Attention Baiting and the Survival of Charlatans*

Ali Kakhbod
UC Berkeley

Dmitry Livdan
UC Berkeley, CEPR

Norman Schürhoff
University of Lausanne, SFI, CEPR

August 26, 2026

VERY PRELIMINARY, NOT FOR CITING

*We thank Bloomberg and StockTwits for providing the data. Ali Kakhbod is with the Walter A. Haas School of Business, University of California, Berkeley, Berkeley, CA 94720. e-mail: akakhbod@berkeley.edu. Dmitry Livdan is with the Walter A. Haas School of Business, University of California, Berkeley, Berkeley, CA 94720. e-mail: livdan@haas.berkeley.edu. Livdan is also a Research Fellow of the CEPR. Norman Schürhoff is with the Faculty of Business and Economics and Swiss Finance Institute, University of Lausanne, CH-1015 Lausanne, Switzerland. e-mail: norman.schuerhoff@unil.ch. Schürhoff is also a Research Fellow of the CEPR. He gratefully acknowledges research support from the Swiss Finance Institute and the Swiss National Science Foundation under SNF project #100018_192584, "Sustainable Financial Market Infrastructure: Towards Optimal Design and Regulation" and has no conflict of interest to disclose.

Attention Baiting and the Survival of Charlatans

Abstract

When attention rather than price allocates demand, competition need not select for quality. We develop a model in which followers learn rationally from noisy outcomes while engagement makes visibility superlinear in audience size. Because beliefs about advice quality adjust slowly while visibility compounds, early attention can lock in on poor advisers. Content slant can make this selection systematically adverse. In data from a social trading platform, distinguishing an expert from a charlatan requires far longer track records than most accounts produce. Using posterior classifications based only on outcomes realized before portfolio formation, stocks with greater charlatan exposure earn 0.90% lower five-factor-adjusted returns over the next 20 trading days, and bullish calls by charlatan accounts earn −0.59%. The relation survives within-stock controls, and reversals are larger when shorting is costly. The results show that stronger preferences for quality improve selection but do not eliminate charlatan lock-in, even with rational learning.

JEL Classification: G12, G14, G41

Key words: attention markets, advice quality, social learning, machine learning, reinforcement, social media

Markets for expertise usually contain institutions that help demand migrate toward quality. Clients pay advisers, investors move capital toward managers who perform, and fees, flows, monitoring, and disclosure gradually discipline low-quality providers (Berk and Van Binsbergen, 2022; Gervais, Lynch, and Musto, 2005; Gervais and Strobl, 2020). Markets organized around attention work differently. Advice is often free, visibility is rewarded by engagement, and audiences must infer quality from noisy realized outcomes. The central allocation problem is therefore dynamic. Quality is learned gradually, while visibility can accumulate immediately. This paper asks what happens when attention compounds faster than evidence about quality arrives.

Financial advice on social media provides a useful setting in which to study this problem because advice quality can be evaluated against subsequent returns. We study active accounts on a large social trading platform whose signed stock opinions can be matched to CRSP outcomes. We call an account an expert when acting on its advice has positive persistent expected signed abnormal returns and a charlatan when that expected value is negative. The labels refer to the economic value of advice rather than to fraud or intent. The relevant question is not simply whether poor advisers are present. It is whether audiences can identify them before engagement and accumulated popularity determine who continues to be seen.

We develop a model in which two selection forces operate on different clocks. Informational selection works through Bayesian beliefs about advice quality. Recommendation returns are noisy, so beliefs move slowly, and even perfect learning can affect follower choice only through a bounded preference for perceived quality. Attention selection works through visibility. A larger audience raises the return to engagement effort, effort raises platform exposure, and additional exposure attracts more audience. This complementarity makes visibility superlinear in audience size. Small differences in initial prominence or early arrival luck can therefore be amplified before track records contain enough information to distinguish good advice from bad.

This difference in speeds produces a selection failure without requiring irrational followers. Under superlinear reinforcement, one account eventually receives almost all new attention, and a charlatan can become the winner with positive probability for every finite preference for quality. In the slow-learning limit, long-run attention reflects initial audience, engagement efficiency, and early luck rather than advice quality. The linear benchmark shows what drives the result. When amplification is only linear, early accidents do not generate the same lock-in, and sufficiently strong informational selection can govern long-run attention.

Attention need not merely fail to select for quality. It can select against it. We allow accounts to choose optimistic, trend-confirming content that raises engagement. Charlatan accounts have stronger incentives to use this content because they sacrifice less reputation value by doing so, while experts have more to lose from degrading a record that can eventually support a reputation for quality. Optimistic slant therefore creates an early attention advantage precisely where advice value is lower. Superlinear reinforcement can preserve that advantage after information improves. A pricing block then maps the resulting composition of visible advice into retail demand and prices.

Optimistic charlatan content generates overpricing and subsequent reversal, with stronger effects when short selling is costly.

The empirical analysis follows this mechanism rather than a taxonomy of social-media users. We estimate latent advice quality from noisy signed recommendation returns and construct expanding-window posterior probabilities that each account is an expert or a charlatan. The principal specifications are outcome-embargoed. At every classification date, each recommendation used to estimate quality has completed its full 20-trading-day outcome window, and the main return tests use only information available before portfolio formation. The posterior probabilities are therefore measurement inputs for the economic mechanism rather than ex post labels chosen to explain subsequent returns.

The first result is that track records accumulate far too slowly to discipline most accounts. Median per-recommendation volatility is 8.6%, while economically relevant expert–charlatan differences in expected returns are only a few percentage points per month. Even under an optimistic Bayesian benchmark, distinguishing the two therefore requires far more observations than most accounts ever generate. The median account produces seven scored recommendations during the three-and-a-half-year sample. Fewer than 4% of accounts reach the detection threshold implied by the largest-weight quality components, and fewer than 1% reach the threshold implied by the typical expert–charlatan gap. Attention is thus allocated in an environment in which reliable information about quality is scarce precisely when audience positions are being established.

The second result clarifies what followers are trying to learn. When we re-estimate advice quality on disjoint halves of the sample, the population distribution reappears but individual realized alpha ranks barely persist. Behavioral strategies and stock habitats are substantially more stable. Moreover, first-half optimism predicts lower advice quality in the second half even after controlling for first-half estimated quality. The persistent object is therefore better understood as an advice strategy with a stable unconditional payoff than as a realized alpha that repeats in every market environment. This distinction strengthens the learning friction: followers must infer persistent strategy through returns that also reflect changing market habitats.

The pricing evidence is formed strictly in real time. Stocks with greater posterior charlatan exposure earn 0.90% lower five-factor-adjusted returns over the next 20 trading days. The relation is not simply a comparison between different kinds of stocks. Within the same stock, a one-standard-deviation increase in charlatan exposure predicts roughly 0.20% lower subsequent returns after stock and time fixed effects and controls for momentum, size, turnover, liquidity, volatility, retail ownership, and shortability. At the account level, bullish recommendations by accounts in the top decile of the prior charlatan posterior earn -0.59% over the following 20 trading days out of sample. Bearish recommendations do not produce a corresponding positive return. Poor advice is therefore not a signal that investors can mechanically invert; its predictive content is concentrated in optimistic calls.

The mechanism tests point in the same direction. Bullish charlatan calls are followed by returns

of −1.98% in hard-to-short stocks and only −0.27% in easy-to-short stocks. High charlatan exposure is associated with greater subsequent volatility, and the coverage–return relation is substantially stronger outside the low-noise segment, where advice quality is harder to infer. Retail order flow comoves strongly with charlatan activity unconditionally, but the incremental relation disappears after controlling for recent returns, overall platform activity, and the visible broadcast. The evidence is therefore consistent with a common optimistic signal reaching investors through the platform rather than with investors responding independently to the identity of charlatan accounts.

A data-guided simulation combines the estimated distribution of advice quality and recommendation noise with the observed relation between activity and audience. Stronger preferences for perceived quality improve selection, but they do not eliminate the possibility that attention locks onto poor advice. Greater amplification offsets part of the gain from better informational selection. The exercise is illustrative rather than structural, but it makes the economic distinction transparent. Policies that improve information about quality strengthen a bounded informational force; policies that reduce amplification alter the mechanism through which early attention becomes persistent.

The paper contributes to the economics of expertise and digital platforms by studying the allocation of demand when quality is learnable but attention compounds before it is learned. Berk and Van Binsbergen (2022) study the persistence of charlatans when quality is difficult to observe, while Acemoglu, Ozdaglar, and Siderius (2024) and Acemoglu, Makhdoumi, Malekian, and Ozdaglar (2025) show how engagement objectives and slow learning can distort digital environments. Our mechanism combines noisy but objectively measurable performance, endogenous content choice, and superlinear audience reinforcement. It also connects this allocation problem to asset prices by showing that the composition of visible advice predicts returns in real time and that the effect varies with limits to arbitrage.

Related work on StockTwits documents heterogeneity in advice quality and its association with audience size (Kakhbod, Kazempour, Livdan, and Schürhoff, 2025). We instead study why competition for attention need not move audiences toward better advice. The analysis also relates to work on aggregate social-media signals (Cookson and Niessner, 2020; Cookson, Lu, Mullins, and Niessner, 2024), information siloing (Cookson, Engelberg, and Mullins, 2023), social learning and homophily (Golub and Jackson, 2012; Currarini, Jackson, and Pin, 2009), extrapolative beliefs (Greenwood and Shleifer, 2014), and performance evaluation (Harvey and Liu, 2018; Crane and Crotty, 2020; Dim, 2022).

Our sample covers July 2013 through the end of 2016, before the Covid-19 and meme-stock episodes that later made social-media-driven trading especially visible. Those episodes are not required for the mechanism. The argument requires only that performance be noisy, visibility respond to engagement, and accumulated audience influence future exposure. Whenever attention can compound faster than quality can be learned, competition for visibility can preserve poor advice rather than eliminate it.

The paper proceeds as follows. Section 1 develops the model. Section 2 describes the data

and measurement. Section 3 quantifies learnability and distinguishes persistent strategies from unstable realized alpha. Section 4 disciplines amplification and endogenous optimism. Section 5 presents the real-time pricing and mechanism tests. Section 6 provides the quantitative illustration and a first look at attention-market design. The appendices contain proofs and the measurement, classification, and mechanism-validation evidence.

1 A Model of Advice in Attention Markets

We begin with a model that isolates the economic mechanism. The model takes as primitives a distribution of advice quality, noisy realized returns from following advice, and an attention technology that determines how audiences find advisers. Later sections estimate these objects directly and test the model's predictions. The model has five blocks. They are (i) types and Bayesian learning from noisy track records, (ii) an engagement choice that makes platform visibility superlinear in audience size, (iii) attention dynamics in the form of a nonlinear Pólya urn, (iv) an endogenous choice of optimistic content slant, and (v) a pricing block in which short-sale-constrained retail investors trade on visibility-weighted sentiment. Blocks (i) through (iii) explain why attention can flow to charlatans, block (iv) explains why popularity can select against skill rather than merely ignoring it, and block (v) maps the biased visible signal into prices. Figure 1 summarizes the two selection forces, and Figure 2 in Section 1.4 maps the resulting attention allocation into prices. The exposition stays at the level of the mechanism. All proofs are in Appendix A, and every quantitative statement about where the data place the model's parameters is deferred to Sections 2 and 3.

1.1 Advice quality and learning

A platform hosts $N \geq 2$ accounts that provide financial advice, indexed by i. Each account has type $\theta_i \in \{\mathrm{E}, \mathrm{C}\}$, where E denotes an expert and C a charlatan, with $\Pr(\theta_i = \mathrm{E}) = P_{\mathrm{E}} \in (0, 1)$. An account observes its own type, while followers learn about the type from its realized track record. We use expert and charlatan to distinguish advice with positive and negative expected value. The labels refer to the economic value of acting on the advice rather than to intent or fraud.

The kth recommendation of account i generates the signed abnormal return

$$r_{i,k} = \alpha_{\theta_i} + \sigma_r \varepsilon_{i,k}, \qquad \varepsilon_{i,k} \overset{iid}{\sim} \mathcal{N}(0, 1), \tag{1}$$

where $\alpha_{\mathrm{E}} > 0 > \alpha_{\mathrm{C}}$. Define the quality gap as $\Delta \equiv \alpha_{\mathrm{E}} - \alpha_{\mathrm{C}} > 0$. A larger quality gap makes type easier to infer, while greater return noise makes realized performance less informative.

The type in equation (1) represents a persistent advice strategy with an unconditional expected payoff. It does not require an account to earn the same realized alpha in every market environment. Section 3.2 shows that behavior is substantially more persistent than realized account-level alpha

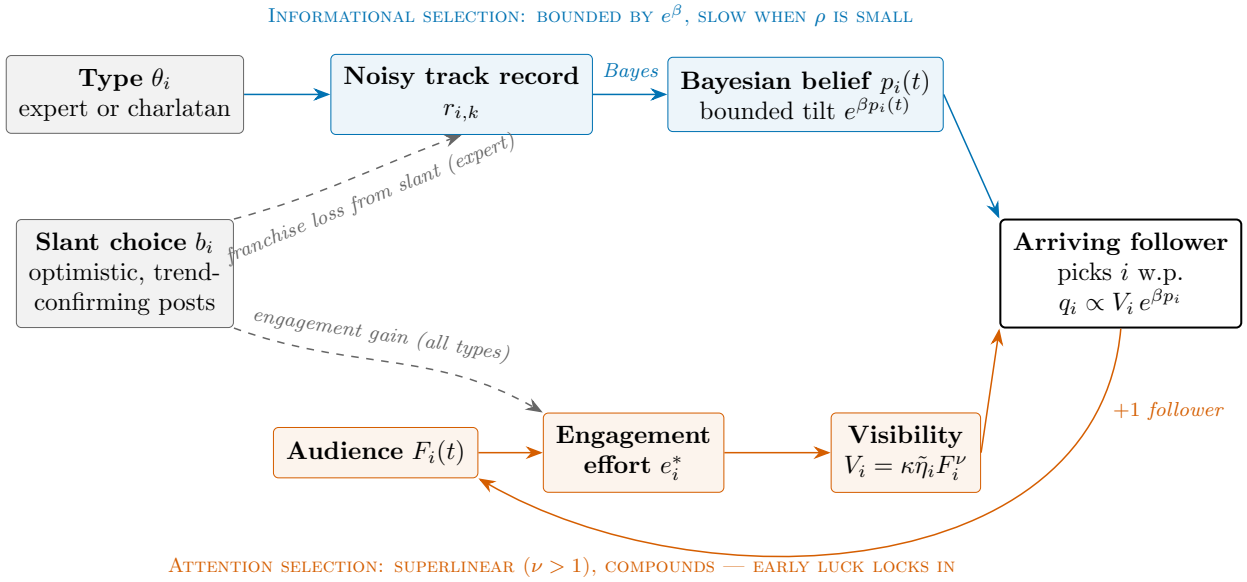


Figure 1: Model Overview: Two Selection Forces in the Market for Advice

The figure summarizes the attention block of the model of Section 1 and shows the two forces that compete for the arriving follower. In the top channel (informational selection), the persistent type θ_i generates a noisy track record $r_{i,k} = \alpha(\theta_i, b_i) + \sigma_r \varepsilon_{i,k}$ (equation 1), followers update the belief $p_i(t)$ slowly at rate $\rho = \Delta^2/(2\sigma_r^2)$, where $\Delta = \alpha_{\mathrm{E}} - \alpha_{\mathrm{C}}$ is the quality gap (Lemma 1), and beliefs enter the follower's choice (equation 5) only through the bounded tilt $e^{\beta p_i} \in [1, e^{\beta}]$. In the bottom channel (attention selection), audience F_i raises the return to engagement effort e_i^* (equation 3), the feed algorithm amplifies engagement into visibility $V_i = \kappa\tilde{\eta}_i F_i^{\nu}$ with $\nu > 1$ and effective efficiency $\tilde{\eta}_i$ (equation 4), and each arriving follower feeds back into audience, a superlinear reinforcement loop that monopolizes attention and can lock in on a charlatan account (Proposition 1). The slant choice b_i (dashed) raises engagement during the early attention phase, when performance histories are too short to reveal quality, so the engagement advantage can become self-reinforcing; a direct reputation franchise later makes expert accounts protect advice quality even though accumulated follower counts are sticky, and superlinear reinforcement links the horizons because an early audience lead is unlikely to be overturned by any bounded later informational tilt, so optimism marks negative advice value (Proposition 2). Figure 2 maps the resulting attention allocation into prices.

across disjoint periods. Equation (1) abstracts from this additional source of variation and therefore gives an optimistic benchmark for learning.[1]

The expected payoff α_{θ_i} is measured from the perspective of a follower who acts at the first post-message price. It can therefore reflect both information in the recommendation and any price movement induced before the follower trades. Section 2.3 discusses this interpretation and estimates the latent distribution of advice quality from realized recommendation returns.

After n recommendations, the public record is $\mathcal{R}_i(n) = (r_{i,1}, \ldots, r_{i,n})$. A Bayesian follower forms

[1] A reduced-form extension is $r_{i,k,t} = \alpha_{\theta_i} + \xi_{i,t} + \sigma_r \varepsilon_{i,k,t}$, where $\xi_{i,t}$ is a mean-zero account-period component. Recommendations within a period average out the idiosyncratic term but not $\xi_{i,t}$. The extension therefore slows learning relative to equation (1). We retain the simpler specification for the formal attention results.

the posterior $p_i(n) = \Pr(\theta_i = \mathrm{E} \mid \mathcal{R}_i(n))$. Under equation (1), log posterior odds satisfy

$$L_i(n) \equiv \ln \frac{p_i(n)}{1 - p_i(n)} = \ln \frac{P_\mathrm{E}}{1 - P_\mathrm{E}} + \frac{\Delta}{\sigma_r^2} \sum_{k=1}^{n} \left(r_{i,k} - \frac{\alpha_\mathrm{E} + \alpha_\mathrm{C}}{2} \right). \tag{2}$$

The amount of information contained in one recommendation is summarized by the learning rate $\rho \equiv \Delta^2/(2\sigma_r^2)$. Under an expert account, each recommendation raises expected log posterior odds by ρ. Under a charlatan account, it lowers them by the same amount.

