Prediction Markets, Bayes and Kalshi

Nicholas G. Polson
Booth School of Business, University of Chicago
5807 South Woodlawn Avenue, Chicago, Illinois 60637
ngp@chicagobooth.edu

August 7, 2026

Abstract

An event contract settles at one dollar or zero, so its price is a probability and the exchange is a market quoted in the currency of Bayes. We give a Bayesian account of trading such contracts, with the CFTC regulated exchange Kalshi as the leading case. Three ingredients are needed. First, a model whose latent parameter is low dimensional relative to the number of listed contracts, so that a few liquid quotes identify the whole board. Second, a posterior that carries parameter uncertainty into every derived contract rather than a plug in estimate. Third, a sizing rule. We show that Kelly betting on an exclusive family of contracts has an explicit water filling solution with a shadow price b that is strictly below one, so that contracts with negative naive expected value can enter the growth optimal portfolio as hedges, and that the optimal growth rate is $\sum_{i\in S} p_i \log(p_i/c_i) + (1-p_S)\log b$, which collapses to the Kullback-Leibler divergence $\mathrm{KL}(p\|c)$ when the book is fair. The Harville (1973) formula is treated as the canonical low dimensional generator of a combinatorially large contract set, with a Jensen argument showing that plug in exacta prices are biased downward. We recall the Skellam calibration of real time soccer odds of Feng et al. (2016) and the model free martingale diagnostics of Aldous (2013), whose halftime price and serious candidates principles give tests of a Kalshi board that require no model at all. Kalshi's fee $\theta\pi(1-\pi)$ with $\theta = 0.07$ creates a no trade band of width $2\theta\pi(1-\pi)$, at most 3.5 cents at the mid, and it changes which positions exist and not merely their size. An ϵ contamination prior that shrinks the posterior toward the price scales the Kelly stake by exactly $1-\epsilon$, which makes fractional Kelly a Bayes rule rather than a heuristic, and a simulation study shows that the shrinkage is worth roughly six times its cost once the trader's input is noisy rather than merely uncertain. For the maker side we derive the adverse selection spread on a binary contract, which is $4\mu p(1-p)$ for an informed order flow fraction μ and therefore has the same quadratic shape as the exchange fee, and we show that a log utility maker's inventory adjusted reservation price is $p_i/(\Lambda W_i)$, falling directly out of the same first order conditions. Three real episodes anchor the framework: the 2025 federal shutdown duration ladder, whose twelve rungs are reproduced to within a cent by a one parameter hazard; the 2024 election board and its cross venue incoherence; and the favourite longshot bias documented on Kalshi by Bürgi et al. (2025), under which contracts below ten cents lose more than sixty percent of the money staked.

Keywords: Prediction markets; Event contracts; Kalshi; Bayesian inference; Harville formula; Plackett-Luce; Skellam process; Kelly criterion; Martingales; e-values; Horseshoe prior; Favourite longshot bias.

MSC: 62C10, 62F15, 60G44, 91G80. **JEL:** C11, D84, G13.

1 Introduction

A Kalshi contract pays one dollar if a stated event occurs and nothing otherwise. If it trades at π, then π is the market's probability for the event, up to risk premia and the cost of carrying collateral.

The exchange is therefore an unusual object for a statistician: an order book quoted directly in the units of a posterior. As Aldous (2013) puts it, prediction markets are essentially the unique setting in which substantial real world data is interpretable as probabilities and as fluctuations of probabilities over time.

The trading problem is a comparison of two probability vectors. On one side is $\pi = (\pi_1, \ldots, \pi_n)$, the quoted prices on a family of contracts. On the other is $p = (p_1, \ldots, p_n)$, the posterior predictive probabilities given a model and an information set. Bayes supplies p, the exchange supplies π, and the Kelly criterion converts the gap into a position whose long run growth rate is the Kullback-Leibler divergence between them (Kelly, 1956; Cover and Thomas, 2006).

Three difficulties separate this picture from practice and they organise the paper.

Dimension. A market on the winner of a twenty runner field lists twenty contracts, but the space of finishing orders has 20! elements. A market on the consumer price index lists a dozen bins, but the underlying quantity is real valued. A market on a soccer match lists three outcomes, while the score matrix is infinite. In each case the listed contracts are a coarse projection of a latent object, and the useful inferential move is to run the projection backwards: recover a low dimensional parameter from the liquid quotes, then push it forward to price the entire board coherently. Harville (1973) performs this move for orderings and Feng et al. (2016) perform it for soccer scores through a Skellam process.

Uncertainty. The projection is estimated. A plug in Harville calculation that inserts point estimates into the exacta formula returns a number with no error bar, and the map is convex, so the number is biased. In our illustration the posterior interval on a single exacta probability spans nearly a factor of three.

Friction. Kalshi's published taker fee is $\theta C\pi(1-\pi)$ dollars on C contracts at price π, rounded up to the next cent, with $\theta = 0.07$ on general markets and a maker schedule near one quarter of that.[1] The quadratic shape is elegant, charging most where the market is least certain, but it is not small: a round trip at the mid costs three and a half cents on the dollar, which exceeds most statistical edges.

Contributions. (i) We formalise event contract families as projections of a latent parameter, with coherence constraints (exclusivity, nesting, monotonicity) imposed as prior support restrictions rather than as post hoc screens. (ii) We give the Bayesian treatment of Harville orderings, including the exponential racing derivation, the Stern gamma family in which Harville is the memoryless boundary case, and a Jensen argument showing the direction of plug in bias. (iii) We solve the Kelly problem on an exclusive family with an overround and a quadratic fee in closed form, obtaining a water filling rule whose inclusion threshold is a shadow price below one, and we record the resulting growth rate. (iv) We show that an ϵ contamination prior toward the market price scales the Kelly stake by $1-\epsilon$, that the resulting wealth process is an e-process for the efficiency null, and that Aldous's serious candidates principle is the same optional sampling argument applied to a whole board. (v) We treat the maker's problem, deriving the binary contract adverse selection spread and the inventory adjusted reservation price, and note that the exchange's quadratic fee has the same shape as the adverse selection it taxes, with the two equal at an informed flow fraction of 3.5 percent. (vi) We give the multi board problem with a budget constraint, in which the per board shadow price is replaced by a common one, and the quadratic expansion that connects growth optimality to mean variance optimisation with unit risk aversion. (vii) We ground the framework in three documented Kalshi episodes and report a simulation study separating the cost of parameter uncertainty from the cost of a noisy input.

[1] Kalshi fee schedule, `https://kalshi.com/docs/kalshi-fee-schedule.pdf`. The July 2026 revision retains the coefficient 0.07 and introduces a category multiplier defaulting to one. Whelan (2024) studies the equilibrium price distortion induced by this schedule.

Related work. Ordering models are due to Harville (1973), Henery (1981), Stern (1990) and Lo and Bacon-Shone (2008). The favourite longshot bias is documented by Ali (1977) and reviewed by Thaler and Ziemba (1988). Prediction markets as aggregators are surveyed by Wolfers and Zitzewitz (2004), with the interpretation caveats of Manski (2006) and the combinatorial designs of Hanson (2003). Aldous (2013, 2012) develops the martingale consequences we use in Section 3. Growth optimal betting begins with Kelly (1956) and Breiman (1961); see MacLean et al. (2011). The betting formulation of testing is due to Shafer (2021), Shafer and Vovk (2019), Grünwald et al. (2024) and Vovk and Wang (2021). Calibration is the criterion of Dawid (1982) and Gneiting and Raftery (2007). Dynamic sports models include Stern (1994), Glickman and Stern (1998), Polson and Stern (2015) and Feng et al. (2016).

2 Event Contracts and the Probability Simplex

2.1 Mechanics and the cost of trading

Let A resolve at a known future date against a named public source. A Yes contract pays $\mathbf{1}_A$; the paired No contract pays $\mathbf{1}_{A^c}$. Since together they pay one dollar with certainty, $\pi^{\text{yes}} + \pi^{\text{no}} = 1$ up to the tick and the spread. Prices live on $\{0.01, \dots, 0.99\}$, which is coarse at the extremes: an event of true probability 1/500 cannot be quoted at its value.

Write $c(\pi)$ for the all in cost of one Yes contract quoted at π,

$$c(\pi) \;=\; \pi + \theta\,\pi(1-\pi), \qquad \theta \in \{0.07 \text{ taker},\ 0.0175 \text{ maker}\}. \tag{1}$$

The fee is symmetric across the two sides, so the No side costs $(1-\pi) + \theta\pi(1-\pi)$ and the two costs sum to $1 + 2\theta\pi(1-\pi)$. The round turn friction is therefore

$$\text{band}(\pi) \;=\; 2\theta\pi(1-\pi) \;\le\; \tfrac{\theta}{2} \;=\; 3.5 \text{ cents}. \tag{2}$$

Quote π	0.05	0.10	0.20	0.30	0.50	0.70	0.90	0.95
Taker fee, cents per contract	0.33	0.63	1.12	1.47	1.75	1.47	0.63	0.33
All in Yes cost $c(\pi)$	0.0533	0.1063	0.2112	0.3147	0.5175	0.7147	0.9063	0.9533
No trade half band $\theta\pi(1-\pi)$	0.0033	0.0063	0.0112	0.0147	0.0175	0.0147	0.0063	0.0033
Fee as percent of stake	6.6	6.3	5.6	4.9	3.5	2.1	0.7	0.4

Table 1: Kalshi taker fee $\theta\pi(1-\pi)$ at $\theta = 0.07$, before rounding up to the next cent. The half band is the smallest absolute deviation of the posterior from the quote at which a taker order has positive expected value. The last row is the one that matters: measured against capital deployed, the fee is worst on longshots, which is exactly where posterior models most often disagree with the market.

Table 1 records the schedule. Per contract the fee peaks at the mid, but as a fraction of capital at risk it peaks at low prices, reaching 6.6 percent at five cents and, once the rounding to a whole cent binds, one hundred percent at one cent. A strategy that systematically buys longshots pays roughly twice the rate of one trading near the mid.

2.2 Families, ladders and coherence

Contracts appear in structured families.

- *Exclusive and exhaustive*, such as the bins of an inflation print or the target range after an FOMC meeting. Then $\sum_{i=1}^n \pi_i = 1 + v$ with $v \ge 0$ the overround, the analogue of the track take.

- *Nested*, such as thresholds. If $A_x = \{Y \geq x\}$ then $x \mapsto \pi(A_x)$ must be non increasing, since the prices trace a survival function.
- *Winner families*, such as a tournament field, where $\sum_i \pi_i = 1 + v$ and where auxiliary contracts on reaching a final or finishing top four impose ordering constraints of the Harville type.

