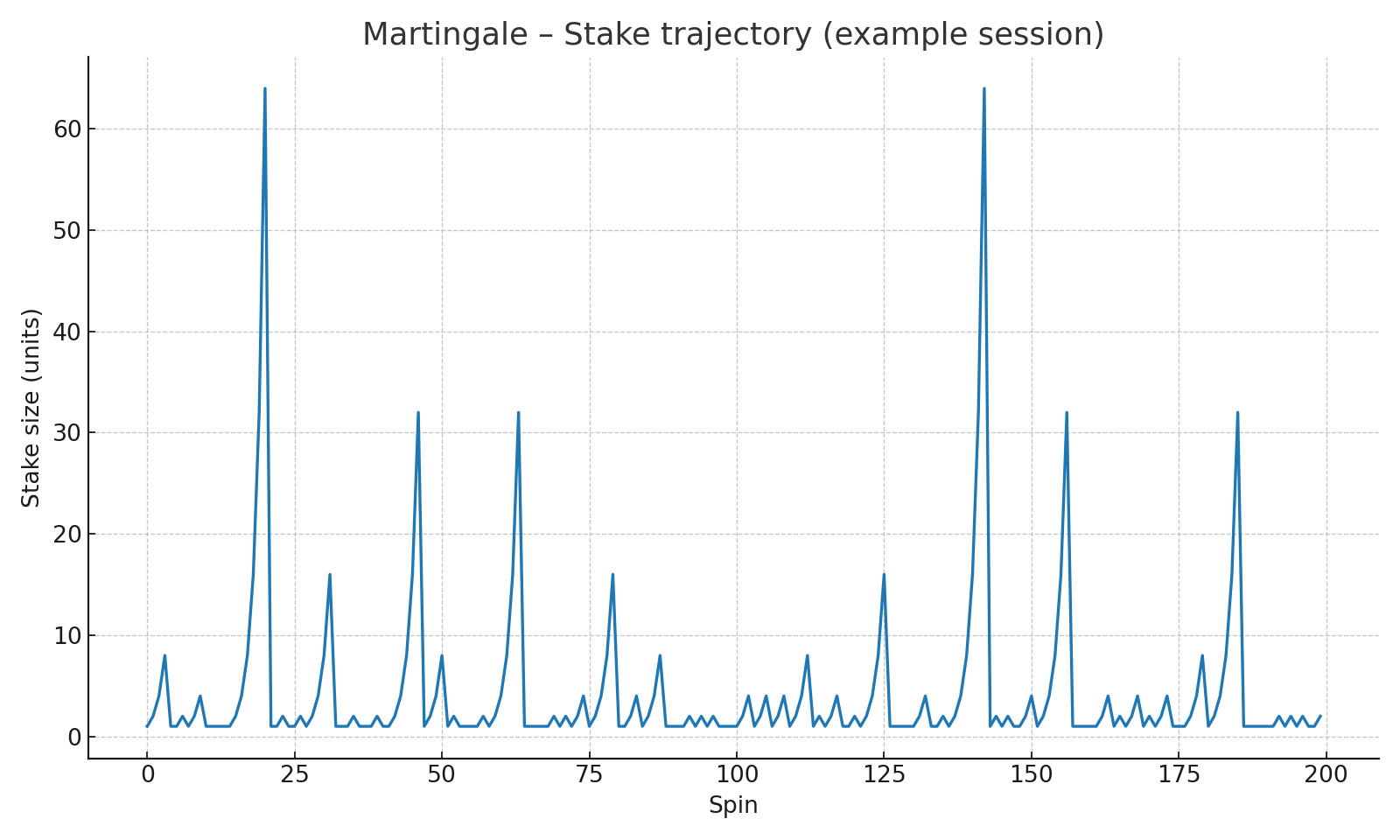
*Figure R.4 – EV convergence (mean±sd) across bet types (EU vs US).*