Lemma 1 (Speed of learning) *Under equation (1), the following statements hold.*

(a) $p_i(n) \to \mathbb{1}\{\theta_i = \mathrm{E}\}$ *almost surely.*

(b) The increment in $L_i(n)$ *has mean* $+\rho$ *for an expert account and* $-\rho$ *for a charlatan account, with variance* 2ρ*.*

(c) Let $\ell_0 = \ln[P_\mathrm{E}/(1 - P_\mathrm{E})]$ *and* $\ell_{\delta_\ell} = \ln[(1 - \delta_\ell)/\delta_\ell]$ *for* $\delta_\ell \in (0, 1/2)$ *with* $|\ell_0| < \ell_{\delta_\ell}$*. Expected log posterior odds reach the corresponding confidence boundary after* $n^*_\mathrm{E}(\delta_\ell) = (\ell_{\delta_\ell} - \ell_0)/\rho$ *for an expert and* $n^*_\mathrm{C}(\delta_\ell) = (\ell_{\delta_\ell} + \ell_0)/\rho$ *for a charlatan. With an even prior, both reduce to* $n^*(\delta_\ell) = \ell_{\delta_\ell}/\rho$*.*

The lemma separates eventual learning from timely learning. Bayesian beliefs are consistent, so the mechanism does not rely on followers remaining permanently mistaken about advice quality. The difficulty is that the number of observations required for a given level of confidence is proportional to σ_r^2/Δ^2. Appendix A shows that the corresponding expected sequential stopping times are weakly longer because posterior odds can overshoot the confidence boundary.

Whether reputation can discipline the allocation of attention is therefore an empirical question. It depends on the size of the quality gap relative to recommendation noise and on the length of the track record followers observe. Section 3.1 estimates these objects and finds that most accounts generate far fewer scored recommendations than the model requires for reliable classification. The next subsection asks what happens when audience can accumulate while this learning is still taking place.

1.2 Attention accumulation and lock-in

Each account begins with an audience $F_i(0) > 0$. New units of attention arrive sequentially. An arrival represents a new follow edge or another unit of persistent audience rather than a person who can follow only one account. Each arriving unit attaches to one account, thereby increasing that account's audience by one.

Accounts can increase their exposure by generating engagement. An account with audience $F_i(t)$ chooses effort $e_i(t) \geq 0$ to maximize

$$\pi_i e_i(t) F_i(t) - c_i \frac{e_i(t)^{1+\chi}}{1+\chi}, \qquad \chi > 0. \tag{3}$$

Effort includes posting, replying, and other activity that generates engagement. The parameter $\pi_i > 0$ is the return from engaging an existing follower and $c_i > 0$ governs the cost of effort. The optimal choice is $e_i^*(t) = [\pi_i F_i(t)/c_i]^{1/\chi}$.

The platform surfaces content in proportion to the engagement it generates. Account i therefore receives exposure

$$V_i(t) = \kappa e_i^*(t) F_i(t) = \kappa \eta_i F_i(t)^\nu, \qquad \eta_i \equiv \left(\frac{\pi_i}{c_i}\right)^{1/\chi}, \qquad \nu \equiv 1 + \frac{1}{\chi} > 1. \tag{4}$$

The scale parameter $\kappa > 0$ is common across accounts. The economically important parameter is ν. Exposure is superlinear in audience because a larger audience raises the return to engagement effort, greater effort raises exposure, and greater exposure attracts additional audience. The complementarity between effort and audience therefore creates reinforcement rather than imposing it directly on follower choice.

This mechanism moves the usual investor-attention argument upstream. Existing work shows that attention can shift investor demand toward salient securities (Barber and Odean, 2008; Da, Engelberg, and Gao, 2011). Here the object receiving attention is itself endogenous. Existing audience changes the effort supplied by an adviser and therefore changes the probability that the adviser enters the consideration set of future followers.

The model begins from a benchmark in which endowed attention advantages contain no information about advice quality.

Assumption 1 (Type-independent initial conditions) *The pair $(\eta_i, F_i(0))$ is independent of θ_i. The triples $(\eta_i, F_i(0), \theta_i)$ are independent across accounts, and the joint distribution of $(\eta_i, F_i(0))$ has bounded support in $(0, \infty)^2$.*

Assumption 1 isolates endogenous selection. Initial audience and endowed engagement efficiency do not favor either type. Section 1.3 introduces a content choice that changes effective engagement efficiency and therefore makes attention selection endogenous to type.

Attention as a reinforced urn. Followers encounter accounts through the platform feed. Conditional on exposure, a follower values account i according to the posterior probability $p_i(t)$ that acting on its advice creates value. Let $\beta \geq 0$ measure the strength of this preference. With standard

extreme-value taste shocks, the probability that the next unit of attention attaches to account i is

$$q_i(t) = \frac{V_i(t)e^{\beta p_i(t)}}{\sum_j V_j(t)e^{\beta p_j(t)}} = \frac{\eta_i F_i(t)^\nu e^{\beta p_i(t)}}{\sum_j \eta_j F_j(t)^\nu e^{\beta p_j(t)}}. \tag{5}$$

Equation (5) defines a generalized Pólya urn. Each account is a color, its current audience F_i is the number of balls of that color, and each arriving follower adds one ball to the color it chooses. Unlike the classical Pólya urn, colors need not have equal attractiveness. Account i is sampled with weight $\eta_i F_i^\nu e^{\beta p_i}$. The factor η_i is its endowed engagement efficiency, F_i^ν is reinforcement from accumulated audience, and $e^{\beta p_i}$ is the follower's preference for expected advice quality.

The urn representation makes the feedback transparent. If an account receives an early follower, its audience rises. A larger audience raises its future sampling weight, which makes another follower more likely, which raises its sampling weight again. When $\nu > 1$, this feedback is superlinear. A small early lead can therefore become self-reinforcing even when it was initially generated by luck.

The urn also makes clear that two selection forces operate on different scales. Informational selection enters through $e^{\beta p_i}$. Because $p_i \in [0, 1]$, the largest possible informational advantage of an expert over a charlatan is the finite factor e^β. Attention selection enters through $\eta_i F_i^\nu$. When $\nu > 1$, the relative advantage generated by audience is not bounded because audience itself evolves with past choices.

This source of path dependence differs from informational herding. In Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992), predecessors' choices matter because they reveal information. In Scharfstein and Stein (1990), reputational concerns can induce agents to imitate one another. Here previous follower choices matter even if they reveal nothing. They increase the number of balls of one color in the attention urn and thereby increase that account's probability of being selected again.

A two-account benchmark gives the deterministic counterpart. Suppose one account is an expert and one is a charlatan, and audience evolves according to $\dot{F}_i(t) = \lambda q_i(t)$. If at some date t_0 the inequality $\eta_C F_C(t_0)^{\nu-1} > e^\beta \eta_E F_E(t_0)^{\nu-1}$ holds, it remains satisfied thereafter regardless of the path of beliefs. The charlatan's audience relative to the expert's then diverges and its share of newly arriving attention converges to one. Appendix A proves the result.

The benchmark identifies the economic threshold. Once accumulated audience generates an advantage larger than the greatest informational premium an expert can ever receive, subsequent learning cannot reverse the allocation. The stochastic urn explains how an account can reach this region. Early follower arrivals are random, and superlinear reinforcement magnifies those small initial differences.

For the exact sequential result, the account's public record is indexed by its own follower counter. Between successive follower arrivals to an account, one additional scored recommendation enters that account's record.[2] Say that an account strongly monopolizes attention if it eventually receives

[2]This timing makes the exponential embedding in Appendix A exact. It is a proof convention rather than a claim

all but finitely many arriving units.

Proposition 1 (Attention lock-in) *Suppose attention is allocated according to equation (5).*

(a) If $\nu > 1$, a unique account strongly monopolizes attention almost surely.

(b) For every finite β, if the realized type profile contains a charlatan account, the probability that a charlatan strongly monopolizes attention is strictly positive.

(c) Under Assumption 1, as $\rho \to 0$ with β fixed, each account's monopoly probability converges to its monopoly probability in the superlinear urn with weights $\eta_i F_i^\nu$. Conditional on $(\eta_i, F_i(0))_{i=1}^N$, the limiting winner is independent of type.

(d) If $\nu = 1$, strong monopoly does not occur. Every account receives infinitely many arrivals. If one account has a unique maximal asymptotic growth rate, its audience share converges to one.

Part (a) is the basic urn result. Superlinear reinforcement eventually selects one color. Almost all subsequent attention flows to the same account. Part (b) shows why rational learning does not guarantee efficient selection. A charlatan can obtain the favorable sequence of early draws needed to become the winning color, and the probability of doing so remains positive for every finite responsiveness to quality.

Part (c) identifies the role of slow learning. When ρ is small, recommendation outcomes move posteriors slowly, so the factors $e^{\beta p_i}$ remain nearly common while the urn begins to reinforce its early draws. The process then approaches a pure attention urn with weights $\eta_i F_i^\nu$. Under Assumption 1, these weights contain no information about type. In the slow-learning limit, the identity of the winner is therefore unrelated to advice quality.

Part (d) shows why superlinearity matters. With $\nu = 1$, every color continues to be drawn infinitely often and early luck cannot generate strong monopoly. Once beliefs converge, the asymptotic growth rate is $\eta_i e^\beta$ for an expert and η_i for a charlatan. A sufficiently high-engagement charlatan can still become the largest account, but a finite early lead cannot overturn a strict difference in long-run growth rates.

The baseline urn therefore produces misallocation but not adverse selection. When learning is slow, attention can lock in on an expert or a charlatan for reasons unrelated to quality. The next subsection adds the missing force. Accounts choose content that changes their effective urn weights. If low-quality accounts have stronger incentives to choose the attention-enhancing strategy, the same reinforcement process that previously ignored quality begins to select systematically against it.

that recommendation and follower arrivals are mechanically linked in the data. The endogenous-content extension in Section 1.3 instead uses block timing.

1.3 Endogenous slant and adverse selection

We now allow accounts to change the weights with which they enter the attention urn. This is the mechanism that turns random lock-in into adverse selection.

Account i chooses a persistent slant intensity $b_i \in \{0, \bar{b}\}$, where $0 < \bar{b} < 1$. The choice $b_i = \bar{b}$ means that a fraction $\bar{b}$ of the account's content uses attention-oriented bait. In the tractable specialization below, bait consists of optimistic calls on recent winners. The empirical analysis measures the broader framing strategy using hype, certainty, price targets, bullish trend references, calls to action, and self-promotion. The model compresses these choices into the single variable b_i.

For this subsection, normalize $\alpha_{\mathrm{E}} = a > 0$ and $\alpha_{\mathrm{C}} = -a$. The normalization is used only to obtain a simple separating equilibrium. The learning and baseline urn results above do not require symmetry.

Slant affects both the attention payoff and advice quality.

Assumption 2 (Payoffs from slant) *The following conditions hold.*

(a) Slant raises engagement. The per-follower reward in equation (3) is $\pi_i g(b)$, where $g(0) = 1 < g(\bar{b})$.

(b) A baited recommendation has expected signed abnormal return $-h$, where $h > a$.

The first condition is the attention-persuasion channel. Bait makes content more effective at generating engagement. Because engagement determines exposure, it changes the weight with which an account competes for the next unit of attention. The second condition gives bait an economic cost. In the specialization used here, bait directs optimistic advice toward recent winners when their subsequent reversal is strongest.

Substituting the engagement reward $\pi_i g(b)$ into equation (3) gives optimal effort $e_i^*(b, F) = [\pi_i g(b) F / c_i]^{1/\chi}$. Exposure and the optimized engagement payoff are

$$V_i(b, F) = \kappa \tilde{\eta}_i(b) F^{\nu}, \quad \Pi_i^*(b, F) = A_i g(b)^{\nu} F^{\nu}, \quad \tilde{\eta}_i(b) \equiv \eta_i g(b)^{1/\chi}, \quad A_i \equiv \frac{\chi}{1+\chi} \pi_i^{\nu} c_i^{-1/\chi}. \tag{6}$$

The key object is $\tilde{\eta}_i(b)$. Without bait, account i enters the attention urn with efficiency η_i. With bait, its efficiency becomes $\eta_i g(\bar{b})^{1/\chi}$. Since $g(\bar{b}) > 1$, bait multiplies every future urn weight generated by a given audience size. Holding beliefs and audience fixed, the account is therefore more likely to receive the next follower at every draw.

Slant also changes the account's track record. A scored recommendation has expected return

$$r_{i,k}(b_i) = \alpha(\theta_i, b_i) + \sigma_r \varepsilon_{i,k}, \qquad \alpha(\theta, b) = (1 - b)\alpha_\theta - bh. \tag{7}$$

An unslanted charlatan has expected return $-a$, while a charlatan choosing $\bar{b}$ has expected return $-a - \bar{b}(h - a)$. Bait therefore lowers charlatan advice value by $\bar{b}(h - a) > 0$.

The question is now which type chooses the larger urn weight. Both types would like the extra engagement generated by bait. They differ in what they lose from producing it. An expert can preserve a valuable reputation by maintaining a good return record. A charlatan eventually receives a poor reputation even without bait and therefore sacrifices less by using the engagement-enhancing strategy.

To separate the period in which attention is won from the period in which reputation becomes informative, attention arrives in two blocks. Early followers observe only a short and noisy performance history. Later evaluators observe a longer record. Let p_i^E and p_i^L denote the corresponding posterior probabilities that account i is an expert.

The early block is itself a reinforced urn. Its sampling probability is

$$q_i^E = \frac{\tilde{\eta}_i(b_i)(F_i^E)^\nu e^{\beta p_i^E}}{\sum_j \tilde{\eta}_j(b_j)(F_j^E)^\nu e^{\beta p_j^E}}. \tag{8}$$

Relative to the baseline urn in equation (5), the only new object is $\tilde{\eta}_i(b_i)$. Content choice therefore enters the attention process as a multiplicative change in the account's urn weight.

The information structure makes the timing economically consequential. Let $\mathcal{I}^*$ denote the information about type contained in one equilibrium recommendation. We consider a sequence in which

$$n_E\mathcal{I}^* \to 0, \qquad n_L\mathcal{I}^* \to \infty. \tag{9}$$

Early records are therefore too short to distinguish types reliably, while later records eventually reveal advice quality. Appendix A gives the corresponding conditions for noisy behavioral information about slant.

The first limit has a particularly simple implication for the urn. During the early attention block, $p_i^E \to P_{\mathrm{E}}$ for every account, so the informational factors $e^{\beta p_i^E}$ converge to the same constant. Common factors cancel from equation (8). The early allocation therefore converges to a pure superlinear urn with weight

$$\eta_i g(b_i)^{1/\chi} F_i^\nu$$

for account i.

This is where bait changes selection. In the separating equilibrium derived below, experts choose $b_{\mathrm{E}}^* = 0$ and charlatans choose $b_{\mathrm{C}}^* = \bar{b}$. Conditional on the same endowed efficiency, initial audience, and early luck, an expert therefore enters the limiting urn with weight $\eta_i F_i^\nu$, while a charlatan enters with the strictly larger weight $\eta_i g(\bar{b})^{1/\chi} F_i^\nu$. Bait does not merely raise contemporaneous engagement. It changes the stochastic process that determines who ultimately receives attention.

To complete the incentive argument, suppose an account also receives a reputation franchise tied to recognized quality. Let $\Phi(p_i^L)$ be increasing in the late posterior and let $\Omega_i > 0$ measure the value of this franchise. It can represent professional opportunities, subscriptions, access to sophisticated clients, or other benefits that depend on being recognized as a high-quality adviser rather than

simply having a large audience.

For experts, assume that the loss of reputation from being classified as a charlatan exceeds the largest possible engagement gain from slant:

$$\Omega_i[\Phi(1) - \Phi(0)] > A_i g(\bar{b})^{\nu} \Delta_{\Sigma}. \tag{10}$$

Here Δ_{Σ} bounds the total weight placed on engagement payoffs across phases. Appendix A gives the exact normalized payoff representation. The condition allows the audience itself to become highly persistent because reputation has value outside the accumulated mass audience.

The separating equilibrium also requires

$$a < h, \qquad \frac{2a}{3a + h} < \bar{b} < \min\left\{1, \frac{2a}{h - a}\right\}. \tag{11}$$

These bounds ensure that the long return record classifies the equilibrium strategies and the relevant deviations in the direction required for separation. Appendix A derives the inequalities.

Proposition 2 (Slant and adverse selection) *Suppose Assumptions 1 and 2, equations (9), (10), and (11) hold along the sequence of economies described above, with the early mass of attention tending to infinity. For all sufficiently large economies, there exists a separating equilibrium with $b^*_{\mathrm{C}} = \bar{b}$ and $b^*_{\mathrm{E}} = 0$. No uniqueness claim is made. In this equilibrium:*

(a) Greater observed slant raises the posterior probability that an account is a charlatan, conditional on its late return record.

(b) Bait raises the charlatan's effective urn efficiency from η_i to $\eta_i g(\bar{b})^{1/\chi}$.

(c) The early allocation converges to a superlinear urn with weights $\eta_i g(\bar{b})^{1/\chi} F_i^{\nu}$ for charlatan accounts and $\eta_i F_i^{\nu}$ for expert accounts. Consequently, if π^E_{C} denotes the limiting probability that the early strong monopolist is a charlatan, then

$$\pi^E_{\mathrm{C}} > 1 - P_{\mathrm{E}}. \tag{12}$$

The probability that the early strong monopolist remains the eventual monopolist converges to one.

(d) Slant lowers charlatan advice quality by $\alpha(\mathrm{C}, \bar{b}) - \alpha(\mathrm{C}, 0) = -\bar{b}(h - a) < 0$.

The key result is part (c). To see why the inequality in equation (12) arises, consider a labeled account with the same endowed efficiency, initial audience, and realization of early urn luck under the two possible types. If the account is an expert, its early urn weight is $\eta_i F_i^{\nu}$. If it is a charlatan, bait multiplies that weight by $g(\bar{b})^{1/\chi} > 1$. Under the common urn coupling, the charlatan version therefore has a strictly higher probability of becoming the strong monopolist.

