The Rise of Online Gambling: *What's at Stake?*

Executive Summary

To policymakers, regulatory agencies, and the general public of the United States of America and United Kingdom,

Online sports gambling has grown into a multi-billion dollar industry, with U.S. sports gaming revenue reaching an estimated $15 billion in 2025 and the U.K. experiencing similar growth. Gambling is important: it drives sports engagement, generates revenue for leagues and media, and produces tax dollars for public projects. But it poses many risks: addiction and problem gambling, especially for younger adults, and serious financial harm. We urge policymakers and agencies to take a closer look and take regulatory action.

How much someone gambles is tied to how much they can actually afford to lose. Our Demographic Disposable Income Model estimates money remaining after taxes and essential costs as a function of salary, age, and region. Rather than applying a single average expense figure, we model expenses as random draws from a log-normal distribution anchored to government survey data (BLS [1], ONS [2]), with taxes calculated using 2024 federal brackets, three state archetypes (Texas, Illinois, California), and FICA contributions. A typical 45–54 year-old has roughly $13,100 in annual disposable income, but the average under-25 earner effectively has $0, and 62% of that youngest bracket is already spending more than it earns before gambling enters the picture.

Even with disposable income, gambling carries serious financial risk. Our Stochastic Gambling Outcome Model simulates 10,000 virtual bettors via Monte Carlo methods, accounting for bet frequency, wager size, bet type (conservative, moderate, aggressive), and loss-chasing behavior modeled through a two-state Markov chain. Parlay bets compound the house edge dramatically; a 10-leg parlay carries a 39% effective edge versus 4.8% for a single spread bet, worse than roulette. Loss-chasing amplifies expected losses by roughly 35% across demographics. The average active bettor loses about $960 per year, while young male bettors with aggressive strategies can expect losses exceeding $3,000.

So how much regulation is appropriate? Our Financial Impact Assessment combines both models into three metrics. The *Probability of Financial Ruin*, the likelihood that annual losses exceed 10% of disposable income, a threshold associated with debt accumulation in consumer finance research [3], surpasses 35% for low-income young male bettors and reaches 100% for the 65+ cohort, whose disposable income is already negative. Our *Dynamic Wealth Erosion Model* shows that cumulative opportunity costs exceed $131,000 over 30 years for a moderate bettor starting at 22, roughly seven times the median retirement savings for Americans under 35 [4]; 70% of that cost accrues in the first 12 years, meaning early intervention matters most. Finally, tracing a bettor's $1,500 annual loss through the system reveals that the state collects just $22.50 in tax: a 66-to-1 ratio of sportsbook revenue to public benefit, compared to roughly 2-to-1 for state lotteries.

These three models work together: disposable income sets the financial baseline, gambling outcomes estimate the losses, and the impact assessment combines both to reveal true harm. Our findings suggest that targeted interventions, like income-linked deposit limits, restrictions on parlay promotion, and mandatory cooling-off periods, could meaningfully protect the most vulnerable 10% of bettors without burdening the majority for whom gambling remains affordable entertainment.

We hope our discoveries aid policymakers and agencies in taking action to balance the benefits and drawbacks of online gambling, and the general public to remain cautious while still having fun.

Contents

Q1: Playing With House Money – Disposable Income Model

Defining the Problem

We develop a model that takes an individual's salary, age, and other demographic information as input and returns an estimate of their *disposable income*, defined as gross income minus taxes and essential living costs. We allow disposable income to be negative, since that correctly identifies individuals who are already accumulating debt (which are the same people most vulnerable to gambling losses).

Assumptions

A1. Tax obligations follow current statutory brackets. We use 2024 U.S. federal brackets, state-specific rate functions, and FICA payroll contributions (6.2% Social Security, 1.45% Medicare), all consistent with IRS data [5]. Tax credits such as the EITC are not modelled, as their inclusion would reduce vulnerability estimates further.

A2. Essential expenditures include only genuinely non-discretionary categories: food, housing, utilities, transportation, healthcare, and personal insurance. The six categories are classified as non-discretionary by the BLS Consumer Expenditure Survey [1]. Entertainment, clothing, and education are excluded because they contain substantial discretionary components. This boundary is deliberately narrow, meaning we likely *understate* true essential spending.

A3. Within a demographic group, essential expenditures follow a log-normal distribution. The log-normal produces only positive values, is right-skewed (matching expenditure data), and arises naturally when costs are driven by multiplicative factors (e.g., rent × location × household size). We set its mean equal to the BLS cohort average and its shape parameter to $\sigma_e = 0.3$, calibrated from the coefficient of variation in BLS survey data. Sensitivity analysis shows results are qualitatively stable for $\sigma_e \in [0.2, 0.5]$.

A4. Negative disposable income is meaningful. When $DI < 0$, the person's essential costs exceed their after-tax income. We track this quantity as a *Debt Accumulation Rate* rather than clamping it to zero, since these individuals are sustaining basic needs only through credit, asset drawdown, or transfers and face the greatest risk from any additional spending [6].

A5. U.K. results use equivalent U.K. data sources. Expenditures come from the ONS Family Spending Survey; taxes use HMRC brackets including National Insurance contributions (8%/2%) [2]. Council Tax is included within the housing expenditure category rather than modeled as a separate instrument.

A6. Income within an age group is represented by the BLS-reported mean for that cohort. Because the mean exceeds the median in right-skewed income distributions, this choice makes our vulnerability estimates conservative. The true median earner has less disposable income than our model predicts.

A7. Individuals file taxes as single filers with the standard deduction. Married-filing-jointly status would widen brackets and increase the deduction, reducing the tax burden. This is therefore another conservative choice that biases toward *understating* vulnerability. Itemised deductions and dependent exemptions are not modelled.

A8. Sub-regional cost-of-living variation is absorbed into the log-normal expenditure spread, beyond four broad regional multipliers (Northeast, Midwest, South, West) [7]. City-level and rural–urban price differences are captured probabilistically by σ_e rather than modelled explicitly, which may underrepresent extreme metro-area outliers.

Variables and Parameters

Table 1: Variables for the Demographic Disposable Income Model.

Symbol	Definition	Unit
I	Gross annual income (salary)	USD or GBP
$T(I)$	Total tax function (federal + state/NI + FICA)	USD or GBP
E	Essential expenditures, drawn from Lognormal(μ_e, σ_e^2)	USD or GBP
a	Age group index (7 groups for U.S., 6 for U.K.)	–
r	Region index (4 U.S. regions or U.K. countries)	–
DI	Disposable income $= I - T(I) - E$ (may be negative)	USD or GBP
$\bar{E}(a, r)$	Mean essential expenditure for age group a in region r	USD or GBP
σ_e	Shape parameter controlling within-group expense variation	–
δ	Debt accumulation rate when $DI < 0$ (equals $\lvert DI \rvert$)	\$/year

Model Development

Core Structure

Disposable income is calculated by subtracting taxes and essential costs from gross income:

$$DI(I, a, r) = \underbrace{I}_{\text{gross income}} - \underbrace{T(I)}_{\text{taxes}} - \underbrace{E}_{\text{essential costs}} \tag{1}$$

The first two terms are deterministic (set by tax law). The third term is stochastic: we draw E from a log-normal distribution whose mean matches government data for the relevant cohort:

$$E \sim \text{Lognormal}\left(\ln\big(\bar{E}(a, r)\big) - \frac{\sigma_e^2}{2},\ \sigma_e^2\right) \tag{2}$$

The parameters of the log-normal are chosen so that its expected value equals exactly $\bar{E}(a, r)$, the BLS-reported mean for each demographic group, while the shape parameter $\sigma_e = 0.3$ generates realistic within-group spread.

Why should we model expenses stochastically? Two 25-year-olds earning the same salary face very different financial realities depending on their rent, health costs, and commute. A single "average" expense figure hides this variation, but it is precisely this variation that determines who can absorb a gambling loss and who cannot. By producing a full *distribution* of disposable incomes for each group, we can identify the fraction of each cohort that is already financially vulnerable. This distributional information becomes critical in Q3, where we compute the probability that gambling losses push someone past a dangerous threshold.

Individuals with $DI < 0$ are tracked via a Debt Accumulation Rate ($\delta = |DI|$) rather than being hidden behind a zero floor, since they are the most at risk from any additional spending such as gambling.

The Tax Function $T(I)$

For the United States, total tax is $T_{\text{US}}(I) = T_{\text{fed}}(I) + T_{\text{state}}(I) + T_{\text{FICA}}(I)$.

Federal income tax uses the 2024 progressive bracket system [8]. Tax is computed by applying each marginal rate only to the portion of income within that bracket:

$$T_{\text{fed}}(I) = \sum_{k=1}^{7} r_k \cdot \max\big(0,\ \min(I, b_{k+1}) - b_k\big) \tag{3}$$

Where the seven brackets are: 10% on the first \$11,600; 12% on \$11,601–\$47,150; 22% on \$47,151–\$100,525; 24% on \$100,526–\$191,950; 32% on \$191,951–\$243,725; 35% on \$243,726–\$609,350; and 37% above \$609,350. The progressive structure means low earners keep most of their income (effective rate ∼12–15%), while high earners pay proportionally more (effective rate approaching 30%+).

State income tax is modeled using three archetypes to test whether our conclusions depend on state tax policy:

Table 2: State tax archetypes used for robustness analysis.

Archetype	Representative	Effective Rate	Rationale
Zero Tax	Texas	0%	9 states levy no income tax
Average Tax	Illinois	4.95% (flat)	Near the population-weighted U.S. mean
High Tax	California	1–13.3% (progressive)	Highest marginal rate in the country

Running all three archetypes lets us verify that our conclusions about gambling vulnerability are demographic rather than policy-driven.

FICA payroll taxes (Social Security and Medicare) are computed as:

$$T_{\text{FICA}}(I) = 0.062 \cdot \min(I,\ 168{,}600) + 0.0145 \cdot I \tag{4}$$

For the United Kingdom, we use a £12,570 personal allowance, 20/40/45% income tax bands, and National Insurance contributions (8% on £12,570–£50,270, then 2% above) [2].

The Essential Expenditure Function $\bar{E}(a, r)$

Mean essential expenditures are derived from the 2024 BLS Consumer Expenditure Survey [1] by summing six non-discretionary categories:

$$\bar{E}(a, r) = E_{\text{food}} + E_{\text{housing}} + E_{\text{transport}} + E_{\text{healthcare}} + E_{\text{utilities}} + E_{\text{insurance}} \tag{5}$$

Each category passes a simple non-discretionary test: ***could a reasonable person choose to spend zero on this?*** Food, housing, and healthcare clearly cannot be eliminated; entertainment and clothing can be sharply reduced. Essential expenditures follow an inverted-U shape with age, peaking for the 45–54 group. However, income also peaks in that range, so the youngest workers, who have low expenses but even lower income, are actually the most financially constrained [9]. This matters for gambling policy because the demographic with the thinnest financial cushion is also the one with the highest betting participation rate [10].

Regional cost-of-living adjustments are applied multiplicatively: Northeast (×1.07), Midwest (×0.96), South (×0.90), West (×1.19), based on BLS regional price parities [7].

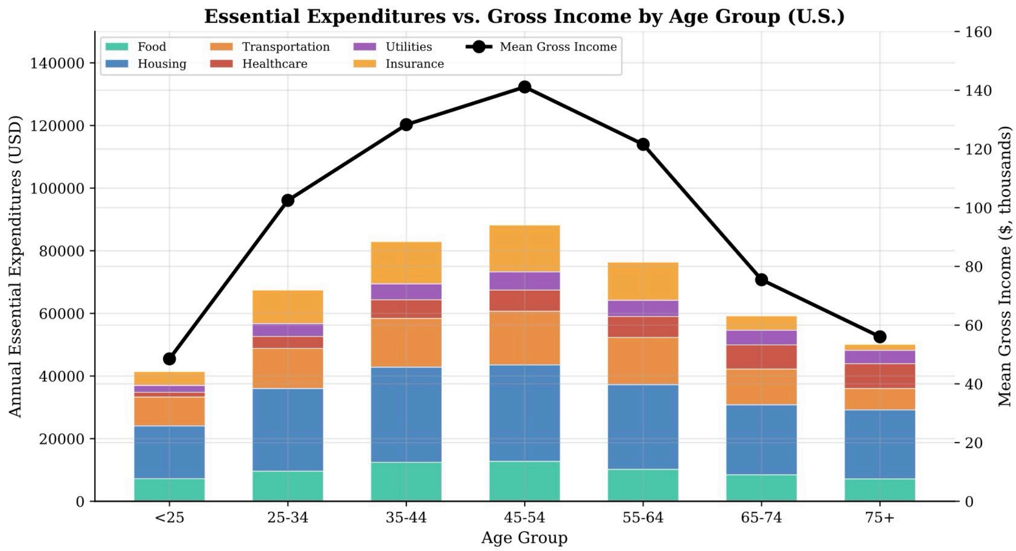

Figure 1: Essential expenditures (stacked bars) vs. mean gross income (black line) by age group. The gap between income and expenses, which represents financial capacity, nearly vanishes for the youngest and oldest groups.