Each structure restricts the set $\mathcal{C}$ of admissible price vectors. It is natural to build the restriction into the model rather than test for it afterwards. Specifying $p = p(\vartheta)$ with $p(\vartheta) \in \mathcal{C}$ for every ϑ makes the posterior predictive coherent by construction. This is the operational content of the claim that Bayes helps: the exchange quotes each contract separately and need not be coherent, whereas a posterior derived from one latent parameter cannot fail to be.

2.3 Removing the overround

Prices on an exclusive family sum to $1 + v > 1$, so they are not probabilities and must be normalised before comparison with a posterior. The choice of normalisation is itself a modelling decision, and the three standard ones embed different assumptions about where the margin sits.

Proportional. Set $\tilde{\pi}_i = \pi_i / B$ with $B = \sum_j \pi_j$. This assumes the margin is levied uniformly in probability, which is the assumption of no favourite longshot bias.

Power, or odds ratio. Set $\tilde{\pi}_i \propto \pi_i^{1/\gamma}$ with γ solving $\sum_i \pi_i^{1/\gamma} = 1$. Values $\gamma < 1$ shrink long prices more than short ones, which is the shape the empirical literature reports.

Shin. Shin (1993) derives the margin endogenously from a bookmaker facing a proportion z of insiders, giving

$$\tilde{\pi}_i \;=\; \frac{\sqrt{z^2 + 4(1-z)\pi_i^2/B} \;-\; z}{2(1-z)}, \tag{3}$$

with z chosen so the $\tilde{\pi}_i$ sum to one. Here z is interpretable as the incidence of informed trading, and (3) reduces to the proportional rule as $z \to 0$.

Quote π_i	0.300	0.250	0.180	0.120	0.110	0.080
Proportional	0.2885	0.2404	0.1731	0.1154	0.1058	0.0769
Power, $\gamma = 0.977$	0.2915	0.2419	0.1728	0.1141	0.1044	0.0753
Shin, $z = 0.008$	0.2913	0.2421	0.1732	0.1142	0.1043	0.0748

Table 2: Three normalisations of the six contract board of Section 11, whose quotes sum to 1.04. Power and Shin agree closely and both raise the favourite and cut the longshot relative to the proportional rule. The gap at the short end, 0.0769 against 0.0750, is about two cents, which is of the same order as the fee and therefore large enough to decide whether a longshot trade exists.

Table 2 applies all three to the board of Section 11. The disagreement between them is concentrated exactly where the fee is proportionally most punitive, at the short prices, so the de-vigging convention is not a cosmetic preprocessing step. It is part of the model, and a Bayesian should carry a prior over γ or z rather than fix one by convention.

3 Martingale Diagnostics for a Board

Before any model is fitted, one can ask whether the board behaves as the theory of conditional expectation requires. Let $\mathcal{F}_t$ be the information available at t and $p_t = \mathbb{P}(A \mid \mathcal{F}_t)$. By the tower property p_t is a martingale converging to $\mathbf{1}_A$. Aldous (2013) observes that if market prices are consensus probabilities then they too should be martingales, and that this hypothesis has sharp, model free, testable consequences. We record three, since each is directly checkable on Kalshi data and none requires the machinery of Sections 4 to 7.

Proposition 1 (Halftime price principle)**.** *Consider a contest between equally matched opponents decided by the sign of* Z_1+Z_2*, where* Z_1 *and* Z_2 *are the point differences in the two halves, independent and identically distributed with continuous distribution function* F *symmetric about zero. Then the price at halftime is uniformly distributed on* $(0,1)$*.*

Proof. $\mathbb{P}(Z_1+Z_2>0 \mid Z_1=z)=\mathbb{P}(Z_2>-z)=F(z)$ by independence and symmetry, so the halftime price is $F(Z_1)$, and $F(Z_1)\sim \mathrm{U}(0,1)$ by the probability integral transform. □

Proposition 2 (Serious candidates principle)**.** *Consider a contest with several entrants, exactly one of which wins, whose prices are continuous path martingales all starting below* $b\in(0,1)$*. Let* N_b *be the number of entrants whose price ever reaches* b*. Then* $\mathbb{E}[N_b]=1/b$*. Under the fee schedule* (1) *the strategy that buys one contract on each entrant when its price first reaches b has expected profit* $1-c(b)\mathbb{E}[N_b]<0$*, so the market free bound* $\mathbb{E}[N_b]\le 1/c(b)$ *is the tradeable version.*

Proof. Buy one contract on each entrant when its price first reaches b. Each purchase is at a stopping time, so by optional sampling the aggregate is a fair bet. The outlay is bN_b. The eventual winner's price converges to one and therefore crosses b, so exactly one held contract pays one dollar. The gain is $1-bN_b$ and has expectation zero, giving $\mathbb{E}[N_b]=1/b$. With fees the outlay is $c(b)N_b$ and the gain $1-c(b)N_b$ has non positive expectation under the null. □

Proposition 3 (Downcrossings)**.** *Under the conditions of Proposition 2 with* $0<a<b<1$*, the expected total number of downcrossings of* $[a,b]$*, summed over entrants, is* $(1-b)/(b-a)$*.*

Proof. Buy at b, sell at a, repeat. Exactly one contract expires at one and each of the $D_{a,b}$ others is sold at a, so the gain is $(1-b)-D_{a,b}(b-a)$, which has expectation zero. □

Aldous (2013) finds all three consistent with data: halftime prices from thirty baseball games track the uniform closely, and for the 2012 Republican nomination the observed counts crossing $b=1/3$ and $b=1/5$ were three and five against theoretical values of three and five, with mild excess at lower thresholds. The lesson for a Kalshi trader is that a board can be tested before it is modelled, and that an excess of entrants crossing a threshold is a direct, calibrated signal that the field is being collectively overpriced.

Proposition 2 also foreshadows the argument of Section 10. It is a betting strategy whose expected gain must vanish under the null, which is precisely the construction of an e-variable.

4 The Bayesian Machinery

4.1 Predictive probabilities are what get traded

Let ϑ be a parameter, $\mathcal{D}$ the data, A the contract event. The object to compare with the price is the posterior predictive

$$p \;=\; \mathbb{P}(A \mid \mathcal{D}) \;=\; \int \mathbb{P}(A \mid \vartheta)\,\pi(\vartheta \mid \mathcal{D})\,d\vartheta, \tag{4}$$

not the plug in $\mathbb{P}(A \mid \hat{\vartheta})$. The two differ whenever $\vartheta \mapsto \mathbb{P}(A \mid \vartheta)$ is non linear, which is the rule for threshold and ordering events, and the direction is systematic.

Lemma 1 (Plug in bias)**.** *If* $h(\vartheta)=\mathbb{P}(A \mid \vartheta)$ *is convex on the support of the posterior then* $\mathbb{P}(A \mid \mathcal{D})\ge \mathbb{P}(A \mid \bar{\vartheta})$ *where* $\bar{\vartheta}=\mathbb{E}[\vartheta \mid \mathcal{D}]$*, with strict inequality unless the posterior is degenerate.*

Proof. Jensen's inequality applied to (4). □

Two consequences. For a tail event $A = \{Y \geq x\}$ with x well beyond the centre of the predictive distribution, the survival function is convex in the location parameter, so integrating over parameter uncertainty fattens the tail and the plug in under prices longshots. This compounds rather than corrects the favourite longshot bias already present in market prices. For the Harville exacta map $h(p) = p_i p_j/(1 - p_i)$ we have $\partial^2 h/\partial p_i^2 = 2p_j/(1 - p_i)^3 > 0$, so the same conclusion applies to ordering contracts, and the bias grows cubically as $p_i \to 1$.

4.2 Shrinkage priors for strengths

Contract families are generated by a vector of latent strengths. For a field of n entrants write $w_i = \exp(\eta_i)$ with

$$\eta_i = \mu + \beta^\top x_i + \delta_i, \qquad \delta_i \mid \lambda_i, \tau \sim N(0, \lambda_i^2 \tau^2), \qquad \lambda_i \sim C^+(0, 1), \quad \tau \sim C^+(0, 1), \tag{5}$$

the horseshoe prior of Carvalho et al. (2010). The structural argument for a global-local prior is that fields are sparse in signal: most entrants sit close to what the covariates x_i predict, and a few are genuinely different, either because of information the covariates miss or because the market has misjudged them. The horseshoe leaves those few unshrunk while pulling the rest onto the regression surface, which is what one wants when the alternative is to overfit a handful of recent results. For logistic and multinomial logistic likelihoods, as in Bradley-Terry (Bradley and Terry, 1952) and Plackett-Luce (Plackett, 1975), Polya-Gamma augmentation (Polson et al., 2013) restores conditional conjugacy and the Gibbs sampler is routine.

4.3 Calibration and scoring rules

A posterior is worth trading only if it is better than the price in a sense that compounds. The relevant criterion is a strictly proper scoring rule, and the logarithmic rule is the one that connects to growth. Writing I for the realised outcome, the Murphy decomposition of the expected score (Murphy, 1973; Gneiting and Raftery, 2007) separates it into uncertainty, resolution and reliability. Uncertainty is a property of the event and not of the forecaster. Reliability measures calibration and is zero for an honest posterior. Resolution measures the sharpness that remains, and it is the only term a trader can improve.

Two practical consequences follow. First, a model can beat the market on reliability while losing on resolution, and such a model is worthless: it is well calibrated and uninformative, which the market already is. Second, posterior predictive checks against realised outcomes should be run per contract family rather than pooled, because a board can be well calibrated on average while systematically wrong on longshots, which is precisely the bias documented in Section 7.

4.4 Edge detection as testing by betting

Whether there is an edge at all is a testing question, and the betting formulation makes it the same object as the trading question. Let H_0 be the hypothesis that the quoted price is the true probability. A wager returning R times its stake is an e-variable for H_0 if $\mathbb{E}_{H_0}[R] \leq 1$; a product of such wagers is an e-process, and its running value is an anytime valid measure of evidence (Shafer, 2021; Vovk and Wang, 2021; Grünwald et al., 2024). Under H_0 the trader's bankroll multiple is exactly such a process. Ville's inequality gives $\mathbb{P}(\sup_t W_t \geq 1/\alpha) \leq \alpha$, so doubling one's money is evidence at level $e = 2$ and a twentyfold increase is $\alpha = 0.05$. This is a sobering calibration: a record that triples capital is not, by itself, distinguishable from luck.