Let ω_{E} and ω_{C} denote the probabilities that a labeled account wins the limiting early urn conditional on being an expert or a charlatan. The preceding comparison gives $\omega_{\mathrm{C}} > \omega_{\mathrm{E}}$. Since accounts are ex ante exchangeable, the probability that the winning account is a charlatan is

$$\pi_{\mathrm{C}}^{E} = \frac{(1 - P_{\mathrm{E}})\omega_{\mathrm{C}}}{(1 - P_{\mathrm{E}})\omega_{\mathrm{C}} + P_{\mathrm{E}}\omega_{\mathrm{E}}} > 1 - P_{\mathrm{E}}. \tag{13}$$

The first inequality is therefore not an assumption about the composition of attention. It is generated by the equilibrium content choice.

The mechanism can now be interpreted in four steps. First, bait raises the engagement return $g(b)$. Second, greater engagement raises effective urn efficiency from η_i to $\eta_i g(b)^{1/\chi}$. Third, a higher urn weight raises the probability of winning the superlinear attention process. Fourth, because charlatans choose bait and experts do not, charlatans win the attention urn more often than their population frequency would imply:

$$\text{bait} \Longrightarrow g(\bar{b}) > 1 \Longrightarrow \tilde{\eta}_{\mathrm{C}} > \eta_{\mathrm{C}} \Longrightarrow \omega_{\mathrm{C}} > \omega_{\mathrm{E}} \Longrightarrow \pi_{\mathrm{C}}^{E} > 1 - P_{\mathrm{E}}.$$

The last step uses the lock-in result from the previous subsection. By the end of the early block, the strong monopolist has accumulated a diverging audience while rival audiences remain bounded in probability. The later informational advantage of an expert remains bounded by e^{β}. Appendix A shows that the probability that the later block overturns the early winner converges to zero. The charlatan overrepresentation created while quality is difficult to observe therefore persists after quality becomes easier to learn.

The reputation franchise explains why the two types choose different urn weights in the first place. A charlatan receives a low late posterior under either content strategy, so protecting its reputation has little marginal value. It chooses bait to obtain the larger engagement payoff and higher urn weight. An expert can preserve a high late posterior by avoiding bait, and condition (10) makes that reputation gain larger than the maximal engagement benefit from slant. Attention persuasion therefore creates adverse selection because the type with lower advice value has the stronger incentive to purchase attention through content.

This mechanism yields three empirical predictions. Baiting should be associated with larger audiences, baiting should predict lower advice value even conditional on directional optimism, and accounts with more reputation outside the mass-audience channel should use less bait. Section 4.2 tests these implications using the pre-specified text-based bait measure. The pricing block next asks what happens when the resulting allocation of visible advice affects investor demand and prices.

1.4 Prices and visible sentiment

The preceding results determine which accounts receive attention. We now ask how the composition of that attention affects prices. The pricing block takes the attention allocation as given and follows

three steps. Account visibility determines the advice investors see, visible advice affects retail demand, and limited arbitrage converts that demand into a temporary price displacement.

Fix a stock j and three dates. Accounts post at date 0, retail investors trade at date 1, and the stock's fundamental abnormal payoff $f_j \sim \mathcal{N}(0, \tau_f^2)$ is realized at date 2. The public pre-period return is $R_{j,0}$.

The link to the attention model is the composition of the feed. For a stock with positive coverage, define the charlatan share of visibility as

$$w_j = \frac{\sum_{i:\theta_i=\mathrm{C}} V_i \mathbb{1}\{i \text{ covers } j\}}{\sum_i V_i \mathbb{1}\{i \text{ covers } j\}}. \tag{14}$$

Propositions 1 and 2 provide the mechanism that can raise w_j above the charlatan population share. The pricing block conditions on w_j and asks how prices vary with the composition of advice investors see.

Expert advice contains information about fundamentals. Charlatan advice reflects the optimistic and extrapolative content generated by the slant mechanism. Expert, charlatan, and visible sentiment are

$$S_j^{\mathrm{E}} = f_j + u_j, \qquad S_j^{\mathrm{C}} = \mu_{\mathrm{C}}(\bar{b}) + \zeta R_{j,0} + e_j, \qquad S_j^{V} = (1 - w_j) S_j^{\mathrm{E}} + w_j S_j^{\mathrm{C}}. \tag{15}$$

The parameter $\mu_{\mathrm{C}}(\bar{b}) > 0$ is the average optimism associated with the equilibrium slant strategy, while $\zeta > 0$ captures extrapolation from past returns. The model therefore allows baiting to affect attention through $g(\bar{b})$ and the content of the resulting advice through $\mu_{\mathrm{C}}(\bar{b})$ and ζ. These are distinct channels. The first determines who is heard. The second determines what investors hear.

The specification builds on the idea that short-sale constraints allow optimistic views to receive disproportionate weight in prices (Miller, 1977; Harrison and Kreps, 1978; Scheinkman and Xiong, 2003; Hong and Stein, 2003). It also follows the empirical evidence that disagreement predicts lower subsequent returns when pessimistic views are difficult to trade (Diether, Malloy, and Scherbina, 2002) and that stock-loan conditions create substantial cross-sectional variation in shorting costs (D'Avolio, 2002).

We formulate the pricing results conditional on the characteristics that determine visibility and arbitrage capacity.

Assumption 3 (Pricing environment) *Conditional on (w_j, BC_j), the variables $R_{j,0}$, f_j, u_j, and e_j are mutually independent with zero means and finite variances τ_R^2, τ_f^2, τ_u^2, and τ_e^2, where $\tau_f^2 > 0$ and $\zeta^2\tau_R^2 + \tau_e^2 > 0$. For comparisons across visibility shares and borrowing-cost states, the conditional distribution of $(R_{j,0}, f_j, u_j, e_j)$ is invariant in (w_j, BC_j).*

The conditioning makes the pricing results partial comparative statics. Stocks with different w_j can differ systematically for many reasons in the data. The model asks what changing the composition of visible advice does while holding the residual return environment fixed.

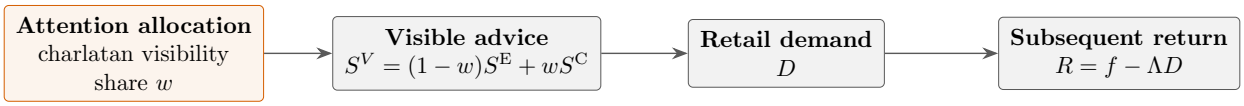


Figure 2: From Attention Allocation to Prices

The attention process determines the charlatan share w of the advice visible about a stock. Greater charlatan visibility makes the visible signal more optimistic. Retail investors trade on that signal, and limited arbitrage converts their demand into a temporary price displacement. Subsequent returns reverse this displacement. The reversal is larger when borrowing costs make arbitrage less effective.

Retail investors observe S_j^V and submit demand

$$D_j = \delta \max\{S_j^V, 0\} - \delta^- \max\{-S_j^V, 0\} + mR_{j,0}, \qquad \delta > 0, \qquad 0 \le \delta^- \le \delta, \qquad m \ge 0. \tag{16}$$

The parameter δ^- governs the ability of retail investors to trade on negative visible sentiment. The case $\delta^- = \delta$ gives symmetric linear demand, while $\delta^- = 0$ gives the binding short-sale constraint used for the directional result below. The term $mR_{j,0}$ allows retail investors to extrapolate directly from past returns.

Arbitrageurs absorb retail demand only imperfectly. Let BC_j denote the stock's borrowing-cost state and let $\Lambda_j = \Lambda(\mathrm{BC}_j) > 0$ with $\Lambda' > 0$. A larger Λ_j means that a given retail imbalance produces a larger price displacement. The subsequent abnormal return is

$$R_j = f_j - \Lambda(\mathrm{BC}_j) D_j. \tag{17}$$

The retail short-sale constraint and limited arbitrage therefore play different roles. The first makes positive and negative visible sentiment affect demand asymmetrically. The second determines how strongly a given demand imbalance moves the price.

The first implication follows directly from equation (15). Conditional on w_j, the difference between visible and expert sentiment has expectation $\mathbb{E}[S_j^V - S_j^{\mathrm{E}} \mid w_j] = w_j \mu_{\mathrm{C}}(\bar{b})$. The attention allocation therefore changes the average signal received by investors even when followers correctly learn the quality of the accounts producing it. The relevant object is the composition of visible advice rather than the population fraction of charlatan accounts.

For completeness, the attention results also imply an aggregate-belief statement.

Proposition 3 (Attention composition and visible beliefs) *Under the conditions of Proposition 1, define* $\bar{\theta}(t) = \sum_i p_i(t) V_i(t) / \sum_k V_k(t)$. *On the event that a charlatan strongly monopolizes attention,* $\bar{\theta}(t) \to 0$. *Under the separating equilibrium of Proposition 2,* $\lim_{m\to\infty} \lim_{t\to\infty} \mathbb{E}_m[\bar{\theta}_m(t)] = 1 - \pi_{\mathrm{C}}^E < P_{\mathrm{E}}$. *Moreover,* $\mathbb{E}[S_j^V - S_j^{\mathrm{E}} \mid w_j] = w_j \mu_{\mathrm{C}}(\bar{b})$.

Proposition 3 distinguishes learning from allocation. Followers can eventually recognize that a dominant account offers poor advice while the account continues to dominate the feed. The

failure is therefore not Bayesian inference. It is the allocation of attention inherited from the earlier reinforcement process.

We next characterize the resulting return predictability. Fix a borrowing-cost state x and write $\Lambda = \Lambda(x)$. Let $\Sigma_V(w) = \mathrm{Var}(S_j^V \mid w_j = w, \mathrm{BC}_j = x)$. For the directional results, fix $s > 0$ and define $m_+(w,x) = \mathbb{E}[R_j \mid S_j^{\mathrm{C}} \geq s, w_j = w, \mathrm{BC}_j = x]$ and $m_-(w,x) = \mathbb{E}[R_j \mid S_j^{\mathrm{C}} \leq -s, w_j = w, \mathrm{BC}_j = x]$.

Proposition 4 (Visibility composition and returns) *Under Assumption 3, the following statements hold.*

(a) Suppose $\delta^- = \delta$ and $m = 0$. In the joint population projection of R_j on $(S_j^{\mathrm{E}}, S_j^{\mathrm{C}})$, the coefficients on charlatan and expert sentiment are $-\Lambda\delta w_j$ and $\tau_f^2/(\tau_f^2 + \tau_u^2) - \Lambda\delta(1 - w_j)$. If $\Lambda\delta(\tau_f^2 + \tau_u^2) < \tau_f^2$, the population coefficient on S_j^V changes sign once, at $\bar{w}(x) \in (0,1)$ defined by

$$(1 - \bar{w}(x))\tau_f^2 = \Lambda(x)\delta\Sigma_V(\bar{w}(x)). \tag{18}$$

The coefficient is positive below $\bar{w}(x)$ and negative above it, and $\bar{w}'(x) < 0$.

(b) Suppose $\delta^- = \delta$. Then

$$\mathbb{E}[R_j \mid w_j, \mathrm{BC}_j = x] = -\Lambda(x)\delta w_j \mu_{\mathrm{C}}(\bar{b}). \tag{19}$$

Expected returns are strictly decreasing in w_j whenever $\mu_{\mathrm{C}}(\bar{b}) > 0$, and strictly decreasing in x whenever $w_j > 0$.

(c) Suppose $\delta^- = 0$ and $m = 0$. For every $w > 0$, $m_+(w,x) < 0$, $m_-(w,x) \leq 0$, and $|m_+(w,x)| > |m_-(w,x)|$. At $w = 0$, the two magnitudes are equal. Moreover, $\lim_{w\uparrow 1} m_-(w,x) = 0$ and

$$\lim_{w\uparrow 1} m_+(w,x) = -\Lambda(x)\delta\mathbb{E}[S_j^{\mathrm{C}} \mid S_j^{\mathrm{C}} \geq s, w_j = 1, \mathrm{BC}_j = x] < 0. \tag{20}$$

(d) Suppose $\delta^- = \delta$. Then $\mathrm{Cov}(D_j, S_j^{\mathrm{C}} \mid w_j, \mathrm{BC}_j = x) = \delta w_j \mathrm{Var}(S_j^{\mathrm{C}} \mid w_j, \mathrm{BC}_j = x) + m\zeta\tau_R^2 \geq 0$. If $w_j < 1$, $\tau_R^2 > 0$, $\mathrm{Var}(S_j^{\mathrm{E}} \mid w_j, \mathrm{BC}_j = x) > 0$, and $\tau_e^2 > 0$, the population coefficient on S_j^{C} in the projection of D_j on $(R_{j,0}, S_j^V, S_j^{\mathrm{C}})$ is zero.

Part (a) shows how the composition of attention changes the information content of the aggregate feed. Expert sentiment can remain informative even after visible sentiment becomes contrarian. The threshold $\bar{w}(x)$ need not equal one half because the two signals differ in both information content and noise. Higher borrowing costs lower the amount of charlatan visibility required for the aggregate signal to become negatively related to subsequent returns. Appendix A gives the complete sign characterization outside the parameter region in which visible sentiment begins positively informative.

Part (b) gives the main cross-sectional prediction. Greater charlatan visibility raises current retail demand through the optimistic component of the feed and therefore produces lower subsequent

returns when the price displacement reverses. The relation strengthens with the cost of arbitrage. Section 5.1 tests this prediction using posterior-weighted message coverage as a proxy for w_j. The empirical measure is not the model's visibility share because platform impressions are not observed.

Part (c) gives the directional prediction. With a binding retail short-sale constraint, bullish charlatan advice generates more buying pressure than bearish charlatan advice generates selling pressure. Bullish calls are therefore followed by larger reversals. As charlatan content comes to dominate visible advice, bearish charlatan sentiment drives retail demand toward zero and its subsequent-return effect disappears. This is why a known charlatan need not provide an economically useful signal that investors can simply invert.

Part (d) distinguishes exposure to charlatan content from a separate charlatan-demand channel. Charlatan sentiment can covary strongly with retail demand because it contributes to the visible signal and because both variables respond to past returns. Once visible sentiment and past returns are included, the model gives charlatan sentiment no additional coefficient. Section 5.4 uses this exclusion as a mechanism test.

Corollary 1 (Borrowing costs and return asymmetry) *Suppose the conditions of Proposition 4(c) hold and $\Lambda' > 0$. For every $w > 0$,*

$$\frac{\partial}{\partial x}\left(|m_+(w,x)| - |m_-(w,x)|\right) > 0. \tag{21}$$

At $w = 0$, the magnitude gap is zero.

The corollary separates the two frictions in the pricing mechanism. The retail constraint creates the asymmetry between bullish and bearish advice. Limited arbitrage determines the magnitude of the resulting price displacement. Higher borrowing costs therefore increase the return gap between bullish and bearish charlatan calls at any positive charlatan visibility share. Section 5.4 tests this comparative statics using stock-level borrowing costs.

1.5 Empirical predictions

The model generates four sets of empirical predictions that follow the sequence of the mechanism. The first concerns whether advice quality can be learned before audience positions are established. The second concerns whether accumulated audience is reinforced through engagement. The third asks whether attention-oriented content systematically favors low-quality accounts. The fourth maps the resulting composition of advice into subsequent returns. These predictions distinguish the informational friction from the attention technology and, within the attention technology, distinguish reinforcement from endogenous persuasion.

Prediction 1 (Slow learning) *The number of recommendations required to distinguish experts from charlatans increases with recommendation noise and decreases with the square of the expert–*

charlatan quality gap. When observed track records are short relative to this detection threshold, advice quality remains difficult to infer while attention is being allocated.

Lemma 1 implies that the relevant learning rate is $\rho = \Delta^2/(2\sigma_r^2)$. For a fixed confidence level, the number of observations required for classification is therefore proportional to σ_r^2/Δ^2. The empirical test compares these model-implied detection thresholds with the track records actually produced on the platform. The prediction does not require quality to remain permanently hidden. It requires only that reliable evidence arrive slowly enough that audience can accumulate before most accounts are statistically distinguishable. The model also predicts a weaker role for informational selection where recommendation outcomes are noisier, although cross-sectional tests of this comparative static are necessarily less direct because return volatility may correlate with other stock characteristics.

Prediction 2 (Attention reinforcement) *Engagement rises with accumulated audience in the direction required for visibility to be superlinear in audience. When this amplification is sufficiently strong, early audience differences become persistent and subsequent quality information need not undo them.*

The model generates visibility $V_i = \kappa \eta_i F_i^{\nu}$ with $\nu > 1$ because a larger audience raises the return to engagement effort and greater engagement raises exposure. The empirical counterpart is therefore a positive relation between accumulated audience and engagement activity that is strong enough to imply superlinear exposure under the model's mapping. This relation is an equilibrium object rather than a causal elasticity of follower count on future visibility. Audience, engagement, account age, survival, and early luck can all be jointly determined. The test is instead intended to discipline whether the amplification mechanism required by the urn is quantitatively plausible.

Prediction 3 (Attention persuasion and adverse selection) *Attention-oriented bait is associated with larger audiences and lower advice value, even conditional on directional optimism. Accounts with a more valuable reputation outside the mass-audience channel use less bait.*

Proposition 2 turns type-neutral lock-in into adverse selection. Bait raises the engagement return $g(b)$ and therefore raises effective urn efficiency from η_i to $\eta_i g(b)^{1/\chi}$. Charlatans have less reputation value to protect and choose the attention-enhancing strategy, whereas experts avoid it when the reputation franchise is sufficiently valuable. The resulting chain is

$$\text{bait} \implies g(\bar{b}) > 1 \implies \widetilde{\eta}_C > \eta_C \implies \omega_C > \omega_E \implies \pi_C^E > 1 - P_E.$$

The empirical tests therefore ask whether promotional framing predicts audience size, whether it predicts lower advice quality after controlling for optimism, and, most importantly, whether bait measured in one part of the sample predicts advice quality measured from disjoint future returns after controlling for prior quality and optimism. The reputation-franchise mechanism yields a

separate prediction that accounts displaying an outside professional or business affiliation use less bait. Because audience and affiliation are endogenous choices, these relations are interpreted as mechanism evidence rather than causal estimates of the effect of changing content.