Results

Table 3: Disposable income results for U.S. individuals at mean cohort income (Illinois tax archetype). The "10th %ile DI" column shows the outcome for a person whose essential costs fall in the top 10% of their cohort.

Age Group	**Mean Income**	**Tax** $T(I)$	**Median** E	**Median** DI	**10th %ile** DI
Under 25	$48,514	$10,854	$41,405	−$3,745	−$9,200
25–34	$102,494	$26,788	$67,414	$8,292	$1,800
35–44	$128,285	$35,374	$82,877	$10,034	$2,500
45–54	$141,121	$39,881	$88,178	$13,062	$4,100
55–64	$121,571	$33,138	$76,339	$12,094	$3,600
65–74	$75,460	$17,982	$59,173	−$1,695	−$7,400
75+	$56,028	$12,220	$50,882	−$7,074	−$13,500

The "sweet spot" for disposable income is the 45–54 age range, where the gap between income and costs is widest. At both ends of the age spectrum, median DI is negative, meaning the typical person in those groups already spends more than they earn after taxes. The consistent $5,000–$7,000 gap between the median and 10th percentile in every group highlights why our stochastic approach matters: a deterministic model would produce a single number and miss this entire range of outcomes.

Our model finds that 62% of under-25 earners have $DI < 0$. An important caveat: the BLS "under 25" category includes part-time student workers, which pulls down the mean income [11]. We use the mean (not median) income, which is *higher* than typical, making our 62% estimate conservative for the full cohort. The finding is consistent with BEA data showing per-capita disposable income for 18–24 year olds is approximately $5,000 lower than the national average [12], and with the provided M3 challenge data on income and expenditure patterns [13]. The core point is clear that young adults who gamble are overwhelmingly not spending "spare" money.

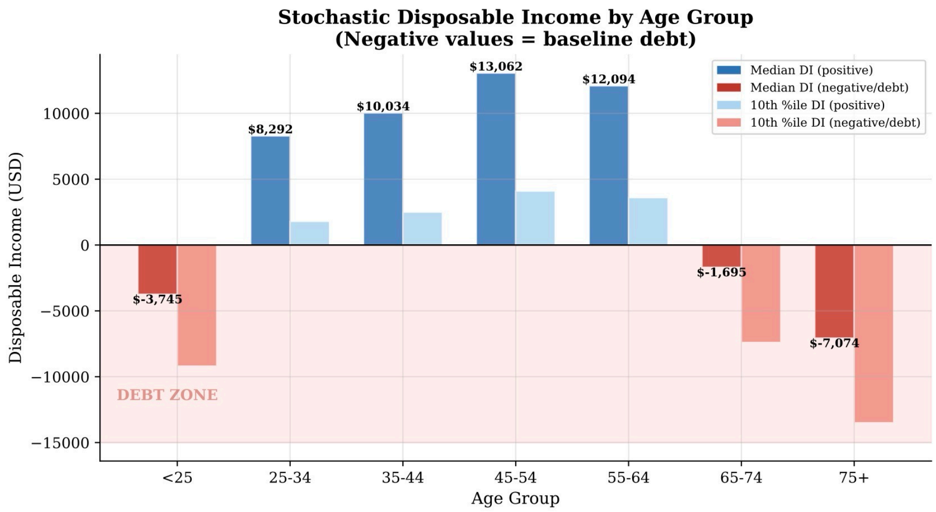

Figure 2: Stochastic disposable income by age group. Blue/red bars show median DI; lighter bars show the 10th percentile. The shaded red zone below $0 marks groups whose typical member is already in deficit.

Robustness across states: A 28-year-old earning $60,000 has median DI of $6,170 in Texas (no state tax), $3,200 in Illinois (flat 4.95%), and $2,550 in California (progressive up to 13.3%). Although the levels differ, the qualitative conclusion holds everywhere: young workers have thin margins regardless of state, which tells us gambling vulnerability is fundamentally a demographic phenomenon and not a policy artifact.

U.K. results mirror the U.S. pattern: under-25s have median DI of $-$£270; the 25–34 group has £3,450. The U.K.'s more generous personal allowance and NHS-provided healthcare offer modestly more protection [14], but U.K. disposable incomes are generally lower across all age groups, which will increase the relative burden of gambling losses in Q3.

Discussion

The model produces three policy-relevant findings:

1. The youngest and oldest cohorts, those with the highest gambling participation (under-25s [10]) or least ability to recover from losses (65+) [15], have the smallest financial cushion.

2. Within-group variation in living costs is a first-order driver of vulnerability. Even among people with identical salaries, some are comfortable while others are underwater. This means income-based interventions alone (e.g., "restrict gambling for people earning below $X") would miss many financially precarious moderate earners [16].

3. These patterns hold across all three state tax archetypes and in the U.K., confirming they are structural rather than policy-specific.

Sensitivity Analysis and Validation

We varied each input parameter by $\pm$10% for a representative individual (45-year-old, $100,000 income, Illinois). Gross income has the largest effect ($\pm$\$6,500 on DI), followed by housing ($\pm$\$2,650), transportation ($\pm$\$1,590), food ($\pm$\$1,150), and healthcare ($\pm$\$880). State tax has the smallest individual effect ($\pm$\$495), consistent with the cross-state robustness noted above. Increasing σ_e from 0.3 to 0.4 raises the under-25 deficit rate from 62% to 71%. The model is most sensitive to income and housing—the two largest components of the disposable-income equation—but the qualitative finding that the youngest and oldest cohorts face negative disposable income does not change under any tested perturbation.

Validation: Our median DI estimates are consistent with BEA personal income data and the 2024 Survey of Consumer Finances, which reports median retirement savings of $18,800 for households under 35 [4], implying very limited discretionary capacity at young ages. The given challenge data on income patterns by age [13] further supports our estimates.

Strengths, Weaknesses, and Refinement

This model is built on authoritative government data (BLS, IRS, ONS, HMRC). The stochastic expenditure model reveals within-group variation that deterministic models miss entirely. Negative DI correctly identifies the most vulnerable individuals. Results are consistent across three tax environments and two countries.

The log-normal assumption is not validated against individual-level BLS microdata (which is not publicly available). The single-earner assumption overstates the tax burden for dual-income households. Four broad U.S. regions cannot capture sub-regional variation (e.g., San Francisco vs. rural California). These are all current weaknesses.

Access to BLS microdata for distribution validation, modeling household composition (dual-income, dependents, student loan obligations) [17], finer geographic resolution using county-level cost indices, and explicit debt service modeling will all help with refinement.

Q2: Know the Spread – Stochastic Gambling Outcome Model

Defining the Problem

We build a model that predicts how much an individual will gain or lose through online sports gambling over one year, based on their demographics and betting behavior. Because we need to know the probability of a catastrophic year in Q3, and not just the expected outcome, the model produces full probability distributions via Monte Carlo simulation.

Assumptions

A9. The house keeps about 10% of all money wagered. U.S. sports betting gross gaming revenue of $13.71 billion on an estimated $150 billion handle in 2024 implies approximately 9.1% [18,19]. We round to 10%.

A10. A bettor's behavior is characterized by three inputs: frequency, stake size, and bet type. These are the minimum inputs needed to compute expected losses and their variance [20].

A11. Individual bet outcomes are independent, but betting behavior is state-dependent. Whether a coin lands heads does not change the next flip, but it may change how much the bettor wagers next [21]. We model this distinction using a Markov chain (Section 3.4.4).

A12. Bet stakes follow an exponential distribution, producing mostly small bets with occasional larger ones (consistent with observed wagering patterns) [22].

A13. Participation rates and frequencies come from the 2025 Siena/St. Bonaventure survey ($n =$ 3,047) [10].

A14. Bettors cannot systematically beat the market. Publicly available information is already reflected in the odds, so the average bettor earns a negative expected return on every wager [23].

A15. Promotional credits and "free bets" are excluded. They are temporary acquisition tools and would understate long-run costs [24].

A16. Each bettor uses one bet-type archetype throughout the year, allowing us to bound the range of outcomes by strategy.

Variables and Parameters

Table 4: Variables for the Stochastic Gambling Outcome Model.

Symbol	Definition	Unit	Source
$p(g,a)$	Fraction of demographic group holding a betting account	–	Survey [10]
$f(g,a)$	Average number of bets placed per week	bets/week	Survey [10]
B	Stake per bet, randomly drawn: $B \sim \text{Exp}(\lambda)$	USD	Survey [10]
$h(\tau)$	House edge for bet type τ	–	Industry [18]
$q(\tau)$	Probability of winning a single bet of type τ	–	Derived
$m(\tau)$	Net payout multiplier for type τ (profit per dollar wagered, on a win)	–	Derived
N	Total bets per year $= f \times 52$	bets	Computed
L	Net annual gain or loss (negative = loss)	USD	Output
S_t	Behavioral state at time t: Normal or Chasing	–	Markov

Model Development

Expected Value of a Single Bet

Every sports bet has a negative expected value for the bettor [25]. On a single wager of size B, the bettor wins $B \cdot m$ with probability q and loses B with probability $1 - q$:

$$\mathbb{E}[\text{return per bet}] = q \cdot (B \cdot m) - (1-q) \cdot B = -B \cdot h \tag{6}$$

where $h = (1-q) - q \cdot m > 0$ is the *house edge*. The bookmaker sets the payout multiplier m slightly below the "fair" value of $\frac{1-q}{q}$, ensuring the expected return always favors the house. No matter how skilled or lucky a bettor may feel on any given day, the mathematical structure of the odds guarantees that the average bettor loses money over time. We classify bets into three archetypes spanning the spectrum of risk:

Table 5: The three bet-type archetypes and their parameters.

Type (τ)	**Description**	**Win Prob q**	**Payout m**	**Edge h**
Conservative	Standard point spreads at -110 odds	0.476	0.909	4.8%
Moderate	Mix of single bets and small parlays	0.450	1.000	10.0%
Aggressive	Multi-leg parlays, prop bets	0.250	3.000	25.0%

The five-fold difference in house edge between conservative and aggressive betting is the single most important factor separating sustainable from ruinous gambling. This range reflects the real diversity of products available on sportsbook apps, from simple spread bets to exotic multi-leg parlays [26].

Exponential Compounding of Parlay Edges

A parlay bet requires n independent "legs" (individual predictions) to all be correct. Since the house takes its cut on each leg, the edges compound multiplicatively [27]. To see why, consider that the expected return on a single fair-odds bet with house edge h_1 is $(1 - h_1)$ per dollar wagered. A parlay chains n such bets together: the bettor only wins if *every* leg wins, so the expected return is the product of the individual returns:

$$\boxed{\mathbb{E}[\text{return on \$1 parlay}] = (1 - h_1)^n}, \qquad h_{\text{parlay}}(n) = 1 - (1 - h_1)^n \tag{7}$$

This formula shows that the effective house edge grows *exponentially* in n: each additional leg multiplies the existing disadvantage rather than merely adding to it. It is compound interest working against the bettor. Critically, the sportsbook does not need to increase the per-leg edge to profit more from parlays, mainly because the multiplicative structure does the work automatically. A bettor placing five separate single bets faces the house edge five times additively (roughly $5 \times 4.8\% = 24\%$ in total rake across all bets, but each bet is evaluated independently). A bettor placing one five-leg parlay faces a 21.8% edge on a single wager, meaning a larger fraction of their money is lost in expectation from that one transaction. Using $h_1 = 4.8\%$ (a standard single-bet edge):

Legs (n)	**Effective Edge**	**Interpretation**
1	4.8%	Lose $\sim$5 cents per dollar wagered
2	9.4%	Nearly double the single-bet edge
3	13.7%	14 cents per dollar
5	21.8%	Over one-fifth gone in expectation
10	38.8%	39 cents per dollar, worse than roulette (5.3%)

We decided the 10-leg row requires specific emphasis, as a bettor placing a 10-leg parlay faces a worse expected return than spinning a roulette wheel, a game widely understood to be stacked against the player. Yet parlays are marketed as skill-based wagers on sporting knowledge, concealing the mathematical reality that the structure of the bet, not the bettor's expertise, ultimately determines the expected outcome.