5 Orderings: the Harville Formula

5.1 Statement and exact derivation

Let the field be $\{1,\dots,n\}$ with strengths $w_i > 0$ and win probabilities $p_i = w_i / \sum_j w_j$. Harville (1973) proposed

$$\mathbb{P}(i,j) \;=\; p_i \,\frac{p_j}{1-p_i}, \qquad \mathbb{P}(i,j,k) \;=\; p_i \,\frac{p_j}{1-p_i}\,\frac{p_k}{1-p_i-p_j} \tag{6}$$

for the exacta and trifecta. The rule is sequential sampling without replacement from the win probabilities: remove the winner, renormalise, repeat. Its axiomatic content is Luce's independence of irrelevant alternatives (Luce, 1959) and it coincides with the Plackett-Luce ranking model (Plackett, 1975).

Lemma 2 (Exponential racing)**.** *If completion times $T_i \sim \mathrm{Exp}(w_i)$ are independent, then the induced law on finishing orders is exactly* (6)*.*

Proof. $\mathbb{P}(T_i = \min_j T_j) = w_i / \sum_j w_j = p_i$. By the memoryless property, conditional on the identity and time of the first finisher, the residual times of the remainder are again independent exponentials with the same rates, so the argument iterates over the reduced field with renormalised probabilities. □

Harville's formula is therefore not an approximation. Under exponential running times it is the exact ordering law, and the entire content of the model is memorylessness.

5.2 Why it fails, and two corrections

Memorylessness says that, conditional on the favourite not winning, the race among the remainder is unchanged. Empirically a strong entrant that fails to win still tends to finish near the front, so the true conditional law favours it for second place more than (6) allows. Harville consequently overstates the probability that a longshot places and understates the place probabilities of favourites.

Running time models. Henery (1981) takes $T_i = \theta_i + \varepsilon_i$ with normal errors and Stern (1990) takes $T_i \sim \mathrm{Gamma}(r, w_i)$, in which Harville is the case $r = 1$ and the normal model is the limit $r \to \infty$. The shape r indexes exactly the departure from memorylessness. Ordering probabilities require an $(n-1)$ dimensional integral, which is a simulation problem: draw T, sort, average indicators. With 10^5 draws the full trifecta table is accurate to three decimals in seconds.

Discounted Harville. Lo and Bacon-Shone (2008) propose the one parameter fix

$$\mathbb{P}(i,j) \;=\; p_i \,\frac{p_j^{\lambda}}{\sum_{k \neq i} p_k^{\lambda}}, \qquad \lambda \approx 0.8, \tag{7}$$

which flattens the second place distribution toward uniform, nests Harville at $\lambda = 1$, and has λ estimable from historical place frequencies.

5.3 Posterior propagation

The plug in use of (6) substitutes $\hat{p}$ and reports one number. The Bayesian version treats p as random: given a posterior on the strength vector, whether from the Dirichlet approximation $p \mid \mathcal{D} \sim \mathrm{Dir}(\kappa \hat{p})$ or from a horseshoe Bradley-Terry fit, draw $p^{(g)}$ for $g = 1,\dots,G$ and average $h_A(p^{(g)})$ over the Harville, Stern or discounted map. The output is a posterior over every ordering contract on the board.

The correction is not academic. In the six entrant field of Section 11 take $\hat{p}_A = 0.34$ and $\hat{p}_B = 0.24$. The plug in exacta for A over B is 0.1236. Under $\mathrm{Dir}(\kappa\hat{p})$ with $\kappa = 50$, roughly the information in fifty comparable events, the posterior mean is 0.1235 with standard deviation 0.0373 and ninety percent interval $[0.069, 0.190]$. Against a quote of eleven cents the plug in reports a two cent edge, while the posterior reports $\mathbb{P}(\text{value} > 0.11) = 0.61$, close to a coin flip. At $\kappa = 200$ the standard deviation halves to 0.0187 and that probability rises to 0.76. Figure 1 plots the two posteriors. Uncertainty is first order for exotics in a way it is not for win contracts, because the Harville map is a ratio and hence convex, and the plug in value sits below the posterior mean by Lemma 1.

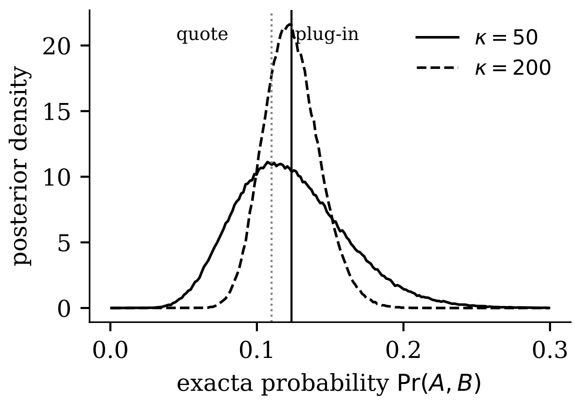

Figure 1: Posterior density of the exacta probability $\Pr(A, B)$ under $p \mid \mathcal{D} \sim \text{Dir}(\kappa\hat{p})$ for $\kappa = 50$ and $\kappa = 200$, against the plug in value 0.1236 and a market quote of 0.11. Quadrupling the effective sample size halves the spread but does not move the centre.

6 Skellam Dynamics and the EPL Odds Market

The closest antecedent to the Kalshi problem is Feng et al. (2016), which supplies the template for recovering a latent generator from a handful of quotes.

Let X_t and Y_t be goals scored by the home and away teams, independent Poisson processes with rates λ_A and λ_B. The score difference $N_t = X_t - Y_t$ is a Skellam process (Skellam, 1946) with

$$\mathbb{P}(N_t = k) \;=\; e^{-(\lambda_A+\lambda_B)t} \left(\frac{\lambda_A}{\lambda_B}\right)^{k/2} I_{|k|}\left(2t\sqrt{\lambda_A\lambda_B}\right), \tag{8}$$

I_k the modified Bessel function of the first kind. The three quoted outcomes are $\mathbb{P}(N_1 > 0)$, $\mathbb{P}(N_1 = 0)$ and $\mathbb{P}(N_1 < 0)$, which after removing the overround give two independent equations in the two unknowns (λ_A, λ_B). Calibration is therefore exactly identified, and once the rates are known (8) prices the whole matrix of correct score contracts. Three liquid quotes generate a full joint law over scores.

Two further features transfer. First, the implied volatility of a game, defined through the diffusion limit of the score difference in the manner of Stern (1994) and Polson and Stern (2015), gives a scalar summary of remaining uncertainty. Second, the model is re-estimated as the game evolves: after a goal or a red card the repricing is fed back through (8) to new rates and a new score matrix. Feng et al. (2016) illustrate this on Everton against West Ham in March 2016 and validate calibration on 1520 matches from 2012 to 2016, finding observed frequencies close to market implied probabilities.

The link to Section 3 is exact and worth stating. Aldous's discrete counterexample to the halftime price principle is a soccer model with independent Poisson scoring at rates $\lambda_1 = \lambda_2$, which is the Skellam setting of (8) with $t = 1/2$. The halftime price is then supported on the few values of F at integer score differences, and its distribution function is a staircase around the uniform diagonal rather than the diagonal itself. Discreteness of the latent count, not any failure of the martingale hypothesis, produces the deviation. For a Kalshi live sports board this is the practical warning: uniformity of the halftime price is a good diagnostic for high scoring sports and a biased one for low scoring sports, where the Skellam model tells you what the correct null distribution actually is.

For live in game contracts the right object is a filtering problem: a latent state, a jump process of scoring events, and a sequence of posteriors, for which sequential Monte Carlo is the natural tool. For high scoring sports the Brownian approximation of Stern (1994) is preferable, with the state space extension of Glickman and Stern (1998) across a season.

7 The Favourite Longshot Bias

Across a century of parimutuel and bookmaker data, low probability outcomes trade above their frequency and high probability outcomes trade below it (Ali, 1977; Thaler and Ziemba, 1988; Ottaviani and Sørensen, 2008). A convenient one parameter description of the distortion is a compression on the logit scale,

$$\operatorname{logit}\pi \;=\; \alpha + \beta \operatorname{logit} p, \qquad \beta \in (0,1), \tag{9}$$

so that quoted probabilities are pulled toward a fixed point. Inverting gives the Bayes corrected probability $p = \operatorname{logit}^{-1}\{(\operatorname{logit}\pi - \alpha)/\beta\}$, and the position implied by (9) has a clean structure.

Proposition 4 (Single crossing)**.** *Under (9) with $\beta < 1$ there is a unique pivot $p^* = \operatorname{logit}^{-1}\{\alpha/(1-\beta)\}$ such that $\pi > p$ for $p < p^*$ and $\pi < p$ for $p > p^*$. The gross Kelly position is therefore short every contract below the pivot and long every contract above it.*

Proof. $\pi < p$ if and only if $\operatorname{logit}\pi < \operatorname{logit} p$, that is $\alpha + \beta \operatorname{logit} p < \operatorname{logit} p$, that is $\operatorname{logit} p > \alpha/(1-\beta)$, since $1-\beta > 0$. □

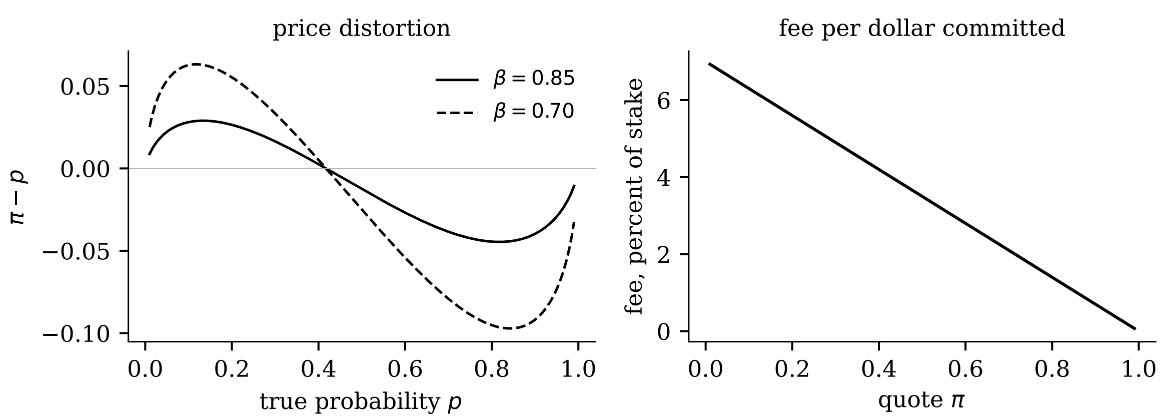

Figure 2: Left: the price distortion $\pi - p$ implied by (9) for two values of β, with the pivot held at $p^* = 0.417$. Longshots are overpriced and favourites underpriced, with the effect concentrated in the tails. Right: the Kalshi taker fee as a percentage of capital committed, $100\,\theta(1-\pi)$, which rises monotonically as the quote falls. The two panels work against each other: the largest apparent edges sit exactly where the fee is proportionally worst.

