

Prediction-Market Seed Capital Recovery from Noise-Dominant Flow

Yankai Chen^{1,2}, Bowei He^{1,2}, Zhuohan Xie², Zhiwei Liu³, Fan Zhang²,
Xueqing Peng⁴, Yan Wang⁴, Lingfei Qian⁴, Xue Liu^{1,2}

¹McGill University ²MBZUAI

³The University of Manchester ⁴The Fin AI

Abstract

Automated prediction markets require sponsors to prefund liquidity before observing order flow, creating a financing challenge at launch. We study whether nonnegative charges conditioned on observable payoff direction can improve recovery of this prefunded capital while limiting their effect on informed participation. We develop Seed Capital Flow (SCF), a direction-conditioned levy, in a stylized binary cost-function market with informed and liquidity-motivated traders. When order composition differs across directions, SCF concentrates the permitted fee burden on the direction with relatively more liquidity-motivated flow, whereas a uniform fee spreads it across both directions. Under a sufficiently tight common retention constraint, this allocation yields higher expected recovery capacity and can make additional liquidity choices financially viable. Synthetic numerical audits examine robustness to alternative flow patterns, stochastic arrivals, and label misspecification. The results characterize a mechanism-design tradeoff rather than an empirical prediction: the market remains prefunded, recovery is expected rather than guaranteed, and the analysis is limited to an opening-cohort setting.

1 Introduction

Automated prediction markets can aggregate the probabilistic signals produced by financial NLP systems (Xie et al., 2026c; Elbadry et al., 2026; Zhou et al., 2026; Wang et al., 2026b), but such a market must first exist, and here they face a cold-start problem. Deeper liquidity makes a cost-function market easier to trade against, but the sponsor must commit a larger worst-case-loss budget before observing any order flow. No fee can create revenue when no potential trader arrives. We instead ask whether fee design can reduce the traffic required to recover prefunded seed capital in expectation.

Uniform fees earn revenue from liquidity-motivated orders but also deter orders that convey information, such as those placed by forecasting models and trading agents (Lin et al., 2025; Peng et al., 2026; Gong et al., 2026). This tension changes when payoff direction predicts order composition. If one side has a higher liquidity-to-informed ratio, a uniform fee spends the informed-participation budget equally across two unequal directions. We call our policy Seed Capital Flow (SCF), a direction-conditioned recovery levy. It gives fee relief to the candidate-information-dominant direction and concentrates the permitted burden on the candidate-noise-dominant direction. This is a Ramsey-style allocation of a common screening budget: the platform observes direction, not trader type.

We analyze this idea in a binary logarithmic market scoring rule with fixed-size orders and one opening cohort. The same primitives determine three objects: the exact finite-trade cost faced by an arriving trader, the participation of informed and liquidity-motivated traders, and the sponsor’s prefunded worst-case-loss budget. Raw surplus distributions and type–direction arrival intensities are fixed before liquidity is chosen. Glosten–Milgrom-style primitives make informed candidate flow symmetric under a common prior, while liquidity-motivated flow is more concentrated in one direction. SCF and the uniform benchmark face the same nonnegativity, fee-cap, and retention rules.

Result. Under a tight retention allowance, the unique SCF optimum exempts the candidate-information-dominant direction and charges the candidate-noise-dominant direction. The optimal uniform policy splits the same permitted burden between both directions. Although the two policies deter the same expected number of informed orders, SCF earns strictly more expected levy revenue at every admissible liquidity level.

The pointwise revenue gap changes the launch boundary. Each mechanism has a minimum potential-arrival scale at which at least one positive liquidity level can recover its seed capital in expectation, and the SCF threshold is strictly lower. Between the two thresholds, some positive liquidity levels are recoverable with SCF, while none are recoverable with any feasible uniform levy. Above the uniform threshold, every uniformly recoverable level is recoverable under SCF, and the inclusion is strict. The separation also survives an informed-order-loss allowance fixed across candidate liquidity levels once the platform imposes a positive minimum depth.

The contribution is deliberately narrow. We identify a sufficient combination of fixed raw Glosten–Milgrom-style flow, a common informed-order-loss rule, and directional asymmetry that makes the weak menu-inclusion comparison strict at every admissible depth. Exact LMSR geometry then converts that strict screening gain into an arrival-threshold and recoverable-set result. Differentiated fees, weak dominance from a larger policy menu, and noise-flow financing are not claimed as new by themselves. A structured simulation audit maps the size of the threshold effect, supplies a wrong-direction negative control, quantifies exact finite-cohort chance thresholds, and probes nonuniform surplus distributions outside the theorem.

2 Related Work

Cost-function prediction markets provide continuous liquidity with bounded sponsor loss (Hanson, 2003; Chen and Pennock, 2007; Abernethy et al., 2013). Profit-charging and liquidity-sensitive market makers use revenue or activity to offset subsidy and adjust liquidity (Othman and Sandholm, 2012; Othman et al., 2013; Li and Vaughan, 2013); recent work studies liquidity pools, general provisioning interfaces, online adaptation, dynamic loss allocation, and quadratic fees (Amini et al., 2026; Bhaskara et al., 2026; Nueve et al., 2026; Moallemi et al., 2026; Nueve and Waggoner, 2025). We instead fix LMSR depth and compare recovery policies under a common informed-order-retention rule.

Our adverse-selection primitives follow the insight that direction changes a market maker’s posterior when informed and uninformed traders coexist (Glosten and Milgrom, 1985). Related models connect those primitives to automated market-maker

curves or optimize prediction-market bid and ask quotes (Nadkarni et al., 2024; Feil and Nendel, 2026). Our one-cohort model has no belief updating, spread choice, or inventory control.

Directional and information-sensitive AMM fees are also established. Prior work studies asymmetric or dynamic charges for toxic arbitrage and noise flow, nonconstant-fee accuracy–loss frontiers, and uninformed-flow financing of equilibrium liquidity (Alexander et al., 2024; Baggiani et al., 2026; Rao and Shah, 2023; Adams et al., 2025). The uniform menu is nested in the SCF policy class, so weak dominance is generic. Our strict result instead uses unequal direction-level liquidity-to-informed ratios, a binding common retention constraint, and the exact LMSR seed requirement. Payoff-representation invariance and split-proof fee design are treated as implementation constraints, not standalone novelty (Frongillo and Waggoner, 2018; Frongillo et al., 2024; Bichuch and Feinstein, 2026).

Financial NLP benchmarks supply the forecasts such markets aggregate (Wang et al., 2026a; Zhang et al., 2026; Xie et al., 2026b,a,d), and autonomous agents form a growing share of traders (Xie et al., 2026e; Safder et al., 2026). A companion mechanism adapts fees to trade span for sequential aggregation (Chen et al., 2026); we instead study financing the launch.

More generally, price-discrimination results already show that a discriminatory menu weakly dominates uniform pricing and characterize when profit gaps can be large (Malueg and Snyder, 2006; Bergemann et al., 2022). Our claim is correspondingly bounded: among the public primary sources reviewed through August 11, 2026, we did not identify the same fixed-raw opening-cohort primitives combined with a common informed-order-loss allowance and a strict LMSR arrival-threshold separation. This scoped search statement is not a priority or firstness claim.

3 Market and Opening-Cohort Model

3.1 Binary payoff space and LMSR capital

There are two mutually exclusive outcomes, Y and N . A contingent claim is a vector $r = (r_Y, r_N) \in \mathbb{R}^2$, and $\mathbf{1} = (1, 1)$. Portfolios r and $r + \alpha\mathbf{1}$ differ only by the sure payoff α . The risky quotient therefore has one coordinate,

$$x(r) = r_Y - r_N. \quad (1)$$

The sign of $x(r)$ defines the economic payoff direction, independently of the order interface. For example, $(z, 0)$ (buy z YES) and $(0, -z)$ (sell z NO) share risky coordinate z and differ only by $z\mathbf{1}$.

For $b > 0$, the binary LMSR cost function is

$$C_b(q) = b \log \left(e^{q_Y/b} + e^{q_N/b} \right). \quad (2)$$

We initialize at $q^0 = (0, 0)$, so the price is $(1/2, 1/2)$. The cost is translation invariant:

$$C_b(q + \alpha \mathbf{1}) = C_b(q) + \alpha. \quad (3)$$

Its worst-case loss from the initial state is

$$K(b) = b \log 2. \quad (4)$$

The outside option $b = 0$ means that the platform does not launch the product.

Fix an indivisible risky lot $z > 0$. The canonical zero-sum representatives are

$$r_h = \left(\frac{z}{2}, -\frac{z}{2} \right), \quad r_n = -r_h. \quad (5)$$

At the initial state, the exact LMSR cost of either representative is

$$D(b) := C_b(q^0 + r_h) - C_b(q^0) = b \log \cosh \left(\frac{z}{2b} \right). \quad (6)$$

Thus both directions have the same finite-trade cost; the benchmark removes LMSR-side asymmetry.