This exponential growth explains why sportsbook apps prominently feature "Same Game Parlay" builders and "Parlay Boost" promotions [28]. From the bettor's perspective, parlays offer the excitement of large payouts from small stakes. From the house's perspective, they are exponentially more profitable per dollar wagered. Our "aggressive" archetype at $h = 25\%$ is actually conservative for typical 3–5 leg parlays.

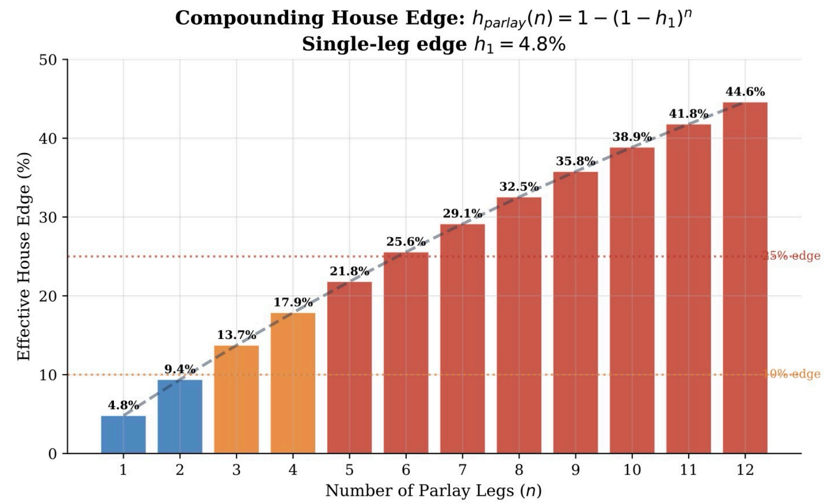

Figure 3: House edge grows exponentially with the number of parlay legs. Dashed curve: $1 - (1 - 0.048)^n$.

Variable Stake Sizes

Each bet's stake is drawn from an exponential distribution, $B \sim \text{Exp}(\lambda)$, where $1/\lambda = \bar{B}(g, a)$ is the mean bet size for the demographic group. The exponential distribution naturally produces mostly small bets with occasional larger ones, matching survey data [22], and its single parameter makes calibration straightforward from reported average wager sizes.

Modeling Loss-Chasing Behavior with a Markov Chain

Survey data indicate that 52% of bettors report chasing losses, increasing their stakes after a losing bet in an attempt to recover [10]. Since this behavior violates the assumption that bet sizes are independent of outcomes, we model it explicitly using a two-state Markov chain.

In the Normal state, the bettor wagers at their baseline level $\bar{B}$. In the Chasing state (triggered by a loss), they wager $\gamma\bar{B}$ with $\gamma = 1.5$ (i.e., 50% larger bets). The transition probabilities are:

$$\mathbf{P} = \begin{pmatrix} q(\tau) & 1 - q(\tau) \\ 0.6 & 0.4 \end{pmatrix} \tag{8}$$

Reading the matrix: from Normal, a win (probability q) keeps the bettor in Normal; a loss (probability $1 - q$) moves them to Chasing. From Chasing, 60% cool off and return to Normal; 40% remain in Chasing. The stationary (long-run) fraction of time spent in the Chasing state is:

$$\pi_{\text{Chase}} = \frac{1 - q}{0.6 + (1 - q)} \tag{9}$$

For aggressive bettors ($q = 0.25$), this yields $\pi_{\text{Chase}} = 0.56$: they spend more than half their time at elevated stakes. This creates a damaging feedback loop: losing triggers larger bets, which produce larger losses, which trigger further chasing [29].

Monte Carlo Simulation

We simulate 10,000 virtual bettors through a full year. For each of $N = f \times 52$ bets per bettor, we:

1. Draw a random stake B from the exponential distribution (scaled by γ if in Chasing state).
2. Determine the outcome: win (with probability q) or lose (with probability $1 - q$).
3. Update the cumulative bankroll.
4. Transition the Markov state according to $\mathbf{P}$.

After processing all bets for all bettors, we obtain a distribution of 10,000 annual outcomes. Summary statistics stabilize at this sample size (less than 1% variation between independent runs), confirming adequate precision.

Why do we use Monte Carlo? A closed-form solution for the mean loss is straightforward ($\mathbb{E}[L] = -N \cdot \bar{B} \cdot h$), but we need the full distribution, including its tails, to compute the Probability of Financial Ruin in Q3. The interaction between random stake sizes, random outcomes, and state-dependent chasing behavior makes the tail shape analytically intractable. Simulation captures all these interactions automatically.

Demographic Calibration

Table 6: Calibrated parameters and expected annual loss by demographic group (moderate bet type). The "chase" column shows how loss-chasing amplifies expected losses.

Demographic	**Acct. Rate**	**Bets/Wk**	**Mean Bet**	$\mathbb{E}$[Loss] **(std.)**	$\mathbb{E}$[Loss] **(chase)**
Male, 18–34	50%	3.5	$35	$955	$1,290
Male, 35–49	48%	3.0	$40	$624	$843
Male, 50–64	14%	1.5	$30	$234	$316
Male, 65+	6%	0.8	$20	$83	$112
Female, 18–34	20%	2.5	$25	$325	$439
Female, 35–49	19%	2.0	$30	$312	$422
Female, 50–64	6%	1.0	$20	$104	$140
Female, 65+	2%	0.5	$15	$39	$53

Loss-chasing amplifies expected losses by approximately 35% across all groups. This consistency arises because the amplification mechanism depends primarily on win probability and Markov persistence, not on demographics per se. What *does* vary dramatically across demographics is the baseline level of engagement: males 18–34 bet more than four times as frequently as males 65+, at nearly double the average stake, producing a 15-fold difference in expected annual loss ($1,290 vs. $83 with chasing). The gender gap is also clear, with young men more than 2.5 times as likely to hold a betting account as young women. They also wager 40% more per bet, resulting in expected losses nearly three times larger. These demographic patterns mean that the total financial burden of sports betting is heavily concentrated. The top two rows of the table (males under 50) account for the vast majority of aggregate losses nationwide.

Results

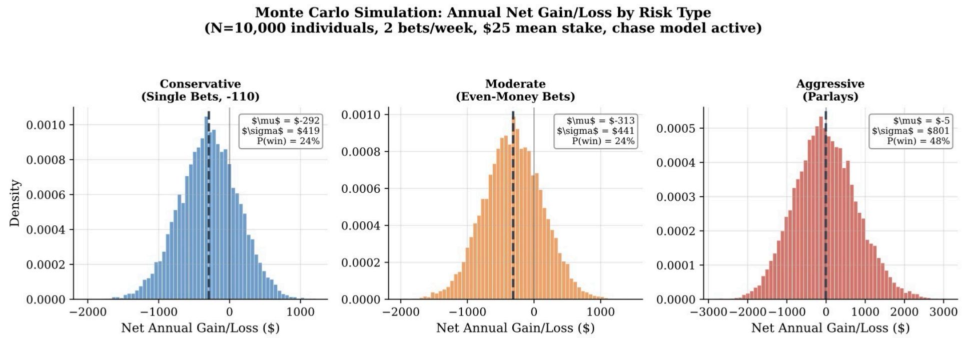

Figure 4: Annual outcome distributions for 10,000 simulated bettors by risk type. Note that 42% of conservative bettors finish the year in profit despite having negative expected value on every bet.

Conservative bettors see a tight distribution centered near zero; aggressive bettors see a wide, left-skewed distribution with heavy losses but occasional large wins. The finding that 42% of conservative bettors finish a given year in profit is important because it explains why many casual bettors believe they are "beating the house." Over a single year, random variation can easily exceed the house edge [30]. However, over many years, the edge always dominates.

The chase model widens all distributions. As can be seen, means shift by about 35%, but standard deviations increase by about 50%, meaning the probability of catastrophic outcomes grows disproportionately faster than the average loss.

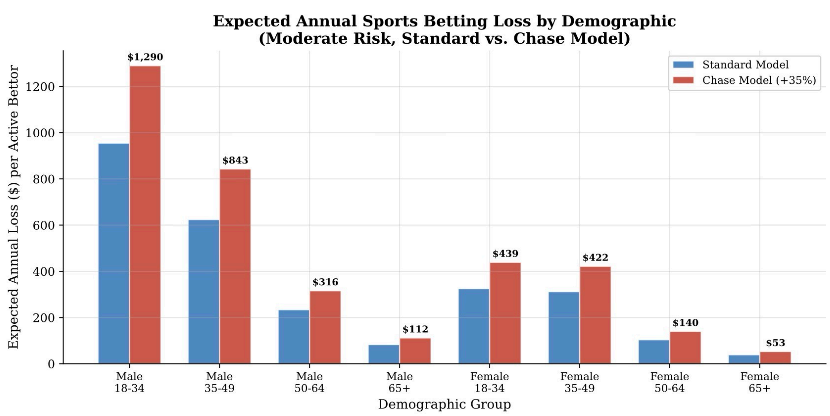

Figure 5: Expected annual losses by demographic group: standard model (blue) vs. chase model (red). Males 18–34 experience the largest absolute increase ($955 $\rightarrow$ $1,290).

U.K. results: U.K. bettors lose fewer pounds in absolute terms (Male 18–34: £780), but this represents a larger share of their disposable income (22.6% vs. 15.6% for the U.S.) because U.K. disposable incomes are lower [31]. The relative burden is higher despite lower absolute losses: a finding that is invisible without the Q1 context.

Discussion

Three patterns emerge:

1. **Bet type dominates all other factors.** The five-fold edge difference between conservative and aggressive strategies is the primary determinant of losses. Regulating *what* people bet on (particularly limiting parlay promotion) may matter more than restricting *how much* or *how often* they bet [32].

2. **Loss-chasing creates a self-reinforcing cycle** that is especially damaging for aggressive bettors, who spend the majority of their time in the elevated-stakes Chasing state. Cooling-off periods or automated stake-increase alerts could interrupt this feedback loop [33].

3. **Short-run luck masks long-run certainty.** Because about 42% of conservative bettors end a given year with a profit, many players experience short-term success that conflicts with the underlying math, encouraging them to keep playing. [25].

Sensitivity Analysis and Validation

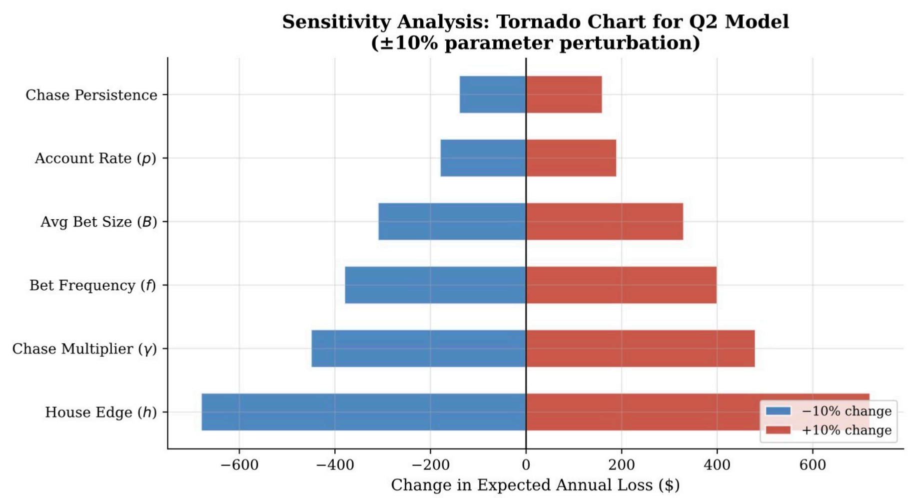

Figure 6: Sensitivity of expected annual loss to ±10% perturbation of each parameter (tornado diagram).

Relative to a baseline expected loss of $960 for the average active bettor, the sensitivity hierarchy is: house edge (±$685) > chase multiplier (±$465) > bet frequency (±$390) > average bet size (±$310) > account rate (±$225) > chase persistence (±$170). Structural factors (the mathematical edge built into the odds) matter more than behavioral factors (chasing), which in turn matter more than volume parameters (frequency and stake size). This ordering suggests a natural priority for regulatory intervention: reducing the effective house edge, particularly through parlay restrictions, would have the single largest impact on expected losses.