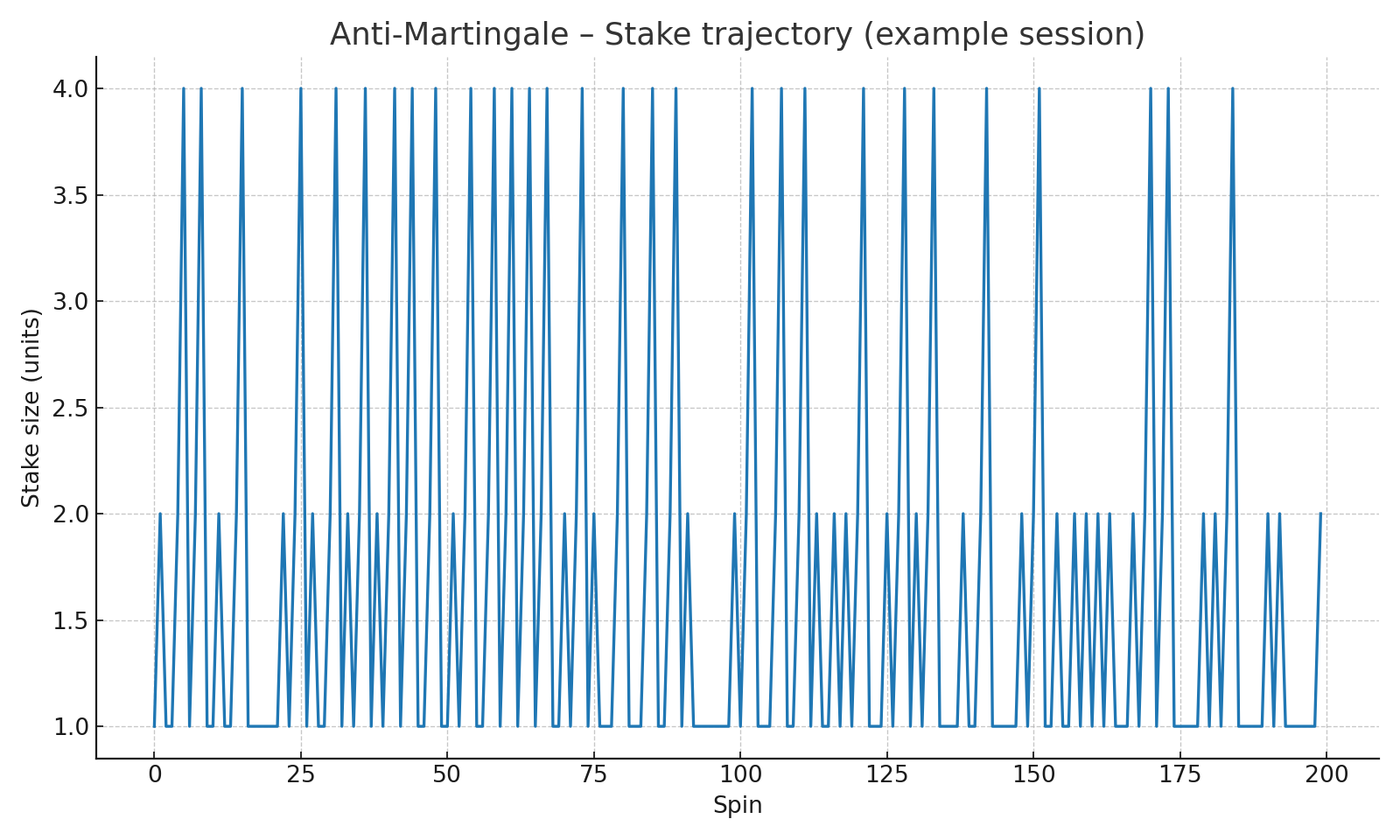
*Figure R.5 – Martingale: ending-wealth distribution under table limits.*



*Figure R.6 – Anti-Martingale (Paroli): ending-wealth distribution (streak cap=3).*



*Figure R.7 – Martingale: stake trajectory (example session).*



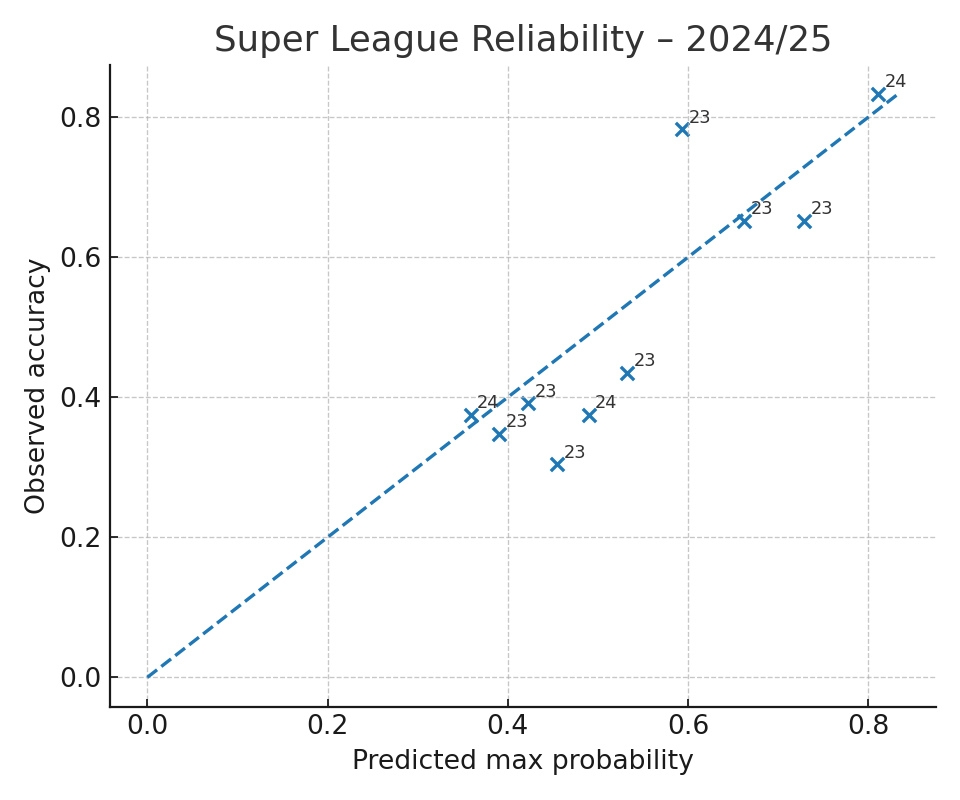
*Figure R.8 – Anti-Martingale: stake trajectory (example session).*

Appendix F — Football (Super League Greece): Market-Implied Probabilities, Utility & Luck