7.1 Evidence from Kalshi

The bias is documented on Kalshi itself. Bürgi et al. (2025) assemble transaction level data on 46,282 contracts drawn from 12,403 distinct events through April 2025, giving over 300,000 prices once both sides of each trade are counted. Their findings are three. Prices are informative and sharpen as expiry approaches. Contracts priced under ten cents lose more than sixty percent of the money staked on them. Contracts priced above fifty cents earn a small but statistically significant positive return. Makers outperform takers, but both sides show the same pattern across the price range.

Two summary points calibrate (9). Reading the sub ten cent bucket at a representative price of five cents with a gross return of -60 percent gives $p = 0.02$ at $\pi = 0.05$; reading the upper range at seventy cents with a return of two percent gives $p = 0.714$ at $\pi = 0.70$. Solving (9) through these two points yields $\beta = 0.79$ and $\alpha = 0.13$, hence a pivot at $p^* = 0.64$.

Quote π	0.03	0.05	0.10	0.30	0.50	0.70	0.90
Implied true p	0.010	0.020	0.050	0.226	0.460	0.714	0.933
Gross return $p/\pi - 1$	-0.66	-0.60	-0.50	-0.25	-0.08	$+0.02$	$+0.04$

Table 3: The logit compression (9) at $\beta = 0.79$, $\alpha = 0.13$, fitted to two summary points from Bürgi et al. (2025). The implied pivot at 0.64 sits above the fifty cent level at which their data first shows positive returns, so this two point calibration overstates the bias in the middle of the range; a serious fit would use the full bucket level table and carry a posterior on (α, β) rather than a point estimate.

The gap deserves stating plainly. A pivot at 0.64 implies that contracts between fifty and sixty four cents lose money, whereas Bürgi et al. (2025) report positive returns above fifty cents. The two point fit is therefore too aggressive in the middle, which is exactly what one expects when a two parameter family is pinned by its tails. It is reported here to make the calibration reproducible, not because the parameters should be trusted.

Take the five cent contract at face value. With $p = 0.02$, selling costs $c' = 0.9533$ after the fee, the Kelly stake is $f^\star = 0.5715$ of bankroll, and the growth rate is 0.0101 nats against 0.0121 without the fee. The trader commits fifty seven percent of capital to earn one percent per event, in a position that is wiped out by a two percent event. Log utility is content with this. Almost no institution would be.

Figure 2 shows why the bias is harder to harvest on an exchange than the size of the distortion suggests. The natural trade is to sell longshots, and the fee on the No side of a cheap contract is small in cents. But the capital committed is large. Take $\pi = 0.08$ against a posterior $p = 0.05$. Without fees the growth rate is 0.0070 nats. Buying No costs $c' = 0.9252$ after the fee, the Kelly stake is $f^\star = 0.332$ of bankroll, and the growth rate falls to 0.0050 nats. The trader ties up a third of his capital to earn half a percent per event, and does so in a position that loses everything staked on a five percent event. Growth optimality is indifferent to this asymmetry, since log utility already prices it, but a trader with any liability constraint is not.

The deeper point is that (9) is a statement about the market, not about the world, so α and β are parameters to be estimated with a posterior like any others. A trader who fixes $\beta = 0.85$ from published sports betting estimates and applies it to a weather board or an inflation ladder has imported a bias correction from a market with a different microstructure, a different clientele, and a different fee schedule.

8 Where Bayesian Logic Helps on Kalshi

Five settings, chosen because in each the exchange lists many contracts driven by few latent quantities.

8.1 Inflation ladders

Kalshi lists inflation prints as ladders of bins or thresholds. Let Y be the release and specify $Y \mid \vartheta \sim F_\vartheta$, for instance a skew normal centred on a nowcast with dispersion calibrated to the historical distribution of surprises. The bin $[a_i, a_{i+1})$ has model price $F_\vartheta(a_{i+1}) - F_\vartheta(a_i)$.

Three gains follow. Only two or three bins near consensus carry liquidity, and the model prices the rest consistently rather than leaving them stale. The ladder must be monotone, since $\{Y \geq a_{i+1}\} \subset \{Y \geq a_i\}$, and monotonicity holds automatically under any F_ϑ but need not hold across the quotes. And the fitted dispersion is the market's implied uncertainty, the analogue of implied volatility.

Example 1 (A violation that is not a trade)**.** Suppose the board quotes $\mathbb{P}(Y \geq 3.1) = 0.58$ and $\mathbb{P}(Y \geq 3.2) = 0.61$. This is incoherent, and selling the second while buying the first collects three cents and can never lose. But the taker fees are $0.07 \times 0.61 \times 0.39 = 1.67$ cents and $0.07 \times 0.58 \times 0.42 = 1.71$ cents, or 3.38 cents together. The arbitrage is destroyed by the schedule. A violation must exceed $\theta[\pi_1(1-\pi_1) + \pi_2(1-\pi_2)]$, about three and a half cents near the mid and nine tenths of a cent for a maker on both legs.

The 2025 shutdown ladder of Section 9 is a worked instance of this structure on live quotes. This is the reason to model rather than to screen. Incoherence screens fire constantly on a real board and almost none of what they find is tradeable. What survives is disagreement between the model and the liquid quotes, not disagreement among the illiquid ones.

8.2 Policy rate paths

The rate board lists, for each meeting, an exclusive family of target ranges, and separately a family on the level at year end. Marginals do not determine the joint law of the path, but a latent factor model does. Write $r_m = r_{m-1} + \Delta_m$ with Δ_m on the grid of twenty five basis point increments and

$$\mathbb{P}(\Delta_m = d \mid \mathcal{F}) \;=\; \frac{\exp\{\gamma_d + \beta_d^\top z_m\}}{\sum_{d'} \exp\{\gamma_{d'} + \beta_{d'}^\top z_m\}}, \tag{10}$$

where z_m collects a policy rule gap, an inertia term in Δ_{m-1}, and a latent hawkishness factor. Fitting to the near dated meetings, which are liquid, prices the far dated ones and the terminal contract jointly. The consistency requirement is that the model implied law of r_M under the fitted path must reproduce the quoted year end family. Systematic discrepancy identifies whether the board's implied path dependence is too weak or too strong, a question on which the individual marginal contracts are silent.

8.3 Daily temperature

Weather is the cleanest case because a good prior is public. Kalshi lists daily maxima at named stations in bins of one or two degrees. Numerical weather prediction ensembles supply forecasts $f_1, \ldots, f_K$, and the Bayesian model averaging calibration of Raftery et al. (2005) takes the predictive density as

$$g(y \mid f_1, \ldots, f_K) \;=\; \sum_{k=1}^{K} \omega_k \, \phi\big(y \mid a_k + b_k f_k, \sigma^2\big)\,, \tag{11}$$

with weights and bias corrections estimated on a rolling window at that station. Raw ensembles are underdispersed, so calibration is not cosmetic; it is what turns a forecast into a probability comparable with a price. Intraday, the observed path is further data and the posterior for the daily maximum conditions on the running maximum and the hours of insolation remaining. Settlement is against a specific station reading, so model risk is confined to the physics rather than the contract definition. This is where a modest repeatable edge is most plausible, and correspondingly where liquidity is thinnest.

8.4 Tournament and award markets

A tournament winner market is a Plackett-Luce ranking market in disguise. Kalshi lists one Yes contract per entrant, often with auxiliary contracts on reaching a final, winning a division, or making a shortlist. Section 5 supplies the coherence conditions. If R_i is the event that i reaches the final and W_i that i wins, then $W_i \subset R_i$, so $\pi(W_i) \leq \pi(R_i)$ and the ratio is the market's implied conditional probability of winning the final. Harville, or better a Stern gamma model, links the two through one strength vector.

The gain is that a single posterior prices the whole board at once. Draw $w^{(g)}$, simulate the bracket, record the frequency of every listed event: winner, finalist, division, top four, head to head. The overround is then visible as the gap between $\sum_i \pi_i$ and one, and mispricings are the residuals of π against p after normalisation. Award markets have the same structure with a far weaker likelihood, since precursor awards are the only data, so the prior does more of the work and shrinkage matters more.

8.5 Correlated elections

State level contracts are marginals of a highly correlated joint law. A one factor probit is the minimal honest model,

$$\mathbb{P}(\text{state } s \text{ flips} \mid Z) = \Phi(a_s + b_s Z), \qquad Z \sim N(0,1), \tag{12}$$

with Z a national swing factor. The state contracts identify the marginals $\Phi\left(a_s/\sqrt{1+b_s^2}\right)$ and a national or seat count contract identifies the loading scale. The joint law is quoted nowhere, so any contract depending on more than one state, including tipping point and majority contracts, must be priced by the model. Correlation misspecification dominates here: independence across states makes a national landslide nearly impossible and radically underprices the tails. A prior on b_s centred on historical covariance with an inflation factor for regime change is the appropriate hedge against exactly the error behind the largest prediction market dislocations of recent cycles.

9 Three Episodes on Kalshi

The framework above is abstract. Three episodes from the exchange make it concrete.

9.1 The 2025 shutdown ladder

The United States federal government shut down on 1 October 2025 and reopened on 12 November, a span of forty three days and the longest on record. Kalshi listed the duration as a ladder of nested threshold contracts, of the form "will the shutdown last more than d days", which is precisely the structure of Section 8.1: the quotes trace a discrete survival function $S(d) = \mathbb{P}(D > d)$, and monotonicity in d is a coherence constraint, not an empirical finding.

On 31 October, with the shutdown in its thirty first day, three rungs traded at roughly

$$S(45) = 0.56, \qquad S(46) = 0.51, \qquad S(60) = 0.18.$$

Rather than read these as three independent opinions, fit a hazard. Let h be the daily probability of resolution conditional on the shutdown still running, so that $S(d) = S(45)(1-h)^{d-45}$ for $d \geq 45$. The rungs at 45 and 60 give $h = 1 - (0.18/0.56)^{1/15} = 0.0729$, and the fitted value at the intermediate rung is $S(46) = 0.519$ against a quote of 0.51. A single parameter reproduces the board to within a cent.