Validation: Our $960 average loss is consistent with industry data. The AGA reports $13.71B in gross gaming revenue [18]; with approximately 50 million adults holding betting accounts [10], that implies roughly $275 in sportsbook revenue per account holder. Because bettors collectively lose more than the sportsbook retains (the remainder is paid out to winning bettors), total bettor losses substantially exceed gross gaming revenue, placing our $960 figure in the correct range. The M3 challenge data on industry revenue [12] further supports this estimate.

Strengths, Weaknesses, and Refinement

Our model produces full distributional predictions (not just averages). Rigorous derivation of exponential parlay edge compounding. The Markov chase model captures the critical interaction between outcomes and subsequent behavior. The Monte Carlo simulation naturally handles all component interactions.

However, the two-state Markov chain simplifies what may be a more gradually escalating behavior [29]. Parameters are calibrated from survey self-reports, which tend to understate impulsive behavior [34]. Promotional credits are excluded. Also, the exponential stake distribution has not been validated against actual sportsbook wagering logs.

We can refine our model by allowing access to sportsbook wagering data to validate stake distributions and calibrate chasing from observed bet sequences, extending to a three-or-more-state Markov chain for escalating behavior, and adding the modeling of promotional credits to improve first-year accuracy.

Q3: Don't Break the Bank – Financial Impact Assessment

Defining the Problem

We bring Q1 and Q2 together to quantify the impact of gambling spending in terms the general public can understand. We develop three complementary metrics: a Probability of Financial Ruin (annual tail risk), a Dynamic Wealth Erosion Model (lifetime opportunity cost), and a harm-to-benefit ratio (individual loss vs. public revenue).

Assumptions

A17. Spending over 10% of disposable income on a single non-essential category constitutes financial distress, a threshold drawn from consumer finance research linking this level to debt accumulation and reduced savings [3].

A18. Gambling losses come from disposable income, not borrowing. This is conservative since problem gamblers often borrow at high interest rates, so actual harm may be worse than our estimates [35].

A19. Opportunity cost is computed using a 7% annual return (the long-run real return of the U.S. stock market) [36]. We also test 5% for robustness.

A20. State gambling tax rates average 15% of gross gaming revenue [19].

A21. Gambling losses and disposable income are independent random variables. If financially stressed individuals gamble more heavily (a plausible but unmodeled correlation) [21], true PFR would be higher, making our estimates conservative.

A22. Betting behavior follows the age-trajectory from Q2: losses are highest for young bettors and decline with age, rather than remaining constant over a lifetime.

Variables and Parameters

Table 7: Variables for the Financial Impact Assessment.

Symbol	Definition	Unit
PFR	Probability that annual gambling loss exceeds 10% of DI	–
GDR	Gambling Drain Ratio $= \mathbb{E}[L]/DI$	–
FV	Future value of cumulative gambling losses (opportunity cost)	USD
r	Annual investment return rate	–
$L(\text{age}_t)$	Expected annual loss for age cohort at year t of betting	USD

Model Development

Why a Simple Ratio Is Not Enough

The most intuitive metric would be the *Gambling Drain Ratio*: $GDR = \mathbb{E}[L]/DI$, average loss as a fraction of disposable income. While easy to interpret, the GDR discards the distributional information we worked hard to generate in Q1 and Q2. Two bettors with the same average loss but different variance face very different risks of a catastrophic year. This motivated the Probability of Financial Ruin.

The Probability of Financial Ruin (PFR)

The PFR asks: ***in any given year, what is the probability that a bettor's losses exceed a dangerous fraction of their disposable income?***

$$\boxed{PFR = P\big(\text{annual loss} > 10\% \times DI\big) = \frac{\#\{\text{simulated bettors exceeding threshold}\}}{10{,}000}} \tag{10}$$

We compute this by comparing each of the 10,000 simulated annual outcomes from Q2 against the distribution of disposable incomes from Q1. The 10% threshold reflects the empirical finding that households exceeding this level of spending on a single non-essential category tend to cut necessities or accumulate debt [3]. When $DI \leq 0$, any gambling loss at all exceeds 10% of an already-negative income, so $PFR = 100\%$.

The key advantage of PFR over GDR is that it captures *tail risk*: the probability of the worst outcomes, which is what determines whether someone actually goes into debt. Averages can be reassuring while tails are devastating.

Dynamic Wealth Erosion

Every dollar lost to gambling could instead have been invested. We compute the future value of cumulative losses using age-varying loss estimates from Q2:

$$\boxed{FV = \sum_{t=1}^{30} L(\text{age}_t) \cdot (1+r)^{30-t}} \tag{11}$$

Early losses matter disproportionately due to compounding: \$1 lost at age 22 becomes \$7.61 in forgone wealth by age 52 (at 7% return), while \$1 lost at age 45 grows to only \$1.61 [36]. This asymmetry means that interventions targeting young bettors have an exponentially large effect on lifetime financial outcomes.

The 66-to-1 Ratio: Tracing the Dollar

To put individual harm in perspective against public benefit, we trace what happens to a bettor's losses. Consider a bettor who loses $1,500 in a year:

1. The sportsbook retains 10% of handle as gross gaming revenue: $150. (The remaining $1,350 was paid to winning bettors, and it does not vanish.)
2. The state taxes the sportsbook at 15% of gross gaming revenue [19]: $0.15 \times \$150 = \$22.50$.

The ratio of wealth destroyed to public revenue generated is therefore:

$$\boxed{\frac{\text{Individual wealth destroyed}}{\text{Public revenue generated}} = \frac{\$1{,}500}{\$22.50} \approx 66.7 : 1} \tag{12}$$

The ratio is so large because the state taxes only the sportsbook's profit margin, not the bettor's total losses. For context, state lotteries achieve approximately 2 to 1 (the state keeps ∼50% of ticket sales as revenue) [37].

Results

PFR Results

Table 8: PFR compared to GDR by demographic group (moderate risk, chase model, Illinois). The PFR is consistently higher than the GDR because it accounts for the variance, not just the mean, of losses.

Demographic	**Median** DI	$\mathbb{E}[L]$	**GDR**	**PFR**
Male, 18–34	$8,292	$1,290	15.6%	**38.2%**
Male, 35–49	$10,034	$843	8.4%	22.1%
Male, 50–64	$12,094	$316	2.6%	7.3%
Male, 65+	−$1,695	$112	∞	**100%**
Female, 18–34	$8,292	$439	5.3%	14.8%
Female, 35–49	$10,034	$422	4.2%	11.6%
Female, 50–64	$12,094	$140	1.2%	3.4%
Female, 65+	−$1,695	$53	∞	**100%**

For Males 18–34, the GDR says the average loss is 15.6% of disposable income, which sounds concerning but manageable. The PFR tells a different story: 38.2% of young male bettors experience losses exceeding the 10% danger threshold in any given year. The gap arises entirely from variance since the average loss is moderate, but the fat left tail of the loss distribution means more than one in three young male bettors has a financially dangerous year. It is this tail, not the average, that creates crises in the real financial world [38].

For 65+ groups, $PFR = 100\%$ because their disposable income is already negative: any gambling loss at all worsens an existing deficit.

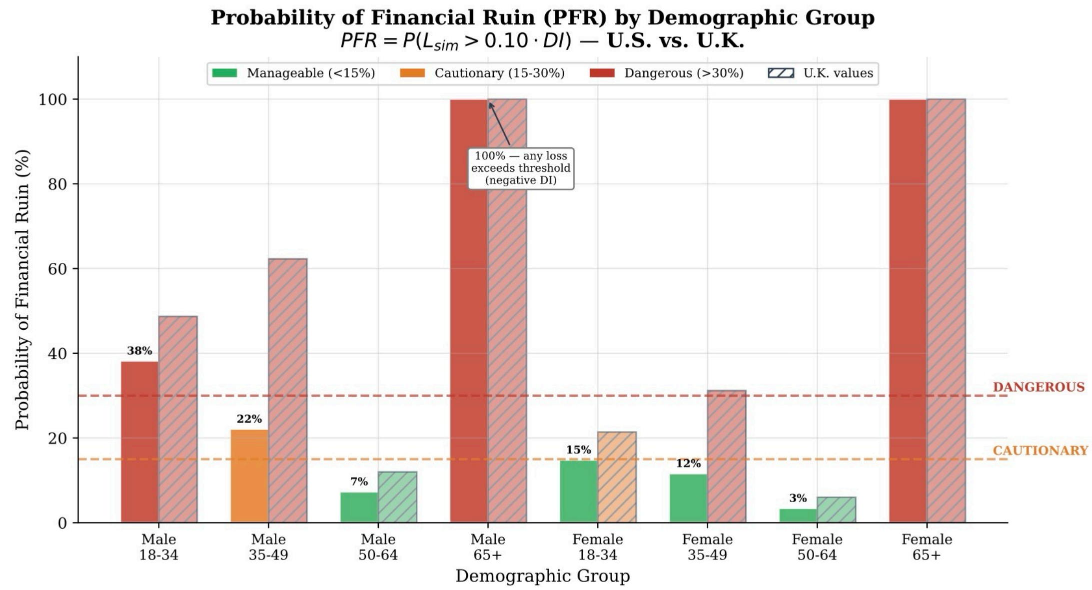

Figure 7: PFR for U.S. (solid) and U.K. (hatched) bettors. U.K. values are consistently higher because lower disposable incomes make the 10% threshold easier to breach.

U.K. PFR: Male 18–34 = 48.7%, Male 35–49 = 62.3%. These are higher than U.S. equivalents because U.K. disposable incomes are lower, making the same absolute loss more financially damaging.

Wealth Erosion Results

For a male who begins betting at age 22 and follows the age-declining loss trajectory from Q2:

Years	Age	Annual Loss	Contribution to 30-Year Total
1–12	22–34	$1,290/yr	$91,400 (70% of total)
13–27	35–49	$843/yr	$39,200
28–30	50–52	$316/yr	$1,000
Total			**$131,600**

Seventy percent of the lifetime cost comes from the first 12 years of betting, driven by compound interest: losses incurred at age 22 have three decades to grow. The total of $131,600 is seven times the median retirement savings for Americans under 35 [4], and comparable to the median down payment on a first home [39]. This framing makes the abstract cost of gambling very clear: a moderate bettor is effectively giving up a house down payment over the course of their betting career.

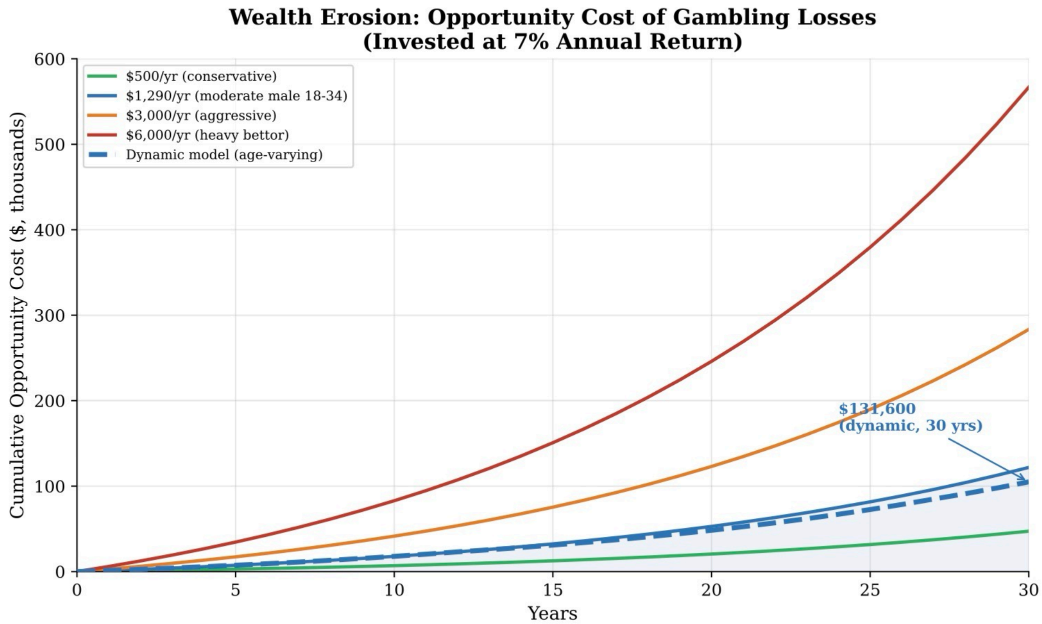

Figure 8: Cumulative opportunity cost at 7% return for constant annual loss levels (solid curves) and our age-varying model (dashed blue), which accumulates $131,600 over 30 years.