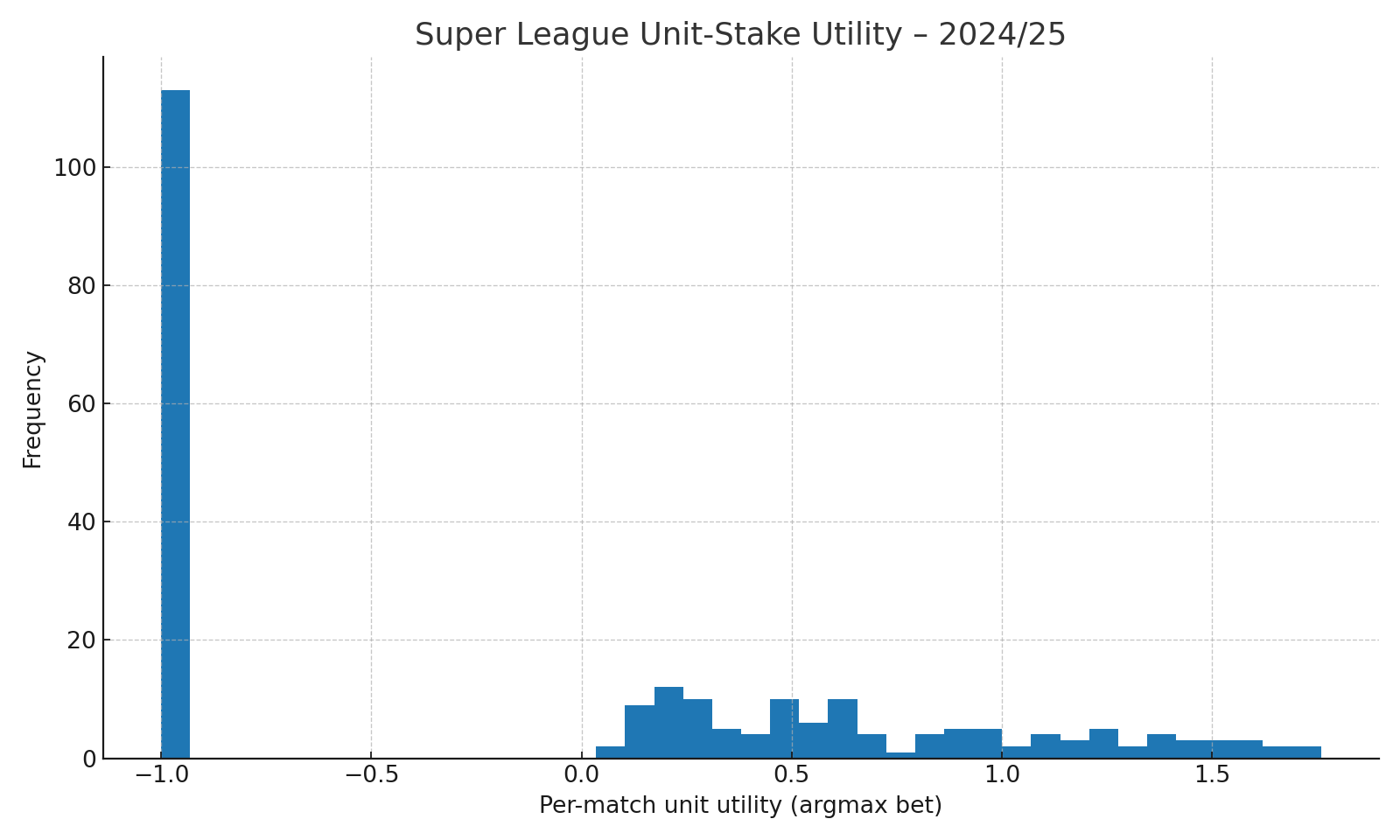
**Super League (Greece) – Market-Implied P, Utility and Luck**

Data source: Football-Data.co.uk (league code G1). We use closing/average odds to derive market-implied probabilities for {Home, Draw, Away}, corrected for overround. The realized outcome’s probability defines surprisal S; entropy H is computed over the 3-way distribution. Utility U is modeled as unit-stake net return on the argmax-probability selection (illustrative baseline).