Three lessons follow, one for each part of this paper.

Threshold d (days)	45	46	50	55	60
Quoted $S(d)$	0.56	0.51	—	—	0.18
Constant hazard fit, $h = 0.0729$	0.560	0.519	0.384	0.263	0.180

Table 4: The shutdown duration ladder on 31 October 2025 and a one parameter geometric hazard fitted to the outer two rungs. The fit prices the unquoted interior rungs and predicts the traded rung at 46 days to within a cent.

Dimension. A ladder of a dozen contracts collapses to one number. The fitted h says the market regarded the political process as memoryless once past day forty five, with an expected residual duration of $1/h \approx 13.7$ further days. That is a statement a trader can hold a view on. A vector of twelve prices is not.

Uncertainty. A Bayesian would put a prior on h, or better allow a time varying hazard, since a memoryless process is a strong claim about a negotiation with known deadlines. The posterior on h then propagates into every rung at once, including the many that were never quoted.

Settlement. The rung "more than 43 days" was defined as the shutdown remaining active at ten in the morning Eastern time on 13 November. The funding bill was signed the previous evening, so that rung resolved No and the effective duration for the board was forty three days exactly. The contract turned on a definitional cutoff measured in hours. No amount of hazard modelling addresses that, which is the content of the remark in Section 15 that settlement risk is not statistical risk.

Ex post, Kalshi traders' mean forecast during the episode was about 41.6 days against a realised 43. The board was slightly short on the mean while pricing a fifty six percent chance of exceeding forty five days a fortnight before the end. Both statements are true, and they illustrate why directional accuracy is a poor summary of a probabilistic forecast.

9.2 The 2024 election board

Kalshi listed federal election contracts for the first time in autumn 2024 after a court ruling against the CFTC's objection. The episode is usually recalled as a success, since the markets favoured the eventual winner while polling aggregates were close to a tie. The systematic evidence is more mixed and more useful.

Clinton and Huang (2025) analyse more than 2,500 political markets across the Iowa Electronic Markets, Kalshi, PredictIt and Polymarket over the final five weeks of the campaign, covering more than two billion dollars of transactions. Directional accuracy on Kalshi was about seventy eight percent. Prices for identical contracts diverged across venues, daily price changes were weakly or negatively autocorrelated, and apparent arbitrage widened rather than narrowed in the final fortnight.

Each finding maps onto something above. Directional accuracy is close to meaningless as a criterion: a perfectly calibrated forecaster whose average confidence is seventy eight percent is wrong twenty two percent of the time by construction, which is why Section 4.4 insists on proper scoring rules and the resolution term rather than a hit rate. Negative autocorrelation in daily changes is a direct violation of the martingale hypothesis of Section 3, and it is testable by the downcrossing count of Proposition 3 without reference to any model. Cross venue divergence is an incoherence in the sense of Section 2.2, and by Proposition 10 it is monetisable only by a participant able to trade both venues, since a divergence invisible to the contracts one can actually buy contributes nothing to $g^\star$.

9.3 Winner families and awards

The data structure in Bürgi et al. (2025) is itself instructive: 46,282 contracts nested inside 12,403 events, where an event is a question such as which nominee wins a given acting award and the contracts are the individual nominees. This is exactly the winner family of Section 8.4. Each event carries an overround $\sum_i \pi_i - 1$, each contract is one coordinate of a Plackett-Luce strength vector, and any auxiliary contract on a shortlist or a category sweep is an ordering event in the sense of Section 5.

Awards are also the case in which the prior does the most work. The likelihood consists of a handful of precursor results, correlated with each other and with the target through voting bodies that overlap. A horseshoe Bradley-Terry fit with precursor indicators as covariates is a reasonable model; a maximum likelihood fit on six observations is not. The practical test of whether the shrinkage was set correctly is Proposition 9: compare the model's log score against the board's, averaged over resolved events, and require the difference to exceed the fee band before treating it as a strategy.

10 Kelly Sizing

10.1 A single contract

Stake a fraction f of bankroll on Yes at all in cost $c = c(\pi)$. Each dollar returns $1/c$ if A occurs and zero otherwise, so expected log growth is

$$g(f) \;=\; p \log\left(1 - f + \frac{f}{c}\right) + (1-p)\log(1-f). \tag{13}$$

Proposition 5 (Binary Kelly with fees)**.** *The growth optimal stake is $f^\star = (p-c)/(1-c)$, positive only if $p > c$. The no trade region is $\pi - \theta\pi(1-\pi) \le p \le \pi + \theta\pi(1-\pi)$, of width $2\theta\pi(1-\pi)$. For a small edge,*

$$g(f^\star) \;=\; \mathrm{KL}\big(p \,\|\, c\big) \;\approx\; \frac{(p-c)^2}{2c(1-c)}. \tag{14}$$

Proof. Write $W_1 = 1 + f(1-c)/c$ and $W_0 = 1 - f$. Then $g'(f) = p(1-c)/(cW_1) - (1-p)/W_0$, and setting this to zero and clearing denominators gives $p(1-c)(1-f) = (1-p)\{c + f(1-c)\}$, that is $p - c = f(1-c)$. Hence $f^\star = (p-c)/(1-c)$, at which $W_1 = p/c$ and $W_0 = (1-p)/(1-c)$, so

$$g(f^\star) = p\log\frac{p}{c} + (1-p)\log\frac{1-p}{1-c} = \mathrm{KL}(p\|c).$$

The stated approximation is the chi square expansion of the divergence about $p = c$. □

The formula reads as edge over the price of the losing side. At $\pi = 0.30$ with $\theta = 0.07$ we have $c = 0.3147$, so $p = 0.34$ gives $f^\star = 0.0369$, under four percent of bankroll for a four cent raw edge. Figure 3 plots the stake and the growth rate as functions of the posterior. Expression (14) is the useful summary: growth is one half the squared standardised edge, so the fee band $|p - \pi| > \theta\pi(1-\pi)$ requires a standardised edge exceeding $\theta\sqrt{\pi(1-\pi)}$, which is 0.035 at the mid, and the gross growth forgone at the boundary is $\theta^2\pi(1-\pi)/2 \approx 6 \times 10^{-4}$ nats.

10.2 Parameter uncertainty, and what actually changes the stake

Suppose p is unknown with posterior Π. Two candidate rules present themselves: plug in $\bar p = \mathbb{E}[p \mid \mathcal{D}]$, or maximise expected growth under the posterior. They agree, and rather more generally than the binary case.

Proposition 6 (Bayesian Kelly)**.** *Let the outcome space be finite with wealth multipliers $W_i(f)$ that do not depend on p. Then $g_p(f) = \sum_i p_i \log W_i(f)$ is affine in p, so $\mathbb{E}_\Pi[g_p(f)] = g_{\bar p}(f)$ for every f and the growth optimal portfolio under posterior uncertainty is the plug in portfolio at the posterior mean.*

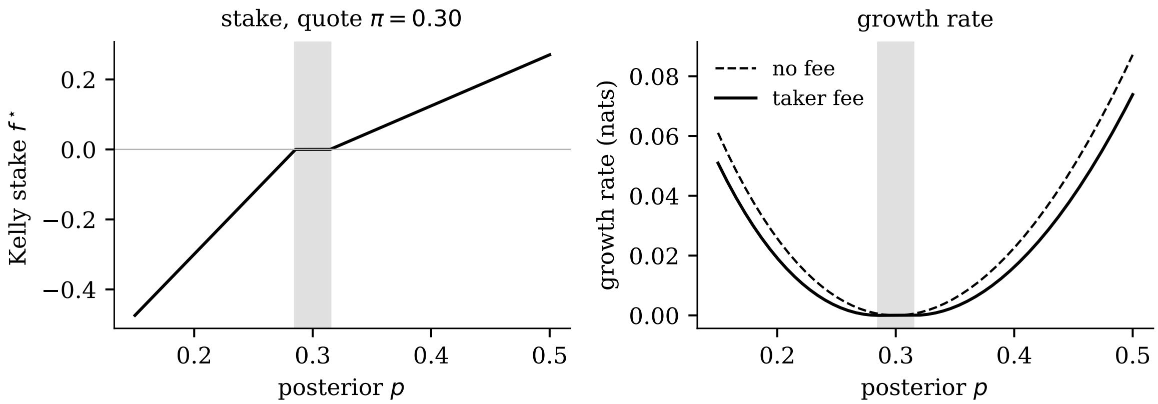

Figure 3: Left: the growth optimal stake against the posterior p at a quote of $\pi = 0.30$, positive for Yes and negative for No. Right: the corresponding growth rate with and without the taker fee. The shaded strip is the no trade band $[\pi - \theta\pi(1-\pi),\ \pi + \theta\pi(1-\pi)]$, of width 2.9 cents at this quote. The fee does not merely lower the curve, it flattens a neighbourhood of the quote to zero.

The result is stronger than it first appears. Because log is additive across periods and because the trader observes the outcome regardless of how he bet, the sequential Bayesian problem separates: maximising $\mathbb{E}_t[\log W_t]$ period by period at the current posterior mean is dynamically optimal, and there is no experimentation premium. Neither parameter uncertainty nor learning is a reason to bet less.

What does change the stake is misspecification, which is not parameter uncertainty at all. Suppose the trader entertains an ϵ contamination (Berger, 1985): with probability $1-\epsilon$ his model is right and with probability ϵ the market is. His operative predictive is $p_\epsilon = (1-\epsilon)\bar{p} + \epsilon\pi$.

Proposition 7 (Fractional Kelly as robust Bayes)**.** *Under ϵ contamination and ignoring fees, $p_\epsilon - \pi = (1-\epsilon)(\bar{p} - \pi)$, hence*

$$f_\epsilon^\star = \frac{p_\epsilon - \pi}{1-\pi} = (1-\epsilon)\, f^\star, \tag{15}$$

and the optimal growth rate falls by the fraction ϵ^2 relative to full Kelly. With fees the stake is $\{(1-\epsilon)(\bar{p}-\pi) - \theta\pi(1-\pi)\}/(1-c)$, so contamination both scales the position and widens the no trade band.

Fractional Kelly is therefore not a prudential haircut. Half Kelly is the exact position of a trader who gives even odds to the proposition that his edge is illusory, and given the e-value calibration of Section 4.3, under which a threefold gain is barely evidence, $\epsilon = 1/2$ is not conservative.