Lemma 1 (LMSR geometry). *The function D is continuous and strictly decreasing on $(0, \infty)$, with*

$$\lim_{b \downarrow 0} D(b) = \frac{z}{2}, \quad \lim_{b \rightarrow \infty} D(b) = 0. \quad (7)$$

Consequently, for every $w_N \in (0, z/2)$, there is a unique $b_0 > 0$ such that $D(b_0) = w_N$.

3.2 Timing and the recovery object

The platform proceeds in two stages. At time zero it chooses $b > 0$, posts the full capital in [equation \(4\)](#), and announces a levy policy. During a specified recovery window, potential fixed-lot orders arrive. The platform calls b *recoverable in expectation* if expected levy receipts during that window are at least $K(b)$. The initial capital remains necessary even when this inequality holds.

Our closed form is a fixed-opening-state cohort model. Participation is evaluated against the initial LMSR schedule in [equation \(6\)](#); the model does not treat [equation \(6\)](#) as the exact cost faced by every order along an endogenous multi-order state path.

The cohort may be viewed as a launch-screening benchmark or as the reduced form of an opening batch. A sequential implementation would have to replace the demand functions below with expectations over the evolving state.

3.3 Latent types and fixed raw primitives

Potential orders have latent type $T \in \{I, N\}$ and observable payoff direction $s \in \{h, n\}$. Type I is informed; type N has an outcome-independent liquidity motive. The labels h and n will denote the candidate-information-dominant and candidate-noise-dominant directions, respectively. The platform observes s , not T .

Direction is a policy-invariant attribute of the candidate order. Conditional on (T, s, W) , the announced levy changes only the execute-versus-abstain decision; a candidate cannot substitute an h order for an n order or vice versa. Thus the type-direction intensities below are primitives rather than equilibrium responses to the fee menu. This restriction is material and is revisited in [section 9](#).

Let W be an order's raw linear advantage before subtracting the LMSR finite-trade cost and the levy. Fix type-specific supports

$$0 < w_N < w_I \leq \frac{z}{2}, \quad d := w_I - w_N > 0. \quad (8)$$

For each type-direction pair, fix a density $c_{T,s} > 0$, independent of b . At arrival scale $m > 0$, the raw expected intensity is

$$m c_{T,s} dW \quad \text{for } W \in [0, w_T], \quad (9)$$

and zero outside the support. These intensities need not sum to one.

For an informed order, direction h can be represented by posterior $\mu = 1/2 + W/z$, and direction n by $\mu = 1/2 - W/z$. Condition [equation \(8\)](#) keeps both posteriors in $[0, 1]$. For a liquidity-motivated order, W is a private trading benefit independent of the realized outcome.

Use the density notation

$$i_s = c_{I,s}, \quad n_s = c_{N,s}, \quad c_s = i_s + n_s. \quad (10)$$

We impose

$$i_h = i_n =: i > 0, \quad n_n > n_h > 0. \quad (11)$$

The first restriction gives a symmetric common-prior informed construction. The second places more liquidity-motivated raw intensity in direction

n . The strict comparison comes from this flow asymmetry, not from directional LMSR curvature.

For $b > b_0$, define the fee headroom

$$a(b) := w_N - D(b) > 0. \quad (12)$$

The corresponding informed-trader headroom is $a(b) + d$. If direction s faces a per-lot levy $\ell_s \in [0, a(b)]$, an order executes exactly when $W \geq D(b) + \ell_s$. The resulting demands are

$$Q_{I,s}(\ell_s; b) = m_i [a(b) + d - \ell_s]. \quad (13)$$

$$Q_{N,s}(\ell_s; b) = m_n n_s [a(b) - \ell_s]. \quad (14)$$

All four type-direction cells remain in their linear regions on the stated levy domain.

At zero levy, the informed share among candidate orders in direction s is

$$\rho_s(b) = \frac{i[a(b) + d]}{i[a(b) + d] + n_s a(b)}. \quad (15)$$

Hence $\rho_h(b) > \rho_n(b)$ for every $b > b_0$. This is why h is called candidate-information-dominant and n candidate-noise-dominant. The classification is made before differential screening.

Among completed orders under levy $\ell < a(b)$, the informed share is

$$\rho_s^E(\ell; b) = \frac{i[a(b) + d - \ell]}{i[a(b) + d - \ell] + n_s [a(b) - \ell]}. \quad (16)$$

It increases strictly with ℓ , because the shorter-support liquidity-motivated type is screened out faster. This compositional statement is not, by itself, a claim about mutual information, forecast accuracy, or price informativeness. Moreover, differential screening can reverse the cross-direction ranking after execution.

4 SCF and the Common Retention Rule

4.1 A payoff-direction levy

For an arbitrary trade r with $x(r) \neq 0$, let $s(r) \in \{h, n\}$ denote the sign-mapped payoff direction. For coefficients $\tau_h, \tau_n \geq 0$, define the payment

$$\text{Pay}_{b,\tau}(q, r) = C_b(q + r) - C_b(q) + \frac{\tau_{s(r)}}{2} x(r)^2. \quad (17)$$

Riskless trades have zero levy. For the fixed risky lot z , write

$$\ell_s = \frac{\tau_s z^2}{2}. \quad (18)$$

The quadratic form in [equation \(17\)](#) applies on arbitrary trade histories; only its fixed-lot value enters the recovery theorem.

Proposition 2 (Representation invariance and safety). *For every $q, r \in \mathbb{R}^2$ and $\alpha, \beta \in \mathbb{R}$,*

$$\begin{aligned} \text{Pay}_{b,\tau}(q + \beta \mathbf{1}, r) &= \text{Pay}_{b,\tau}(q, r), \\ \text{Pay}_{b,\tau}(q, r + \alpha \mathbf{1}) &= \text{Pay}_{b,\tau}(q, r) + \alpha. \end{aligned} \quad (19)$$

Thus economically equivalent implementations have identical state-contingent net profits. Moreover, a finite history under [equation \(17\)](#) admits no pathwise arbitrage, and nonnegative levies do not increase the LMSR worst-case loss.

The proposition is an implementation requirement rather than a novelty claim. In particular, buy YES and sell NO cannot be charged differently when they represent the same risky payoff after a cash adjustment. Because levies are nonnegative, the lower charge on direction h is fee relief rather than a rebate.

4.2 Expected recovery revenue

Using [equations \(13\) and \(14\)](#), expected SCF receipts are

$$\begin{aligned} \Gamma(\ell_h, \ell_n; b, m) &= \sum_{s \in \{h, n\}} \ell_s [Q_{I,s} + Q_{N,s}] \\ &= m \sum_{s \in \{h, n\}} i d \ell_s \\ &\quad + m \sum_{s \in \{h, n\}} c_s \ell_s [a(b) - \ell_s]. \end{aligned} \quad (20)$$

The revenue is strictly concave under

$$0 \leq \ell_s \leq a(b). \quad (21)$$

Retention cost is the expected number of informed orders screened out relative to zero levy. With a uniform informed density,

$$L_s(\ell_s; b) = m i \ell_s. \quad (22)$$

This completion metric does not directly measure information incorporated into prices.

The tight-loading condition is

$$\bar{\delta} := \frac{i(n_n - n_h)}{2(i + n_n)} > 0, \quad 0 < \delta < \bar{\delta}. \quad (23)$$

At each candidate $b > b_0$, both mechanism classes face the same allowance

$$\bar{L}(b) = m \delta a(b). \quad (24)$$

Equivalently, the permitted loss is a constant fraction of common fee headroom. Although pointwise fair, the rule varies with b ; see [section 7](#).

The SCF recovery capacity, with D denoting its directional menu, is

$$\begin{aligned} \Gamma_D(b; m) &:= \max_{\ell_h, \ell_n} \Gamma(\ell_h, \ell_n; b, m) \\ \text{s.t. } i(\ell_h + \ell_n) &\leq \delta a(b), \\ 0 &\leq \ell_h, \ell_n \leq a(b). \end{aligned} \quad (25)$$

The uniform benchmark chooses the best feasible common levy:

$$\begin{aligned} \Gamma_U(b; m) &:= \max_{\ell} \Gamma(\ell, \ell; b, m) \\ \text{s.t. } 2i\ell &\leq \delta a(b), \quad 0 \leq \ell \leq a(b). \end{aligned} \quad (26)$$

Thus both mechanisms face the same retention allowance. Weak dominance is mechanical because the uniform class is the diagonal subset of the SCF policy class. The substantive question is whether strict dominance holds at every depth and expands the recoverable set.