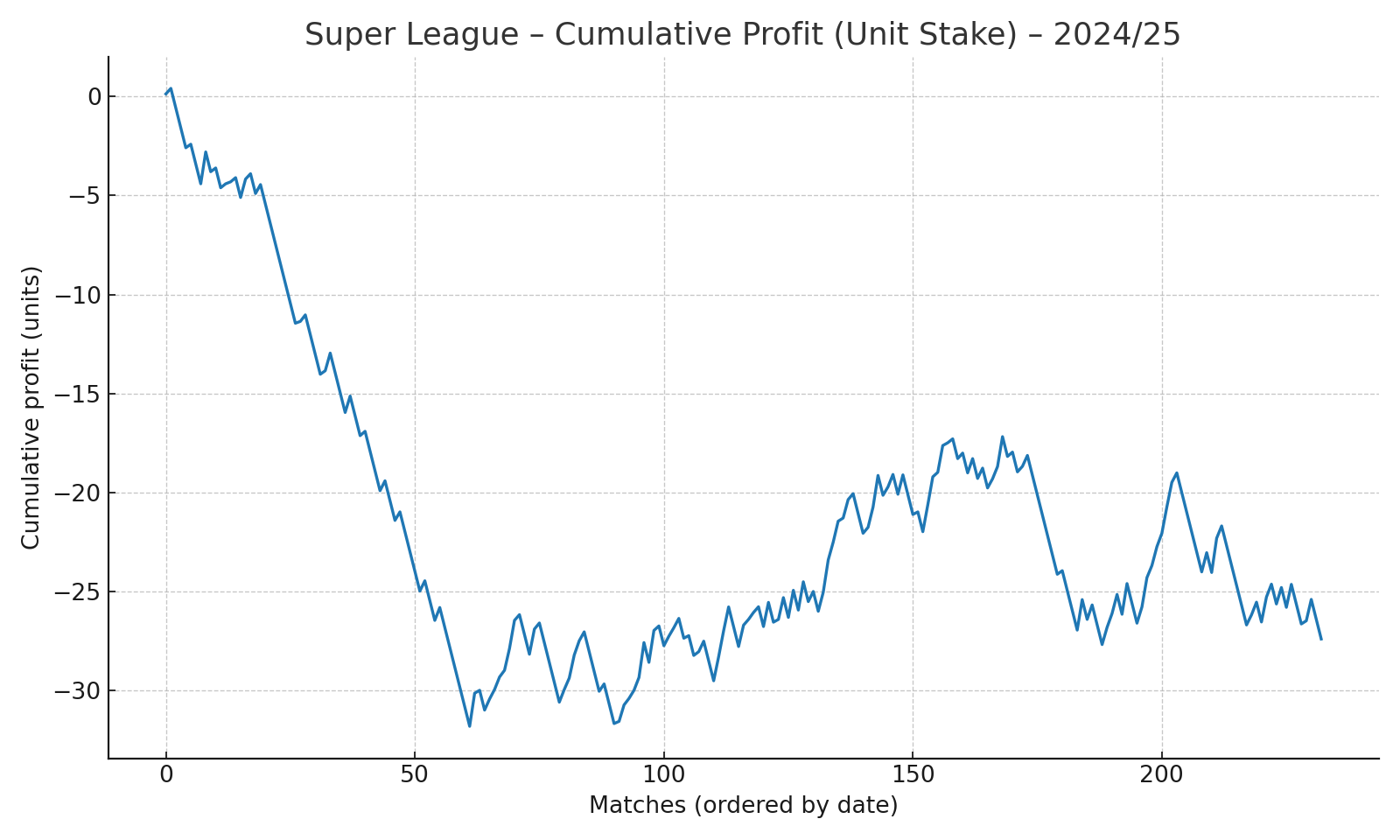
|  |  |
| --- | --- |
| Metric | Value |
| Season (latest) | 2024/25 |
| Matches | 233 |
| Brier score (3-way) | 0.5812 |
| Log loss (3-way) | 0.9794 |
| Mean unit utility (per match) | -0.1176 |



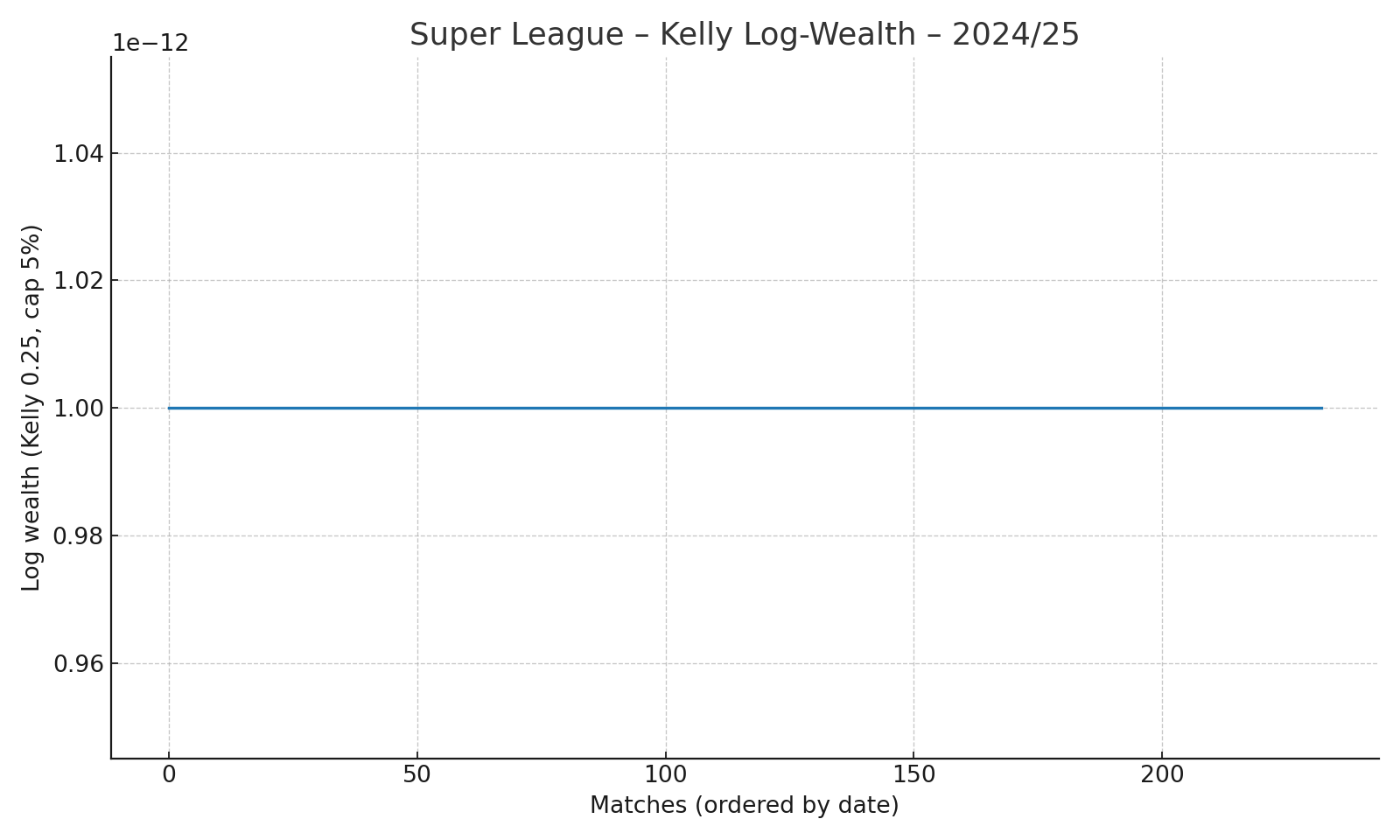
*Figure F.1 – Reliability of argmax market probabilities (latest season).*