The genuine limitations of Proposition 6 lie elsewhere: early liquidation at a random price makes W_i depend on the future price path rather than on the terminal event alone; market impact makes c a function of size, so the multipliers depend on f non linearly; and the relevant bankroll is rarely the account balance.

10.3 An exclusive family

Take n exclusive and exhaustive outcomes with all in costs c_i and posterior probabilities p_i. Let $f_i \geq 0$ be the fraction staked on i and $f_0 = 1 - \sum_i f_i$ the reserve. If i occurs, wealth is multiplied by $W_i = f_0 + f_i/c_i$, so

$$\max_{f\geq 0,\ \sum_i f_i \leq 1}\ \sum_{i=1}^{n} p_i \log\left(f_0 + \frac{f_i}{c_i}\right). \tag{16}$$

Proposition 8 (Water filling)**.** *Order the outcomes by p_i/c_i decreasing. The solution to* (16) *is supported on a prefix S of that order, with*

$$f_i = p_i - b\,c_i \quad (i \in S), \qquad f_0 = b = \frac{1-p_S}{1-c_S}, \qquad p_S = \sum_{i\in S} p_i, \quad c_S = \sum_{i\in S} c_i, \tag{17}$$

where S is the smallest prefix with $f_i > 0$ for all $i \in S$ and $p_j/c_j \le b$ for all $j \notin S$. The wealth multipliers are $W_i = p_i/c_i$ on S and $W_i = b$ off it, and the optimal growth rate is

$$g^\star \;=\; \sum_{i\in S} p_i \log\frac{p_i}{c_i} \;+\; (1-p_S)\log b. \tag{18}$$

Proof. Write $W_i = f_0 + f_i/c_i$ with $f_0 = 1 - \sum_j f_j$, so $\partial W_i/\partial f_i = 1/c_i - 1$ and $\partial W_j/\partial f_i = -1$ for $j \ne i$. Then $\partial g/\partial f_i = p_i/(c_iW_i) - S^\dagger$ where $S^\dagger = \sum_j p_j/W_j$. The Karush-Kuhn-Tucker conditions give $p_i/(c_iW_i) = S^\dagger$ for $f_i > 0$ and $p_i/(c_if_0) \le S^\dagger$ otherwise, hence $W_i = p_i/(c_iS^\dagger)$ on S and $W_i = f_0$ off it. Substituting into the definition of $S^\dagger$ gives $S^\dagger = c_S S^\dagger + (1-p_S)/f_0$, and the budget identity $\sum_{i\in S}(p_i/S^\dagger - c_if_0) = 1 - f_0$ then forces $S^\dagger = 1$. Hence $f_0 = (1-p_S)/(1-c_S) = b$, $f_i = p_i - bc_i$, and $W_i = p_i/c_i$ on S. Substituting into the objective gives (18). The exclusion condition $p_j/(c_jf_0) \le 1$ is $p_j/c_j \le b$. Concavity of the objective makes the stationary point the global maximum. □

Two consequences deserve emphasis.

First, when the book is fair, $\sum_i c_i = 1$, then $b = 1$, S is everything, $f_i = p_i$, and (18) reduces to

$$g^\star \;=\; \sum_i p_i \log\frac{p_i}{c_i} \;=\; \mathrm{KL}(p\,\|\,c). \tag{19}$$

Bet your beliefs, and the growth rate is the divergence of the posterior from the price, measured in nats per event (Kelly, 1956; Cover and Thomas, 2006). This is the cleanest statement of why a Bayesian belongs in an event contract market: the compensation for a better calibrated posterior is exactly its distance from the consensus. The general formula (18) is the overround adjusted version, in which the trader also earns $(1-p_S)\log b < 0$, the cost of the states he has chosen not to cover.

Second, the inclusion threshold is b, not one. Since $b < 1$ whenever the trader has any edge, contracts with $p_i < c_i$, that is with negative naive expected value, can enter the optimal portfolio. They raise the wealth multiplier in states where the concentrated positions lose, and log utility pays for that insurance. Screening contract by contract on positive expected value is not the growth optimal rule.

10.4 Growth as a scoring rule advantage

Identity (19) has an interpretation that makes the value of a posterior directly measurable.

Proposition 9 (Growth is log score advantage)**.** *On a fair exclusive board, the optimal growth rate equals the expected difference in logarithmic score between the posterior and the price,*

$$g^\star \;=\; \mathbb{E}_p[\log p_I - \log c_I]\,, \qquad I \sim p. \tag{20}$$

Proof. Immediate from (19), since $\mathbb{E}_p[\log p_I - \log c_I] = \sum_i p_i \log(p_i/c_i)$. □

The practical value is that the right hand side is estimable from historical forecasts without ever placing a trade. A backtest of the model's log score against the board's log score, averaged over resolved contracts, is an unbiased estimate of the growth rate the strategy would have achieved, and it requires no assumptions about execution. A model whose log score advantage is smaller than the fee band is not a strategy.

10.5 Incomplete boards

An exclusive family is a complete market on its outcome space, which is why (19) is available. Most interesting boards are incomplete: the exchange lists marginals, such as individual states in an election or individual meetings in a rate path, but not the joint events a model has views on. Let Ω be the underlying state space, let the traded contracts be indicators $\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_m}$ with all in costs $c_1, \dots, c_m$, and let

$$\mathcal{M} \;=\; \{q \in \Delta(\Omega) : q(A_k) = c_k,\; k = 1, \dots, m\} \tag{21}$$

be the set of probability measures consistent with the quotes. On an exclusive family $\mathcal{M}$ is a single point; in general it is a convex polytope.

Proposition 10 (Duality). *The maximal growth rate attainable by trading the listed contracts is*

$$g^\star \;=\; \min_{q \in \mathcal{M}} \mathrm{KL}(p \,\|\, q), \tag{22}$$

and the minimiser q^* is the I-projection of p onto $\mathcal{M}$, an exponential tilt of p by a linear combination of the traded indicators.

This is the log optimal case of the general minimal distance martingale measure duality (Goll and Rüschendorf, 2001). Three readings are useful. Growth is the divergence to the nearest consistent measure rather than to any particular one, so a model can disagree violently with the market about the joint law and earn nothing if the disagreement is invisible to the traded marginals. The optimal q^* has exponential family form, so the traded contracts are the sufficient statistics of the trade. And listing a new contract can only weaken the constraint set, hence weakly increase $g^\star$: every additional listed event is a potential increase in achievable growth, which is the precise sense in which combinatorial market design (Hanson, 2003) adds value to an informed participant.

For the correlated election model of Section 8.5 this says that a view on the national correlation factor Z is monetisable only to the extent that the board lists a contract whose payoff depends on more than one state.

10.6 Drawdown and evidence

In the diffusion approximation, staking a fraction $\rho \in (0, 1]$ of full Kelly gives log wealth with drift $(\rho - \rho^2/2)\Lambda^2$ and volatility $\rho\Lambda$, where Λ is the Sharpe ratio of the full Kelly portfolio. Brownian first passage then gives

$$\mathbb{P}\left(\inf_t W_t \le x\right) \;=\; x^{(2-\rho)/\rho}, \qquad 0 < x < 1. \tag{23}$$

Full Kelly carries a one half chance of ever halving the bankroll, half Kelly $0.5^3 = 12.5$ percent, quarter Kelly $0.5^7 = 0.8$ percent, while retaining 100, 75 and 44 percent of the maximum growth rate respectively. By Proposition 7 the choice of ρ is a statement about ϵ, and both are statements about trust in the model.

Under the null that the market is correct, $\mathbb{E}[W_i] \le 1$ with equality when fees vanish, so W_t is a non negative supermartingale started at one and Ville's inequality applies. The bankroll multiple is an anytime valid e-value against market efficiency (Shafer and Vovk, 2019; Shafer, 2021), and Proposition 2 is the same construction applied across a board rather than through time.

11 A Worked Example

Consider a six entrant winner market, equally a race, a tournament field, or an award shortlist. The board quotes $\pi = (0.30, 0.25, 0.18, 0.12, 0.11, 0.08)$ with $\sum_i \pi_i = 1.04$, a four percent overround. A horseshoe Bradley-Terry posterior returns $p = (0.38, 0.22, 0.19, 0.09, 0.08, 0.04)$.

Entrant	A	B	C	D	E	F
Quote π_i	0.300	0.250	0.180	0.120	0.110	0.080
All in cost c_i	0.3147	0.2631	0.1903	0.1274	0.1169	0.0852
Posterior p_i	0.380	0.220	0.190	0.090	0.080	0.040
Ratio p_i/c_i	1.208	0.836	0.998	0.706	0.684	0.469
Stake f_i, no fee ($b = 0.778$)	0.1467	0.0256	0.0500	0	0	0
Stake f_i, with fee ($b = 0.869$)	0.1066	0	0.0247	0	0	0
Multiplier W_i, with fee	1.208	0.869	0.998	0.869	0.869	0.869

Table 5: Growth optimal allocation on a six contract exclusive family. Without fees the shadow price is $b = 0.778$ and three contracts are held, including B whose ratio 0.836 is below one. With the taker fee the shadow price rises to $b = 0.869$, B drops out, and deployed capital falls from 22.2 to 13.1 percent of bankroll. Multipliers satisfy $W_i = p_i/c_i$ on the support and $W_i = b$ off it, as Proposition 8 requires.

Growth rates from (18) are $g^\star = 0.0192$ nats per event without fees and 0.0108 with them. Fees consume forty four percent of the growth rate while leaving the raw edge untouched, which is the most important number in the table. Over 250 independent events, full Kelly compounds by $\exp(250 \times 0.0108) \approx 14.9$ and half Kelly by $\exp(250 \times 0.0082) \approx 7.8$, with the drawdown profile improved as above. Both figures assume the posterior is correct at every event, which is what Proposition 7 exists to discipline.

Contract B in the fee free row repays attention. Its ratio is 0.836, so it loses money in expectation, yet it enters because the threshold is $b = 0.778$ rather than one: the trader is heavily long A and needs wealth in the states where A loses. Once the fee lifts the threshold to 0.869 the hedge is no longer worth its cost. Fee schedules do not merely scale positions, they change which positions exist.