Prediction 4 (Visible advice and subsequent returns) *Greater charlatan exposure predicts lower subsequent returns. The reversal is concentrated in bullish charlatan advice and is stronger when short selling is costly. Conditional on past returns and the aggregate visible signal, charlatan sentiment has no separate effect on retail demand in the model's linear-demand benchmark.*

Propositions 3 and 4 connect the allocation of attention to prices. A larger charlatan visibility share w_j makes visible advice more optimistic because $E[S_j^V - S_j^E \mid w_j] = w_j \mu_C(\bar{b})$. With symmetric linear retail demand, expected subsequent return satisfies $E[R_j \mid w_j, BC_j = x] = -\Lambda(x)\delta w_j \mu_C(\bar{b})$. Greater charlatan visibility therefore predicts lower subsequent returns, and the effect becomes stronger as borrowing costs raise the price impact of a given retail imbalance.

The directional prediction is sharper. When retail investors cannot respond symmetrically to negative sentiment, bullish charlatan advice generates more buying pressure than bearish charlatan advice generates selling pressure. Bullish charlatan calls should therefore be followed by larger reversals, and the bullish–bearish return gap should increase with borrowing costs. The model also distinguishes exposure to charlatan content from a separate demand channel. Charlatan sentiment can covary strongly with retail trading because it contributes to the visible signal and because both respond to past returns. Once past returns and visible sentiment are included, however, the linear-demand benchmark gives charlatan sentiment no additional coefficient.

The empirical analysis follows this sequence. Section 3 measures the speed at which quality can be learned. Section 4 tests the reinforcement and endogenous-bait mechanisms. Section 5 then uses real-time posterior classifications to ask whether the composition of posted advice predicts subsequent returns and whether the relation varies with direction and limits to arbitrage. Because platform impressions are not observed, the empirical charlatan-coverage measure is a proxy for the model's visibility share rather than a direct observation of w_j. The tests therefore evaluate the predictions of the mechanism without treating message coverage as identical to platform exposure.

2 Data and Measurement

We use data from StockTwits, a free platform on which investors share stock-specific views and follow other users. Our objective is to measure the economic value of the advice produced on the platform and, because realized performance is noisy, the latent distribution from which that value is drawn. We first construct user-stock-day sentiment from individual messages, match these signals to subsequent stock returns, and estimate each account's signed abnormal performance. We then use an empirical-Bayes deconvolution to separate persistent advice quality from sampling noise.

StockTwits contains both active users who produce stock-specific content and passive users who primarily consume it. We study the active population. Their posts range from technical commentary and investment ideas to explicit directional recommendations, and we do not require that a user formally present herself as an adviser or intend to persuade other investors.[3]

2.1 Platform data, sentiment, and returns

StockTwits is organized around short messages about individual securities. Users can post their own views, follow other accounts, and interact with the content they see. The platform reports more than six million registered users, although only a small fraction actively produces content. Unlike curated investment services, StockTwits does not screen contributors ex ante. Entry is therefore broad, and informative analysis can coexist with advice of much lower economic value.

Data sources. We combine four sources. Bloomberg provides StockTwits messages from July 13, 2013, when its collection begins, through January 1, 2017. We collect account-level information directly from StockTwits through February 2, 2018.[4] We match messages to CRSP stock returns and use factor returns from Kenneth French's data library to construct abnormal returns. Markit provides daily stock-level measures of short interest and borrowing costs used in the mechanism tests.

Bloomberg's StockTwits coverage contains one interruption. It ends in late April 2014 and resumes in late December 2014, leaving May through November 2014 unobserved. The gap appears to reflect Bloomberg's data collection rather than selective coverage of particular messages or market conditions. We therefore treat it mechanically. Daily and account-level samples omit the missing period, monthly tests use only covered months, and the out-of-sample portfolio analysis excludes formation windows that overlap the interruption.

Bloomberg collects social-media messages from both StockTwits and X, formerly Twitter. We use StockTwits because Bloomberg provides both a standardized sentiment score and the identity of the poster. Poster identities are unavailable for X, which prevents us from constructing account-level histories of advice quality there.

For each StockTwits message, Bloomberg reports its timestamp, text, stock ticker, username, relevance and confidence scores, topic codes, and a social-sentiment classification produced by a supervised machine-learning algorithm (Bloomberg, 2016). For the nth message by user i about stock j on day t, the sentiment score is $SocSent_{i,j,t,n} \in \{-1, 0, 1\}$. The Bloomberg feed contains 74 million messages across the two social platforms, including 18.3 million StockTwits messages.

[3]FINRA and IOSCO use a similarly broad notion of online financial advice, encompassing individuals who share investment-related content ranging from financial education to specific recommendations and promotions. Passive, non-posting users have been studied as peer interactions by Cookson, Engelberg, and Mullins (2023). Our analysis instead concerns the accounts that produce content. We later impose activity and audience screens and, in the stock-level tests, place greater weight on accounts whose messages receive more exposure.

[4]The platform has not provided us with API access after this date.

Table 1: Summary Statistics

	N	Mean	S.D.	Min	p10	p50	p90	Max
Activity	29,391	530	4,333	9.32	31.00	137.00	970.00	615,145
Followers	22,017	1,040	8,988	0.00	1.00	10.00	135.00	489,704
Measured alpha $\tilde{\alpha}_i$	29,391	−0.60	7.47	−55.22	−7.67	−0.23	6.00	58.21
S.E. of $\tilde{\alpha}_i$	29,391	4.28	4.72	0.00	0.84	2.75	9.46	56.71
t-stat of $\tilde{\alpha}_i$	29,387	−0.92	122.14	−20,824	−2.07	−0.10	1.71	1,088

Notes: This table reports account-level activity, follower counts, measured alpha $\tilde{\alpha}_i$, its standard error, and its t-statistic. Activity and follower counts are obtained from StockTwits. Activity is lifetime posting volume; for accounts pooled across duplicate screen-name matches, it is averaged across the pooled profiles, so its minimum need not be an integer. Measured alpha is the average signed Fama–French five-factor-adjusted return over the 20 trading days following an account's messages. Alphas and standard errors are in percentage points.

Among the StockTwits messages, 9.8% are negative, 76.2% neutral, and 14.0% positive. The distribution is more balanced across positive and negative views than the self-declared sentiment studied in Dim (2022) and Cookson, Engelberg, and Mullins (2023).

News data. Bloomberg also provides the public-news feed used to benchmark social sentiment. For each story, we observe its release time, headline, stock ticker, source, and a machine-generated sentiment score (Bloomberg, 2016). The score $NewsSent_{j,t,n} \in \{-1, 0, 1\}$ is defined analogously to social sentiment. Among 38 million news stories, 12% are negative, 59% neutral, and 29% positive. StockTwits messages are therefore considerably more likely to be neutral and somewhat less likely to be negative than public news.

User-level data. Table 1 summarizes account activity, audience size, and measured advice quality. We use the StockTwits API to obtain lifetime posting activity and follower counts after the return sample ends.[5] Lifetime activity averages 530 posts, with a median of 137. The follower distribution is highly skewed. The average account has 1,040 followers, while the median has only 10.

Sample construction and filtering. Appendix Table D.6 traces the analysis samples from the raw Bloomberg and StockTwits records. The latent-quality estimation uses 29,387 accounts with at least two return-matched directional observations and nondegenerate sampling errors. Of these, 22,017 have follower counts, and 22,931 have complete strategy covariates. The empirical design therefore separates the broad population of messages from the subset of accounts for which persistent advice quality can be estimated.

[5] We collect 139,401 accounts as of February 2, 2018 and match 105,535 of them to Bloomberg usernames. Because many accounts post very little, we pool profiles with fewer than 20 lifetime messages or reposts when resolving duplicate screen-name matches. A user's StockTwits history can extend beyond the Bloomberg sample, so some matched users have fewer than 20 messages during our return sample. We can estimate measured alpha for 29,391 accounts, of which 22,017 also have follower counts.

The measurement is robust to activity screens, minimum track-record requirements, alternative return horizons, corrections for overlapping returns, and heavier-tailed sampling error. Under the previous latent-quality specification, the estimated negative-alpha share ranges from 40.7% to 41.7% across the activity filters reported in Appendix B, and requiring longer scored histories does not eliminate the negative-alpha population. These exercises are useful because the latent-quality estimates enter the model as inputs for the type distribution and the speed of learning, rather than as a descriptive taxonomy of StockTwits users.

User-stock-day sentiment scores. We first match Bloomberg screen names to StockTwits accounts. Bloomberg records the screen name displayed on the platform, whereas the StockTwits account data contain the underlying username. The two coincide in most cases, but screen names need not be unique. When a Bloomberg identity cannot be mapped uniquely to a StockTwits account, we either pool the corresponding profiles or remove the ambiguous match.

We next align each message with market prices. A message posted during trading hours is assigned to that trading day. A message posted after the close, on a weekend, or on a market holiday is assigned to the next trading day. Thus, every message is associated with the first closing price available after it is posted.

For a given user-stock-day, let $SocSent^{+}_{i,j,t} = \sum_{n=1}^{N_{i,j,t}} \mathbb{1}\{SocSent_{i,j,t,n} = 1\}$ and $SocSent^{-}_{i,j,t} = \sum_{n=1}^{N_{i,j,t}} \mathbb{1}\{SocSent_{i,j,t,n} = -1\}$ denote the numbers of positive and negative messages. We aggregate them into the bounded directional score

$$SocSent_{i,j,t} = \max\left\{-1, \min\left(1, SocSent^{+}_{i,j,t} - SocSent^{-}_{i,j,t}\right)\right\}. \tag{22}$$

A user-stock-day is therefore positive when positive messages dominate, negative when negative messages dominate, and zero when the two offset or all messages are neutral.

The return-matched sample contains 6,649,510 user-stock-days. The average numbers of positive and negative messages are 0.37 and 0.19, with maxima of 158 and 169. The corresponding neutral count, $SocSent^{0}_{i,j,t} = \sum_{n=1}^{N_{i,j,t}} \mathbb{1}\{SocSent_{i,j,t,n} = 0\}$, averages 1.37 and reaches a maximum of 448. After applying equation (22), the mean user-stock-day sentiment score is 0.13 with a standard deviation of 0.60.

Vendor sentiment and independent validation. Bloomberg's sentiment score is externally constructed and never trained on the return outcomes used in this paper. As an initial check, we compare it with an independent rule-based classifier applied directly to message text. Across 15.5 million messages with both measures, the two classifications agree 67.8% of the time and have a correlation of 0.37. Direct sign reversals are rare; most disagreements concern whether a message is directional or neutral. Appendix B reports the confusion matrix and the construction of this benchmark. Because message direction determines which subsequent return enters measured alpha, however, agreement with a second automated classifier is not sufficient for the final measurement

design. We therefore build an independent text-classification pipeline that can reconstruct the signed-return stack and, at the same time, measure the framing choices that the model associates with baiting for attention.

2.2 Independent text classification and the bait index

The independent classifier serves two purposes. First, it removes the proprietary Bloomberg score from the construction of recommendation direction. Second, it measures how a message is framed rather than only whether it is bullish or bearish. The distinction matters because Proposition 2 concerns a content choice that raises attention without raising advice value.

We assign nine labels to each message. Two of them describe the recommendation itself, including its directional sentiment and whether the message contains current advice. Six describe framing characteristics that enter the bait measure: hype, certainty, a numerical price target, bullish reference to a recent trend, a call to action, and self-promotion. A final affiliation label records whether the account presents a professional role or outside business and is kept separate as a proxy for reputation outside the platform. The ontology is fixed before any return, posterior, or portfolio outcome is joined to the labels. Appendix B.2 reports the definitions, prompt, illustrative examples, and validation protocol.

We first apply a locally run open-weight language model with deterministic decoding to a fixed stratified sample of 400,000 messages and then train a lightweight classifier on those labels to score the full corpus of 18,348,209 StockTwits messages.[6] The stratified sample deliberately oversamples vendor-positive, vendor-negative, vendor-neutral, and unscored messages so that the classifier sees the regions in which alternative directional labels are most likely to disagree. Population moments are computed on the full corpus rather than on this stratified training sample.

Human coding provides the validation benchmark. Two coders independently label a stratified sample of 1,500–2,000 messages, with adjudication of disagreements. We compare human agreement, the teacher model, the distilled classifier, and the Bloomberg score against the same adjudicated labels, reporting population-reweighted accuracy, class-specific precision and recall, Cohen's kappa, and opposite-sign errors. The independent classifier replaces the vendor score in the headline measurement stack only after passing this validation gate.

Baiting for attention. The framing labels provide a direct measure of the endogenous slant in Proposition 2. For account i in period t, let each framing component denote its average across the

[6] The teacher model is `qwen3:30b` served locally through Ollama with temperature zero, reasoning traces disabled, and a constrained JSON output schema. Neither the teacher nor the distilled classifier observes returns, posterior quality, follower outcomes, or any other economic outcome.

account's messages in that period. We define

$$Bait_{i,t} = \frac{1}{6}\left[\frac{hype_{i,t}}{3} + certainty_{i,t} + target_{i,t} + trend_{i,t}\mathbb{1}\{sentiment_{i,t} = 1\} + action_{i,t} + promo_{i,t}\right]. \tag{23}$$

Each component lies on the unit interval after normalization, so $Bait_{i,t} \in [0, 1]$. Bullishness by itself does not increase the index. The trend component enters only when the message is bullish because the model's baiting mechanism concerns extrapolation from recent winners rather than any reference to a past price move.

The equal weights in equation (23) are fixed before outcomes are examined. They therefore summarize a pre-specified framing bundle rather than an index chosen to maximize a return or audience relation. We report a first-principal-component index estimated from the label covariance matrix as an outcome-free robustness check. The baseline averages over all messages because content can compete for attention even when it does not contain a current recommendation; an advice-only version provides a second robustness check. The individual components are reported separately so that any relation with attention or quality can be traced to the underlying framing choices rather than to the composite alone.

2.3 Measuring advice quality

We measure advice quality by the return obtained from acting in the direction of an account's message. Because an individual track record is noisy, observed performance cannot be interpreted directly as persistent quality. We therefore first estimate each account's signed abnormal return and its sampling uncertainty, and then use an empirical-Bayes deconvolution to recover the latent distribution of advice value and the posterior probability that a given account has positive or negative expected performance.

The object we measure is implementable advice value. A follower enters at the first closing price after the message and holds over the specified return horizon. For a sufficiently visible account, this return can reflect both information in the recommendation and price pressure generated by the message itself. An informative post that moves the price before the follower enters can attenuate measured alpha, while attention-driven overreaction can push measured alpha below zero. We therefore interpret latent quality as the expected payoff from acting on the account's observed advice at the implementable entry price.

Measured alpha. For user i and return horizon H, we estimate

$$SocSent_{i,j,t} \cdot AbnRet_{j,t+1,t+H} = \tilde{\alpha}_i + \epsilon_{i,j,t+1,t+H}. \tag{24}$$

The estimation uses all user-stock-days with $SocSent_{i,j,t} \neq 0$. The dependent variable is positive when the subsequent abnormal return moves in the direction of the account's view and negative

when it moves against it. We compute abnormal returns from the Fama–French one-, three-, and five-factor models over horizons of 1, 2, 5, 10, and 20 trading days. The results are similar across horizons and factor specifications.[7]

Table 1 reports the 20-day five-factor specification used in the baseline analysis. The mean measured alpha is -0.60% per month, and the median is -0.23%. The dispersion in raw estimates is large relative to their precision. The mean standard error is 4.28%, and the median is 2.75%. Nevertheless, 18.9% of accounts have measured alphas significant at the 5% level and 24.8% at the 10% level, considerably more than would arise if all accounts had zero expected advice value. The next step separates this cross-sectional heterogeneity from sampling noise.

2.4 Latent advice quality

Individual track records are too noisy to classify accounts directly from measured alpha. We therefore treat measured performance as a noisy signal of latent advice quality and recover the cross-sectional distribution of that quality by deconvolution. For account i,

$$\tilde{\alpha}_i = \alpha_i + \epsilon_i, \qquad \epsilon_i \sim \mathcal{N}(0, \tilde{\sigma}_i^2). \tag{25}$$

The sampling uncertainty $\tilde{\sigma}_i$ is account-specific and reflects the length and volatility of the account's observed recommendation history.

Latent advice quality is drawn from a flexible two-sided distribution,

$$F_\alpha = \pi_{\mathrm{E}} F_{\mathrm{E}} + \pi_{\mathrm{C}} F_{\mathrm{C}}, \qquad \mathrm{supp}(F_{\mathrm{E}}) \subset (0, \infty), \qquad \mathrm{supp}(F_{\mathrm{C}}) \subset (-\infty, 0). \tag{26}$$

An account is therefore an expert when $\alpha_i > 0$ and a charlatan when $\alpha_i < 0$. These are economic categories defined by expected advice value rather than by the numerical components used to approximate the two distributions. Section 3.2 asks whether α_i should be interpreted literally as a permanent individual trait or more broadly as the unconditional payoff to a persistent advice strategy.

Following performance-evaluation methods and their applications to online investment advice (Harvey and Liu, 2018; Crane and Crotty, 2020; Dim, 2022; Kakhbod, Kazempour, Livdan, and Schürhoff, 2025), we approximate F_{E} and F_{C} with mixtures of K one-sided exponential components on their respective half-lines. The components are sieve terms that allow each economic type to contain substantial heterogeneity. They are not additional types. The distribution of observed $\tilde{\alpha}_i$ is the convolution of this latent distribution with the account-specific Gaussian sampling error in equation (25). We estimate component means and weights by maximum likelihood. Internet

[7] The same estimator can be implemented as a pooled regression with user indicators, which produces identical point estimates. Day-clustered standard errors from the pooled specification are used in the corresponding robustness exercise in Internet Appendix Table F.8. Dependence in the estimated latent-quality distribution is addressed separately with the month-block bootstrap reported in Table 2.

Appendix E gives the likelihood and computational details.

The baseline uses three sieve components on each side. Table 2 reports the estimates. On the expert side, component means are 8.52%, 1.58%, and 0.12% per month, with population weights of 4.7%, 22.5%, and 19.5%. On the charlatan side, the corresponding means are −0.43%, −1.96%, and −10.75%, with weights of 25.3%, 21.3%, and 6.8%. Aggregating within each half-line gives an expert population share of $\pi_E = 46.6\%$ and a charlatan share of $\pi_C = 53.4\%$.