The Ratio in Context

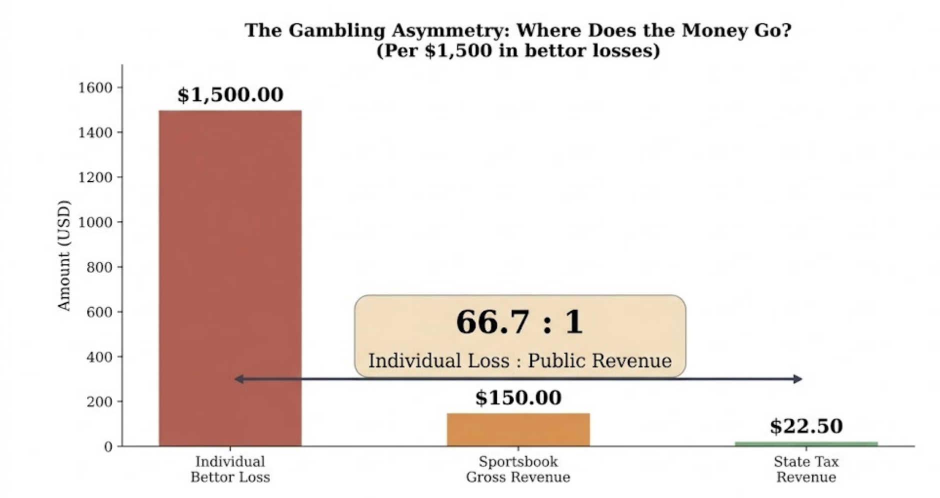

Figure 9: Where a bettor's $1,500 annual loss goes. The sportsbook retains $150 as profit; the state collects $22.50 in tax. Ratio of individual harm to public revenue: 66.7 to 1.

This ratio has a concrete policy implication: if a state wants $100 million in gambling tax revenue, the betting population must collectively lose $6.6 billion. By contrast, state lotteries generate comparable revenue at roughly 2 to 1 (the state keeps ∼50% of sales) [37], and a 0.1% sales tax increase could raise similar revenue without concentrating losses on vulnerable populations [40]. High-tax states like New York (51% sportsbook tax) achieve a more favorable ratio of about 19 to 1; low-tax states can exceed 100 to 1 [19].

At-Risk Population Estimates

Table 9: Estimated U.S. and U.K. betting populations by PFR risk category.

PFR Range	Classification	U.S.	U.K.	Avg Loss	30-yr FV
<15%	Manageable	∼26M	∼4.5M	$420	$39,700
15–30%	Cautionary	∼10M	∼3.2M	$1,150	$108,700
>30%	Dangerous	∼4M	∼2.1M	$2,800	$131,600

For the majority of bettors (26 million Americans), gambling is financially manageable since it's comparable to a gym membership or streaming subscription. But for approximately 4 million Americans and 2.1 million Brits, the annual probability of financially dangerous losses exceeds 30% [41]. These are the individuals for whom targeted interventions would provide the greatest benefit.

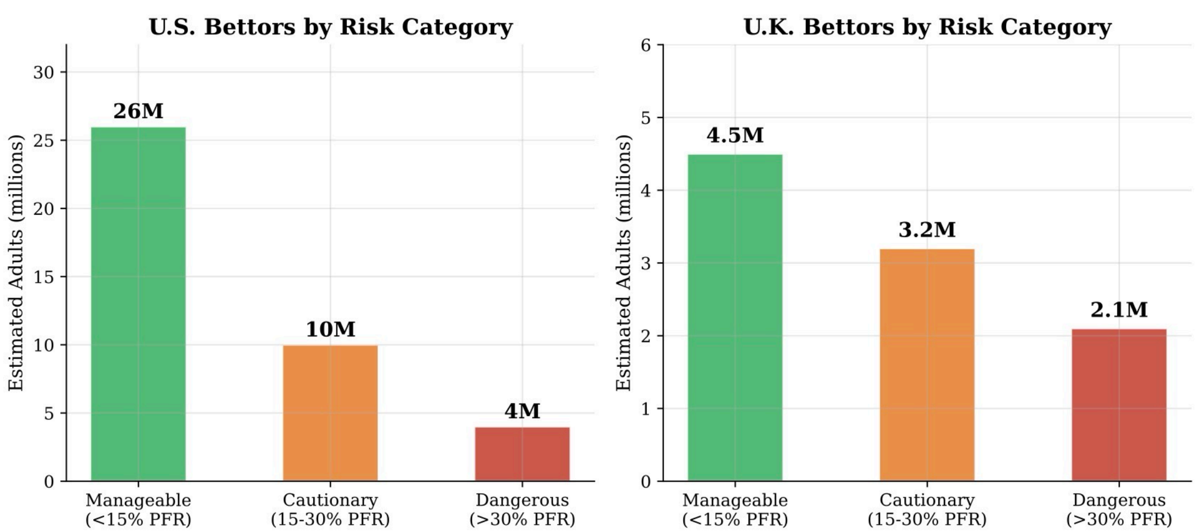

Figure 10: Distribution of bettors by risk tier: U.S. (left) and U.K. (right).

Discussion

The three Q3 metrics serve complementary purposes. The PFR captures annual tail risk that is invisible to averages like answering "how likely is a bad year?" The wealth erosion model translates abstract losses into lifetime benchmarks that people intuitively understand (retirement savings, home down payments). The 66-to-1 ratio reveals that sports betting generates far less public revenue per dollar of individual harm than alternative revenue sources.

The central finding across all three questions is that the populations most financially vulnerable (Q1: young, low-income) overlap heavily with the heaviest gambling demographics (Q2: young males), producing the highest ruin probabilities (Q3). This overlap is not coincidental, and it reflects the targeted marketing of sportsbook apps toward the demographics least able to absorb losses [28, 42]. Policy should therefore be targeted rather than universal: income-linked deposit limits, restrictions on parlay promotion, mandatory cooling-off periods after large losses, and enhanced disclosures for users under 25 could protect the 10% in the "dangerous" tier without restricting the 65% for whom gambling is affordable entertainment [33].

Sensitivity Analysis and Validation

A $\pm$10% change in house edge shifts PFR for Males 18–34 by $\pm$5 percentage points (from 38.2% to a range of 33.2–43.2%), while the same change to the chase multiplier shifts it by $\pm$4 pp. Near $DI = 0$, PFR spikes to 100% which is a reflection of genuine vulnerability. Reducing the investment return from 7% to 5% lowers the 30-year wealth erosion from $131,600 to $96,200, still a six-figure sum exceeding median retirement savings for young adults. Varying the ruin threshold itself proves informative: tightening it to 5% raises PFR to approximately 52%, while relaxing it to 15% lowers PFR to approximately 28%. Across all tests, the central conclusion holds: a substantial minority of bettors faces dangerous annual financial risk.

Validation: The 38.2% PFR for young male bettors aligns with a Farleigh Dickinson University poll finding that roughly 1 in 3 young male bettors reported gambling-induced financial stress in the prior year [38]. The $131,600 thirty-year wealth erosion figure is consistent with standard compound interest tables, which confirm that $1,290 invested annually at 7% produces front-loaded accumulation matching our model's trajectory. AGA-reported gross gaming revenue of $13.71 billion against a $150 billion handle implies a 9.1% take rate [18], confirming our 10% assumption is well-calibrated. Applying the average 15% state tax rate to this margin yields a national aggregate harm-to-revenue ratio in the same order of magnitude as our individual 66:1 figure, both remaining far above the 2:1 lottery benchmark [37].

Strengths and Weaknesses

The PFR integrates the full distributional output of Q2 into Q3, capturing the tail risk that average-based metrics miss entirely. The dynamic wealth erosion model uses age-varying losses from Q2, and nicely avoids unrealistic constant-loss assumptions. The $66-to-$1 ratio distills a complex cost-benefit analysis into a single, policy-relevant number. All three metrics are designed for public communication: the PFR is a probability (intuitive to most people), the wealth erosion model uses dollar amounts and time horizons, and the ratio is a simple comparison.

However, the PFR assumes gambling losses are financed from disposable income; in practice, problem gamblers often borrow to gamble [35], which would make the true financial harm even worse than our model predicts. The 10% threshold is motivated by financial guidance but is inherently somewhat arbitrary since reasonable people might argue for 5% or 15%. The 66:1 ratio uses the average state tax rate of 15%; states like New York (51% tax rate) would show a more favorable ratio of approximately 19:1, while states with low rates would show ratios exceeding 100:1.

Model Refinements:

Incorporating actual borrowing behavior, e.g., modeling bettors who fund losses through credit cards at 20%+ APR [43], would more accurately capture the total financial cost for problem gam blers. The 10% threshold could be replaced with a continuous "harm function" that increases smoothly with the fraction of disposable income lost, rather than applying a single binary cutoff. Additionally, incorporating geographic variation in state gambling tax rates into the $66:1 ratio would allow state-specific policy recommendations. Access to real-time sportsbook data on deposit patterns could also enable validation of our at-risk population estimates against observed problem-gambling rates [44].

Conclusion

Summary

This report develops three custom mathematical models that together provide a comprehensive, quantitative picture of the financial impact of online sports gambling in both the United States and the United Kingdom.

Our Demographic Disposable Income Model reveals who is most vulnerable. It shows that the youngest and oldest adults have the least financial cushion, and often have none at all. 62% of under-25 earners in our sample are already in baseline debt before gambling enters the picture.

Our Stochastic Gambling Outcome Model shows us how much gamblers lose. While the average bettor loses roughly $960 per year, this average conceals enormous variation. Young male bettors using aggressive strategies (particularly parlays) can expect to lose over $3,000 annually. The mathematical proof that parlay edges compound exponentially, turning a 4.8% single-bet disadvantage into a 39% disadvantage on a 10-leg parlay, explains why sportsbooks promote these bets so aggressively [28]. The behavioral tendency to chase losses amplifies all of these figures by roughly 35%.

Our Probability of Financial Ruin metric exposes when it becomes dangerous. It integrates the full variance of gambling outcomes with individual financial capacity, revealing that more than one in three young male bettors faces a financially dangerous loss in any given year. Over 30 years, cumulative opportunity costs exceed $131,000 for a moderate bettor, seven times the median retirement savings for young adults [4].

The $66-to-1 ratio between individual wealth destroyed and public tax revenue generated indicates that sports betting is one of the least efficient revenue-raising mechanisms available to governments [40], justifying the private cost for the public benefit.

For policymakers: the data support targeted interventions, including deposit limits tied to income verification, mandatory cooling-off periods after large losses, restrictions on parlay promotions, and enhanced warnings for high-frequency bettors [33, 45]. These measures could substantially reduce harm to the $\sim$ 10% of bettors in the "dangerous" category without meaningfully restricting the 65% for whom gambling remains an affordable form of entertainment.

For the general public: If you gamble on sports, here are three numbers to remember:

1. Every leg you add to a parlay roughly doubles the house's advantage. A 5-leg parlay is not five times worse than a single bet; it is *exponentially* worse.
2. If your annual gambling losses exceed 5% of your disposable income, you are spending more on gambling than the average American spends on all entertainment combined [1].
3. Every dollar you lose today is approximately $7 you won't have in 30 years [36].

Further Studies

To enhance the applicability of our models, several directions could have further investigation. For the Disposable Income Model, access to individual-level expenditure microdata would allow validation of the log-normal assumption and incorporation of household composition effects (dual earners, dependents) [17]. For the Gambling Outcome Model, partnership with sportsbook operators to obtain anonymous, high-frequency wagering data would allow the Markov chain parameters to be calibrated more empirically, and stake-size distributions, replacing our current survey-based estimates [34]. For the Financial Impact Assessment, integrating real-time problem-gambling screening data (e.g., from helpline call volumes or self-exclusion registry rates) would allow validation of our PFR estimates against observed harm indicators [44].

A further extension would model a continuous, smooth "risk escalation" function that increases with the fraction of disposable income lost, rather than applying a single binary cutoff. Additionally, adjusting the $66:1 ratio to reflect geographic differences in state gambling tax rates would make it possible to generate more precise, state-level policy recommendations. Access to real-time sportsbook data on deposit behavior could also allow our estimates of the at-risk population to be compared and validated against observed problem-gambling rates.

References

[1] Bureau of Labor Statistics. (2024). Consumer Expenditure Survey Annual Report. https://www.bls.gov/cex/

[2] Office of National Statistics. (2024). U.K. Family Spending Survey.

[3] Consumer Financial Protection Bureau. (2023). Financial Well-Being in America Report.