*Figure F.2 – Distribution of per-match unit utility (latest season).*



*Figure F.3 – Cumulative unit-stake profit (latest season).*



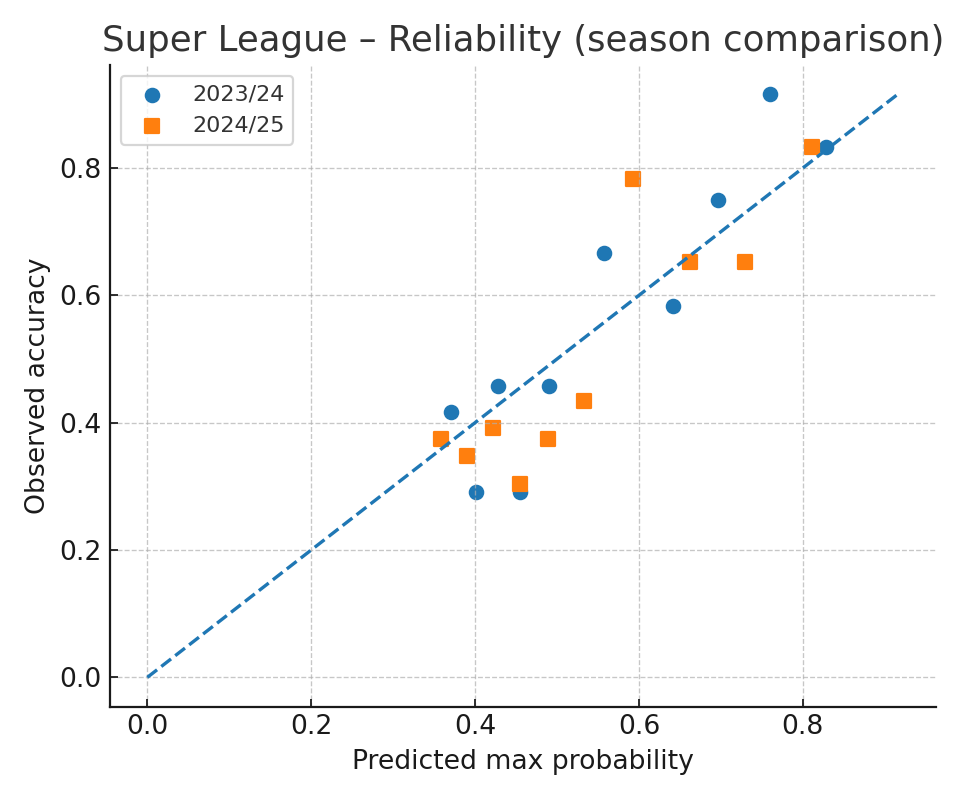
*Figure F.4 – Kelly log-wealth (λ\_K=0.25, cap 5%/match).*