On the ordering contracts, the Harville map applied to the posterior gives $\mathbb{P}(A, B) = 0.124$, $\mathbb{P}(B, A) = 0.107$, $\mathbb{P}(A, C) = 0.088$, and trifectas $\mathbb{P}(A, B, C) = 0.050$ against $\mathbb{P}(C, B, A) = 0.028$. The discounted correction (7) at $\lambda = 0.81$ moves $\mathbb{P}(A, B)$ to 0.113, a nine percent reduction, and raises the chance that a favourite finishes behind a longshot. On a board quoted in one cent ticks with a one and a half cent fee, a nine percent model disagreement is roughly the entire tradeable signal. The choice between Harville and its corrections is not a refinement, it is the trade.

12 Making Rather Than Taking

Every calculation so far has assumed the trader crosses the spread. Kalshi's maker fee is one quarter of the taker fee, $\theta_M = 0.0175$, so the no trade half band shrinks from 1.75 cents at the mid to 0.44 cents. Since Bürgi et al. (2025) find that makers earn more than takers on Kalshi at every point of the price range, the maker's problem deserves its own treatment. It is not free money: the maker is paid the spread and charged adverse selection.

12.1 Adverse selection on a binary contract

Adapt the sequential trade model of Glosten and Milgrom (1985). A single contract on event A has consensus probability p. A fraction μ of arriving orders come from traders who know whether A will occur; the remaining $1 - \mu$ are uninformed and buy or sell with equal probability. The informed buy Yes if and only if A is true. A competitive maker quotes prices equal to his posterior given the direction of the order.

Proposition 11 (Adverse selection spread)**.** *Write $B = (1-\mu)/2$. The competitive ask and bid are*

$$a \;=\; \frac{p(1+\mu)/2}{B + p\mu}, \qquad b \;=\; \frac{pB}{B + (1-p)\mu}, \tag{24}$$

so that

$$a - p \;=\; \frac{\mu\, p(1-p)}{B + p\mu}, \qquad p - b \;=\; \frac{\mu\, p(1-p)}{B + (1-p)\mu}, \tag{25}$$

and for small μ the spread is $a - b \approx 4\mu\, p(1-p)$.

Proof. $\mathbb{P}(\text{buy} \mid A) = \mu + (1-\mu)/2 = (1+\mu)/2$ and $\mathbb{P}(\text{buy} \mid A^c) = (1-\mu)/2 = B$. Bayes gives $a = \mathbb{P}(A \mid \text{buy}) = p(1+\mu)/2 \,/\, \{p(1+\mu)/2 + (1-p)B\}$, and the denominator simplifies to $B + p\mu$. Subtracting p and collecting terms gives $a - p = p\{(1+\mu)/2 - B - p\mu\}/(B + p\mu) = \mu p(1-p)/(B + p\mu)$. The bid is symmetric. As $\mu \to 0$ both denominators tend to $1/2$. $\square$

The shape is the point. Adverse selection costs the maker $4\mu p(1-p)$ per round turn and the exchange charges the taker $2\theta\pi(1-\pi)$. Both are quadratic in the price with a maximum at the mid, which means the fee schedule mimics the economics of the market it taxes rather than distorting them. Equating the two at $p = 1/2$ gives $\mu = \theta/2 = 0.035$: a maker breaks even against the taker's fee burden when three and a half percent of arriving order flow is informed. Below that the maker's edge exceeds the taker's cost, which is consistent with the maker premium in the data.

12.2 Inventory and the reservation price

A maker already holding a position should not quote around p. The correct reference is the price at which he is indifferent to one more contract, and the Karush-Kuhn-Tucker conditions of Proposition 8 supply it directly.

Proposition 12 (Reservation price)**.** *Let a log utility trader hold stakes producing wealth multipliers $W_1, \ldots, W_n$ across the outcomes of an exclusive family, and write $\Lambda = \sum_j p_j/W_j$. He is indifferent to a marginal contract on outcome i at cost*

$$c_i^{\mathrm{res}} \;=\; \frac{p_i}{\Lambda\, W_i}. \tag{26}$$

At zero inventory $W_i \equiv 1$, $\Lambda = 1$ and $c_i^{\mathrm{res}} = p_i$: quote the posterior. Long exposure to i raises W_i and lowers c_i^{res}, so the maker shades his bid away from a position he already holds.

Proof. From the proof of Proposition 8, $\partial g/\partial f_i = p_i/(c_i W_i) - \Lambda$, which vanishes at $c_i = p_i/(\Lambda W_i)$. $\square$

This is the binary contract analogue of the inventory adjusted reservation price of Avellaneda and Stoikov (2008), derived here without an approximation because the outcome space is finite and log utility is exactly solvable. A maker running a full board therefore has a vector of reservation prices, each tilted by his exposure to that particular outcome, and the tilt is largest exactly where his wealth multiplier is largest. It is also the reason a maker on a winner family cannot quote each contract independently: the reservation price for contract i depends on Λ, which depends on the positions in all the others.

13 Many Boards at Once

A trader does not face one board. Kalshi lists thousands of contracts across weather, macroeconomics, politics and sport, most of them resolving independently.

13.1 Growth is additive, capital is not

For independent events, expected log growth adds:

$$g_{\text{total}} \;=\; \sum_{k=1}^{K} g_k. \tag{27}$$

Nothing in Proposition 8 couples the boards through the objective. The coupling is the budget. Solving each board in isolation gives stakes $f^{(k)}$ with total commitment $\sum_k \|f^{(k)}\|_1$, and this exceeds one as soon as K is moderate. In the worked example of Section 11 each board absorbs 13.1 percent of bankroll, so roughly seven simultaneous boards of that quality saturate the account.

Proposition 13 (Common shadow price)**.** *With independent boards and a single budget $\sum_k \|f^{(k)}\|_1 \le 1$, the optimum solves each board's problem with its own b_k replaced by a common multiplier: outcome i on board k enters if $p_{ki}/c_{ki} > b^*$, with b^* the unique value making total commitment equal to one. Boards whose best ratio falls below b^* are dropped entirely.*

The economic reading is that a binding budget converts a per board threshold into a market wide one. A contract that is worth holding in isolation is not worth holding when better contracts elsewhere compete for the same dollar, which is the sense in which breadth is a substitute for edge.

13.2 The quadratic approximation and Markowitz

For small edges the log can be expanded. Writing the portfolio return as $R = \sum_j f_j(\mathbf{1}_{A_j}/c_j - 1)$ and using $\log(1+R) \approx R - R^2/2$,

$$g(f) \;\approx\; f^\top \tilde{\mu} \;-\; \tfrac{1}{2} f^\top \Sigma f, \qquad \tilde{\mu}_j = \frac{p_j - c_j}{c_j}, \qquad \Sigma_{jl} = \frac{\mathbb{P}(A_j \cap A_l) - p_j p_l}{c_j c_l}, \tag{28}$$

whose maximiser is $f^\star \approx \Sigma^{-1}\tilde{\mu}$. Growth optimal betting is mean variance optimisation with unit risk aversion, the covariance taken under the posterior and the returns normalised by price. Two remarks. For a single binary contract $\Sigma_{jj} = p_j(1-p_j)/c_j^2$, so $f^\star \approx c_j(p_j - c_j)/\{p_j(1-p_j)\}$, which agrees with Proposition 5 to first order in the edge. And on an exclusive family Σ is singular, since the indicators sum to one, which is why the exact solution of Proposition 8 takes the water filling form rather than an inverse covariance form. The approximation is useful for the cross board problem, where the exclusivity constraints are local and Σ is block structured, and it makes clear that the relevant input is the posterior covariance across contracts, an object the exchange never quotes.

14 Model Averaging and a Simulation Study

14.1 Averaging over ordering laws

Section 5 offered three ordering models: Harville, the Stern gamma family, and the discounted rule of Lo and Bacon-Shone (2008). A trader need not choose. Let $\mathcal{M}_1, \ldots, \mathcal{M}_J$ carry posterior weights ω_j and produce contract probabilities $p^{(j)}$. Proposition 6 settles what to do.

Proposition 14 (Bet the mixture, not the mixture of bets)**.** *The growth optimal position under model uncertainty uses the mixed probability $\bar{p} = \sum_j \omega_j p^{(j)}$. For a single binary contract the map $p \mapsto f^\star(p)$ is affine, so the mixture of the individual Kelly stakes coincides with the Kelly stake at $\bar{p}$. On an exclusive family the water filling map is piecewise linear with kinks at the points where the support set S changes, so the two differ, and averaging the portfolios is strictly suboptimal whenever the models disagree about S.*

In the worked example the Harville exacta for A over B is 0.124 and the discounted value at $\lambda = 0.81$ is 0.113. Equal weights give 0.1185, which against an eleven cent quote is an edge of eighty five basis points, well inside the fee band. Model averaging here does not produce a trade; it correctly reports that there is none. That is the useful outcome. A trader who picks whichever ordering model happens to favour the position he wants will always find one.

14.2 Does fractional Kelly actually help?

Proposition 7 is a statement about a trader who knows he might be wrong. Proposition 6 says that a trader whose posterior is honest should not shrink. The two are reconciled by asking what happens when the trader treats a noisy estimate as if it were a posterior mean, which is the empirically common failure.

We simulate the six contract board of Section 11 for 500 events and 600 independent paths. Outcomes are drawn from the true p. In the first regime the trader knows p. In the second his belief at each event is an independent draw $\tilde{p} \sim \mathrm{Dir}(150\,p)$, so it is unbiased but noisy, with a standard deviation of about two cents on the favourite, and he stakes as though $\tilde{p}$ were certain. Stakes are computed by Proposition 8 at the fee adjusted costs and scaled by $\rho \in \{1, \frac{1}{2}\}$.

Model	ρ	median $\log W_{500}$	fifth percentile	ninety fifth percentile	$\mathbb{P}(W_{500} < 1)$
Correct	1	5.61	0.15	10.79	0.05
Correct	$\frac{1}{2}$	4.08	1.42	6.68	0.01
Noisy	1	0.37	−6.02	7.86	0.47
Noisy	$\frac{1}{2}$	2.57	−0.93	6.40	0.11

Table 6: Terminal log wealth after 500 events, 600 paths. When the model is correct, full Kelly dominates in the median and half Kelly merely buys a tighter distribution, as the theory requires. When the trader acts on an unbiased but noisy estimate, full Kelly loses money on nearly half of all paths and half Kelly is better on every statistic reported.