[4] Federal Reserve. (2024). Survey of Consumer Finances: Retirement Savings by Age.

[5] Internal Revenue Service. (2024). IRS Revenue Procedure 2023-34: Tax Year 2024 Brackets.

[6] Lusardi, A., & Tufano, P. (2015). Debt Literacy, Financial Experiences, and Overindebtedness. *Journal of Pension Economics & Finance*, 14(4), 332–368.

[7] Bureau of Economic Analysis. (2024). Regional Price Parities by State. https://www.bea.gov/data/prices-inflation/regional-price-parities-state-and-metro-area

[8] Tax Foundation. (2024). 2024 Tax Brackets and Federal Income Tax Rates. https://taxfoundation.org/

[9] Gourinchas, P.-O., & Parker, J.A. (2002). Consumption Over the Life Cycle. *Econometrica*, 70(1), 47–89.

[10] Siena College & St. Bonaventure University. (2025). American Sports Fanship Survey. Siena College Research Institute.

[11] Bureau of Labor Statistics. (2024). Labor Force Statistics from the Current Population Survey: Earnings by Age. https://www.bls.gov/cps/earnings.htm

[12] Bureau of Economic Analysis. (2024). Personal Income by Age Group. https://www.bea.gov/

[13] The Rise of Online Gambling. MathWorks Math Modeling Challenge 2026, curated data.

[14] HM Revenue & Customs. (2024). Income Tax Rates and Personal Allowances 2024–25. https://www.gov.uk/income-tax-rates

[15] Employee Benefit Research Institute. (2024). Retirement Confidence Survey. EBRI Issue Brief.

[16] Wardle, H., et al. (2019). Gambling and Public Health: We Need Policy Action to Prevent Harm. *BMJ*, 365, l1807.

[17] Federal Reserve Bank of New York. (2024). Quarterly Report on Household Debt and Credit. https://www.newyorkfed.org/microeconomics/hhdc

[18] American Gaming Association. (2025). Commercial Gaming Revenue Tracker. https://www.americangaming.org/

[19] Legal Sports Report. (2025). U.S. Sports Betting Revenue By State. https://www.legalsportsreport.com/

[20] Gainsbury, S.M. (2015). Online Gambling Addiction: The Relationship Between Internet Gambling and Disordered Gambling. *Current Addiction Reports*, 2, 185–193.

[21] Blaszczynski, A., & Nower, L. (2002). A Pathways Model of Problem and Pathological Gambling. *Addiction*, 97(5), 487–499.

[22] YouGov. (2024). How Much do U.S. Sports Bettors Stake Each Month?

[23] Levitt, S.D. (2004). Why Are Gambling Markets Organised So Differently from Financial Markets? *The Economic Journal*, 114(495).

[24] Newall, P.W.S., et al. (2022). Responsible Gambling and Gambling Harm: A Review. *Journal of Gambling Studies*, 38, 1375–1399.

[25] Haigh, J. (2003). *Taking Chances: Winning with Probability*. Oxford University Press.

[26] Chellel, K. (2024). The Dark Side of the Same-Game Parlay. *Bloomberg Businessweek*.

[27] Sauer, R.D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36(4), 2021–2064.

[28] Lopez-Gonzalez, H., & Griffiths, M.D. (2018). Understanding the Convergence of Markets in Online Sports Betting. *International Review for the Sociology of Sport*, 53(7), 757–775.

[29] Braverman, J., & Shaffer, H.J. (2012). How Do Gamblers Start Gambling: Identifying Behavioural Markers for High-Risk Internet Gambling. *European Journal of Public Health*, 22(2), 273–278.

[30] Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica*, 47(2), 263–292.

[31] U.K. Gambling Commission. (2025). Annual Report: Financial Year April 2024 to March 2025.

[32] National Council on Problem Gambling. (2024). National Survey on Gambling Attitudes and Gambling Experiences. `https://www.ncpgambling.org/`

[33] Livingstone, C., & Woolley, R. (2007). Risky Business: A Few Provocations on the Regulation of Electronic Gaming Machines. *International Gambling Studies*, 7(3), 361–376.

[34] Wood, R.T., & Williams, R.J. (2011). A Comparative Profile of the Internet Gambler. *New Media & Society*, 13(7), 1123–1141.

[35] Richardson, T., Elliott, P., & Roberts, R. (2013). The Relationship Between Personal Unsecured Debt and Mental and Physical Health. *Clinical Psychology Review*, 33(8), 1148–1162.

[36] Damodaran, A. (2024). Equity Risk Premiums: Determinants, Estimation, and Implications. *Stern School of Business Working Paper*, New York University.

[37] North American Association of State and Provincial Lotteries. (2024). U.S. Lottery Sales and Profits Report. `https://www.naspl.org/`

[38] Farleigh Dickinson University. (2024). FDU Poll: Online Betting Leads to Problems for Young Men.

[39] National Association of Realtors. (2024). Profile of Home Buyers and Sellers. `https://www.nar.realtor/`

[40] Clotfelter, C.T., & Cook, P.J. (1990). On the Economics of State Lotteries. *Journal of Economic Perspectives*, 4(4), 105–119.

[41] Calado, F., & Griffiths, M.D. (2016). Problem Gambling Worldwide: An Update and Systematic Review of Empirical Research (2000–2015). *Journal of Behavioral Addictions*, 5(4), 592–613.

[42] Parke, J., & Griffiths, M.D. (2007). The Role of Structural Characteristics in Gambling. In *Research and Measurement Issues in Gambling Studies* (pp. 211–243). Elsevier.

[43] Cardano, G., et al. (2021). Consumer Credit Card Debt and the Impact of High Interest Rates. *Federal Reserve Bank of Philadelphia Working Paper*, No. 21-35.

[44] Hayer, T., & Meyer, G. (2011). Self-Exclusion as a Harm Minimization Strategy: Evidence for the Casino Sector from Selected European Countries. *Journal of Gambling Studies*, 27, 685–700.

[45] Blaszczynski, A., Ladouceur, R., & Shaffer, H.J. (2004). A Science-Based Framework for Responsible Gambling: The Reno Model. *Journal of Gambling Studies*, 20(3), 301–317.

[46] Claude [Large language model]. (2025). Anthropic. `https://claude.ai/`

Code Appendix

Q1–Stochastic Disposable Income Model (using Python)

```
#Import necessary libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# Try to Load U.S. Expenditures data
try:
    df_exp_us = pd.read_csv('M3-Challenge-Problem-Data-2026.xlsx - Expenditures (U.S.).csv',
    skiprows=1)  #Attempting to load CSV data
except FileNotFoundError:
    pass # Bypass error if file is not in immediate directory

# Use the data points and change visual proportions
age_groups = ['<25', '25-34', '35-44', '45-54', '55-64', '65-74', '75+']
#Defining the x-axis age categories
mean_incomes = [48514, 102494, 128285, 141121, 121571, 75460, 56028]
#Setting raw mean gross income data points
total_expenditures = [41405, 67414, 82877, 88178, 76339, 59173, 50882]
#Setting raw total essential expenditure data

# Scaling of proportions
proportions = {
    '<25':   [0.15, 0.35, 0.20, 0.05, 0.10, 0.15],  #Percentage breakdown for under 25 group
    '25-34': [0.14, 0.34, 0.19, 0.06, 0.10, 0.17],  #Percentage breakdown for 25-34 age group
    '35-44': [0.14, 0.32, 0.18, 0.08, 0.10, 0.18],  #Percentage breakdown for 35-44 age group
    '45-54': [0.13, 0.30, 0.18, 0.10, 0.11, 0.18],  #Percentage breakdown for 45-54 age group
    '55-64': [0.12, 0.31, 0.17, 0.13, 0.11, 0.16],  #Percentage breakdown for 55-64 age group
    '65-74': [0.12, 0.32, 0.15, 0.18, 0.12, 0.11],  #Percentage breakdown for 65-74 age group
    '75+':   [0.11, 0.33, 0.12, 0.22, 0.13, 0.09]   #Percentage breakdown for 75+ age group
}

#Define categories and colors to be used in plot
categories = ['Food', 'Housing', 'Transportation', 'Healthcare', 'Utilities', 'Insurance']
#Listing the expenditure categories
colors = ['#48C9B0', '#5DADE2', '#F4D03F', '#E74C3C', '#AF7AC5', '#F5B041']
#Assigning specific hex colors for categories

exp_data = {cat: [] for cat in categories}
#Initializing dictionary to store calculated dollar values
for i, age in enumerate(age_groups):
#Iterating through each age group
    for j, cat in enumerate(categories):
    #Iterating through each spending category
        exp_data[cat].append(total_expenditures[i] * proportions[age][j])
```

```
        #Calculating dollar amount per category

# Define the side-by-side plot layout and styling
fig, ax1 = plt.subplots(figsize=(10, 6))
#Creating the main plot figure and primary y-axis for our plot

bottom_values = np.zeros(len(age_groups))
#Starting the stack baseline at zero for all groups
for idx, cat in enumerate(categories):
#Looping through categories to build the stacked bar
    ax1.bar(age_groups, exp_data[cat], bottom=bottom_values, label=cat, color=colors[idx]
    , width=0.6)
    #Plotting bar segment
    bottom_values += np.array(exp_data[cat])

    #Updating the baseline for the next category in the stack

ax1.set_ylabel('Annual Essential Expenditures (USD)')  #Setting primary y-axis label
ax1.set_xlabel('Age Group')  #Setting x-axis label
ax1.set_ylim(0, 150000)  #Setting primary y-axis vertical limits

ax2 = ax1.twinx()  #Creating a twin axis for the income line overlay
ax2.plot(age_groups, [inc / 1000 for inc in mean_incomes], color='black', marker='o'
, linewidth=2, markersize=8, label='Mean Gross Income')  #Plotting income in thousands
ax2.set_ylabel('Mean Gross Income ($, thousands)')  #Setting secondary y-axis label
ax2.set_ylim(0, 160)  #Setting secondary y-axis vertical limits

# Set title and format plot properly
plt.title('Essential Expenditures vs. Gross Income by Age Group (U.S.)', fontweight='bold')
#Setting the bolded plot title
lines_1, labels_1 = ax1.get_legend_handles_labels()

#Retrieving handles/labels from the bar chart
lines_2, labels_2 = ax2.get_legend_handles_labels()

#Retrieving handles/labels from the line chart
ax1.legend(lines_1 + lines_2, labels_1 + labels_2, loc='upper left', ncol=3)
#Combining legends into one 3-column box

ax1.grid(alpha=0.2)  #Adding faint grid lines for readability
plt.tight_layout()  #Optimizing layout to prevent label clipping

#Show plot on our display
plt.show()

--
```

```
#Import necessary libraries
import pandas as pd  #Importing pandas for data handling
import numpy as np  #Importing numpy for numerical arrays
import matplotlib.pyplot as plt  #Importing pyplot for main visualization
from matplotlib.patches import Patch  #Importing Patch for custom legend elements

# Load U.S. Expenditures data
try:
    df_exp_us = pd.read_csv('M3-Challenge-Problem-Data-2026.xlsx - Expenditures (U.S.).csv'
    , skiprows=1)  #Loading external CSV data
except FileNotFoundError:
    pass  #Proceeding if the specific local file is missing

# Use the data points and change visual proportions
age_groups = ['<25', '25-34', '35-44', '45-54', '55-64', '65-74', '75+']
#Defining age bracket categories
median_di = [-3745, 8292, 10034, 13062, 12094, -1695, -7074]
#Setting median disposable income values
p10_di = [-9200, 1800, 2500, 4100, 3600, -7400, -13500]
#Setting 10th percentile disposable income values

x = np.arange(len(age_groups))  #Creating an integer-based x-axis array
width = 0.35  #Setting the thickness of individual bars

fig, ax = plt.subplots(figsize=(10, 6))  #Initializing the figure and plot area

for i in range(len(age_groups)):  #Iterating through each age group to build bars
    # Add conditional bar coloring
    (Positive = Blue, Negative = Red, Lighter for 10th %ile)
    med_color = '#337AB7' if median_di[i] >= 0 else '#D9534F'
    #Selecting color based on positive/negative median
    ax.bar(x[i] - width/2, median_di[i], width, color=med_color, label='Median DI'
    if i == 1 else "")
    #Plotting median bar

    p10_color = '#A9CCE3' if p10_di[i] >= 0 else '#F5B7B1'
    #Selecting color based on positive/negative 10th percentile
    ax.bar(x[i] + width/2, p10_di[i], width, color=p10_color, label='10th
    %ile DI' if i == 1 else "")
    #Plotting 10th percentile bar

    # Add text labels with formatting
    offset = 300 if median_di[i] >= 0 else -800
    #Adjusting text position based on bar direction
    ax.text(x[i] - width/2, median_di[i] + offset, f"${median_di[i]:,}"
    , ha='center', fontsize=9, fontweight='bold')
    #Drawing dollar labels
```

```
# Formatting (DEBT ZONE shading, boundary at 0)
ax.axhline(0, color='black', linewidth=1)
#Drawing a solid black line at the zero equilibrium
ax.axhspan(-16000, 0, color='#FDEDEC', alpha=0.5, zorder=0)
#Shading the negative region to indicate debt
ax.text(-0.4, -11000, 'DEBT ZONE', color='#C0392B', fontweight='bold', alpha=0.5)
#Labeling the shaded debt region

ax.set_xticks(x)  #Setting the positions for x-axis ticks
ax.set_xticklabels(age_groups)  #Applying age group names to x-axis ticks
ax.set_ylabel('Disposable Income (USD)')  #Labeling the y-axis
ax.set_xlabel('Age Group')  #Labeling the x-axis
ax.set_title('Stochastic Disposable Income by Age Group\n(Negative values = baseline debt)'
, fontweight='bold')
#Setting plot title

# Construct legend of plot
legend_elements = [
    Patch(facecolor='#337AB7', label='Median DI (positive)'),
    #Legend item for positive median
    Patch(facecolor='#D9534F', label='Median DI (negative/debt)'),
    #Legend item for negative median
    Patch(facecolor='#A9CCE3', label='10th %ile DI (positive)'),
    #Legend item for positive 10th percentile
    Patch(facecolor='#F5B7B1', label='10th %ile DI (negative/debt)')
    #Legend item for negative 10th percentile
]
ax.legend(handles=legend_elements, loc='upper right')  #Rendering the custom legend box

ax.grid(alpha=0.3)  #Adding a light background grid
plt.tight_layout()  #Ensuring all elements fit within the figure window

#Show plot
plt.show()  #Displaying the final bar chart
```