**SPI Features & Kelly Staking (Optional Enhancements)**

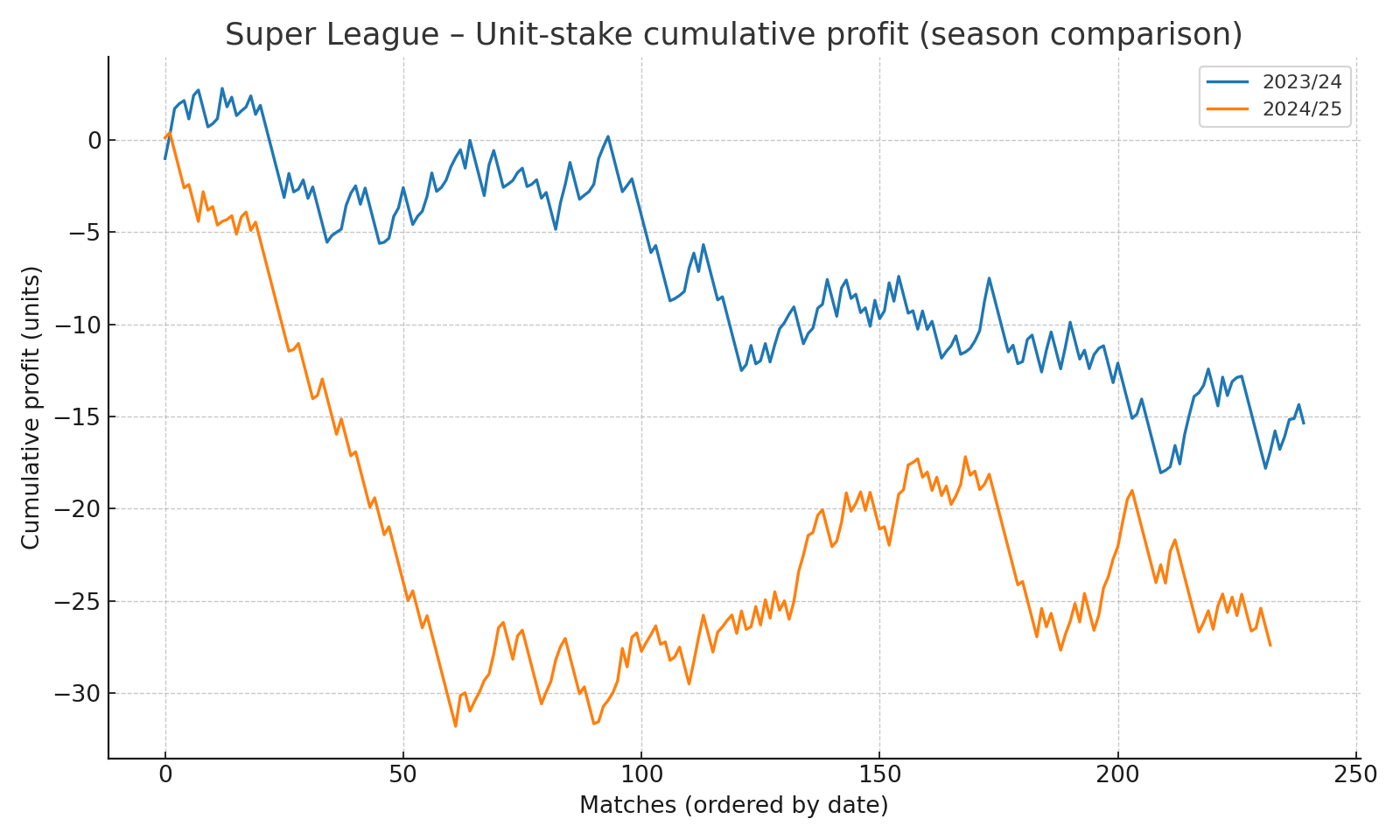
If FiveThirtyEight SPI probabilities are available, blend or recalibrate market probabilities with SPI to minimize 3-way log-loss. Kelly staking with fractional λ\_K and per-match caps evaluates bankroll growth; block-bootstrap over matchdays yields uncertainty bands.

**Comparative Season Analysis – 2023/24 vs 2024/25**

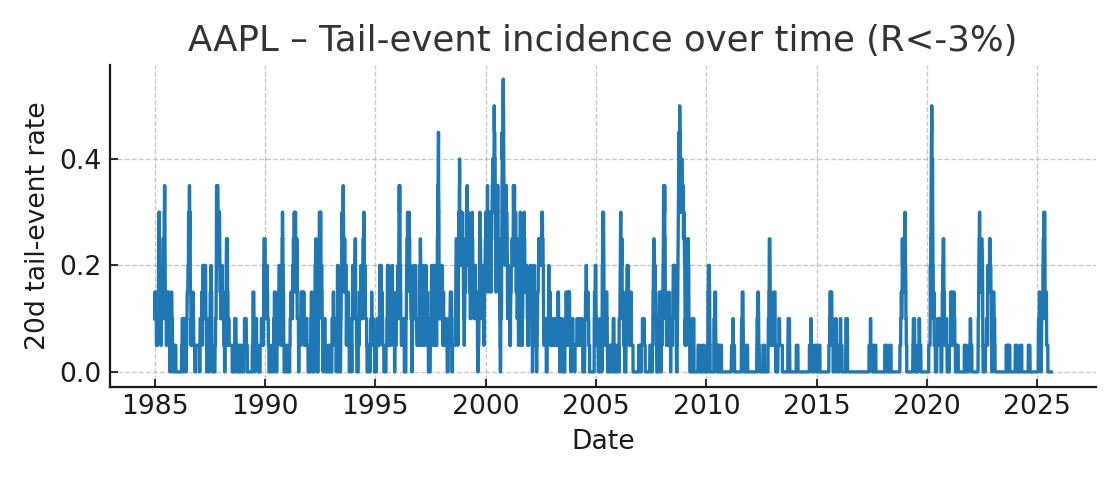
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Season | Matches | Brier (3-way) | Log-loss (3-way) | Mean unit utility |
| 2023/24 | 240 | 0.5306 | 0.9004 | -0.0640 |
| 2024/25 | 233 | 0.5812 | 0.9794 | -0.1176 |