5 Seed-Capital Recovery Thresholds

We now solve equations (25) and (26) and connect the revenue comparison to prefunded seed capital. Use the abbreviations

$$I := 2i, \quad C := c_h + c_n. \quad (27)$$

Theorem 3 (Informed-order-retention-constrained liquidity-set expansion). *Suppose equations (8), (11) and (23) hold. For every $b > b_0$, the unique optimal SCF levies are*

$$\ell_h^D(b) = 0, \quad \ell_n^D(b) = \frac{\delta a(b)}{i}, \quad (28)$$

whereas the unique optimal uniform levy is

$$\ell^U(b) = \frac{\delta a(b)}{2i}. \quad (29)$$

Let $\Gamma_M(b; m) = mG_M(b)$ for $M \in \{D, U\}$. The optimized recovery functions are

$$G_D(b) = \delta d a(b) + c_n a(b)^2 \left(\frac{\delta}{i} - \frac{\delta^2}{i^2} \right). \quad (30)$$

$$G_U(b) = \delta d a(b) + C a(b)^2 \left(\frac{\delta}{I} - \frac{\delta^2}{I^2} \right). \quad (31)$$

For every $b > b_0$,

$$0 < G_U(b) < G_D(b). \quad (32)$$

Normalize capacity by seed capital and define

$$\Psi_M(b) := \frac{G_M(b)}{b \log 2}, \quad S_M := \max_{b > b_0} \Psi_M(b). \quad (33)$$

Both maxima are finite, positive, and attained at interior liquidity levels:

$$0 < S_U < S_D < \infty. \quad (34)$$

The expected seed-capital recovery sets are

$$\mathcal{B}_M(m) := \{b > b_0 : mG_M(b) \geq b \log 2\}. \quad (35)$$

Then:

(i) if $0 < m < 1/S_D$, then $\mathcal{B}_D(m) = \mathcal{B}_U(m) = \emptyset$;

(ii) if $1/S_D \leq m < 1/S_U$, then $\mathcal{B}_D(m) \neq \emptyset$ and $\mathcal{B}_U(m) = \emptyset$;

(iii) if $m \geq 1/S_U$, then $\mathcal{B}_U(m) \subsetneq \mathcal{B}_D(m)$.

Why SCF loads one side. The retention constraint in equation (25) is

$$\ell_h + \ell_n \leq \frac{\delta a(b)}{i}. \quad (36)$$

For a fixed total levy burden, the term $id(\ell_h + \ell_n)$ in equation (20) is constant. The remaining revenue is more productive in direction n , because $c_n > c_h$. Condition $\delta < \bar{\delta}$ is strong enough that even after assigning the full allowed burden to n , the marginal return there remains above the marginal return from starting to charge h . Strict concavity then gives equation (28). Under a common levy, the same constraint becomes $2i\ell \leq \delta a(b)$, giving equation (29) as the benchmark.

Why the gap is strict. The uniform solution spends half of the common screening allowance in each direction. SCF spends all of it where the liquidity-motivated density is higher. Both lose exactly $m\delta a(b)$ expected informed orders, but their liquidity-motivated fee bases differ. The strict inequality in equation (32) therefore arises from $n_n > n_h$, a binding common constraint, and strict concavity, not merely from nesting the uniform menu in the SCF policy class.

From pointwise revenue to launchability. The seed-capital-normalized ratio Ψ_M tends to zero at both ends of its domain. As $b \downarrow b_0$, fee headroom and revenue vanish. As $b \rightarrow \infty$, $a(b) \rightarrow w_N$, so $G_M(b)$ remains bounded while the seed $b \log 2$ diverges. Continuity yields an interior maximum. At a maximizer of Ψ_U , pointwise strictness implies $\Psi_D > \Psi_U$, and hence $S_D > S_U$. The three threshold cases then follow from

$$b \in \mathcal{B}_M(m) \iff m\Psi_M(b) \geq 1. \quad (37)$$

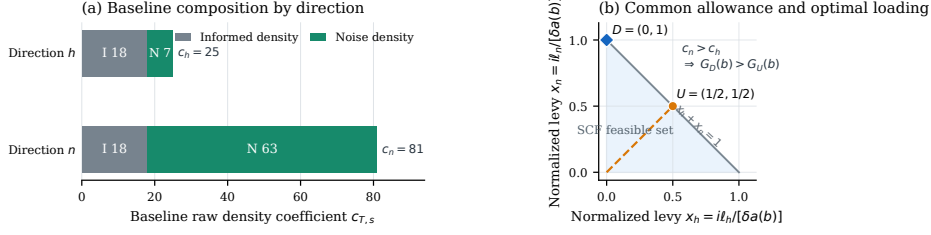


Figure 1: Direction-conditioned levy allocation under a common retention budget. (a) Informed raw density is symmetric, whereas noise is concentrated in direction n , giving $c_h = 25$ and $c_n = 81$. (b) With $x_s = i\ell_s/[\delta a(b)]$, SCF chooses over the shaded simplex while uniform is restricted to its diagonal. The optima $D = (0, 1)$ and $U = (1/2, 1/2)$ exhaust the same allowance, but directional loading yields $G_D(b) > G_U(b)$ for every $b > b_0$.

Robustness to a liquidity-invariant allowance.

The pointwise allowance in equation (24) varies with b . Appendix B instead fixes the allowed informed-order loss per unit of arrival scale at Λ on a domain $b \geq \underline{b} > b_0$. If $0 < \Lambda < \bar{\delta}a(\underline{b})$, the SCF solution still loads the entire burden onto n , strictly dominates the optimal uniform levy pointwise, and preserves all three arrival-threshold conclusions. The positive liquidity floor is necessary because fee headroom vanishes as $b \downarrow b_0$.

Proposition 4 (Symmetric benchmark). *Maintain the linear-surplus opening-cohort model, but replace equations (11) and (23) with $i_h = i_n =: i > 0$, $n_h = n_n =: n > 0$, and a direction-symmetric retention allowance. Specifically, for some $\bar{L}(b) \geq 0$, let SCF face $mi(\ell_h + \ell_n) \leq \bar{L}(b)$ with identical caps in both directions, and let the uniform program face $2mi\ell \leq \bar{L}(b)$. Then the SCF objective and feasible set are invariant to exchanging h and n . The unique SCF optimizer is diagonal, and the optimized SCF and uniform recovery capacities coincide.*

Remark 5 (Candidate versus executed composition). Theorem 3 classifies direction n using the zero-levy candidate posterior in equation (15). Because $\rho_s^E(\ell; b)$ rises with the levy, the higher charge can make completed orders in direction n more informed on average and can even reverse the original ranking. The theorem does not require the ex-post ranking to remain fixed.

Remark 6 (What is optimal). The word *optimal* always means optimal within the capped class of non-negative per-lot levies in equations (25) and (26), for the opening-cohort revenue objective and the stated retention allowance. It does not mean globally optimal market design.

6 Experiments

The theorem is analytical; the experiments are synthetic comparative statics and falsification checks (Liu et al., 2026), not empirical evidence. For each parameter vector we globally maximize $\Psi_M(b) = G_M(b)/(b \log 2)$ over $b > b_0$ using a 512-point adaptive log grid and golden-section refinement. The primary outcome is the proportional reduction in the minimum arrival scale,

$$R := 1 - \frac{m_D^*}{m_U^*} = 1 - \frac{S_U}{S_D}. \quad (38)$$

Positive R favors SCF. No main-model maximum is on a numerical boundary, and the baseline agrees with a separate 20,001-point grid within 3.5×10^{-9} in normalized capacity. Parameter-draw percentiles below describe the chosen synthetic design measure; they are not confidence intervals or estimates of a market population.

Baseline. We set $z = 1$, $(w_I, w_N) = (0.45, 0.30)$, $(i_h, i_n) = (18, 18)$, $(n_h, n_n) = (7, 63)$, and $\delta = \bar{\delta}/2 = 3.111111$. This gives

$$b_0 = 0.304759, \quad S_D^{-1} = 1.503569, \quad (39) \\ R = 22.84\%, \quad S_U^{-1} = 1.948531.$$

At $m = 1.726050$, SCF crosses the recovery target while the uniform policy does not (figure 2a). One unit of m contains 37.2 expected raw candidate orders under this calibration, so the two thresholds correspond to 55.93 and 72.49 candidates before participation screening; this conversion does not turn m into observed traffic.

Independent program and boundary audits.

The large sweeps inside the theorem domain reuse its closed-form capacity and therefore cannot independently validate the optimizer. We separately solve the raw two-levy programs in equations (25)

Audit	Setting	Result
Raw program	20 designs	$\max \Delta G = 2.22 \times 10^{-16}$
Symmetry	$A = 0$	$R \doteq 0$
Side swap	$A = \pm 0.8$	same R , levies swap
Loading	$q : .99 \rightarrow 1.25$	n -only $\rightarrow$ split
Recovery sets	low/mid/high m	$(0, 0), (1, 0), (1, 1)$

Table 1: Audits that can contradict the closed-form implementation or its boundary interpretation. Recovery-set entries count nonempty SCF and uniform sets, respectively. Full rows and residuals are in the supplement.

and (26), without calling the closed-form capacity routine, over 20 designs spanning four calibrations and five fee-headroom levels. A complete 1,001-point feasible-edge search with local refinement matches the closed-form levies within 3.34×10^{-16} , capacities within 2.22×10^{-16} , and the common loss budget within 2.22×10^{-16} .