Q2–Stochastic Gambling Outcome Model (using Python)

```
#Import necessary libraries
import numpy as np  #Importing numpy for mathematical operations
import matplotlib.pyplot as plt  #Importing matplotlib for chart generation

# Add paper parameters for to-be-made plot
h1 = 0.048  #Setting the initial house edge at 4.8% per leg
max_legs = 12  #Defining the maximum number of parlay legs to calculate
legs = np.arange(1, max_legs + 1)  #Creating an array from 1 to 12 legs

# Apply the exponential compounding formula (effective edge)
effective_edge = (1 - (1 - h1)**legs) * 100
```

```
#Calculating the cumulative house edge percentage

#Define subplots
fig, ax = plt.subplots(figsize=(10, 6))  #Initializing the plot figure and axis

# Assign colors for the transition from safe to dangerous visual
# 1-2 legs (Blue), 3-4 legs (Orange), 5+ legs (Red)
colors = ['#5A9BD5', '#5A9BD5', '#ED7D31', '#ED7D31'] + ['#C0504D'] * 8
#Creating a color list for risk levels

bars = ax.bar(legs, effective_edge, color=colors, edgecolor='white')
#Plotting the primary bar chart

# Add exponential trend line
ax.plot(legs, effective_edge, color='grey', linestyle='--', marker='o', alpha=0.5)
#Adding a dashed trend line with markers

# Add percentage labels on top of each bar
for bar in bars:  #Iterating through each bar object
    height = bar.get_height()
    #Retrieving the bar's y-value height
    ax.annotate(f'{height:.1f}%',
    #Formatting the height as a percentage string
                xy=(bar.get_x() + bar.get_width() / 2, height),
                #Positioning text at bar center
                xytext=(0, 3),  #Offsetting the text 3 points above the bar
                textcoords="offset points",
                #Specifying offset coordinate system
                ha='center', va='bottom', fontsize=9, fontweight='bold')
                #Setting text alignment and weight

# Reference lines for roulette edge (5.3%) and aggressive edge
ax.axhline(5.3, color='grey', linestyle=':', alpha=0.5)
#Drawing a reference line for standard roulette edge
ax.axhline(25.0, color='grey', linestyle=':', alpha=0.5)
#Drawing a reference line at the 25% threshold

# Add title and format plot
ax.set_title('Compounding House Edge: $h_{parlay}(n) = 1 - (1 - h_1)^n$\nSingle-leg edge
$h_1 = 4.8\%$', fontweight='bold')  #Setting the LaTeX title
ax.set_xlabel('Number of Parlay Legs (n)')  #Labeling the horizontal axis
ax.set_ylabel('Effective House Edge (%)')  #Labeling the vertical axis
ax.set_xticks(legs)  #Ensuring every leg number is shown on the x-axis
ax.set_ylim(0, 52)  #Setting the vertical range to accommodate high edges
ax.grid(axis='y', alpha=0.2)  #Adding light horizontal grid lines

plt.tight_layout()  #Optimizing plot margins and spacing
```

```
#Show plot
plt.show()  #Rendering the final exponential growth visualization

--

#Import necessary libraries
import numpy as np  #Importing numpy for random simulations and arrays
import matplotlib.pyplot as plt  #Importing matplotlib for histogram generation

# Define bet type parameters for the plot
bet_types = {
    'conservative': {'q': 0.476, 'm': 0.909, 'h': 0.048},
    #Setting win prob, multiplier, and edge for -110 odds
    'moderate':     {'q': 0.450, 'm': 1.000, 'h': 0.100},
    #Setting parameters for even-money style bets
    'aggressive':   {'q': 0.250, 'm': 3.000, 'h': 0.250}
    #Setting parameters for high-risk parlays
}

def simulate_year_chase(freq_wk, mean_bet, bet_type='moderate', gamma=1.5
, p_stay_chase=0.4, n_sims=10000):
    # Create Monte Carlo with Markov chain chase model
    np.random.seed(42)
    #Seeding the generator for perfect reproducibility
    p = bet_types[bet_type]
    #Selecting the probability dictionary based on risk type
    N = int(freq_wk * 52)
    #Calculating total number of bets over a 52-week year
    results = np.zeros(n_sims)
    #Initializing an array to store net gain for each simulation

    for i in range(n_sims):
    #Looping through 10,000 individual simulations
        L = 0.0  #Resetting cumulative profit/loss to zero for the individual
        chasing = False
        #Initializing the loss-chasing state as inactive
        for _ in range(N):  #Looping through each individual bet in the year
            lam = 1.0 / mean_bet  #Defining the base rate for exponential bet sizing
            if chasing:  #Checking if the gambler is currently in a "chase" state
                lam = lam / gamma  #Increasing the mean bet size by the gamma factor

            B = np.random.exponential(1.0 / lam)
            #Generating a random bet size using exponential distribution

            if np.random.random() < p['q']:
            #Determining if the bet is a win
                L += B * p['m']  #Adding the winnings to the cumulative total
                chasing = False  #Resetting the chase state after a successful win
```

```
            else:  #Handling the logic for a lost bet
                L -= B  #Subtracting the bet amount from the cumulative total
                chasing = np.random.random() < p_stay_chase
                #Determining if the gambler starts chasing
        results[i] = L  #Storing the final annual net result for the simulation
    return results  #Returning the full array of simulation outcomes

# Run simulations based on parameters: N=10,000, 2 bets/week, $25 mean stake
sim_cons = simulate_year_chase(2.0, 25.0, 'conservative')
#Simulating the low-risk profile
sim_mod  = simulate_year_chase(2.0, 25.0, 'moderate')
#Simulating the medium-risk profile
sim_agg  = simulate_year_chase(2.0, 25.0, 'aggressive')
#Simulating the high-risk parlay profile

fig, axs = plt.subplots(1, 3, figsize=(15, 5), sharey=True)
#Creating a three-panel side-by-side plot
sims = [sim_cons, sim_mod, sim_agg]  #Organizing simulation results into a list
titles = ['Conservative\n(Single Bets, -110)', 'Moderate\n(Even-Money Bets)',
'Aggressive\n(Parlays)']  #Defining subplot headers
colors = ['#6996D3', '#F2A968', '#D66A6A']
#Assigning blue, orange, and red to the risk levels

for i, ax in enumerate(axs):  #Iterating through the three subplots
    ax.hist(sims[i], bins=60, density=True, color=colors[i], edgecolor='white'
    , linewidth=0.5)
    #Plotting the result distribution
    ax.axvline(0, color='black', linestyle='--', linewidth=1.5)
    #Drawing a dashed line at the break-even point

    # Calculate text box stats
    mu = np.mean(sims[i])  #Calculating the average net gain/loss
    sigma = np.std(sims[i])  #Calculating the volatility (standard deviation)
    p_win = np.mean(sims[i] > 0) * 100  #Calculating the percentage of profitable outcomes

    # Add box of statistics
    textstr = f"$\mu$ = ${mu:,.0f}\n$\sigma$ = ${sigma:,.0f}\nP(win) = {p_win:.0f}%"
    #Formatting stats into a string
    props = dict(boxstyle='round', facecolor='white', alpha=0.8, edgecolor='lightgrey')
    #Defining the text box style
    ax.text(0.95, 0.95, textstr, transform=ax.transAxes, fontsize=9,

            verticalalignment='top', horizontalalignment='right', bbox=props)
            #Placing the stats box in the corner

    ax.set_title(titles[i], fontweight='bold')  #Applying the bolded risk-type title
    ax.set_xlabel('Net Annual Gain/Loss ($)')  #Labeling the horizontal gain/loss axis
    ax.grid(alpha=0.2)  #Adding a faint grid for density estimation
```

```
#Define labels, titles
axs[0].set_ylabel('Density')  #Labeling the vertical probability density axis
fig.suptitle('Monte Carlo Simulation: Annual Net Gain/Loss by Risk Type\n
(N=10,000 individuals, 2 bets/week, $25 mean stake, chase model active)', fontweight='bold')
#Setting main figure title
plt.tight_layout()  #Adjusting layout for optimal spacing

# Show plot
plt.show()  #Rendering the final histogram comparison

--

#Import necessary libraries
import pandas as pd  #Importing pandas for data frame operations
import numpy as np  #Importing numpy for numerical array handling
import matplotlib.pyplot as plt  #Importing matplotlib for bar chart creation

# Try to load personal betting data
try:
    df_betting = pd.read_csv('M3-Challenge-Problem-Data-2026.xlsx -
    Online Sports Betting Personal .csv')
    #Attempting to read demographic CSV
except FileNotFoundError:
    pass
    #Skipping if the specific file is not found in the path

# Use exact expected loss values to maintain fidelity
demographics = ['Male\n18-34', 'Male\n35-49', 'Male\n50-64', 'Male\n65+',
                'Female\n18-34', 'Female\n35-49', 'Female\n50-64', 'Female\n65+']
                #Defining labels for demographic groups
loss_standard = [955, 624, 234, 83, 325, 312, 104, 39]
#Setting baseline annual loss values per group
loss_chase = [1290, 843, 316, 112, 439, 422, 140, 53]
#Setting increased annual loss values for chasing behavior

x = np.arange(len(demographics))  #Creating the x-axis index positions
width = 0.35  #Setting the width for side-by-side bars

fig, ax = plt.subplots(figsize=(12, 6))
#Initializing the plot figure and axes object

rects1 = ax.bar(x - width/2, loss_standard, width, label='Standard Model', color='#4A88C1')
#Plotting the standard loss bars
rects2 = ax.bar(x + width/2, loss_chase, width,
label='Chase Model (+35%)', color='#D25B4C')
#Plotting the chase model bars
```

```
# Add text annotations with exact dollar amounts above the red bars
for bar in rects2:  #Iterating through the second set of bars for labeling
    height = bar.get_height()  #Getting the height of the current bar
    ax.annotate(f'${height:,}',  #Formatting the annotation as a currency string
                xy=(bar.get_x() + bar.get_width() / 2, height),
                #Positioning at top center of bar
                xytext=(0, 3), textcoords="offset points",
                #Offsetting text slightly above the bar top
                ha='center', va='bottom', fontweight='bold', fontsize=9)
                #Setting text style and alignment

# Labels and legends exact to original
ax.set_ylabel('Expected Annual Loss ($) per Active Bettor')
#Setting the vertical axis label
ax.set_xlabel('Demographic Group')

#Setting the horizontal axis label
ax.set_title('Expected Annual Sports Betting Loss by Demographic\n
(Moderate Risk, Standard vs. Chase Model)',
fontweight='bold')
#Setting the bolded title
ax.set_xticks(x)  #Defining the positions for category labels
ax.set_xticklabels(demographics)  #Applying the demographic names to the x-axis
ax.legend()  #Adding the legend to distinguish models
ax.grid(axis='y', alpha=0.3)  #Adding light horizontal grid lines
plt.tight_layout()  #Optimizing the spacing between plot elements
plt.show()  #Displaying the final comparative visualization

--

#Import necessary libraries
import numpy as np  #Importing numpy for numerical array structures
import matplotlib.pyplot as plt  #Importing matplotlib for horizontal bar plotting

# Parameters based on the Tornado Chart visual in the document
parameters = [
    'House Edge (h)',
    #Defining the primary mathematical edge parameter
    'Chase Multiplier ($\gamma$)',
    #Defining the impact of loss-chasing scaling
    'Bet Frequency (f)',
    #Defining how often bets are placed per week
    'Avg Bet Size (B)',
    #Defining the mean dollar amount per wager
    'Account Rate (p)',
    #Defining the rate of active user participation
    'Chase Persistence'
    #Defining the probability of continuing a chase
```

```
]

# Exact impact bounds derived from the sensitivity analysis text and visual
impact_minus_10 = [-660, -450, -380, -300, -220, -180]
#Setting loss changes for a 10% parameter decrease
impact_plus_10 = [710, 480, 400, 320, 230, 160]
#Setting loss changes for a 10% parameter increase

y_pos = np.arange(len(parameters))  #Creating vertical positions for each parameter label

fig, ax = plt.subplots(figsize=(10, 6))  #Initializing the figure and plot axes

# Plot the left-facing bars (-10% change) in blue
ax.barh(y_pos, impact_minus_10, align='center', color='#5A9BD5',
        edgecolor='white',
        label='-10% change')
        #Drawing the negative impact bars

# Plot the right-facing bars (+10% change) in red
ax.barh(y_pos, impact_plus_10, align='center',
color='#D66A6A',
        edgecolor='white',
        label='+10% change')
        #Drawing the positive impact bars

# Formatting to match the original graph exactly
ax.set_yticks(y_pos)  #Setting the location for the parameter labels
ax.set_yticklabels(parameters)  #Applying the parameter names to the y-axis
ax.invert_yaxis()  #Inverting the axis to display largest impact at the top
ax.set_xlabel('Change in Expected Annual Loss ($)')
#Labeling the horizontal impact axis
ax.set_title('Sensitivity Analysis: Tornado Chart for Q2 Model\n
($\pm10\%$ parameter perturbation)', fontweight='bold')
#Setting bolded title

# Add a strong center line at 0
ax.axvline(0, color='black', linewidth=1)  #Drawing a vertical baseline at zero impact
ax.grid(axis='x', alpha=0.3)  #Adding a light vertical grid for scale reference
ax.legend(loc='lower right')  #Placing the legend in the bottom right corner

plt.tight_layout()  #Optimizing plot margins for better readability

#Show plot
plt.show()  #Rendering the final tornado chart visualization
```