*Figure F.5 – Reliability comparison (2023/24 vs 2024/25).*



*Figure F.6 – Unit-stake cumulative profit comparison (2023/24 vs 2024/25).*

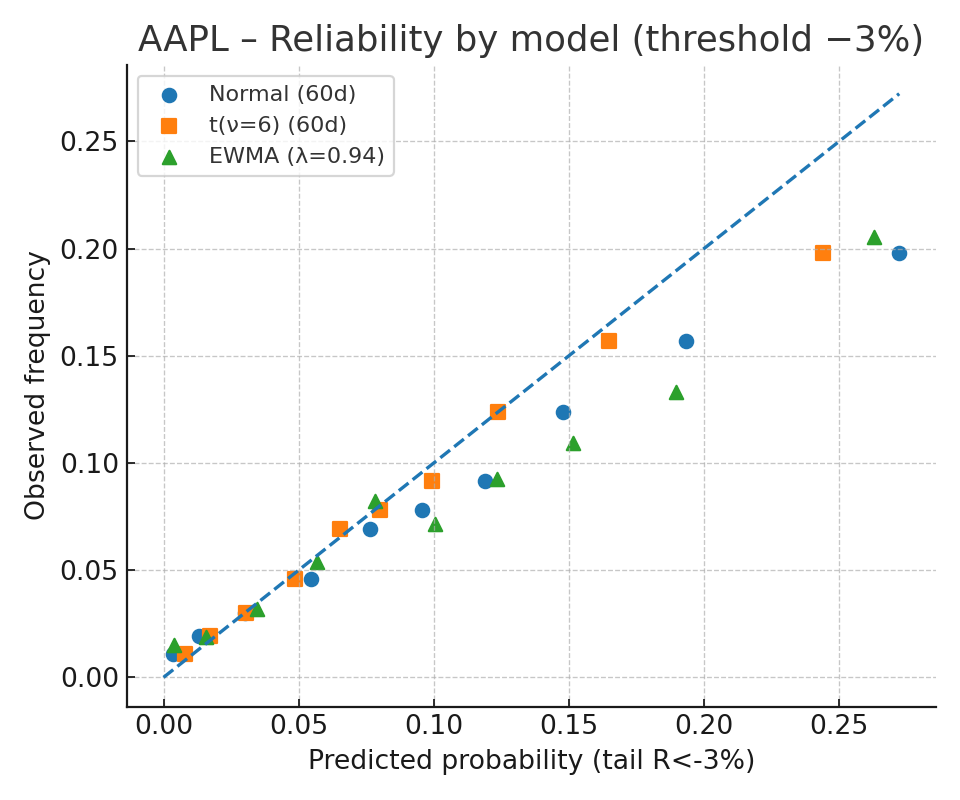
**

**Appendix G — Equity Tail-Event Calibration: SPX (R<−2%) & AAPL (R<−3%), Normal vs t vs EWMA**

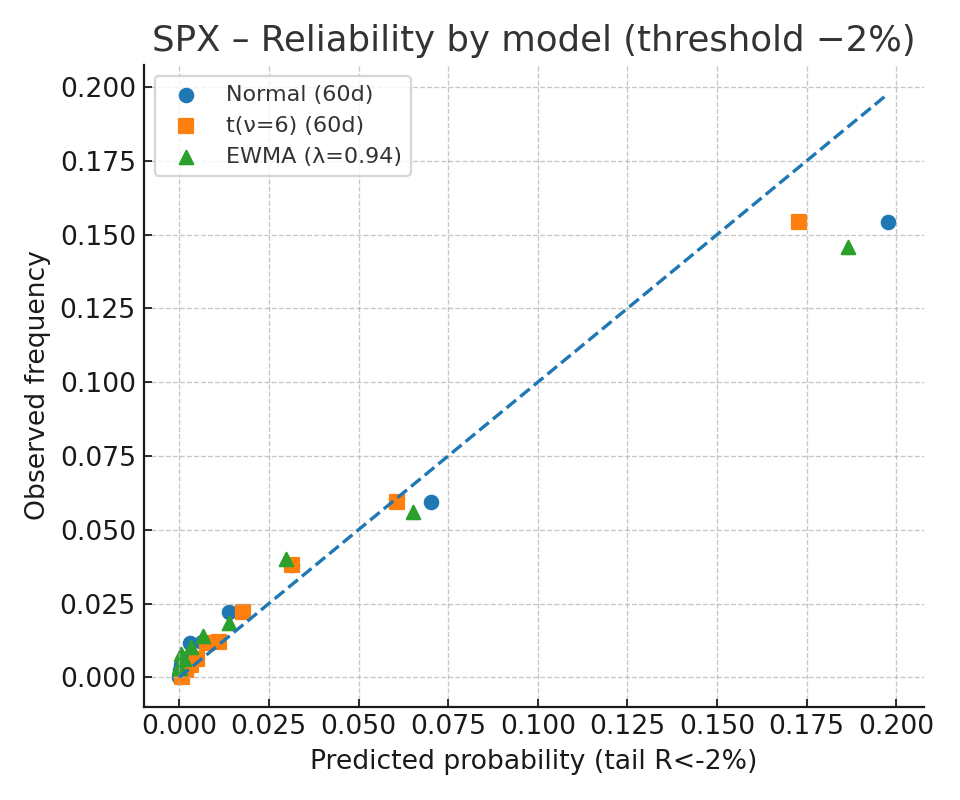
**Stocks Mini-Case – SPX (R<-2%) vs AAPL (R<-3%), Normal vs t(ν=6) vs EWMA**

We add SPX heavy-tail and time-varying volatility estimators for symmetry. Models: 60-day Normal; 60-day Student‑t (ν=6, variance-matched); EWMA Normal (λ=0.94, mean≈0). Thresholds: SPX R<−2%, AAPL R<−3%.