We then keep the baseline normalized loss budget but include signed noise-flow asymmetry $A \in \{-0.8, 0, 0.01, 0.2, 0.8\}$. At $A = 0$, the numerical optimum is diagonal and $R = 4.44 \times 10^{-16}$; swapping $A = 0.8$ to -0.8 leaves $R = 22.84\%$ unchanged and swaps the charged side. At $A = 0.01$, the gain is only 0.0073%. Finally, holding $A = 0.8$ fixed and moving the allowance share from $q = 0.99$ to $q = 1.25$ changes the optimum from one-sided to split loading. The gain remains non-negative because the optimized directional menu contains the uniform menu, but the one-sided prescription is not extended beyond its stated tight-loading condition.

Recoverable sets and structured sweep. Root refinement directly recovers the three cases of theorem 3. At $m = 1.353212$, both sets are empty. At $m = 1.726050$, the SCF set is $[0.724392, 1.828707]$ and the uniform set is empty. At $m = 2.143384$, the SCF and uniform sets are $[0.591848, 2.664094]$ and $[0.753474, 1.625315]$, respectively; the latter is strictly contained in the former. Maximum endpoint residual is 1.34×10^{-15} .

For comparative statics, define flow asymmetry $A = (n_n - n_h)/(n_n + n_h)$ and retention tightness $q = \delta/\bar{\delta}$. A 64×61 grid over $A \in [0.02, 0.95]$ and $q \in [0.05, 0.95]$, holding mean noise density fixed, produces a smooth phase diagram (figure 2b). All 3,904 cells have $R > 0$, as the theorem requires, while the magnitude shrinks toward symmetry.

Finite-cohort recovery probability. Expected intensities do not specify realized order counts. Let the four type-direction cells be independent Poisson point processes over the recovery window. Thinning by participation and locking the levies before orders realize gives $V_M = \ell_M \text{Pois}(m\lambda_M(b))$, with $\mathbb{E}[V_M] = mG_M(b)$. For each target probability γ , each mechanism chooses b ex ante to minimize the required m . At $\gamma = 0.95$, the thresholds are 2.0494 for SCF and 2.4436 for uniform, a 16.13% reduction (figure 3). The reduction falls from 23.40% at $\gamma = 0.50$ to 13.40% at $\gamma = 0.99$.

Independent Poisson flow is thin-tailed. As a stress test, multiply all four cell intensities in a window by a common $U \sim \text{Gamma}(k, k)$, preserving their means while inducing clustering. At the 95% target, $k = 10$ raises the SCF and uniform thresholds to 2.9706 and 3.7256; $k = 2$ raises them to 8.5794 and 11.0436. Thus common shocks sharply worsen absolute launch traffic, while the relative reductions remain 20.27% and 22.31%. This check covers a shared multiplicative shock, not arbitrary directional dependence.

Implementation coverage and distributional stress. A 324-case factorial and 10,000 fixed-seed parameter draws vary w_N/z , the normalized surplus gap, mean noise-to-informed density, A , and q over the theorem’s admissible interior. All 10,324 cases retain the strict ordering, with R ranging from 0.016% to 43.04%. Because strictness is implied by the theorem and the draw distribution is artificial, this is implementation coverage rather than independent or statistical evidence.

To move outside the uniform-surplus assumption, we replace each type’s normalized surplus density by Beta(1, 3), Beta(1, 1), Beta(2, 2), or Beta(3, 1), preserving cell mass, informed symmetry, and the common absolute loss allowance. Tail-certified multistart reoptimization in all 16 pairs retains $S_D > S_U$, with $R = 0.030$ –38.35%. The one-sided pattern survives in only 10 cases; high-surplus informed flow with center-concentrated noise is the near-counterexample. This finite search is not a theorem extension.

Wrong-label negative control. The theorem assumes the noise-dominant direction is known. Let π be the probability that a recovery window uses the correct direction mapping and let $r = \delta/i$. The

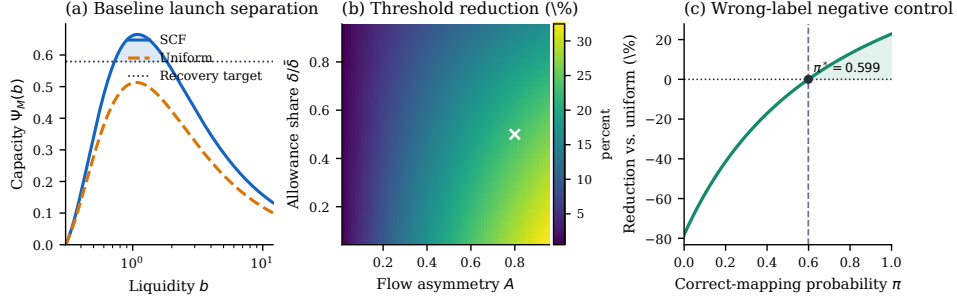


Figure 2: Synthetic audit. (a) Baseline seed-normalized recovery curves; shading marks liquidity levels recoverable only under SCF at the displayed arrival scale. (b) Threshold reduction over flow asymmetry and retention tightness; the cross marks the baseline. (c) Exact wrong-label negative control.

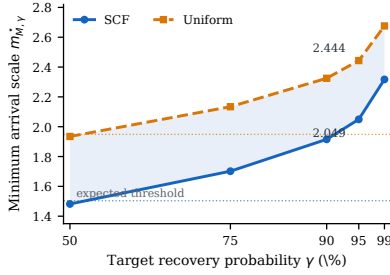


Figure 3: Exact finite-cohort thresholds under independent Poisson thinning. Each mechanism chooses b before the opening cohort realizes. Dotted lines mark the corresponding expected-receipt thresholds.

exact baseline break-even is

$$\pi^* = \frac{C(r/2 - r^2/4) - c_h(r - r^2)}{(c_n - c_h)(r - r^2)} = 0.598881. \quad (40)$$

SCF beats uniform only when $\pi > \pi^*$ (figure 2c). Complete reversal raises the baseline arrival threshold by 78.17% relative to uniform. The value 59.89% is not universal: across the 10,324 admissible designs, π^* has median 61.17%, 5th–95th percentiles 51.36–72.77%, and range 50.34–74.47%. Here π is a window-level mapping probability, not per-order classifier accuracy.

Fixed allowance and fragmentation. For four positive minimum depths and three tightness levels, the liquidity-invariant allowance of theorem 7 preserves a 17.78%–27.08% threshold reduction. In contrast, splitting a quadratically charged lot into k equal pieces leaves only $1/k$ of the levy, supporting the indivisible-lot domain but not Sybil robustness.

7 Discussion

What kind of liquidity is added? The mechanism does not manufacture reserves after a market

has opened. The sponsor supplies the full LMSR seed before trading. Conditional on a positive potential opening cohort, the fee allocation can make more candidate values of b recoverable in expectation. This is a financing or launchability result: if $m = 0$, levy revenue is zero under every policy and no positive b is recoverable.

Pointwise fairness and cross- b normalization.

At each candidate b , both mechanisms face the same allowance $\bar{L}(b) = m\delta a(b)$, but this allowed number of screened informed orders varies with depth. Proposition 7 shows that the launch separation survives a fixed per-arrival allowance on a domain bounded away from the zero-headroom boundary. Neither result establishes robustness to every information-quality or retention objective.

8 Conclusion

SCF does not solve a literal no-trade cold start. It can nevertheless change the financing boundary of a prefunded prediction market when potential order flow is positive and direction predicts order composition. In the binary LMSR opening-cohort model, a tight common informed-order-retention rule makes it optimal to exempt the candidate-information-dominant direction and concentrate the permitted charge on the candidate-noise-dominant direction. Relative to the best uniform levy under the same rule, SCF has strictly higher expected recovery capacity at every admissible liquidity level and lowers the launch threshold.

The contribution links a common screening constraint to an exact capital threshold; fees, directional pricing, and noise-flow revenue are not new by themselves. A natural extension is a sequential model with an information-quality objective.

9 Limitations

The analysis is limited to a fixed opening cohort in a binary LMSR. The retention constraint counts informed orders rather than signal quality and does not establish sequential equilibrium, price accuracy, or endogenous direction choice. Ex-ante direction labels may be misspecified, and no empirical classifier is evaluated. Recovery is expected rather than pathwise; the Poisson and shared-shock checks omit directional clustering and serial dependence. The quadratic levy is vulnerable to trade splitting. Other settings require separate analysis.