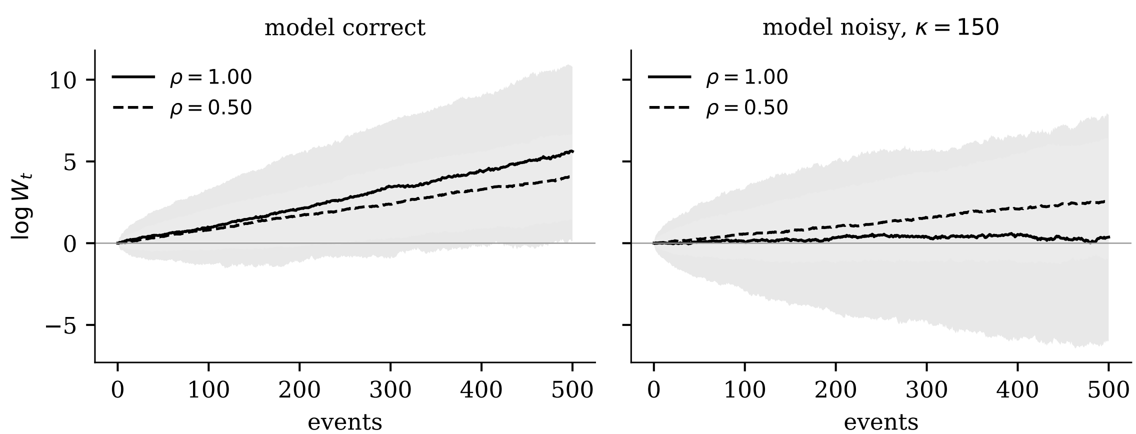

Figure 4: Median log wealth with fifth and ninety fifth percentile bands over 500 events. Left: the trader's probabilities are correct, and full Kelly ($\rho = 1$, solid) beats half Kelly ($\rho = \frac{1}{2}$, dashed) in the median while carrying a wider band. Right: the trader's probabilities are unbiased but noisy, and the ordering reverses.

Table 6 and Figure 4 give the result. With a correct model, full Kelly earns a median log wealth of 5.61 against 4.08 for half Kelly, and the shrinkage costs growth exactly as Proposition 7 predicts.

With a noisy model the ranking inverts: full Kelly ends below its starting bankroll on forty seven percent of paths and its median gain collapses to 0.37, while half Kelly retains 2.57 and loses on eleven percent of paths.

Nothing here contradicts Proposition 6. The simulated trader is not using a posterior mean; he is using a single draw and treating it as certain, which is what overfitting to a short history looks like in practice. Shrinking that draw halfway toward the price is, by Proposition 7, the behaviour of someone who assigns even odds to the market being right, and it recovers most of the loss. The lesson is not that Kelly is too aggressive. It is that the input is usually worse than the trader believes, and that the correct response is to fix the input, with fractional Kelly as the insurance one buys against failing to.

15 Discussion

What Bayes buys. Coherence under a low dimensional generator. One posterior prices every contract on a board at once, respects nesting, exclusivity and monotonicity by construction, fills illiquid quotes from liquid ones, attaches uncertainty to derived contracts such as exactas and multi state election payoffs, and yields a sizing rule whose growth rate is a divergence from the price. None of this is available to a trader reasoning contract by contract.

What it does not buy. Settlement risk is not statistical risk: contracts resolve against a named source, and modelling that source's definition, revision policy and timing is a legal exercise. Adverse selection concentrates at the release instant for economic contracts, where the entire information content arrives in one second and resting liquidity is withdrawn beforehand, so an edge that exists at t^- is not executable at t. Position limits and depth mean the Kelly fractions above may exceed what the book absorbs, in which case the relevant problem carries a capacity constraint. And the regulatory perimeter for sports contracts remains contested across states, a source of variance no model represents.

Directions. Sequential Monte Carlo for live in game and intraday weather boards, where the filtering structure of Feng et al. (2016) applies directly. Generative Bayesian computation for posterior simulation where the likelihood is implicit, as it is for bracket and path dependent contracts available only through a simulator. And conformal calibration layered on the posterior predictive, converting the calibration assumption behind a ladder into a distribution free coverage statement.

16 Conclusion

Event contracts make the Bayesian bargain explicit. The exchange quotes a probability, the model produces a probability, and Kelly converts the gap into a stake whose growth rate is the divergence between the two. Harville shows how a combinatorially large contract space is generated by a vector of strengths, and how badly a plug in calculation misstates both the level and the uncertainty of the result. The Skellam calibration of Feng et al. (2016) performs the same reversal for scores and supplies the dynamic version. The martingale principles of Aldous (2013) show that a board can be tested before it is modelled. The fee schedule $\theta\pi(1-\pi)$ shows that any resulting edge must clear up to three and a half cents before it is an edge at all. And an ϵ contamination prior toward the market shows that fractional Kelly, long defended as prudence, is simply what a Bayesian does once he takes seriously the possibility that the price is right and he is not.

Two asymmetries deserve a final word. The first is between making and taking. The maker pays one quarter of the fee and, by Proposition 11, faces an adverse selection cost of the same quadratic shape, so the exchange has built a schedule that mirrors the economics it taxes; the empirical maker premium on Kalshi is the visible consequence. The second is between uncertainty and error. Proposition 6 says that an honest posterior should not be shrunk, and the simulation of Section 14 says that a noisy point estimate should be shrunk hard. Almost every trader believes he is in the first situation. The evidence of Section 7, in which contracts under ten cents return less than forty cents on the dollar, suggests that most are in the second.

References

Aldous, D. J. (2012). *On Chance and Unpredictability: Twenty Lectures on the Links between Mathematical Probability and the Real World*. Lecture notes, University of California, Berkeley.

Aldous, D. J. (2013). Using prediction market data to illustrate undergraduate probability. *American Mathematical Monthly*, 120(7), 583–593.

Ali, M. M. (1977). Probability and utility estimates for racetrack bettors. *Journal of Political Economy*, 85(4), 803–815.

Avellaneda, M. and Stoikov, S. (2008). High-frequency trading in a limit order book. *Quantitative Finance*, 8(3), 217–224.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer.

Bradley, R. A. and Terry, M. E. (1952). Rank analysis of incomplete block designs. *Biometrika*, 39(3/4), 324–345.

Breiman, L. (1961). Optimal gambling systems for favorable games. *Proceedings of the Fourth Berkeley Symposium*, 1, 65–78.

Bürgi, C., Deng, W. and Whelan, K. (2025). Makers and takers: the economics of the Kalshi prediction market. CEPR Discussion Paper 20631; George Washington University CER Working Paper 2026-001.

Carvalho, C. M., Polson, N. G. and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, 97(2), 465–480.

Clinton, J. D. and Huang, T. (2025). Assessing the accuracy and efficiency of 2024 political prediction markets. Working paper, Vanderbilt University.

Cover, T. M. and Thomas, J. A. (2006). *Elements of Information Theory*, 2nd ed. Wiley.

Dawid, A. P. (1982). The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77(379), 605–610.

Feng, G., Polson, N. G. and Xu, J. (2016). The market for English Premier League (EPL) odds. *Journal of Quantitative Analysis in Sports*, 12(4), 167–178.

Glickman, M. E. and Stern, H. S. (1998). A state-space model for National Football League scores. *Journal of the American Statistical Association*, 93(441), 25–35.

Glosten, L. R. and Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1), 71–100.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378.

Grünwald, P., de Heide, R. and Koolen, W. M. (2024). Safe testing. *Journal of the Royal Statistical Society, Series B*, 86(5), 1091–1128.

Goll, T. and Rüschendorf, L. (2001). Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics*, 5(4), 557–581.

Hanson, R. (2003). Combinatorial information market design. *Information Systems Frontiers*, 5(1), 107–119.

Harville, D. A. (1973). Assigning probabilities to the outcomes of multi-entry competitions. *Journal of the American Statistical Association*, 68(342), 312–316.

Henery, R. J. (1981). Permutation probabilities as models for horse races. *Journal of the Royal Statistical Society, Series B*, 43(1), 86–91.

Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology*, 12(4), 595–600.

Kelly, J. L. (1956). A new interpretation of information rate. *Bell System Technical Journal*, 35(4), 917–926.

Lo, V. S. Y. and Bacon-Shone, J. (2008). Approximating the ordering probabilities of multi-entry competitions by a simple method. In *Handbook of Sports and Lottery Markets*, Elsevier, 51–65.

Luce, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis*. Wiley.

MacLean, L. C., Thorp, E. O. and Ziemba, W. T., eds. (2011). *The Kelly Capital Growth Investment Criterion: Theory and Practice*. World Scientific.

Manski, C. F. (2006). Interpreting the predictions of prediction markets. *Economics Letters*, 91(3), 425–429.

Ottaviani, M. and Sørensen, P. N. (2008). The favorite-longshot bias: an overview of the main explanations. In *Handbook of Sports and Lottery Markets*, Elsevier, 83–101.

Plackett, R. L. (1975). The analysis of permutations. *Applied Statistics*, 24(2), 193–202.

Polson, N. G., Scott, J. G. and Windle, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *Journal of the American Statistical Association*, 108(504), 1339–1349.

Polson, N. G. and Stern, H. S. (2015). The implied volatility of a sports game. *Journal of Quantitative Analysis in Sports*, 11(2), 145–153.

Raftery, A. E., Gneiting, T., Balabdaoui, F. and Polakowski, M. (2005). Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, 133(5), 1155–1174.

Shafer, G. (2021). Testing by betting: a strategy for statistical and scientific communication. *Journal of the Royal Statistical Society, Series A*, 184(2), 407–431.

Shafer, G. and Vovk, V. (2019). *Game-Theoretic Foundations for Probability and Finance*. Wiley.

Shin, H. S. (1993). Measuring the incidence of insider trading in a market for state-contingent claims. *Economic Journal*, 103(420), 1141–1153.

Skellam, J. G. (1946). The frequency distribution of the difference between two Poisson variates belonging to different populations. *Journal of the Royal Statistical Society*, 109(3), 296.

Stern, H. (1990). Models for distributions on permutations. *Journal of the American Statistical Association*, 85(410), 558–564.

Stern, H. (1994). A Brownian motion model for the progress of sports scores. *Journal of the American Statistical Association*, 89(427), 1128–1134.

Thaler, R. H. and Ziemba, W. T. (1988). Anomalies: parimutuel betting markets, racetracks and lotteries. *Journal of Economic Perspectives*, 2(2), 161–174.

Vovk, V. and Wang, R. (2021). E-values: calibration, combination and applications. *Annals of Statistics*, 49(3), 1736–1754.

Whelan, K. (2024). How do prediction market fees affect prices and participants? Working paper, University College Dublin.

Wolfers, J. and Zitzewitz, E. (2004). Prediction markets. *Journal of Economic Perspectives*, 18(2), 107–126.