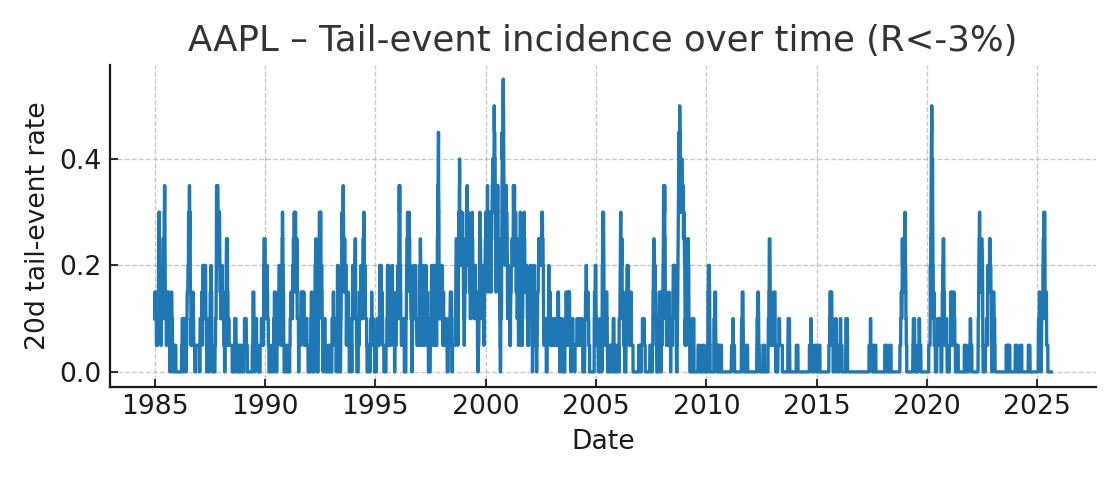
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Ticker | Threshold | Model | Obs | Event rate | Brier | Log-loss |
| SPX | -2% | Normal(60d) | 39,451 | 0.0311 | 0.02815 | 0.11617 |
| SPX | -2% | t(ν=6, 60d) | 39,451 | 0.0311 | 0.02796 | 0.11298 |
| SPX | -2% | EWMA(λ=0.94) | 39,260 | 0.0305 | 0.02785 | 0.12115 |
| AAPL | -3% | Normal(60d) | 10,256 | 0.0824 | 0.07286 | 0.26441 |
| AAPL | -3% | t(ν=6, 60d) | 10,256 | 0.0824 | 0.07223 | 0.26153 |
| AAPL | -3% | EWMA(λ=0.94) | 10,065 | 0.0814 | 0.07264 | 0.26669 |



*Figure M.1 – AAPL (R<−3%): reliability across Normal, t(ν=6), and EWMA.*



*Figure M.2 – SPX (R<−2%): reliability across Normal, t(ν=6), and EWMA.*



*Figure M.3 – AAPL (R<−3%): 20-day tail-event incidence over time.*

### *Conclusions*

We introduced a unified “Luck” functional that fuses probability, utility, surprisal, and information collapse, and proved normalization and convergence properties (GC/ULLN) enabling consistent ERM. Applications across lotteries, football markets, roulette, and slots demonstrate that the framework is domain-agnostic yet empirically testable: calibration curves align with theory, and variance-driven phenomena (e.g., table limits or high-volatility slots) appear exactly where the axioms predict.

### Practical Implications

For applied modelers, the framework yields: (i) transparent calibration targets, (ii) robustness via Lipschitz constraints, and (iii) decision metrics that remain comparable across domains with different base rates and paytables.

### Limitations & Future Work

Our results depend on sales reporting quality (lotteries), seasonal demand (sports), operator RTP settings (slots), and simple return dynamics (equities). Future work: richer utility UUU (risk-averse/CRRA, prospect-theory), heavy-tail & regime-switching for returns, multi-market information flows (cross-entropy regularizers), and non-parametric calibration under partial observability.

### Data & Code Availability

All code and data to reproduce figures and tables (lottery tiers, football odds, roulette/slots simulations, SPX/AAPL analyses) will be released at [repository/DOI placeholder]. A reproducibility README includes versions, seeds, and run scripts.

### CRediT Author Statement

* Conceptualization: [Name]
* Methodology: [Name]
* Software: [Name]
* Validation: [Name]
* Formal Analysis: [Name]
* Investigation: [Name]
* Data Curation: [Name]
* Writing – Original Draft: [Name]
* Writing – Review & Editing: [Name]
* Visualization: [Name]
* Supervision: [Name]
* Project Administration: [Name]
* Funding Acquisition: [Name]

### Declaration of Generative AI Use

Parts of the writing, figure scripting, and document assembly used an AI assistant (ChatGPT, OpenAI). The authors reviewed and take full responsibility for the content. No AI tools were listed as co-authors, and all sources and computations were verified by the authors.

### Declaration of Competing Interest

The authors declare no competing financial interests or personal relationships that could have appeared to influence the work reported.

### Funding

This research received no specific grant from funding agencies in the public, commercial, or not-for-profit sectors.  
(If funded, replace with the grant and number.)

### Acknowledgments

We thank [names/affiliations] for helpful discussions and feedback. Any remaining errors are our own.

### Ethics Statement

This study does not involve human participants, personal data, or animal subjects. All analyses use public or synthetic datasets.

### Disclaimer (Research-Only)

This article is for research and education. It is not gambling or investment advice. Casino games have negative expected value; sports and markets involve financial risk; past performance does not guarantee future results.

### Trademarks

Starburst (NetEnt) and Book of Dead (Play’n GO) are trademarks of their respective owners. Slot models herein are surrogate, RTP/hit-rate–calibrated approximations, not reverse-engineered paytables.

### Supplementary/Appendices

* Appendix B – Explanatory Notes on Core Definitions and Results (with mini-examples).
* Appendix C – Lottery (OPAP Joker): Mechanics & Jackpot Frequency.
* Appendix D – Online Slots: Starburst (Low-Vol) vs Book of Dead (High-Vol).
* Appendix E – Roulette: Theory, La Partage / En Prison, Systems & Variance.
* Appendix F – Football (Super League): Market-Implied PPP, Utility & Luck.
* Appendix G – Equities: SPX (R<−2%) & AAPL (R<−3%), Normal vs t vs EWMA.