References

- Jacob Abernethy, Yiling Chen, and Jennifer Wortman Vaughan. 2013. [Efficient market making via convex optimization, and a connection to online learning](#). *ACM Transactions on Economics and Computation*, 1(2):1–39.
- Austin Adams, Ciamac C. Moallemi, Sara Reynolds, and Dan Robinson. 2025. [am-AMM: An auction-managed automated market maker](#). *Preprint*, arXiv:2403.03367.
- Abe Alexander, Jesse Moestaredjo, Mart Heuvelmans, and Lars Fritz. 2024. [Role of fee choice in revenue generation of AMMs: A quantitative study](#). *Preprint*, arXiv:2406.12417.
- Hamed Amini, Maxim Bichuch, and Zachary Feinstein. 2026. [Decentralized prediction markets and sports books](#). *Mathematical Finance*.
- Leonardo Baggiani, Martin Herdegen, and Leandro Sánchez-Betancourt. 2026. [Optimal dynamic fees in automated market makers](#). *Preprint*, arXiv:2506.02869.
- Dirk Bergemann, Francisco Castro, and Gabriel Y. Weintraub. 2022. [Third-degree price discrimination versus uniform pricing](#). *Games and Economic Behavior*.
- Adithya Bhaskara, Rafael Frongillo, Elias Lindgren, and Maneesha Papireddygar. 2026. [A general theory of liquidity provisioning for prediction markets](#). *Preprint*, arXiv:2311.08725. Current public version; to appear at EC 2026.
- Maxim Bichuch and Zachary Feinstein. 2026. [Axioms for automated market makers](#). *Operations Research*.
- Yankai Chen, Rassul Magauin, Bowei He, Anuar Aimoldin, Sirui Song, Bin Xiao, Zangir Iklassov, and Xue Liu. 2026. Trade-adaptive aggregation of probabilistic financial forecasts. In *Proceedings of the Financial Technology and Natural Language Processing Workshop (FinNLP) @ EMNLP 2026*. Association for Computational Linguistics.
- Yiling Chen and David M. Pennock. 2007. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*, pages 49–56.
- Rania Elbadry, Sarfraz Ahmad, Ahmed Heakl, Dani Bouch, Momina Ahsan, Muhra AlMahri, Marwa El-said Khalil, Yuxia Wang, Salem Lahlou, Sophia Ananiadou, Veselin Stoyanov, Jimin Huang, Xueqing Peng, Preslav Nakov, and Zhuohan Xie. 2026. [SAHM: A benchmark for Arabic financial and shari’ah-compliant reasoning](#). In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 34509–34536, San Diego, California, United States. Association for Computational Linguistics.
- Dominik Feil and Max Nendel. 2026. [Optimal market making in prediction markets](#). *Preprint*, arXiv:2607.17991.
- Rafael Frongillo, Maneesha Papireddygar, and Bo Waggoner. 2024. [An axiomatic characterization of CFMMs and equivalence to prediction markets](#). In *15th Innovations in Theoretical Computer Science Conference*, volume 287, pages 51:1–51:21.
- Rafael Frongillo and Bo Waggoner. 2018. [An axiomatic study of scoring rule markets](#). In *9th Innovations in Theoretical Computer Science Conference*, volume 94, pages 15:1–15:20.
- Lawrence R. Glosten and Paul R. Milgrom. 1985. [Bid, ask and transaction prices in a specialist market with heterogeneously informed traders](#). *Journal of Financial Economics*, 14(1):71–100.
- Jie Gong, Maowei Jiang, Zhiwei Liu, Yang Qiao, Wenxi Wu, Mengxi Xiao, Enze Zhang, Ziyang Kuang, Yankai Chen, Caishuang Huang, Meng Zhou, Xiku Du, Xue Liu, Guojun Xiong, Min Peng, Qianqian Xie, and Sophia Ananiadou. 2026. [ShiJianBench: From dialogue to decision for long-horizon evaluation of investment advisors](#). *arXiv preprint arXiv:2608.01204*.
- Robin Hanson. 2003. [Combinatorial information market design](#). *Information Systems Frontiers*, 5(1):107–119.
- Xiaolong Li and Jennifer Wortman Vaughan. 2013. [An axiomatic characterization of adaptive-liquidity market makers](#). In *Proceedings of the 14th ACM Conference on Electronic Commerce*, pages 657–674.
- Sida Lin, Yankai Chen, Yiyang Qi, Chenhao Ma, Bokai Cao, Yifei Zhang, Xue Liu, and Jian Guo. 2025. CSPO: Cross-market synergistic stock price movement forecasting with pseudo-volatility optimization. In *Proceedings of the ACM Web Conference 2025 (WWW)*. ACM.
- Zhiwei Liu, Qiqi Qiang, Jie Gong, Runteng Guo, Haoyang Liu, Ziyang Kuang, Sichen Hui, Jingchi Yang, Wanying He, Maowei Jiang, Zhuohan Xie, Yankai Chen, Xue Liu, Min Peng, Qianqian Xie, and Sophia Ananiadou. 2026. [A survey of simulation in](#)

- finance: Data, methods, applications, and evaluation. SSRN preprint.
- David A. Malueg and Christopher M. Snyder. 2006. [Bounding the relative profitability of price discrimination](#). *International Journal of Industrial Organization*, 24(5):995–1011.
- Ciamac C. Moallemi, Dan Robinson, and Brian Zhu. 2026. [Uniform-loss automated market making for prediction markets](#). *Preprint*, arXiv:2607.17428.
- Viraj Nadkarni, Sanjeev Kulkarni, and Pramod Viswanath. 2024. [Adaptive curves for optimally efficient market making](#). In *6th Conference on Advances in Financial Technologies*, volume 316, pages 25:1–25:24.
- Enrique Nueve, Bao Nguyen, Rafael Frongillo, and Bo Waggoner. 2026. [Adaptive liquidity in prediction markets via online learning](#). *Preprint*, arXiv:2605.09599.
- Enrique Nueve and Bo Waggoner. 2025. [Smooth quadratic prediction markets](#). *Preprint*, arXiv:2505.02959.
- Abraham Othman, David M. Pennock, Daniel M. Reeves, and Tuomas Sandholm. 2013. [A practical liquidity-sensitive automated market maker](#). *ACM Transactions on Economics and Computation*, 1(3):1–25.
- Abraham Othman and Tuomas Sandholm. 2012. [Profit-charging market makers with bounded loss, vanishing bid/ask spreads, and unlimited market depth](#). In *Proceedings of the 13th ACM Conference on Electronic Commerce*, pages 790–807.
- Xueqing Peng, Zhuohan Xie, Yupeng Cao, Haohang Li, Lingfei Qian, Yan Wang, Vincent Jim Zhang, Huan He, Xuguang Ai, Linhai Ma, Ruoyu Xiang, Yueru He, Yi Han, Shuyao Wang, Yuqing Guo, Mingyang Jiang, Yilun Zhao, Youzhong Dong, Xiaoyu Wang, et al. 2026. [Herculean: An agentic benchmark for financial intelligence](#). *arXiv preprint arXiv:2605.14355*.
- Rithvik Rao and Nihar Shah. 2023. [Triangle fees](#). *Preprint*, arXiv:2306.17316.
- Muhammad Usman Safder, Ayesha Gull, Rania Elbadry, Fan Zhang, Yankai Chen, Xueqing Peng, Xue Liu, Preslav Nakov, and Zhuohan Xie. 2026. [FinPersona-Bench: A benchmark for longitudinal psychometric stability of autonomous financial agents](#). *arXiv preprint arXiv:2606.31522*.
- Yan Wang, Yi Han, Lingfei Qian, Yueru He, Xueqing Peng, Dongji Feng, Zhuohan Xie, Vincent Jim Zhang, Rosie Guo, Fengran Mo, Jimin Huang, Yankai Chen, Xue Liu, and Jian-Yun Nie. 2026a. [Conv-FinRe: A conversational and longitudinal benchmark for utility-grounded financial recommendation](#). In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*. ACM.
- Yan Wang, Keyi Wang, Shanshan Yang, Jaisal Patel, Jeff Zhao, Fengran Mo, Xueqing Peng, Lingfei Qian, Jimin Huang, Guojun Xiong, Yankai Chen, Víctor Gutiérrez-Basulto, Xiao-Yang Liu, Xue Liu, and Jian-Yun Nie. 2026b. [Finauditing: A financial taxonomy-structured multi-document benchmark for evaluating LLMs](#). In *Proceedings of the 49th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*. ACM.
- Zhuohan Xie, Yuyang Dai, Rania Elbadry, Vanshikaa Jani, Georgi Georgiev, Dimitar Dimitrov, Fan Zhang, Xueqing Peng, Lingfei Qian, Jimin Huang, Jiahui Geng, Yankai Chen, Ye Yuan, Haolun Wu, Yuxia Wang, Ivan Koychev, Veselin Stoyanov, Mingzi Song, Yu Chen, and 2 others. 2026a. [Overview of FinMMEval 2026 task 1: Multilingual financial multiple-choice question answering](#). In *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2026)*, CEUR Workshop Proceedings. CEUR-WS.org.
- Zhuohan Xie, Rania Elbadry, Fan Zhang, Georgi Georgiev, Xueqing Peng, Lingfei Qian, Jimin Huang, Dimitar Dimitrov, Vanshikaa Jani, Yuyang Dai, Jiahui Geng, Yuxia Wang, Ivan Koychev, Veselin Stoyanov, and Preslav Nakov. 2026b. [The CLEF-2026 FinMMEval lab: Multilingual and multimodal evaluation of financial AI systems](#). In *Advances in Information Retrieval*, volume 16486 of *Lecture Notes in Computer Science*, pages 267–276. Springer.
- Zhuohan Xie, Daniil Orel, Rushil Thareja, Dhruv Sahnan, Hachem Madmoun, Fan Zhang, Debopriyo Banerjee, Georgi Nenkov Georgiev, Xueqing Peng, Lingfei Qian, Jimin Huang, Jinyan Su, Aaryamonvikram Singh, Rui Xing, Rania Elbadry, Chen Xu, Haonan Li, Fajri Koto, Ivan Koychev, and 6 others. 2026c. [FinChain: A symbolic benchmark for verifiable chain-of-thought financial reasoning](#). In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14529–14553, San Diego, California, United States. Association for Computational Linguistics.
- Zhuohan Xie, Xueqing Peng, Georgi Georgiev, Dimitar Dimitrov, Yuyang Dai, Rania Elbadry, Vanshikaa Jani, Lingfei Qian, Fan Zhang, Jimin Huang, Jiahui Geng, Yankai Chen, Ye Yuan, Haolun Wu, Yuxia Wang, Ivan Koychev, Veselin Stoyanov, Mingzi Song, Yu Chen, and 2 others. 2026d. [Overview of FinMMEval 2026 task 2: Multilingual financial short-answer question answering](#). In *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2026)*, CEUR Workshop Proceedings. CEUR-WS.org.
- Zhuohan Xie, Lingfei Qian, Georgi Georgiev, Dimitar Dimitrov, Yuyang Dai, Rania Elbadry, Vanshikaa Jani, Xueqing Peng, Fan Zhang, Jimin Huang, Jiahui Geng, Yankai Chen, Ye Yuan, Haolun Wu, Yuxia Wang, Ivan Koychev, Veselin Stoyanov, Mingzi Song, Yu Chen, and 2 others. 2026e. [Overview of](#)