The flexible two-sided specification also clarifies the role of accounts whose estimated performance lies close to zero. An earlier specification imposed a separate point mass at zero in addition to two components on each side. Allowing a third component on each half-line improves both AIC and BIC without such a mass. The additional components locate substantial probability near zero, at 0.12% on the expert side and −0.43% on the charlatan side. The appropriate conclusion is therefore narrow. The data do not require a distinct zero-quality population once heterogeneity on the positive and negative sides is modeled flexibly.

This distinction is important for the economics. An account with an uninformative track record need not have zero true quality. It can instead have a posterior close to one half because the available observations reveal little about the sign of α_i. Under the baseline $K = 3$ specification, 45% of accounts have charlatan posteriors between 0.4 and 0.6, compared with 35% under $K = 2$. Appendix Figure B.1 plots the posterior distribution. Uncertainty about type is therefore represented as uncertainty in beliefs, which is the object that enters the learning mechanism of Section 1.

The likelihood is evaluated in logs to avoid numerical underflow and converges to the same optimum from all random starting values.[8] Bootstrap standard errors for the $K = 3$ two-sided estimates will be obtained from the month-block procedure used for the earlier specification.

Robustness. The economic conclusions do not depend on a particular sieve representation. Internet Appendix F estimates the competing specifications using a common pipeline and retains the earlier zero-mass model as a benchmark. Posterior rankings are nearly unchanged. The expert posterior under the new specification has a Spearman correlation of 0.98 with the earlier positive-alpha posterior, the charlatan posterior has a correlation of 0.97 with the earlier negative-alpha posterior, and posterior mean quality correlates at 0.9994 across the two specifications. The change in specification therefore affects how near-zero accounts are represented without materially changing the ranking of accounts that drives the empirical analysis.

The estimates are also robust to how overlapping returns and return horizons are treated. Newey–West corrections, non-overlapping outcome windows, and calendar-day clustering produce charlatan shares at least as large as the baseline estimate. Posterior rankings remain informative across horizons. Expert probabilities have Spearman correlations of 0.61–0.67 across adjacent return horizons and 0.27 between the 1- and 20-day horizons. Appendix Table D.7 reports the full

[8] Evaluating the convolution density in levels can underflow for accounts with large sampling errors, flattening the likelihood and creating spurious numerical optima.

Table 2: Estimating the Distribution of Latent Advice Quality

Panel A: Baseline two-sided sieve ($K = 3$ per side)	μ_k (%)	π_k (%)
Expert component 3	8.52	4.7
Expert component 2	1.58	22.5
Expert component 1	0.12	19.5
Charlatan component 1	−0.43	25.3
Charlatan component 2	−1.96	21.3
Charlatan component 3	−10.75	6.8
Expert share π_{E}		46.6
Charlatan share π_{C}		53.4
N	29,387	
Log Likelihood	−87,589.2	
AIC	175,200	
BIC	175,292	

Panel B: Specification comparison	AIC	BIC
Two-sided, $K = 3$ per side (baseline; LL −87,589.2)	175,200	175,292
Zero mass plus $K = 2$ per side (old baseline; LL −87,609.9)	175,236	175,302
Two-sided, $K = 2$ per side (LL −87,671.8)	175,358	175,416

Notes: This table reports empirical-Bayes estimates of the latent distribution of advice quality. Measured alpha is the average signed Fama–French five-factor-adjusted return over the 20 trading days following an account's messages. The baseline specification approximates the positive and negative halves of the latent-quality distribution with three exponential sieve components each and imposes no point mass at zero. Panel A reports component means μ_k and population weights π_k. The components approximate heterogeneity within the expert and charlatan populations and are not economic types. Panel B compares the baseline with the earlier specification containing a point mass at zero and with a two-component-per-side specification. Means and weights are in percentage points and percentages, respectively. Bootstrap standard errors from the month-block procedure are to be added.

comparison.

Real-time estimation. The pricing tests require a stronger restriction than robustness of the full-sample estimator. At each portfolio-formation date, the classification must use only investment outcomes that an investor could already have observed. We therefore re-estimate the latent-quality distribution recursively in expanding windows.

At the end of quarter q, the estimation sample contains only recommendations whose full 20-trading-day return window has closed by that date. Messages posted during the final 20 trading days of the quarter therefore enter the next estimation vintage rather than the current one. Winsorization thresholds are also recomputed separately within each vintage. No posterior used in the real-time tests contains a recommendation outcome that had not yet been realized when the posterior was formed.[9]

The full-sample estimates are used to describe the latent-quality distribution and discipline the model's learning parameters. The expanding-window estimates are used whenever a posterior enters

[9]The embargo removes approximately the final month of recommendations from each estimation vintage. Embargoed and unembargoed posteriors have correlations ranging from 0.92 in the earliest window to above 0.98 thereafter.

a subsequent-return test. Section 3.2 separately examines whether measured quality persists across genuinely disjoint periods rather than across nested estimation windows.

Sieve order in real time. We use the same $K = 3$ specification in the expanding-window estimator as in the full sample. All 14 real-time vintages converge, including the earliest three-quarter window with only 1,887 classified accounts. This choice follows a rule fixed before the pricing results were examined: use $K = 3$ throughout if the early vintages remain numerically stable and use $K = 2$ only if thin samples require greater parsimony.

In practice, the posterior rankings are almost identical across the two choices. The Spearman correlation between $K = 2$ and $K = 3$ charlatan posteriors has a minimum of 0.992 and a median of 0.998 across vintages. Posterior mean quality correlates at least 0.9995 in every vintage. We therefore report the $K = 3$ vintages in the baseline real-time tests and use $K = 2$ as a sieve-order robustness check.

Posterior measures of advice quality. Let $\mathcal{H}_{i,t}$ denote the recommendation outcomes of account i that have been realized by date t. We summarize the resulting posterior with three objects,

$$p_{i,t} = \Pr(\alpha_i > 0 \mid \mathcal{H}_{i,t}), \qquad c_{i,t} = \Pr(\alpha_i < 0 \mid \mathcal{H}_{i,t}) = 1 - p_{i,t}, \qquad q_{i,t} = \mathbb{E}[\alpha_i \mid \mathcal{H}_{i,t}]. \tag{27}$$

The probabilities $p_{i,t}$ and $c_{i,t}$ measure uncertainty about the sign of advice quality, while $q_{i,t}$ also incorporates its expected magnitude. In the full-sample estimates, $\mathcal{H}_{i,t}$ is replaced by the complete account history and the time subscript is omitted.

Operationally, let $f_i(\tilde{\alpha}_i)$ denote the estimated mixture density of measured alpha. Posterior mean quality is

$$q_{i,t} = \frac{\int_{\mathbb{R}} \alpha \, \phi(\tilde{\alpha}_i; \alpha, \tilde{\sigma}_i) \, dF_\alpha(\alpha)}{f_i(\tilde{\alpha}_i)}, \tag{28}$$

where $\phi(\tilde{\alpha}_i; \alpha, \tilde{\sigma}_i)$ is the $\mathcal{N}(\alpha, \tilde{\sigma}_i^2)$ density. The expert and charlatan probabilities are obtained by integrating the same posterior over the positive and negative halves of the latent distribution.

Because $p_{i,t} + c_{i,t} = 1$, every account carries a complete two-type posterior rather than being assigned mechanically to a third "uncertain" category. We use posterior probabilities when the economic prediction is naturally stated in terms of type and $q_{i,t}$ when the magnitude of expected advice value is more informative. Hard classifications are reserved for transparent portfolio exercises, such as the top decile of the charlatan posterior.

2.5 Text-informed priors

The baseline posterior learns quality from realized investment outcomes alone. Followers also observe content, and the model predicts that framing can contain information about type before a

long return history accumulates. We therefore ask how much observable content can improve the starting point of the learning problem.

We allow the prior probability of expertise to depend on pre-specified content characteristics x_i, so that $\pi_{\mathrm{E}}(x_i) = \Lambda(x_i'\gamma)$, where Λ is the logistic function. The latent expert and charlatan distributions are unchanged. Text shifts only the prior odds that an account is drawn from one side of the quality distribution; realized returns continue to determine the likelihood.

The optimism measure provides a transparent benchmark. A one-standard-deviation increase in the share of positive messages lowers the logit prior for expertise by 0.32. The concomitant-variable specification improves the fit relative to the return-only model, with a likelihood-ratio statistic of 179.5 and AIC and BIC of 175,180 and 175,246. A split-sample version that measures optimism on messages disjoint from the returns entering the likelihood gives a coefficient of -0.29 and a likelihood-ratio statistic of 18.4.

The bait labels provide the preferred content specification because they separate directional optimism from the framing bundle that Proposition 2 treats as an attention choice. We therefore repeat the concomitant-prior exercise with the pre-specified bait components, keeping direction itself separate. If baiting contains information about type beyond optimism, followers can improve their prior by observing how advice is presented, not merely whether it is bullish. If it does not, optimism remains the sufficient content signal and the richer framing measure serves to identify the engagement channel.

Content can improve priors without solving the learning problem. Under the optimism-only specification, the mean-evidence threshold moves from 124 recommendations under an unconditional prior to 111 for an account at the 90th percentile of optimism and 153 at the 10th percentile. Framing changes where posterior odds begin; it does not change the information contained in each realized recommendation. Section 3.1 shows that this distinction is quantitatively important because return histories remain short relative to the evidence needed to identify quality.

3 Learning About Advice Quality

The latent-quality estimates in Table 2 provide the empirical inputs for the model's learning problem. They imply that 46.6% of accounts have positive expected advice value and 53.4% have negative expected advice value. In this section, we use the estimated quality distribution for two purposes. We first ask how long a track record must be before a follower can reliably distinguish good advice from bad. We then ask what, if anything, persists across disjoint periods when realized account-level alpha itself is unstable. Appendix B reports additional classification, activity-filter, directional-content, and return-chasing exercises.

3.1 Slow learning from realized performance

The relevant friction is not that advice quality is permanently hidden. It is that realized returns reveal it slowly. Suppose account i produces signed abnormal returns $r_{i,1}, \ldots, r_{i,n}$ with mean α_i and per-recommendation volatility $\hat{\sigma}_{r,i}$. The standard error of its average track-record return is $\hat{\sigma}_{r,i}/\sqrt{n}$. Learning therefore depends on the amount of return noise relative to economically meaningful differences in advice quality.

Definition 1 (Expert-charlatan detectability benchmark) *Consider an expert and a charlatan whose recommendation returns are Gaussian with means $\alpha_{\mathrm{E}} > 0 > \alpha_{\mathrm{C}}$, quality gap $\Delta = \alpha_{\mathrm{E}} - \alpha_{\mathrm{C}}$, and common per-recommendation volatility σ_r. Under an even prior, the expected log posterior odds of a Bayesian observer reach the confidence boundary $\ln[(1-\delta)/\delta]$ after*

$$n^*(\delta) = \frac{2\sigma_r^2 \ln \frac{1-\delta}{\delta}}{\Delta^2} \tag{29}$$

recommendations. A fixed-sample one-sided test that separates the two means with size δ and power $1-\delta$ instead requires $n \geq n^F(\delta) = 4z_{1-\delta}^2 \sigma_r^2/\Delta^2$.

The Bayesian threshold is deliberately favorable to learning. It assumes that the follower knows the two candidate drifts and updates correctly after every recommendation. It measures the point at which expected log odds cross the confidence boundary; the expected sequential stopping time is weakly longer because the posterior can overshoot that boundary. The threshold is the empirical counterpart of Lemma 1. A conventional fixed-sample test requires still more observations and, at $\delta = 0.05$, raises the required history by roughly 80%. Both criteria scale with σ_r^2/Δ^2, so the substantive question is how large the relevant expert-charlatan gap is in the data.

We define the typical gap before examining the resulting threshold as the distance between the medians of the positive and negative latent-quality distributions,

$$\Delta_{\mathrm{typ}} = \mathrm{median}(\alpha \mid \alpha > 0) - \mathrm{median}(\alpha \mid \alpha < 0). \tag{30}$$

This definition avoids tying the learning calculation to any particular sieve component. Table 3 also reports two alternative gaps that summarize different parts of the estimated distribution, while Figure 3 traces the full relationship between the quality gap and the number of observations required for learning.

Calibration. For each account, we recover per-recommendation volatility as $\hat{\sigma}_{r,i} = \tilde{\sigma}_i\sqrt{N_i}$. Its cross-sectional median is 8.6%, with a mean of 10.7% and 10th and 90th percentiles of 3.0% and 20.9%. We evaluate the detection thresholds at the median volatility.

The magnitudes are large relative to observed track records. The gap between the largest-weight sieve components on the two sides is 2.01 percentage points per month, which implies a Bayesian

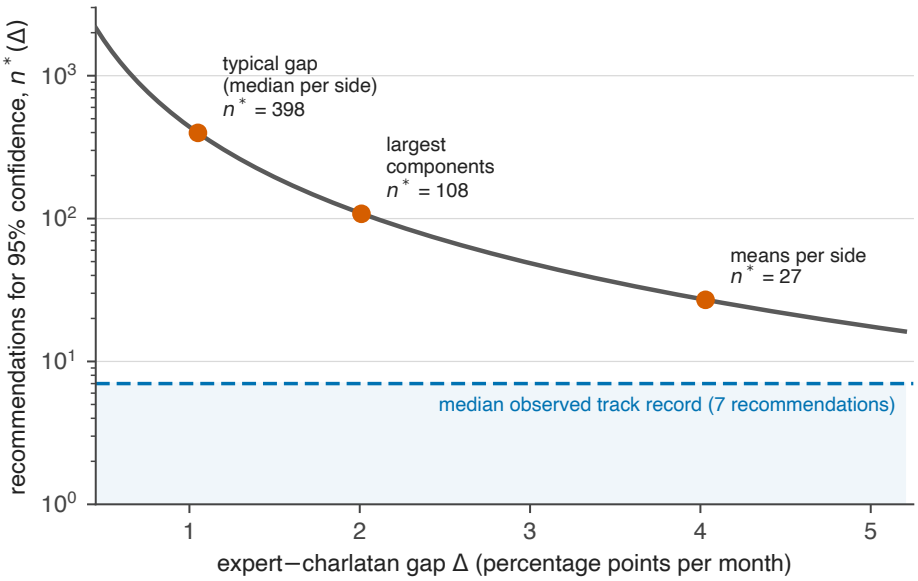


Figure 3: The Detection Threshold as a Function of the Quality Gap

Notes: The figure plots the mean-evidence Bayesian detection threshold $n^*(\Delta) = 2\sigma_r^2 \ln 19/\Delta^2$ from equation (29) at $\delta = 0.05$ and the median per-recommendation volatility $\hat{\sigma}_r = 8.6\%$. The marked points correspond to the three quality gaps in Table 3: the median expert-charlatan gap ($\Delta = 1.05$, $n^* = 398$), the gap between the largest-weight sieve components on the two sides ($\Delta = 2.01$, $n^* = 108$), and the gap between the mean latent alpha on the two sides ($\Delta = 4.03$, $n^* = 27$). The dashed line marks the median observed track record of seven scored recommendations.

threshold of 108 scored recommendations. Only 3.9% of accounts ever produce that many. Using the mean latent alpha on each side produces the largest gap, 4.03 percentage points, because it places substantial weight on the tails. Even under this favorable comparison, the threshold is 27 recommendations and only 16.6% of accounts reach it. The typical gap in equation (30) is 1.05 percentage points and implies 398 recommendations, reached by only 0.8% of accounts.

The contrast with actual histories is stark. The median account produces only seven scored recommendations during the three-and-a-half-year sample, or roughly two per year. A message is scored only if it has non-neutral sentiment, can be matched to a CRSP security, and occurs during the return sample, which explains why this number is much smaller than the median of 137 lifetime posts in Table 1. Even the most favorable gap definition requires several times the median observed history, while the economically more representative comparisons require one to two orders of magnitude more observations. Fixed-sample significance criteria lengthen these horizons further.

The precise numerical threshold will be updated after the independent text-classification pipeline is frozen because that procedure changes both the number of scored recommendations and the estimated latent-quality distribution. The economic object is fixed by equation (30). The relevant question is whether observed track records approach the amount of evidence required to separate positive from negative advice value.

Table 3: The Speed of Learning About Advice Quality

Definition of the gap Δ	Δ (%)	$n^*(0.05)$	% reaching n^*
Largest-weight component per side	2.01	108	3.9
Mean alpha per side (tail-inflated)	4.03	27	16.6
Median alpha per side (Δ_{typ})	1.05	398	0.8

Notes: This table evaluates Definition 1 using the latent-quality estimates in Table 2 and the median per-recommendation volatility $\hat{\sigma}_r = 8.6\%$. The three rows use, respectively, the gap between the largest-weight sieve components on the two sides, the difference between mean latent alpha on the expert and charlatan sides, and the pre-specified typical gap Δ_{typ} between their medians. n^* is the mean-evidence Bayesian threshold at $\delta = 0.05$. It is an optimistic learning benchmark because the expected sequential stopping time weakly exceeds it. The final column reports the share of accounts whose observed number of scored recommendations reaches the corresponding threshold. The median account has seven scored recommendations. The final value of Δ_{typ} and its threshold will be updated after the independent text-classification pipeline is frozen.

Limits of reputation. The tails of the quality distribution are easier to learn. The outermost sieve components have monthly means of 8.52% and −10.75%, so an observer can distinguish these extreme accounts after only a small number of recommendations. Together, however, the two components account for only 11.5% of the population. For the much larger interior of the distribution, realistic histories contain too little information to separate persistent advice quality from return noise.

Reputation can therefore discipline the clearest cases while remaining weak for most accounts. This distinction is central to the model. Attention begins accumulating immediately, while reliable evidence about quality arrives only gradually. It also yields a cross-sectional prediction that we return to below. Where recommendation outcomes are noisier, informational selection should be weaker and attention misallocation more severe.