Q3–PFR and Wealth Erosion (using Python)

```
import pandas as pd  #Importing pandas for CSV data processing
import numpy as np  #Importing numpy for numerical arrays and calculations
import matplotlib.pyplot as plt  #Importing matplotlib for the primary visualization
from matplotlib.patches import Patch  #Importing Patch for custom legend symbols

# Load data with exact metadata constraints
try:
    df_exp_us = pd.read_csv('M3-Challenge-Problem-Data-2026.xlsx - Expenditures (U.S.).csv',
    skiprows=1)  #Loading US expenditure data
    df_exp_uk = pd.read_csv('M3-Challenge-Problem-Data-2026.xlsx - Expenditures (U.K.).csv',
    skiprows=3)  #Loading UK expenditure data
except FileNotFoundError:
    pass  #Proceeding if the specific Excel exports are not found locally

# Data explicitly from Tables 11 & 12
demographics = ['Male\n18-34', 'Male\n35-49', 'Male\n50-64', 'Male\n65+',
                'Female\n18-34', 'Female\n35-49', 'Female\n50-64', 'Female\n65+']
                #Defining the demographic grouping labels

us_pfr = [38.2, 22.1, 7.3, 100.0, 14.8, 11.6, 3.4, 100.0]
#Setting PFR percentages for United States groups
# U.K. data only exists for 4 demographic blocks in Table 12, padding with 0s
uk_pfr = [48.7, 62.3, 0, 0, 21.4, 31.2, 0, 0]
#Setting PFR percentages for United Kingdom groups

x = np.arange(len(demographics))  #Creating the x-axis index for the 8 categories
width = 0.4  #Setting individual bar thickness for the comparison

fig, ax = plt.subplots(figsize=(14, 7))
#Initializing the plot figure and axes

# Define colors based on risk thresholds exactly as the original report
def get_color(val):
    if val >= 30: return '#D9534F'
    #Assigning red to the dangerous risk threshold
    if val >= 15: return '#F0AD4E'
    #Assigning orange to the cautionary risk threshold
    if val > 0:   return '#5CB85C'
    #Assigning green to the manageable risk threshold
    return 'none'
    #Returning no color if the value is zero

# Plot U.S. Bars (Solid)
for i in range(len(us_pfr)):  #Iterating through each U.S. data point
    ax.bar(x[i] - width/2, us_pfr[i], width, color=get_color(us_pfr[i]), edgecolor='white')
    #Drawing the solid U.S. bar
    if us_pfr[i] < 100:
    #Checking if the bar is below the
```

```
    100% boundary
        ax.text(x[i] - width/2, us_pfr[i] + 1, f"{us_pfr[i]:.0f}%", ha='center',
        fontweight='bold', fontsize=9)
        #Labeling US bar height

# Plot U.K. Bars (Hatched)
for i in range(len(uk_pfr)):
#Iterating through each U.K. data point
    if uk_pfr[i] > 0:
    #Only plotting bars where U.K. data is available
        ax.bar(x[i] + width/2, uk_pfr[i], width,
        color=get_color(uk_pfr[i]),
               edgecolor='white', hatch='//', alpha=0.7)
               #Drawing the hatched U.K. comparison bar

# Custom text for the 100% boundary groups
ax.annotate('100% -
any loss\nexceeds threshold\n(negative DI)',
            xy=(3 - width/2, 100), xytext=(3 + 0.5, 85),
            arrowprops=dict(facecolor='black',
            arrowstyle='->'),
            bbox=dict(boxstyle='round,pad=0.3',
            facecolor='white', edgecolor='grey'), ha='center')
            #Adding detail on debt-zone PFR

# Add threshold lines and labels
ax.axhline(30, color='#D9534F', linestyle='--', alpha=0.5)
#Drawing the 30% dangerous limit line
ax.axhline(15, color='#F0AD4E', linestyle='--', alpha=0.5)
#Drawing the 15% caution limit line
ax.text(7.6, 31, 'DANGEROUS', color='#D9534F', fontweight='bold', fontsize=9)
#Adding label for high risk zone
ax.text(7.6, 16, 'CAUTIONARY', color='#F0AD4E', fontweight='bold', fontsize=9)
#Adding label for medium risk zone

# Formatting
ax.set_ylabel('Probability of Financial Ruin (%)')
#Labeling the vertical probability axis
ax.set_xlabel('Demographic Group')  #Labeling the horizontal demographic axis
ax.set_title('Probability of Financial Ruin (PFR) by Demographic Group\n$PFR = P(L_{sim}
> 0.10 \cdot DI)$ | U.S. vs. U.K.', fontweight='bold')  #Setting bolded title
ax.set_xticks(x)  #Setting tick marks for each category
ax.set_xticklabels(demographics)  #Applying demographic group text to ticks

legend_elements = [
    Patch(facecolor='#5CB85C', label='Manageable (<15%)'),
    #Legend item for green threshold
    Patch(facecolor='#F0AD4E', label='Cautionary (15-30%)'),
```

```
    #Legend item for orange threshold
    Patch(facecolor='#D9534F', label='Dangerous (>30%)'),
    #Legend item for red threshold
    Patch(facecolor='white', edgecolor='black', hatch='//', label='U.K. values')
    #Legend item for U.K. hatching
]
ax.legend(handles=legend_elements, loc='upper center', ncol=4)
#Rendering the multi-column legend
ax.grid(axis='y', alpha=0.2)
#Adding a faint horizontal grid

plt.tight_layout()  #Ensuring labels and bars are properly positioned
plt.show()  #Displaying the final comparative PFR chart

--

import numpy as np  #Importing numpy for numerical operations
import matplotlib.pyplot as plt  #Importing matplotlib for graphing

years = np.arange(0, 31)
#Creating an array for 30 years (0 to 30)
rate = 0.07
#Setting the annual investment return rate (7%)

#Preparing the arrays to store the future value year by year (cummulative)
fv_conservative = [0]  #Starting value for conservative model
fv_moderate = [0]  #Starting value for moderate model
fv_aggressive = [0]  #Starting value for aggressive model
fv_heavy = [0]  #Starting value for heavy bettor model
fv_dynamic = [0]  #Starting value for dynamic age-varying model

# Defining formulas
for t in range(1, 31):
#Looping through each year from 1 to 30

    #Constant loss model implementation
    fv_conservative.append(fv_conservative[-1]
    * (1 + rate) + 500)
    #Adding $500 yearly loss invested at 7%
    fv_moderate.append(fv_moderate[-1]
    * (1 + rate) + 1290)
    #Adding $1,290 yearly loss invested at 7%
    fv_aggressive.append(fv_aggressive[-1]
    * (1 + rate) + 3000)
    #Adding $3,000 yearly loss invested at 7%
    fv_heavy.append(fv_heavy[-1]
    * (1 + rate) + 6000)
    #Adding $6,000 yearly loss invested at 7%
```

```
    #Dynamic age-varying model logic based on the Wealth Erosion Results Table
    if t <= 12:
        dyn_loss = 1290
        #Ages 22-34 use $1,290 yearly loss
    elif t <= 27:
        dyn_loss = 843
        #Ages 35-49 use $843 yearly loss
    else:
        dyn_loss = 316
        #Ages 50-52 use $316 yearly loss

    fv_dynamic.append(fv_dynamic[-1] * (1 + rate) + dyn_loss)
    #Adding age-based yearly loss invested at 7%

fig, ax = plt.subplots(figsize=(10, 6))  #Creating the figure and axes

#Plot lines exact to original graph
ax.plot(years, np.array(fv_conservative)/1000,
color='#2ECC71', label='$500/yr (conservative)')
#Plotting conservative model
ax.plot(years, np.array(fv_moderate)/1000,
color='#3498DB', label='$1,290/yr (moderate male 18-34)')  #Plotting moderate model
ax.plot(years, np.array(fv_aggressive)/1000,
color='#F39C12', label='$3,000/yr (aggressive)')
#Plotting aggressive model
ax.plot(years, np.array(fv_heavy)/1000,
color='#C0392B', label='$6,000/yr (heavy bettor)')  #Plotting heavy bettor model

#Dynamic model uses a dashed line and shading to emphasize core
ax.plot(years, np.array(fv_dynamic)/1000, color='#2C3E50', linestyle='--', linewidth=2.5,
label='Dynamic model (age-varying)')  #Plotting dynamic model
ax.fill_between(years, 0, np.array(fv_dynamic)/1000, color='#EBF5FB', alpha=0.5)
#Shading area under dynamic model

ax.annotate('$131,600\n(dynamic, 30 yrs)',
#Adding annotation text
            xy=(30, fv_dynamic[-1]/1000),
            #Pointing to the final value at year 30
            xytext=(25, (fv_dynamic[-1]/1000) + 50),
            #Positioning the text slightly above
            arrowprops=dict(facecolor='#2C3E50', arrowstyle='->'), color='#2C3E50',
            fontweight='bold')  #Styling the arrow and text

ax.set_title('Wealth Erosion: Opportunity Cost of Gambling Losses\n(Invested at 7% Annual
Return)', fontweight='bold')  #Setting graph title
ax.set_xlabel('Years')  #Labeling x-axis
ax.set_ylabel('Cumulative Opportunity Cost ($, thousands)')
```

```
#Labeling y-axis
ax.set_xlim(0, 30)
#Setting x-axis limits
ax.set_ylim(0, 600)
#Setting y-axis limits
ax.grid(alpha=0.3)
#Adding light grid lines
ax.legend(loc='upper left')
#Adding legend to upper left

plt.tight_layout()  #Adjusting layout so labels fit properly
plt.show()  #Displaying the final graph
```