FinMMEval 2026 task 3: Live financial decision-making agents. In *Working Notes of the Conference and Labs of the Evaluation Forum (CLEF 2026)*, CEUR Workshop Proceedings. CEUR-WS.org.

Fan Zhang, Mingzi Song, Rania Elbadry, Yankai Chen, Shaobo Wang, Yixi Zhou, Xunwen Zheng, Yueru He, Yuyang Dai, Georgi Nenkov Georgiev, Ayesha Gull, Muhammad Usman Safder, Fan Wu, Liyuan Meng, Fengxian Ji, Junning Zhao, Xueqing Peng, Jimin Huang, Yu Chen, and 3 others. 2026. [FinReporting: An agentic workflow for localized reporting of cross-jurisdiction financial disclosures](#). In *Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 3: System Demonstrations)*, pages 728–735, San Diego, California, United States. Association for Computational Linguistics.

Yixi Zhou, Fan Zhang, Yu Chen, Haipeng Zhang, Preslav Nakov, and Zhuohan Xie. 2026. [FinCARDS: Card-based analyst reranking for financial document question answering](#). In *Findings of the Association for Computational Linguistics: ACL 2026*, pages 24836–24852, San Diego, California, United States. Association for Computational Linguistics.

A Proofs

A.1 Proof of the LMSR-geometry lemma

Put $x = z/(2b) > 0$. Differentiating [equation \(6\)](#) gives

$$D'(b) = \log \cosh x - x \tanh x.$$

As a function of x , the right-hand side has derivative

$$-x \operatorname{sech}^2 x < 0$$

and limit zero at $x = 0$. It is therefore negative for every $x > 0$, so D is strictly decreasing in b . Using $\log \cosh x = x - \log 2 + o(1)$ as $x \rightarrow \infty$ gives

$$\lim_{b \downarrow 0} D(b) = \frac{z}{2}.$$

Using $\log \cosh x = x^2/2 + o(x^2)$ as $x \downarrow 0$ gives

$$D(b) = \frac{z^2}{8b} + o(b^{-1})$$

and hence $D(b) \rightarrow 0$ as $b \rightarrow \infty$, completing the uniqueness argument. $\square$

A.2 Proof of Proposition 2

Translation invariance of C_b gives

$$C_b(q + \beta \mathbf{1} + r) - C_b(q + \beta \mathbf{1}) = C_b(q + r) - C_b(q).$$

Because $x(r + \alpha \mathbf{1}) = x(r)$, the levy is unchanged by either a state translation or a riskless overlay. Translating the traded claim gives

$$C_b(q + r + \alpha \mathbf{1}) - C_b(q) = C_b(q + r) - C_b(q) + \alpha,$$

which proves the two payment identities. The terminal payoff of the overlay also rises by α in each outcome, so state-contingent net profit is unchanged.

For pathwise no arbitrage, consider a finite history $q_{t+1} = q_t + r_t$. Exact cost differences telescope:

$$\begin{aligned} \sum_t \operatorname{Pay}_{b,\tau}(q_t, r_t) &= C_b(q_T) - C_b(q_0) \\ &\quad + \sum_{t: x(r_t) \neq 0} \frac{\tau_s(r_t)}{2} x(r_t)^2 \\ &\geq C_b(q_T) - C_b(q_0). \end{aligned}$$

Define the net position, total levy, and minimum component by

$$v = q_T - q_0, \quad v_{\min} = \min\{v_Y, v_N\}.$$

$$L := \sum_{t: x(r_t) \neq 0} \frac{\tau_s(r_t)}{2} x(r_t)^2 \geq 0.$$

Let

$$\begin{aligned} p_j(q_0) &:= \frac{e^{q_{0,j}/b}}{e^{q_{0,Y}/b} + e^{q_{0,N}/b}} \\ &\in (0, 1), \quad j \in \{Y, N\}. \end{aligned}$$

If $v_Y \neq v_N$, then

$$\begin{aligned} C_b(q_0 + v) - C_b(q_0) &= b \log \left(p_Y(q_0) e^{v_Y/b} \right. \\ &\quad \left. + p_N(q_0) e^{v_N/b} \right) \\ &> v_{\min}, \end{aligned}$$

where strictness follows because both LMSR state prices are positive. The aggregate net payoff in an outcome attaining $v_{\min}$ is therefore strictly negative after including L . If instead $v_Y = v_N = v_{\min}$, translation invariance gives

$$C_b(q_0 + v) - C_b(q_0) = v_{\min},$$

so aggregate net payoff equals $-L \leq 0$ in both outcomes. Thus no finite history has nonnegative net payoff in every outcome and strictly positive net payoff in at least one outcome.

Finally, starting from $q^0 = 0$, the worst-case loss in outcome Y is

$$\begin{aligned} &\sup_q \{q_Y - [C_b(q) - C_b(0)]\} \\ &= b \log 2 + \sup_q \{q_Y - C_b(q)\} \\ &= b \log 2, \end{aligned}$$

where the last supremum is zero and is approached as $q_Y - q_N \rightarrow \infty$. The same calculation holds in outcome N . Every nonnegative levy adds receipts and hence cannot increase this loss. $\square$

A.3 Proof of the main theorem

Fix $b > b_0$ and abbreviate $a = a(b)$. Divide expected revenue by m , and write $x = \ell_h$, $y = \ell_n$. The SCF objective is

$$F(x, y) = id(x + y) + c_h x(a - x) + c_n y(a - y). \quad (41)$$

Its Hessian is

$$\nabla^2 F = -2 \begin{pmatrix} c_h & 0 \\ 0 & c_n \end{pmatrix},$$

so F is strictly concave. The retention constraint is

$$x + y \leq E, \quad E := \frac{\delta a}{i}. \quad (42)$$

Condition [equation \(23\)](#) implies

$$\frac{\delta}{i} < \frac{n_n - n_h}{2(i + n_n)} < \frac{1}{2},$$

so $0 < E < a/2$. Both partial derivatives of F are positive throughout the feasible triangle:

$$\begin{aligned} \frac{\partial F}{\partial x} &= id + c_h(a - 2x) > 0, \\ \frac{\partial F}{\partial y} &= id + c_n(a - 2y) > 0. \end{aligned}$$

The constraint in [equation \(42\)](#) must therefore bind.

Substitute $y = E - x$ and differentiate:

$$\frac{d}{dx} F(x, E - x) = a(c_h - c_n) + 2c_n E - 2(c_h + c_n)x.$$

At $x = 0$, this derivative is negative exactly when

$$\delta < \frac{i(c_n - c_h)}{2c_n} = \frac{i(n_n - n_h)}{2(i + n_n)} = \bar{\delta}.$$

It decreases further with x , so strict concavity yields the unique SCF maximum $x = 0$, $y = E$ in [equation \(28\)](#).

Under a uniform levy ℓ , revenue per unit of arrival scale is

$$F_U(\ell) = 2id\ell + C\ell(a - \ell),$$

and feasibility requires $0 \leq \ell \leq E/2$. Because $E/2 < a/4$,

$$F'_U(\ell) = 2id + C(a - 2\ell) > 0$$

over the feasible interval. The unique uniform maximum is $\ell = E/2$, proving [equation \(29\)](#).