3.2 Persistent strategies, unstable realized alpha

Slow learning matters only if there is something persistent to learn. The model represents that object as account type, but the empirical counterpart need not be a fixed realized alpha that repeats from one market environment to another. We therefore distinguish persistence of measured performance from persistence of the strategies that generate it. Expanding windows establish that the posterior measures used in the real-time tests evolve smoothly as information accumulates. Disjoint samples provide the more demanding test because they share no recommendation outcomes.

Disjoint-sample evidence. We split the sample at June 2015 and estimate advice quality independently in the two periods. The analysis includes 6,101 accounts with at least two scored recommendations in each half. The current version of Table 4 still reports the previous zero-mass specification and will be replaced by the two-sided $K = 3$ estimates. Its rank-correlation exercises, which are the important objects for the argument below, are insensitive to this change in mixture representation.

Table 4: Disjoint-Sample Persistence

Panel A: Independent half-sample mixture estimates				
	First half ($N = 14{,}109$)		Second half ($N = 20{,}976$)	
	μ_k (%)	π_k (%)	μ_k (%)	π_k (%)
Expert component 2	7.88	4.9	8.48	5.5
Expert component 1	1.35	29.2	1.62	23.2
Zero mass (old spec)	0.00	21.9	0.00	26.8
Charlatan component 1	−1.05	32.5	−1.42	37.3
Charlatan component 2	−6.02	11.4	−12.10	7.2

Panel B: Transition matrix across halves (row %, $N = 6{,}101$ users)			
	Expert	Zero mass	Charlatan
Expert (first half)	23.7	14.9	61.4
Zero mass (first half)	23.8	35.3	40.9
Charlatan (first half)	23.6	13.4	63.0
Unconditional second-half share	23.6	15.1	61.3

Panel C: Rank correlations across halves and benchmarks		
	All users	$n \geq 10$ per half
Measured alpha	0.015	0.063
Pr(expert)	0.030	
$\mathbb{E}[\alpha_i \mid \tilde{\alpha}_i]$	0.020	
Benchmark, simulated permanent types	0.377	0.536
Placebo, odd/even split within periods	0.204	0.314
Fraction positive (behavior)	0.298	0.462
Activity (behavior)	0.453	
Top-decile repeat, Pr(expert) / Pr(charlatan)	11.6% / 10.6%	

Notes: This table re-estimates the latent-quality model independently in the two halves of the sample. Account-level alpha and its standard error are computed separately from recommendations through June 2015 and after June 2015, so the two estimation samples contain no common outcomes. Panel A reports independent mixture estimates. Panel B reports transitions between maximum-posterior classifications for the 6,101 accounts with at least two scored recommendations in each half. Panel C reports Spearman correlations across the two halves and comparison benchmarks. The permanent-type benchmark draws latent alphas from the estimated distribution and applies each account's observed recommendation counts and sampling errors in the two periods. The odd-even placebo splits recommendations within the same calendar environments. The behavior rows report persistence in optimism and posting activity. The current Panel A and Panel B use the previous specification containing a zero point mass and will be replaced by the two-sided $K = 3$ estimates.

The striking feature of the time split is how little individual realized performance recurs. The Spearman correlation in measured alpha is 0.015 across the two halves and rises only to 0.063 among accounts with at least ten recommendations in each period. Expert posteriors correlate at 0.03. Likewise, only about 11% to 12% of first-half top-decile accounts remain in the same posterior tail in the second half, barely above the 10% rate implied by independence.

This instability is too large to attribute simply to short histories. We simulate the same experiment under permanent latent quality while preserving each account's observed recommendation counts and sampling errors. The resulting expected cross-half alpha correlations are 0.38 for the full sample and 0.54 for accounts with at least ten recommendations in each half. An odd-even split of recommendations within the same calendar environments also produces much greater persistence,

with correlations between 0.20 and 0.31. In contrast, separating the observations in calendar time reduces the correlations to between 0.02 and 0.06.

The natural interpretation is that measured performance contains a substantial account-period component. Users repeatedly operate in particular styles and stock habitats, but the payoff to those habitats changes across market environments. Additional recommendations within the same period can average away recommendation-level noise without eliminating this common component. This is precisely the distinction captured by the user-period shock introduced in Section 1.

What persists. Behavior is considerably more stable than realized performance. The fraction of positive messages has a cross-half correlation of 0.30, rising to 0.46 among users with at least ten scored recommendations in each period. Posting activity correlates at 0.45. There is also some persistent component in realized advice value. Accounts in the bottom decile of first-half alpha t-statistics underperform those in the top decile by 0.35% over the next 20 trading days in the second half, although this spread is estimated imprecisely.

A simple variance decomposition reaches the same conclusion. The covariance of half-sample alphas attributes 0.57 of variance to the persistent component, compared with total non-sampling variance of approximately 6.0 in each half. The corresponding monthly standard deviations are 0.75% for persistent alpha and 2.3% for the account-period component. Persistent heterogeneity therefore exists, but it accounts for only about one tenth of the non-sampling variation in realized performance.[10]

Weak persistence in individual alpha does not imply that prior information has no economic value. The real-time tests aggregate many noisy account histories. Portfolios formed from prior realized outcomes earn -0.59% over the next 20 trading days in Table 8, while the embargoed exposure sort in Section 5.1 produces a -0.90% spread. The relevant question is therefore not whether realized alpha repeats account by account, but whether persistent observable behavior helps identify future advice value.

From persistent behavior to future advice value. Persistence in behavior is economically relevant only if the persistent behavior is informative about subsequent advice quality. We therefore construct behavior from the first half of the sample and ask whether it predicts quality estimated independently from second-half returns. The two-sided latent-quality model is fit separately in each half, and no second-half observation enters either a first-half regressor or its quality control. Table 5 reports the current benchmark results for 6,096 accounts active in both periods.

Optimism provides the clearest bridge in the existing classification. A one-standard-deviation increase in the first-half share of positive messages predicts second-half posterior mean quality that is 0.088 percentage points lower, with a t-statistic of -3.60. It also raises the second-half

[10]For accounts active in both halves, the variance of the half-sample mean alpha net of sampling variance identifies the sum of persistent and account-period variation, while the covariance across halves identifies the persistent component under independence of the account-period shocks.

Table 5: First-Half Behavior and Second-Half Advice Quality

First-half regressor	q_2	c_2 (pp)	α_2
Optimism (share positive)	−0.088	+0.77	−0.192
	(−3.60)	(+2.54)	(−2.61)
Return chasing	+0.021	≈ 0	+0.163
	(+0.79)	(−0.48)	(+1.94)
Activity (log scored recommendations)	+0.143		
	(+3.85)		
Breadth (log # tickers)	−0.136		
	(−3.70)		
First-half quality control	Yes	Yes	Yes
N		6,096	

Notes: This table relates second-half advice quality to behavior measured exclusively in the first half of the sample for 6,096 accounts with at least two scored recommendations in each period. The sample is split at June 2015, and the two-sided latent-quality model is estimated separately in the two halves. The outcomes are second-half posterior mean quality q_2, the second-half charlatan posterior c_2 in percentage points, and second-half raw signed alpha α_2. Optimism is the share of positive first-half messages, return chasing follows the paper's preferred construction, activity is the log number of scored recommendations, and breadth is the log number of distinct tickers. Each specification controls for the corresponding first-half quality measure. Coefficients are reported for a one-standard-deviation increase in the regressor, with HC1 t-statistics in parentheses. The return-chasing estimates use the paper-exact construction; the remaining rows use the companion specification, with differences across implementations confined to the third decimal.

charlatan posterior by 0.77 percentage points and predicts second-half raw signed alpha that is 0.192 percentage points lower. These relations are incremental to first-half estimated quality. In the posterior-mean specification, the coefficient on first-half quality itself is 0.033 with a t-statistic of 1.19.

Return chasing does not have the same forward content. Its coefficient is 0.021 for second-half posterior quality, approximately zero for the charlatan posterior, and weakly positive for subsequent raw alpha. Activity predicts higher future quality, with a coefficient of 0.143 and a t-statistic of 3.85, while broader stock coverage predicts lower future quality, with a coefficient of −0.136 and a t-statistic of −3.70. These results imply that the persistent content signal is not simply recent-winner chasing.

The bait index gives a more direct test of the strategy interpretation. We repeat the same disjoint-sample regression with first-half $Bait_{i,1}$ as the regressor and second-half $q_{i,2}$, $c_{i,2}$, and raw signed alpha as outcomes. The pre-specified specification controls for first-half quality and first-half optimism, so baiting must predict future advice value beyond the direction of the account's messages. This is the strongest empirical counterpart of type as a persistent strategy in the model: framing is measured before the returns used to evaluate subsequent quality and is never constructed from those returns.

Taken together, the evidence suggests that the persistent object is not realized account's alpha in a particular sample window. Behavioral choices and stock habitats recur much more strongly, and optimism already carries information about future advice value that first-half estimated quality

does not. The bait test asks whether this persistence extends to the broader promotional framing described by the model. Realized returns remain useful for learning, but they combine persistent strategy with substantial market-environment variation. Followers must therefore infer that strategy through both recommendation-level noise and account-period shocks. This is the empirical counterpart of type in Section 1: a persistent advice strategy associated with an unconditional payoff, rather than a fixed realized alpha that repeats in every market environment.

4 Attention Reinforcement and Endogenous Slant

The model generates adverse selection through two distinct forces. Audience reinforcement determines how quickly an early attention advantage can compound, while endogenous content choice determines whether the accounts that benefit from this process are systematically different in quality. We therefore ask two questions. Does engagement increase with accumulated audience in the direction required by the reinforcement mechanism? And which observable content choices are associated both with greater audience and with lower subsequent advice value? The underlying strategy regressions are reported in Appendix C. Broader descriptive patterns in posting behavior and audience size are documented in Kakhbod, Kazempour, Livdan, and Schürhoff (2025).

4.1 Audience accumulation and amplification

Large audiences do not appear to reveal superior advice quality. Follower counts are measured in February 2018, after the return sample ends. The posterior probability of positive alpha falls from 30.3% among accounts with fewer than 10 followers to 23.3% among accounts with more than 1,000, a cross-sectional pattern also documented in Kakhbod, Kazempour, Livdan, and Schürhoff (2025). We do not give this gradient a causal interpretation. Audience can affect engagement, engagement can affect audience, and both may reflect account age, survival, or early luck. The narrower point is that realized prominence does not sort strongly toward the quality followers would ultimately like to identify.

The model's reinforcement mechanism concerns how engagement scales with audience. From equation (4), optimal effort satisfies $e_i^* \propto F_i^{\nu-1}$. This motivates the reduced-form relation

$$\ln \#\text{tweets}_i = a_0 + (\nu - 1)\ln(1 + F_i) + u_i. \tag{31}$$

Across 39,183 accounts, the estimated slope is 0.394 with a standard error of 0.005 and an R^2 of 0.22. Interpreted through the model, this corresponds to $\nu = 1.39$.

The estimate is not a structural estimate of ν. Audience size is jointly determined with posting activity and can be correlated with engagement efficiency, account age, survival, and earlier attention shocks. We therefore use the slope only to discipline the quantitative magnitude of reinforcement. Its positive value is consistent with the complementarity in the model: accounts with

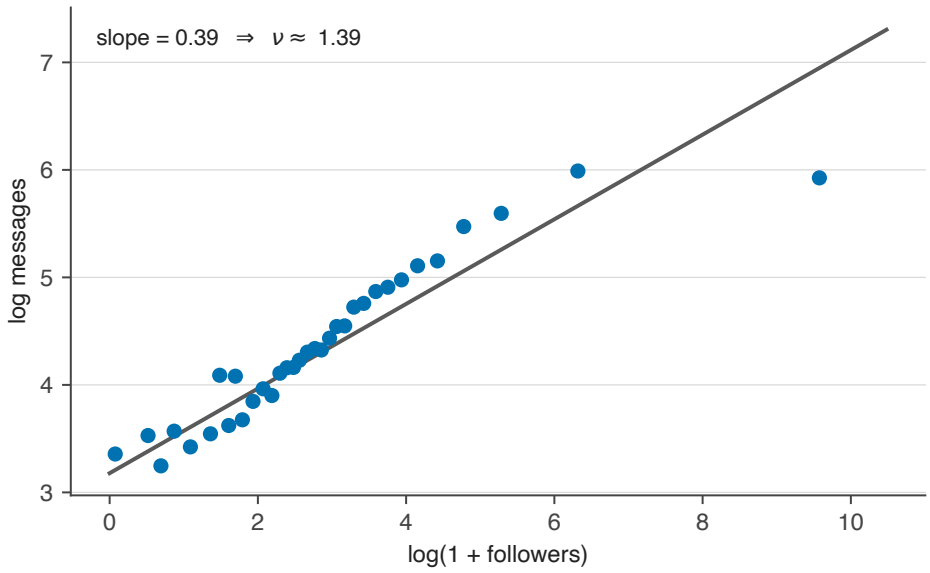


Figure 4: Activity versus Audience Size

Notes: The figure shows a binscatter with 50 equal-count bins of log message activity against log(1 + followers). The regression direction follows the first-order condition underlying equation (31). The sample contains 39,183 accounts with positive activity and non-missing follower counts. Followers are measured in February 2018 and winsorized at the 99.5th percentile. The fitted slope is 0.394 with a robust standard error of 0.005. Under the model's interpretation, the slope corresponds to $\nu - 1$.

larger audiences produce disproportionately more engagement activity, which is the force that can make visibility increase more than proportionally with accumulated audience.[11]

4.2 Baiting for attention

Reinforcement can make attention path dependent without making it adversely selected. Proposition 2 adds the second force: low-quality accounts have stronger incentives to choose content that raises engagement because they have less reputation value to protect. The empirical counterpart therefore requires a strategy that is associated with both audience and lower advice value.

Optimism provides the first indication of this channel. In the multivariate quality regressions, moving from entirely negative to entirely positive content raises the charlatan posterior by 7.63 percentage points and lowers the expert posterior by 5.16 percentage points. Optimism is also positively associated with audience size. In the follower regression, the fraction of positive messages has a coefficient of 0.48, while log activity has an elasticity of 0.76. These are equilibrium associations rather than causal effects of changing message tone, but they show that the content dimension most clearly associated with poor advice is also associated with larger audiences.

The bait index in equation (23) tests the mechanism more directly. It is constructed only from

[11] The corresponding slope is 0.37 when activity is measured over 2015–2016 and 0.24 over 2013–2014. Differences across windows can reflect account growth, selection, and measurement error as well as changes in the underlying relationship.

promotional framing and keeps bullishness itself outside the index. We use it in three pre-specified comparisons. First, we relate audience size to baiting conditional on activity, optimism, account age, and the other strategy controls. Second, we relate contemporaneous posterior quality to baiting while controlling for optimism, so the index must distinguish framing from directional sentiment. Third, we use the disjoint-sample design of Section 3.2 to ask whether first-half baiting predicts second-half advice quality after controlling for first-half quality and optimism. The third test is particularly important because neither the bait measure nor the quality outcome shares return observations.

The individual framing labels make the composite result interpretable. We report hype, certainty, price targets, bullish trend references, calls to action, and self-promotion separately in a companion specification. The equal-weight index remains the primary test because its weights were fixed without outcomes; the component regressions show which behaviors account for any relation with attention and quality. A principal-component index estimated only from the label covariance matrix provides an additional outcome-free robustness check.

The affiliation label provides a separate test of the reputation-franchise mechanism. It is excluded from the bait index because affiliation is not promotional framing. Instead, we ask whether accounts that display an outside professional or business affiliation use less bait, conditional on current audience, activity, prior performance, and stock habitat. A negative relation would be consistent with the model's prediction that an outside reputation franchise raises the cost of sacrificing advice quality for engagement. We treat this comparison as supporting evidence rather than as causal identification because affiliation is chosen and may proxy for other account characteristics.

Return chasing remains useful for interpreting the mechanism but is not itself the attention strategy. Accounts that recommend recent winners are more likely to have negative contemporaneous advice value, particularly for bullish calls, yet return chasing enters the follower regression negatively and does not predict lower quality in a subsequent disjoint period. It therefore describes part of the performance footprint of baited optimism rather than the source of the audience advantage. The distinction is useful because the model bundles optimistic slant and trend confirmation for tractability, while the text classifier allows their roles to be separated.

Appendix C reports the full strategy regressions, and Appendix B.2 reports the text-classification and validation details. The main mechanism test is deliberately narrow: attention-oriented framing must be associated with audience and with lower advice value, and the disjoint-sample specification asks whether that relation survives when the framing and payoff measurements are separated in time.

5 Real-Time Pricing Evidence

We next ask whether the allocation of advice across stocks contains information about subsequent prices. The empirical design proceeds from aggregate exposure to progressively sharper tests. Sec-

tion 5.1 constructs stock-level exposure using account posteriors based only on previously realized outcomes. Section 5.2 asks whether the relation survives within a stock as the composition of its coverage changes. Section 5.3 moves to individual recommendations made after an account has been classified. Sections 5.4 and 5.5 then examine the model's limits-to-arbitrage and learnability predictions.

5.1 Real-time exposure and subsequent returns

Proposition 4(b) predicts lower subsequent returns when a larger share of the advice investors see comes from charlatan accounts. We do not observe platform impressions and therefore cannot measure the model's visibility share w_j directly. Instead, we construct a stock-month measure of the quality composition of posted coverage. Charlatan exposure is

$$C_{j,m} = \frac{\sum_i c_{i,m} N_{i,j,m}}{\sum_i N_{i,j,m}}, \tag{32}$$

where $N_{i,j,m}$ is the number of classified messages posted by account i about stock j during month m, and $c_{i,m}$ is the account's charlatan posterior formed from outcomes observed before the formation period. Because $p_{i,m} + c_{i,m} = 1$, every classified message enters the denominator. Thus, $C_{j,m}$ measures the posterior charlatan composition of message coverage. It is an empirical proxy for the model's visibility share under the maintained interpretation that relative posting volume is informative about relative exposure.

We also construct a continuous measure based on posterior expected quality,

$$Q_{j,m} = \frac{\sum_i q_{i,m} N_{i,j,m}}{\sum_i N_{i,j,m}}. \tag{33}$$

The two variables summarize the same underlying posterior information in different ways. $C_{j,m}$ emphasizes the probability that coverage comes from negative-value accounts, while $Q_{j,m}$ retains the estimated magnitude of advice quality.

At the close of each month, we sort stocks with at least five classified messages into quintiles by exposure. Returns are measured during the following month, so every message entering the characteristic is posted before the first return used to evaluate the portfolio. Account posteriors are also predetermined. Messages posted during calendar quarter Q are weighted by outcome-embargoed posteriors estimated through quarter $Q-1$, using only recommendations whose complete 20-trading-day outcome windows had closed by the classification date.

Table 6 reports the results. Stocks in the highest quintile of charlatan exposure earn 0.90 percentage points lower five-factor-adjusted returns over the next 20 trading days than stocks in the lowest quintile. The conventional t-statistic is -4.23 and the Newey–West statistic with three monthly lags is -3.14. Sorting on posterior expected quality produces the mirror image. The high-minus-low $Q_{j,m}$ portfolio earns 0.94%, with a conventional t-statistic of 4.83 and a Newey–West

Table 6: Cross-Sectional Pricing by Charlatan Exposure

	Q5−Q1 (20d)	t	NW(3)
Panel A: Baseline real-time sorts ($K = 3$ vintages)			
Charlatan exposure $C_{j,m}$	-0.90^{***}	-4.23	-3.14
Expected-quality exposure $Q_{j,m}$	$+0.94^{***}$	$+4.83$	$+4.33$
Panel B: Sieve-order robustness ($K = 2$ vintages)			
Charlatan exposure $C_{j,m}$	-0.758^{***}	-4.10	-3.37
Expected-quality exposure $Q_{j,m}$	$+1.055^{***}$	$+4.95$	
Panel C: Value-weighted ($K = 2$ vintages)			
Charlatan exposure $C_{j,m}$	$+0.11$	$+0.43$	
Expected-quality exposure $Q_{j,m}$	-0.23	-0.64	

Notes: At the close of month m, stocks with at least five classified messages are sorted into quintiles by charlatan exposure $C_{j,m}$ from equation (32) or posterior expected-quality exposure $Q_{j,m}$ from equation (33). Returns are cumulative Fama–French five-factor-adjusted returns over stock-days in month $m+1$, so all messages used to construct the characteristic precede the evaluated returns. Messages in calendar quarter Q are weighted by outcome-embargoed posteriors estimated through quarter $Q-1$. The time series contains 31 monthly spreads. Panel A uses the baseline $K = 3$ vintages. Panel B reports $K = 2$ estimates. For the $K = 2$ charlatan-exposure spread, a moving-block bootstrap gives a 95% interval of $[-1.18, -0.36]$, 24 of 31 monthly spreads are negative, and a wild bootstrap gives $p = 0.0006$. Panel C value-weights the quintile portfolios and currently uses the $K = 2$ vintages. Returns are in percentage points.

statistic of 4.33. The result therefore does not depend on expressing latent quality as a discrete type probability.

The posterior representation also matters little. Using $K = 2$ rather than $K = 3$ components per side produces a charlatan-exposure spread of -0.758%, with a conventional t-statistic of -4.10 and a Newey–West statistic of -3.37. A moving-block bootstrap gives a 95% interval of $[-1.18, -0.36]$, 24 of the 31 monthly spreads are negative, and the wild-bootstrap p-value is 0.0006. The corresponding expected-quality spread is 1.055% with a t-statistic of 4.95. These similarities are consistent with the near-perfect rank correlations between the $K = 2$ and $K = 3$ real-time posteriors documented in Section 2.4.

The effect is concentrated outside the largest stocks. Under the currently available $K = 2$ vintages, value-weighted quintiles produce a charlatan-exposure spread of 0.11% with a t-statistic of 0.43 and a quality-exposure spread of -0.23% with a t-statistic of -0.64. Thus, the equal-weighted result describes the typical covered stock rather than an aggregate market-capitalization effect. The $K = 3$ value-weighted specification is being recomputed.

The earlier analysis used the ratio $W = A/(A + S)$, which compared charlatan-weighted with expert-weighted coverage under the previous three-type architecture. Because that measure discards part of the classified flow and no longer matches the two-type model as naturally as $C_{j,m}$, we retain it only as an appendix robustness measure. The associated design-sensitivity exercises will be re-estimated using $C_{j,m}$ and the $K = 3$ real-time vintages.

5.2 Within-stock evidence

The portfolio sort compares different stocks, so a fixed stock habitat could account for part of the relation. Charlatan accounts may simply specialize in securities whose subsequent returns are low for reasons unrelated to contemporaneous advice. We therefore ask a sharper question. Does the same stock earn lower subsequent returns when the composition of its own coverage becomes more charlatan-heavy?

We estimate

$$R^{20}_{j,m+1} = \alpha_j + \gamma_m + \beta C_{j,m} + X'_{j,m}\theta + \varepsilon_{j,m+1}, \tag{34}$$

where α_j and γ_m are stock and calendar-month fixed effects. Exposure is standardized, so β measures the change in 20-day abnormal returns associated with a one-standard-deviation increase in charlatan exposure. All controls are measured at or before formation. The panel contains 2,566 stocks over 31 formation months, and standard errors are two-way clustered by stock and month.[12]

Table 7 reports the specification sequence. With month fixed effects alone, a one-standard-deviation increase in charlatan exposure predicts 0.356 percentage points lower subsequent returns, with a t-statistic of -4.89. Adding stock fixed effects reduces the coefficient to -0.171 with a t-statistic of -2.80. Thus, a meaningful part of the cross-sectional relation reflects persistent differences in the stocks charlatan-heavy accounts cover.

The within-stock component nevertheless remains economically and statistically meaningful. Adding momentum over one week, one month, and twelve-minus-one months, together with size, turnover, Amihud illiquidity, idiosyncratic volatility, retail trading share, and the Markit borrowing-cost measure, produces a coefficient of -0.196 with a t-statistic of -3.45. Replacing charlatan exposure with posterior expected quality reverses the sign as predicted: the coefficient is 0.163 with a t-statistic of 2.72. An interaction with an indicator for newly covered stock-months is close to zero, so the relation does not appear to arise mechanically when an account begins covering a stock.

Momentum sorts provide a complementary check on the most immediate alternative explanation. Charlatan accounts disproportionately discuss recent winners, so exposure could simply reproduce familiar return reversal. Within both one-month and twelve-minus-one-month momentum terciles, however, the charlatan-exposure spread remains negative. The corresponding momentum-adjusted Q5–Q1 spreads are -0.94% with a t-statistic of -4.45 and -0.74% with a t-statistic of -4.27. Recent returns therefore explain neither the cross-sectional sort nor the within-stock relation.

The current within-stock estimates use the $K = 2$ real-time vintages. Because the $K = 2$ and $K = 3$ posterior rankings are nearly identical, they provide a useful check on the mechanism, but the final table should use the baseline $K = 3$ vintages throughout.

[12] With only 31 month clusters, the two-way variance estimate uses the Cameron–Gelbach–Miller adjustment when the estimated covariance matrix is not positive semidefinite. One-way clustered statistics are similar.

Table 7: Within-Stock Charlatan Exposure and Subsequent Returns

	(1)	(2)	(3)	(4)	(5)
Exposure (per 1 s.d.)	-0.356^{***}	-0.171^{***}	-0.196^{***}	$+0.163^{***}$	-0.187^{***}
t-statistic	(-4.89)	(-2.80)	(-3.45)	$(+2.72)$	(-3.16)
New-coverage interaction					≈ 0 $(t = 0.35)$
Exposure measure	$C_{j,m}$	$C_{j,m}$	$C_{j,m}$	$Q_{j,m}$	$C_{j,m}$
Month FE	Yes	Yes	Yes	Yes	Yes
Stock FE	No	Yes	Yes	Yes	Yes
Full controls	No	No	Yes	Yes	Yes
N	48,472	48,372	45,287	45,287	44,472

Notes: This table reports panel regressions of 20-day Fama–French five-factor-adjusted returns in month $m + 1$ on standardized stock-month exposure. Columns (1)–(3) use charlatan exposure $C_{j,m}$, column (4) uses posterior expected-quality exposure $Q_{j,m}$, and column (5) interacts $C_{j,m}$ with an indicator for newly covered stock-months. Controls include momentum over one week, one month, and twelve-minus-one months, size, turnover, Amihud illiquidity, idiosyncratic volatility, retail trading share, and the Markit DCBS borrowing-cost score, all measured at or before formation. Standard errors are clustered by month in column (1) and two-way by stock and month elsewhere, with the Cameron–Gelbach–Miller adjustment where required. The current estimates use $K = 2$ real-time vintages; the baseline $K = 3$ re-estimation is pending.

5.3 Out-of-sample directional calls

The stock-level exposure measures average posterior uncertainty across many accounts. We next ask whether an account classification formed from prior realized outcomes predicts the returns following that account's subsequent recommendations.

At the end of each quarter, we estimate $c_{i,t}$ from the outcome-embargoed expanding-window vintage. Every return used in that posterior has therefore been realized before the account makes the recommendations evaluated in the next quarter. We then identify accounts in the tweet-weighted top decile of the charlatan posterior and evaluate the stocks they subsequently recommend.

The current construction uses decile cutoffs calculated from the distribution of tweet-stock-day observations within the evaluation quarter. It therefore contains no future return information, but the breakpoint applied to an early-quarter recommendation can depend on messages posted later in the same quarter. The account posterior itself is strictly predetermined; the portfolio cutoff is not. A version using only prior-quarter information to fix the breakpoint is therefore the appropriate fully predetermined specification and is being recomputed. The top-decile rule itself was fixed before the two-type estimates were examined.

Table 8 reports the eleven evaluation quarters under the current construction. Pooling daily portfolio returns across these periods and averaging within calendar month, stocks receiving bullish recommendations from high-charlatan-posterior accounts earn -0.59% over the next 20 trading days. The conventional t-statistic is -2.91 and the Newey–West statistic with three monthly lags is -2.45. Calendar-time aggregation gives the truncated 2014Q2 period weight in proportion to the number of observed days rather than treating it as a complete quarter.

The result is again insensitive to sieve order. Under $K = 2$ vintages, the corresponding return is -0.556%, with a conventional t-statistic of -2.94 and a Newey–West statistic of -2.48. The

Table 8: Out-of-Sample Portfolio Returns

	Charl bullish	Charl contrarian	N obs
2013Q4	−1.293	1.111	241,006
2014Q1	−0.315	−0.876	363,058
2014Q2 (truncated)	0.862	0.815	55,440
2015Q1	−0.346	0.381	325,047
2015Q2	−0.817	−1.049	514,090
2015Q3	−0.144	−0.322	590,619
2015Q4	−0.823	0.472	531,416
2016Q1	0.699	0.985	516,240
2016Q2	−0.344	−0.399	621,912
2016Q3	−0.655	−0.113	776,218
2016Q4	−1.956	0.264	892,345
Calendar-time monthly mean	−0.588***		
t-stat	−2.91		

Notes: This table reports 20-day Fama–French five-factor-adjusted returns using the baseline two-sided $K = 3$ real-time posteriors. At each quarter-end, account posteriors are estimated from recommendations whose complete 20-trading-day outcome windows have already closed. Portfolios are evaluated using recommendations posted during the following quarter. Charlatan bullish contains stocks receiving positive recommendations from accounts in the tweet-weighted top decile of the charlatan posterior. Charlatan contrarian is long charlatan-bearish and short charlatan-bullish. Under the current implementation, posterior probabilities are fully predetermined, but tweet-weighted decile breakpoints use the evaluation quarter's message composition; the prior-quarter-breakpoint version is being recomputed. The 2014Q3 evaluation period contains no covered messages because of the Bloomberg outage, 2014Q4 falls below the pre-specified 100-observation minimum, and 2014Q2 is truncated. The headline statistic pools daily returns across evaluation periods and averages them within calendar months. Its Newey–West t-statistic with three monthly lags is −2.45. Returns are in percentage points.

monthly return is negative in 24 of 31 months, and the wild-bootstrap p-value is 0.0047.

The magnitude is smaller than the −0.74% obtained under the earlier three-type classification. This attenuation follows naturally from the change in measurement. Under the two-type system, accounts that previously received substantial posterior probability on zero alpha must instead carry some probability on each side, making the upper tail of $c_{i,t}$ less extreme. We leave the pre-specified top-decile cutoff unchanged rather than selecting a more favorable threshold after observing the new estimates. Continuous regressions on $c_{i,t}$ and $q_{i,t}$ and the full posterior-decile profile provide complementary specifications.

The directional asymmetry is also informative. Bearish recommendations from high-charlatan-posterior accounts do not generate a corresponding positive return pattern, so the negative performance of bullish recommendations is not consistent with a symmetric signal that investors could simply reverse. It instead matches the one-sided optimism in Proposition 4(c).

Full-sample classifications produce the same directional asymmetry with larger return magnitudes. Because those classifications and the recommendations being evaluated share observations, however, the exercise is descriptive rather than out-of-sample. We therefore leave the in-sample portfolios untabulated and place the evidential weight on the predetermined posterior tests above.

5.4 Limits to arbitrage and retail demand

The pricing model yields two additional predictions. If optimistic charlatan advice contributes to temporary price pressure, subsequent reversals should be larger when short selling is more costly. If retail investors respond to the common signal visible on the platform, charlatan activity should have little incremental relation with retail demand once recent returns and aggregate platform activity are held fixed. The estimates in this subsection still use the previous three-type classification and should therefore be viewed as provisional until they are re-estimated with the two-sided $K = 3$ posteriors.

Short-sale constraints. We first examine whether the return reversal following bullish charlatan recommendations varies with the cost of trading against them. Each day, stocks are sorted into terciles using the trailing five-day average of Markit's cost-of-borrow score. We compare the easy-to-short and hard-to-short terciles using the same out-of-sample portfolio construction as in Table 8. Formation windows affected by the 2014 Bloomberg interruption are excluded.

The difference is economically large. Bullish charlatan recommendations are followed by 20-day abnormal returns of -1.98% in hard-to-short stocks and -0.27% in easy-to-short stocks. A pooled regression using the 46,509 tagged stock-days in the two extreme terciles implies a hard-minus-easy difference of -1.30 percentage points, with a date-clustered t-statistic of -3.79.

The evidence is less informative for bearish recommendations. Their subsequent returns are -0.24% in easy-to-short stocks and -1.96% in hard-to-short stocks, but both estimates are imprecise. More importantly, the bullish-minus-bearish difference does not vary significantly with shortability. The result therefore supports the narrower prediction that bullish reversals become larger when shorting is costly. It does not by itself establish the full directional asymmetry in Corollary 1. The corresponding test for expert bearish recommendations is also underpowered because few observations satisfy both the Markit coverage and account-classification requirements.

Retail flow. We next ask how charlatan activity relates to retail trading. For each stock-day, define posterior-weighted charlatan and expert volume as

$$CharlatanVol_{j,t} = \sum_i \Pr(\text{charlatan}_i)|SocSent_{i,j,t}|,$$

and

$$ExpertVol_{j,t} = \sum_i \Pr(\text{expert}_i)|SocSent_{i,j,t}|.$$

We relate these measures to the retail order imbalance of Boehmer, Jones, Zhang, and Zhang (2021) using

$$OIB_{j,t+h} = a + b_1 CharlatanVol_{j,t} + b_2 ExpertVol_{j,t} + b_3 \ln(1 + TweetVol_{j,t}) + b_4 Ret_{j,t} + \mu_j + e_{j,t+h}, \tag{35}$$

for $h \in \{1, 2, 5\}$. The final specification includes stock fixed effects, and standard errors are clustered by stock and date.

The unconditional relation is strong. Next-day retail order imbalance averages 56.6 on stock-days in the top decile of charlatan volume and 15.8 on all other stock-days, a difference of 40.9 with a t-statistic of 13.23. On the common regression sample, the univariate coefficient on charlatan volume is 5.17. It falls to 2.04 after controlling for expert volume, becomes statistically insignificant after aggregate tweet volume is added, and falls further to 0.60 after controlling for lagged returns and stock fixed effects. Charlatan volume is also insignificant at the two- and five-day horizons.

The same pattern appears when platform activity is decomposed into total volume and the charlatan share. Total message volume carries most of the relation with retail order flow, while the charlatan share has little within-stock incremental effect. Using outcome-embargoed prior-quarter posteriors produces an unconditional charlatan-volume coefficient of 5.60 with a t-statistic of 7.19, which falls to 0.87 with a t-statistic of 0.33 in the fixed-effects specification. The unconditional relation is therefore not an artifact of using future recommendation outcomes to classify accounts.

We also examine the exclusion in Proposition 4(d) more directly. In a regression of next-day retail imbalance on the day's return, aggregate visible sentiment, and charlatan sentiment, with stock and date fixed effects, the coefficient on visible sentiment is 2.98 with a t-statistic of 2.73. The coefficient on charlatan sentiment is 0.39 with a t-statistic of 0.35 and a 95% confidence interval of $[-1.79, 2.58]$. Using outcome-embargoed posteriors gives a charlatan-sentiment coefficient of 1.89 with a t-statistic of 1.48 and a confidence interval of $[-0.62, 4.40]$.

These estimates are consistent with the model's exclusion restriction, but they should not be read as evidence that the incremental effect is literally zero. Visible and charlatan sentiment are mechanically related, which reduces the variation available to distinguish their coefficients, and the confidence intervals remain economically nontrivial. The more robust conclusion is that the strong unconditional association between charlatan activity and retail buying is largely absorbed by recent returns and the overall platform signal.

This pattern fits an environment in which charlatan accounts and retail investors respond to some of the same recent price movements and in which investors subsequently react to the aggregate signal carried by the platform. There is little evidence that charlatan identity has an additional effect on retail demand once these common components are accounted for. The positive coefficient on expert volume in some full-sample specifications is not predicted by the model. Under real-time classification, it is smaller and only marginally significant.