Substituting the two optimizers into [equation \(20\)](#) gives [equations \(30\)](#) and [\(31\)](#). Both are positive because $a, d, \delta > 0$ and

$$0 < \frac{\delta}{i} < \frac{1}{2}, \quad 0 < \frac{\delta}{I} < \frac{1}{4}.$$

The uniform optimizer corresponds to a feasible diagonal point in the SCF program. The SCF optimizer is unique and non-diagonal. Strict concavity therefore implies

$$G_D(b) > G_U(b).$$

This proves [equation \(32\)](#).

We next establish the seed-capital-normalized claims. Both G_D and G_U are continuous on (b_0, ∞) . By [theorem 1](#), $a(b) \downarrow 0$ as $b \downarrow b_0$, so

$$\lim_{b \downarrow b_0} \Psi_M(b) = 0.$$

As $b \rightarrow \infty$, $a(b) \rightarrow w_N$. Each $G_M(b)$ therefore approaches a finite constant, whereas $b \log 2 \rightarrow \infty$, so

$$\lim_{b \rightarrow \infty} \Psi_M(b) = 0.$$

Because Ψ_M is continuous and positive in the interior, it attains a finite positive maximum at an interior point. Let b_U^* maximize Ψ_U . Pointwise strictness gives

$$S_D \geq \Psi_D(b_U^*) > \Psi_U(b_U^*) = S_U,$$

proving [equation \(34\)](#).

The recoverability condition is

$$b \in \mathcal{B}_M(m) \iff m\Psi_M(b) \geq 1.$$

Below $1/S_D$, neither mechanism reaches the target. If $1/S_D \leq m < 1/S_U$, attainment of S_D makes the SCF recovery set nonempty, whereas the uniform maximum remains below the target.

Suppose $m \geq 1/S_U$ and set $q = 1/m$. Then $0 < q \leq S_U < S_D$. Pointwise strictness immediately gives $\mathcal{B}_U(m) \subseteq \mathcal{B}_D(m)$. The uniform superlevel set

$$A_U = \{b > b_0 : \Psi_U(b) \geq q\}$$

is nonempty and compact because Ψ_U is continuous and tends to zero at both endpoints. Let $\bar{b} = \sup A_U$. Continuity gives $\Psi_U(\bar{b}) = q$, while strictness gives $\Psi_D(\bar{b}) > q$. Immediately to the right of $\bar{b}$, continuity leaves $\Psi_D > q$, whereas the definition of the right endpoint gives $\Psi_U < q$. Thus some point lies in $\mathcal{B}_D(m) \setminus \mathcal{B}_U(m)$, so the inclusion is strict. $\square$

B Fixed Loss-Allowance Robustness

Proposition 7 (Liquidity-invariant per-arrival loss allowance). *Maintain the assumptions of [theorem 3](#). Fix a minimum candidate liquidity $\underline{b} > b_0$ and restrict attention to $b \in [\underline{b}, \infty)$. Replace [equation \(24\)](#) with*

$$\bar{L}_\Lambda(m) = m\Lambda, \quad 0 < \Lambda < \bar{\delta} a(\underline{b}), \quad (43)$$

which is constant across candidate liquidity levels per unit of arrival scale. Here Λ is an informed-order-loss allowance per unit of arrival scale, not an aggregate allowance independent of m or a fixed retained fraction. For every $b \geq \underline{b}$, the unique optima are

$$\begin{aligned} \ell_h^{D,\Lambda}(b) &= 0, & \ell_n^{D,\Lambda}(b) &= \frac{\Lambda}{i}, \\ \ell^{U,\Lambda}(b) &= \frac{\Lambda}{2i}. \end{aligned} \quad (44)$$

Their per-unit-arrival recovery functions are

$$G_D^\Lambda(b) = d\Lambda + c_n \frac{\Lambda}{i} \left(a(b) - \frac{\Lambda}{i} \right), \quad (45)$$

$$G_U^\Lambda(b) = d\Lambda + C \frac{\Lambda}{2i} \left(a(b) - \frac{\Lambda}{2i} \right). \quad (46)$$

They satisfy $0 < G_U^\Lambda(b) < G_D^\Lambda(b)$ for every $b \geq \underline{b}$. For each mechanism, set

$$\begin{aligned} S_M^\Lambda &:= \max_{b \geq \underline{b}} \frac{G_M^\Lambda(b)}{b \log 2}, \\ \mathcal{B}_M^\Lambda(m) &:= \{b \geq \underline{b} : mG_M^\Lambda(b) \geq b \log 2\}. \end{aligned}$$

Then $0 < S_U^\Lambda < S_D^\Lambda < \infty$, and all three arrival-threshold and strict-set conclusions of [theorem 3](#) hold after replacing $(S_M, \mathcal{B}_M)$ by $(S_M^\Lambda, \mathcal{B}_M^\Lambda)$.

B.1 Proof of Proposition 7

By [theorem 1](#), $D(b)$ is strictly decreasing, so $a(b) = w_N - D(b)$ is strictly increasing. Hence, for every $b \geq \underline{b}$,

$$\Lambda < \bar{\delta} a(\underline{b}) \leq \bar{\delta} a(b).$$

Fix such a b , abbreviate $a = a(b)$, and set

$$E_\Lambda := \frac{\Lambda}{i}.$$

The SCF retention constraint is $x + y \leq E_\Lambda$. As in the proof of [theorem 3](#), the bound on Λ implies $0 < E_\Lambda < a/2$, both partial derivatives of [equation \(41\)](#) are positive on the feasible triangle, and the retention constraint binds.

After substituting $y = E_\Lambda - x$, the derivative at $x = 0$ is

$$a(c_h - c_n) + 2c_n E_\Lambda < 0,$$

where the inequality is equivalent to $\Lambda < \bar{\delta} a$. The derivative decreases with x , so strict concavity gives the unique SCF optimum $(x, y) = (0, E_\Lambda)$. The uniform objective is strictly increasing on $[0, E_\Lambda/2]$, giving the unique common levy $E_\Lambda/2$. This proves [equation \(44\)](#). Substitution into [equation \(20\)](#) yields [equations \(45\) and \(46\)](#).

The uniform optimum is a feasible diagonal point of the SCF problem, whereas the unique SCF optimum is non-diagonal. Strict concavity therefore gives $0 < G_U^\Lambda(b) < G_D^\Lambda(b)$ pointwise on the restricted domain. The two seed-capital-normalized ratios are continuous and positive on $[\underline{b}, \infty)$, and they tend to zero as $b \rightarrow \infty$ because their numerators remain bounded while $b \log 2$ diverges. Consequently both positive finite maxima are attained, and evaluating the SCF ratio at a uniform maximizer gives $S_D^\Lambda > S_U^\Lambda$. Attainment of both maxima gives the first two arrival-threshold cases, including the equality at $1/S_D^\Lambda$. For the strict-inclusion case, let $q = 1/m \leq S_U^\Lambda$ and define

$$A_U^\Lambda(q) = \left\{ b \geq \underline{b} : \frac{G_U^\Lambda(b)}{b \log 2} \geq q \right\}.$$

Let $\bar{b}$ be the finite right endpoint of this nonempty compact set. Continuity gives $G_U^\Lambda(\bar{b})/(\bar{b} \log 2) = q$, whereas pointwise strictness gives $G_D^\Lambda(\bar{b})/(\bar{b} \log 2) > q$. Immediately to the right, the SCF ratio remains above q while the uniform ratio is below it. Hence $\mathcal{B}_U^\Lambda(m) \subsetneq \mathcal{B}_D^\Lambda(m)$. $\square$

B.2 Proof of Proposition 4

Under the symmetric benchmark, exchanging h and n leaves the SCF objective and feasible set unchanged. Strict concavity gives a unique optimizer. The exchanged optimizer must therefore equal the original one, implying $\ell_h^D = \ell_n^D$. This diagonal solution is feasible for the uniform problem with the same objective value, so the two capacities coincide. $\square$

C Additional Calculations and Extensions

C.1 Executed-order posterior

Use the shorthand

$$H_s(\ell; b) := i[a(b) + d - \ell] + n_s[a(b) - \ell].$$

The derivative of [equation \(16\)](#) is

$$\frac{\partial \rho_s^E(\ell; b)}{\partial \ell} = \frac{in_s d}{H_s(\ell; b)^2} > 0.$$

At zero levy, the difference $\rho_h(b) - \rho_n(b)$ is positive because $n_n > n_h$. No analogous cross-direction ranking is asserted after applying different levies.

C.2 Mechanism-independent screening

The allocation part of [theorem 3](#) does not depend on the logarithmic form by itself. Consider a base price-plus-fee mechanism B with:

- (a) a common opening-state participation cost $D_B(b)$ for both fixed-lot payoff directions;
- (b) a prefunded capital requirement $K_B(b) > 0$;
- (c) a nonnegative recovery levy that does not increase that capital requirement; and
- (d) the same fixed raw type-direction primitives as in [section 3](#).

On the domain where

$$a_B(b) := w_N - D_B(b) > 0,$$

the constrained levy programs yield

$$\ell_h^D = 0, \quad \ell_n^D = \frac{\delta a_B(b)}{i}, \quad \ell^U = \frac{\delta a_B(b)}{2i},$$

and the same strict pointwise ordering after replacing a by a_B in [equations \(30\)](#) and [\(31\)](#). If, in addition, the ratios

$$\Psi_M^B(b) = \frac{G_M^B(b)}{K_B(b)}$$

are continuous, vanish at the relevant domain end-points, and attain positive interior maxima, the threshold and strict-set arguments in [theorem 3](#) carry over verbatim.

This reduction explains how the screening logic may be examined with CAPM- or SQPM-style base fees: their base fee enters D_B , while their capital guarantee enters K_B . It is not a claim that those mechanisms automatically satisfy every condition above. Their participation, safety, information-incorporation, and fragmentation properties must be verified separately.

C.3 Why unrestricted splitting defeats a quadratic recovery levy

For a risky trade with $x(r) = z$, the quadratic levy is $\tau_s z^2/2$. If the trader can split it into k consecutive pieces of risky size z/k without any aggregation rule, total additional levy is

$$k \frac{\tau_s}{2} \left(\frac{z}{k}\right)^2 = \frac{\tau_s z^2}{2k} \rightarrow 0.$$

Thus a positive anonymous quadratic recovery levy cannot simultaneously preserve its stated fixed-lot revenue and be robust to arbitrary fragmentation.

D Reproducibility

The anonymous supplementary material contains the reference implementation. The deterministic audit can be reproduced with:

```
python3 -m unittest discover \
    -s tests -v
python3 run_sanity.py \
    --output-dir results
python3 run_experiments.py \
    --output-dir results
```

The frozen result files generate the paper figures:

```
MPLBACKEND=Agg python3 make_figures.py
```

D.1 Frozen experiment designs

The independent program audit reconstructs demand and revenue directly from [equations \(13\)](#), [\(14\)](#) and [\(20\)](#). For each fixed b , it enumerates the complete feasible edge of the directional loss-budget set with 1,001 points, locally refines the best basin, and separately maximizes the uniform objective over its exact feasible interval. It does not call the theorem's closed-form capacity function. The 20 cases cross four calibrations (baseline, near symmetry, low loading, and $q = 0.99$) with five headroom levels. The maximum absolute levy, capacity, and budget-residual errors are 3.34×10^{-16} , 2.22×10^{-16} , and 2.22×10^{-16} , respectively.

The boundary controls use the same raw-program solver. Signed asymmetry holds mean noise density and the baseline normalized loss budget fixed, including the symmetry null and a label permutation. The loading audit holds baseline asymmetry fixed while moving the budget across the theorem's tight-loading boundary. Selected results appear in [table 2](#); every reported levy is normalized by $a(b_D^*)$.

For the recoverable-set audit, an 8,193-point adaptive log grid locates every sign change of

Audit	Level	R (%)	ℓ_h/a	ℓ_n/a	Allocation
Signed A	-0.80	22.836	.17284	.00000	h -only
Signed A	0	0.000	.08642	.08642	split
Signed A	0.01	0.007	.08369	.08915	split
Signed A	0.20	2.829	.03180	.14104	split
Signed A	0.80	22.836	.00000	.17284	n -only
Loading q	0.99	13.802	.00000	.34222	n -only
Loading q	1.25	8.964	.06604	.36606	split
Loading q	2.00	2.023	.26415	.42721	split

Table 2: Raw-program assumption-boundary controls. A fully optimized directional menu weakly dominates uniform even outside the one-sided-loading condition because it contains the uniform menu.

m	Policy	Lower b	Upper b
1.353212	SCF	—	—
1.353212	Uniform	—	—
1.726050	SCF	0.724392	1.828707
1.726050	Uniform	—	—
2.143384	SCF	0.591848	2.664094
2.143384	Uniform	0.753474	1.625315

Table 3: Direct numerical recovery-set audit. The maximum root residual is 1.34×10^{-15} .

$m\Psi_M(b) - 1$, after which each endpoint is refined by bisection. Table 3 reports all three theorem regimes. A dash means the set is empty.

The expected-receipt suite uses the 512-point adaptive log grid and golden-section refinement stated in section 6. The structured grid has 64 asymmetry values and 61 allowance shares. The factorial and random designs contain 324 and 10,000 cases, respectively; the latter uses seed 20260812. All parameter vectors and unrounded outputs are stored as CSV files.

For the finite-cohort experiment, write $V_M = \ell_M \text{Pois}(m\lambda_M(b))$. Under the theorem’s levies, the charged-order intensities per unit m are

$$\begin{aligned}\lambda_D &= i(a + d - \ell_D) + n_n(a - \ell_D), \\ \lambda_U &= 2i(a + d - \ell_U) + (n_h + n_n)(a - \ell_U),\end{aligned}\quad (47)$$

where $\ell_D = \delta a/i$ and $\ell_U = \delta a/(2i)$. Define $k_M(b) := \lceil K(b)/\ell_M(b) \rceil$. Independent Poisson thinning of the opening state gives

$$\Pr[V_M \geq K(b)] = \Pr[\text{Pois}(m\lambda_M(b)) \geq k_M(b)]. \quad (48)$$

For each γ , define

$$m_{M,\gamma}^* = \inf_{b > b_0} \frac{\mu_\gamma(k_M(b))}{\lambda_M(b)}, \quad (49)$$

where $\mu_\gamma(k)$ is the least Poisson mean attaining upper-tail probability γ . The implementation

γ	SCF m^*	Uniform m^*	Reduction
0.50	1.48195	1.93460	23.40%
0.75	1.70165	2.13404	20.26%
0.90	1.91531	2.32479	17.61%
0.95	2.04940	2.44362	16.13%
0.99	2.31707	2.67570	13.40%

Table 4: Exact Poisson chance-recovery thresholds. Each mechanism separately chooses b before realized arrivals.

Arrival model	SCF m^*	Uniform m^*	Reduction
Independent Poisson	2.04940	2.44362	16.13%
Gamma-Poisson, $k = 10$	2.97057	3.72561	20.27%
Gamma-Poisson, $k = 2$	8.57941	11.04360	22.31%

Table 5: Exact 95% chance-recovery thresholds under a shared multiplicative arrival shock. Smaller k means stronger overdispersion.

enumerates the integer discontinuities of K/ℓ_M through 512 charged orders. Indeed, K/ℓ_M has its unique minimum at $b_x = z/[2 \operatorname{atanh}(2w_N/z)]$, while $\lambda_M(b)$ increases; within each integer cell, its right root therefore minimizes $\mu_\gamma(k)/\lambda_M(b)$. The excluded region is certified using $\mu_\gamma(513)/\lambda_M^\infty$. Even at $\gamma = 0.50$, these lower bounds are 22.49 for SCF and 14.88 for uniform, far above the incumbent thresholds. The resulting values are in table 4.

For the common-shock stress test, let $U \sim \text{Gamma}(k, k)$ be drawn once per recovery window and replace every conditional charged-order intensity by $mU\lambda_M(b)$. Marginally, the charged count is negative binomial with mean $m\lambda_M(b)$ and variance $m\lambda_M(b) + m^2\lambda_M(b)^2/k$. We invert its exact recurrence and use the same integer-cell enumeration and excluded-region certificate as in the Poisson calculation. Results for the 95% target are in table 5.

The Poisson assumption also makes the screened informed-order counts distributionally comparable. SCF screens $\text{Pois}(m\ell_D) = \text{Pois}(m\delta a)$, while the sum across the two uniformly charged directions is $\text{Pois}(2m\ell_U) = \text{Pois}(m\delta a)$. This is equality in marginal distribution, not pathwise equality.

For the distributional stress test, let $g_{\alpha,\beta}$ be a Beta density on $[0, 1]$ and replace the raw intensity by $m c_{T,s} g_{\alpha_T,\beta_T}(W/w_T) dW$. Its integral remains $m c_{T,s} w_T$, and $\text{Beta}(1, 1)$ reproduces the main model. If F_T is the associated CDF, the executed mass per unit m is

$$c_{T,s} w_T \left[1 - F_T \left(\frac{D(b) + \ell_s}{w_T} \right) \right], \quad (50)$$

Informed shape	Noise-surplus shape			
	L	U	C	H
L	18.489	32.384	31.077	38.345
U	6.463	22.836	21.874	30.436
C	3.133	16.535	16.249	25.504
H	0.203	0.689	0.030	3.748

Table 6: Threshold reduction R (percent) outside the theorem under Beta surplus shapes. L, U, C, and H denote Beta(1, 3), Beta(1, 1), Beta(2, 2), and Beta(3, 1), respectively.

and the informed loss in direction s is

$$iw_I \left[F_I \left(\frac{D(b) + \ell_s}{w_I} \right) - F_I \left(\frac{D(b)}{w_I} \right) \right]. \quad (51)$$

The optimizer allows the common loss budget to be slack, refines every grid-detected one-dimensional local maximum after reparameterizing levies by informed loss, and then searches liquidity on a tail-certified finite log interval. A production grid of 65 inner by 193 outer points is checked against 129 by 385 points; the largest absolute change in R is 4.89×10^{-15} .

The outputs are formula checks and synthetic stress tests, not observations from a live market.