Table 9: Retail Trading Response to Tweet Activity

Panel A: Coefficient path for $OIB_{j,t+1}$					
	(1)	(2)	(3)	(4)	(5)
$CharlatanVol_{j,t}$	5.169***	2.042**	1.075	1.295	0.603
	(0.618)	(0.980)	(1.233)	(1.232)	(1.634)
$ExpertVol_{j,t}$		6.360***	5.322***	4.950***	6.258
		(1.600)	(1.737)	(1.740)	(4.264)
Log tweet volume	No	No	Yes	Yes	Yes
Lagged return	No	No	No	Yes	Yes
Stock FE	No	No	No	No	Yes
R^2	0.001	0.001	0.001	0.002	0.031
N	180,116	180,116	180,116	180,116	179,991

Panel B: Full specification at longer horizons				
	$OIB_{j,t+2}$		$OIB_{j,t+5}$	
	OLS	FE	OLS	FE
$CharlatanVol_{j,t}$	−0.629	−0.505	0.333	0.536
	(1.237)	(1.289)	(1.236)	(1.744)
$ExpertVol_{j,t}$	8.773***	10.164**	6.192***	6.310*
	(1.772)	(4.008)	(1.786)	(3.721)
N	166,401	166,260	158,269	158,119

Event study: Mean $OIB_{j,t+1}$ by charlatan tweet volume	
High charlatan (D10)	56.6
Other stock-days	15.8
Difference	40.9 ($t = 13.23$)

Notes: Panel A traces the coefficient on charlatan message volume as expert volume, aggregate message volume, lagged returns, and stock fixed effects are added sequentially. Panel B reports the corresponding full specifications for retail order imbalance at two- and five-day horizons. The event study compares stock-days in the top decile of charlatan volume with all remaining stock-days. The final specifications use standard errors clustered by stock and date.

5.5 Heterogeneity by statistical learnability

The learning model yields cross-sectional comparative statics. Because $\rho = \Delta^2/(2\sigma_r^2)$, a noisier return environment makes advice quality harder to infer from a given number of recommendations. If slow learning contributes to attention misallocation, the return relation associated with charlatan exposure should therefore be stronger where recommendation quality is more difficult to evaluate.

We proxy for this environment using prior-year idiosyncratic volatility, market capitalization, and retail trading share. Idiosyncratic volatility is the standard deviation of daily Fama–French five-factor-adjusted returns over the preceding 12 months, requiring at least 100 observations. These variables do not directly measure the account-specific learning rate. Volatility can also mechanically increase realized return spreads, and small, volatile, retail-heavy stocks differ along many dimensions unrelated to learning. We therefore view these tests as supportive comparative-static evidence rather than as separate identification of the learning mechanism. The estimates currently use the previous coverage ratio and three-type classification and will be replaced by the $C_{j,m}$ specification under the $K = 3$ real-time posteriors.

Stock-level evidence. Monthly cross-sectional regressions relate subsequent 20-day abnormal returns to standardized charlatan exposure, standardized idiosyncratic volatility, their interaction, market capitalization, and retail trading share. Under the current specification, the average coefficient on charlatan exposure is -0.40 with a Fama–MacBeth t-statistic of -3.42. The interaction with idiosyncratic volatility is negative as predicted but imprecise, at -0.10 with a t-statistic of -0.94.

The portfolio comparison is more pronounced. Splitting the charlatan-exposure sort by idiosyncratic-volatility terciles gives a Q5–Q1 return spread of -0.52% in the lowest-volatility tercile, compared with -1.83% and -1.63% in the middle and highest terciles. The pattern is not monotonic across all three groups, but the spread is roughly three times as large outside the low-volatility segment.

The corresponding account-level exercise is less informative because most active users cover portfolios of stocks rather than a narrow volatility segment. Re-estimating the latent-quality model across stock-noise exposure terciles produces charlatan shares of 36.3%, 48.2%, and 44.0%. The quality-audience gradients are also similar across the three groups. Most of the useful variation in statistical noise therefore appears at the stock level rather than across accounts.

Taken together, these results are consistent with the prediction that the pricing relation is weaker where outcomes are easier to evaluate. They do not isolate statistical learnability from the other characteristics of small and volatile stocks. For that reason, the heterogeneity evidence plays a supporting role relative to the real-time exposure and within-stock results.

6 Quantitative Discipline

We close the empirical analysis with a quantitative illustration of the interaction between learning and reinforcement. The exercise is not a structural estimate of the full endogenous-slant model. Its purpose is narrower. We ask how frequently poor advice can dominate when the simulation is disciplined by the observed distribution of advice quality, recommendation noise, and the empirical relation between activity and audience.

Each simulated platform contains 100 accounts. Engagement efficiency is drawn from a lognormal distribution calibrated to observed activity, recommendation noise is taken from Section 3.1, and the amplification parameter is set to the reduced-form value $\nu = 1.39$ from Section 4.1. Latent quality in the current simulation is still drawn from the previous specification and will be replaced by the two-sided $K = 3$ distribution before the exercise is final.

Under the current calibration, one account eventually dominates every simulated platform. When followers place no weight on perceived quality, $\beta = 0$, the dominant account is a charlatan in 44% of simulations. Raising β to 1 lowers this probability to 26%, and raising it to 2 lowers it to 19%. Better use of quality information therefore improves the allocation of attention substantially, but under this calibration it does not eliminate the possibility that a poor adviser becomes dominant.

Amplification works in the opposite direction. Holding $\beta = 2$ fixed, increasing ν from 1.1 to 2 raises the charlatan-dominance probability from roughly 10% to 35%. When followers ignore quality, the corresponding probability remains close to the underlying charlatan population share because type does not affect attention. The simulation therefore illustrates the distinction at the center of the model. Better information strengthens selection on quality, while stronger reinforcement increases the value of early audience advantages.

Because both the quality distribution and the empirical estimate of amplification enter the exercise in reduced form, these magnitudes should not be interpreted as structural policy effects. Their role is to show that empirically plausible learning noise and audience reinforcement can be quantitatively important at the same time.

6.1 Attention-market design

The model distinguishes interventions that improve information from interventions that change how attention accumulates. Better disclosure of track records or ranking rules that place greater weight on estimated quality make followers more responsive to the evidence available about an account. In the model, these changes strengthen informational selection but leave the accumulation technology unchanged. They can reduce the probability that attention settles on poor advice without removing the history dependence created by superlinear reinforcement.

Changes to amplification operate on a different margin. A more linear relation between existing audience and subsequent visibility, weaker reliance on popularity signals, or depreciation of stale audience advantages reduces the extent to which early attention determines future exposure. These policies alter the reinforcement mechanism itself rather than the amount of information followers possess.

This distinction matters because improving the information available to followers and changing the allocation rule are not substitutes. When performance is noisy, even rational followers require long histories before quality becomes clear. A platform that simultaneously allows early popularity to compound can therefore commit substantial attention before that information arrives. The design problem is not only how much users know about quality, but how much of the future attention allocation is determined while they are still learning it.

7 Conclusion

Attention markets need not select for quality even when followers are rational and eventually learn. The problem is a mismatch in speeds. Advice quality is revealed through noisy realized returns, so Bayesian beliefs adjust slowly and influence follower choice through a bounded informational premium. Attention can move much faster. Audience raises the return to engagement, engagement raises visibility, and visibility attracts additional audience. When this feedback is superlinear, early

prominence can compound before quality becomes statistically distinguishable. Endogenous content choice can then turn history dependence into adverse selection. Charlatan accounts have stronger incentives to produce optimistic, attention-generating content, while expert accounts preserve the value of a reputation for quality.

The evidence is consistent with each part of this mechanism. Typical track records are far too short to distinguish reliably between positive- and negative-value advice. Across disjoint periods, realized account alpha is unstable, but behavioral strategies and stock habitats persist, and early optimism predicts lower advice quality later. Attention therefore accumulates around persistent strategies whose economic value is difficult to infer from short histories. Consistent with the amplification channel, activity rises with audience size, while optimism is the content choice most clearly associated with both larger audiences and lower expected advice value.

The consequences extend to prices. Using only recommendation outcomes available before portfolio formation, stocks with greater charlatan exposure earn 0.90% lower abnormal returns over the next 20 trading days. The relation survives within stock after controlling for momentum, liquidity, volatility, retail ownership, shortability, and other characteristics. Bullish recommendations by accounts classified as charlatans using only prior outcomes earn -0.59% out of sample, while bearish recommendations contain little corresponding positive information. Reversals are larger when shorting is costly and weaker among the largest, most easily arbitraged stocks. Retail trading responds primarily to the same recent returns and visible platform signal rather than to charlatan identity itself, consistent with the idea that attention operates through the broadcast investors observe.

The broader lesson is about the design of markets for information. Improving the supply of expert advice is not sufficient if the mechanism allocating attention can lock demand onto poor advice before quality is learned. Stronger preferences for quality help, but they operate through a bounded informational advantage. Reducing the amplification of accumulated popularity changes the feedback mechanism itself. In attention markets, the allocation rule is therefore part of the economic institution. Whether useful information reaches investors depends not only on who produces it, but also on how quickly evidence about quality arrives relative to how quickly attention compounds.

References

Acemoglu, Daron, Ali Makhdoumi, Azarakhsh Malekian, and Asuman Ozdaglar (Mar. 2025). "When Big Data Enables Behavioral Manipulation". In: *American Economic Review: Insights* 7.1, pp. 19–38.

Acemoglu, Daron, Asuman Ozdaglar, and James Siderius (Nov. 2024). "A Model of Online Misinformation". In: *The Review of Economic Studies* 91.6, pp. 3117–3150.

Banerjee, Abhijit V. (1992). "A Simple Model of Herd Behavior". In: *The Quarterly Journal of Economics* 107.3, pp. 797–817. DOI: 10.2307/2118364.

Barber, Brad M. and Terrance Odean (2008). "All That Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors". In: *The Review of Financial Studies* 21.2, pp. 785–818. DOI: 10.1093/rfs/hhm079.

Berk, Jonathan B. and Jules H. Van Binsbergen (2022). "Regulation of Charlatans in High-Skill Professions". In: *The Journal of Finance* 77.2, pp. 1219–1258.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades". In: *Journal of Political Economy* 100.5, pp. 992–1026. DOI: 10.1086/261849.

Bloomberg (2016). "EDF Research Note — Social Sentiment". In: *DOCS 2080750*.

Boehmer, Ekkehart, Charles M. Jones, Xiaoyan Zhang, and Xinran Zhang (2021). "Tracking retail investor activity". In: *The Journal of Finance* 76.5, pp. 2249–2305.

Cookson, J. Anthony, Joseph E. Engelberg, and William Mullins (2023). "Echo Chambers". In: *The Review of Financial Studies* 36.2, pp. 450–500.

Cookson, J. Anthony, Runjing Lu, William Mullins, and Marina Niessner (2024). "The Social Signal". In: *Journal of Financial Economics* 158, p. 103870.

Cookson, J. Anthony and Marina Niessner (2020). "Why Don't We Agree? Evidence from a Social Network of Investors". In: *The Journal of Finance* 75.1, pp. 173–228.

Crane, Alan and Kevin Crotty (2020). "How skilled are security analysts?" In: *The Journal of Finance* 75.3, pp. 1629–1675.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin (2009). "An Economic Model of Friendship: Homophily, Minorities, and Segregation". In: *Econometrica* 77.4, pp. 1003–1045.

D'Avolio, Gene (2002). "The Market for Borrowing Stock". In: *Journal of Financial Economics* 66.2–3, pp. 271–306. DOI: 10.1016/S0304-405X(02)00206-4.

Da, Zhi, Joseph Engelberg, and Pengjie Gao (2011). "In Search of Attention". In: *The Journal of Finance* 66.5, pp. 1461–1499. DOI: 10.1111/j.1540-6261.2011.01679.x.

Diether, Karl B., Christopher J. Malloy, and Anna Scherbina (2002). "Differences of Opinion and the Cross Section of Stock Returns". In: *The Journal of Finance* 57.5, pp. 2113–2141. DOI: 10.1111/0022-1082.00490.

Dim, Chukwuma (2022). "Should retail investors listen to social media analysts? evidence from text-implied beliefs". In: *Working Paper*.

Gervais, Simon, Anthony W. Lynch, and David K. Musto (2005). "Fund Families as Delegated Monitors of Money Managers". In: *The Review of Financial Studies* 18.4, pp. 1139–1169. DOI: 10.1093/rfs/hhi031.

Gervais, Simon and Günter Strobl (2020). "Transparency and Talent Allocation in Money Management". In: *The Review of Financial Studies* 33.8, pp. 3889–3924. DOI: 10.1093/rfs/hhz116.

Giannini, Robert, Paul Irvine, and Tao Shu (2018). "Nonlocal disadvantage: An examination of social media sentiment". In: *The Review of Asset Pricing Studies* 8.2, pp. 293–336.

Golub, Benjamin and Matthew O. Jackson (2012). "How Homophily Affects the Speed of Learning and Best-Response Dynamics". In: *The Quarterly Journal of Economics* 127.3, pp. 1287–1338.

Greenwood, Robin and Andrei Shleifer (2014). "Expectations of Returns and Expected Returns". In: *The Review of Financial Studies* 27.3, pp. 714–746.

Harrison, J. Michael and David M. Kreps (1978). "Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations". In: *Quarterly Journal of Economics* 92.2, pp. 323–336. DOI: 10.2307/1884166.

Harvey, Campbell R. and Yan Liu (2018). "Detecting Repeatable Performance". In: *The Review of Financial Studies* 31.7, pp. 2499–2552.

Hong, Harrison and Jeremy C. Stein (2003). "Differences of Opinion, Short-Sales Constraints, and Market Crashes". In: *Review of Financial Studies* 16.2, pp. 487–525. DOI: 10.1093/rfs/hhg006.

Kakhbod, Ali, Seyed Kazempour, Dmitry Livdan, and Norman Schürhoff (Mar. 2025). "Finfluencers". Working paper.

Miller, Edward M. (1977). "Risk, Uncertainty, and Divergence of Opinion". In: *Journal of Finance* 32.4, pp. 1151–1168. DOI: 10.1111/j.1540-6261.1977.tb03317.x.

Scharfstein, David S. and Jeremy C. Stein (1990). "Herd Behavior and Investment". In: *American Economic Review* 80.3, pp. 465–479.

Scheinkman, José A. and Wei Xiong (2003). "Overconfidence and Speculative Bubbles". In: *Journal of Political Economy* 111.6, pp. 1183–1219. DOI: 10.1086/378531.

Appendix

A Model Derivations and Proofs

Proof of Lemma 1 Fix an account i and suppress the account subscript when no confusion can arise. Conditional on type $\theta \in \{\mathrm{E}, \mathrm{C}\}$, the observations are independent with density $\sigma_r^{-1}\phi((r - \alpha_\theta)/\sigma_r)$. The factor σ_r^{-1} cancels from every likelihood ratio. Bayes' rule therefore gives posterior odds equal to prior odds times the product of recommendation-level likelihood ratios. For a single observation r, the log-likelihood ratio in favor of E is $\ln\phi((r-\alpha_\mathrm{E})/\sigma_r) - \ln\phi((r-\alpha_\mathrm{C})/\sigma_r) = -[(r-\alpha_\mathrm{E})^2 - (r-\alpha_\mathrm{C})^2]/(2\sigma_r^2)$. Expanding the difference of squares gives $(r-\alpha_\mathrm{C})^2 - (r-\alpha_\mathrm{E})^2 = 2(\alpha_\mathrm{E}-\alpha_\mathrm{C})r - (\alpha_\mathrm{E}^2 - \alpha_\mathrm{C}^2) = 2\Delta[r - (\alpha_\mathrm{E}+\alpha_\mathrm{C})/2]$. Hence, the recommendation-level log-likelihood ratio is $\Delta[r-(\alpha_\mathrm{E}+\alpha_\mathrm{C})/2]/\sigma_r^2$. Summing over the first n recommendations and adding the prior log odds $\ell_0 = \ln[P_\mathrm{E}/(1-P_\mathrm{E})]$ yields equation (2).

We next derive the distribution of one increment of the log posterior odds. Under $\theta = \mathrm{E}$, write $r_k = \alpha_\mathrm{E} + \sigma_r\varepsilon_k$. Then $r_k - (\alpha_\mathrm{E}+\alpha_\mathrm{C})/2 = \Delta/2 + \sigma_r\varepsilon_k$, so $L(k) - L(k-1) = \Delta^2/(2\sigma_r^2) + (\Delta/\sigma_r)\varepsilon_k = \rho + \sqrt{2\rho}\,\varepsilon_k$. The increment therefore has mean ρ and variance 2ρ. Under $\theta = \mathrm{C}$, $r_k - (\alpha_\mathrm{E}+\alpha_\mathrm{C})/2 = -\Delta/2 + \sigma_r\varepsilon_k$, which gives $L(k) - L(k-1) = -\rho + \sqrt{2\rho}\,\varepsilon_k$. This proves the drift and innovation-variance statement. Under symmetric normalization $\alpha_\mathrm{E} = a$ and $\alpha_\mathrm{C} = -a$, $\Delta = 2a$, and therefore $\rho = \Delta^2/(2\sigma_r^2) = 2a^2/\sigma_r^2$.

For posterior consistency, let $s_\theta = 1$ for $\theta = \mathrm{E}$ and $s_\theta = -1$ for $\theta = \mathrm{C}$. The preceding calculation gives $L(n) = \ell_0 + s_\theta\rho n + \sqrt{2\rho}\sum_{k=1}^n \varepsilon_k$. By the strong law of large numbers, $n^{-1}\sum_{k=1}^n \varepsilon_k \to 0$ almost surely, and hence $L(n)/n \to s_\theta\rho$ almost surely. Because $\rho > 0$, $L(n) \to +\infty$ under E and $L(n) \to -\infty$ under C. Since $p(n) = e^{L(n)}/[1+e^{L(n)}]$, it follows that $p(n) \to 1$ under E and $p(n) \to 0$ under C. This proves the posterior consistency.

We now derive the mean-evidence thresholds. Let $\ell_{\delta_\ell} = \ln[(1-\delta_\ell)/\delta_\ell] > 0$. The condition $|\ell_0| < \ell_{\delta_\ell}$ places the prior strictly between the upper and lower confidence boundaries. Under an expert account, $\mathbb{E}_\mathrm{E}L(n) = \ell_0 + \rho n$. Solving $\ell_0 + \rho n = \ell_{\delta_\ell}$ gives $n_\mathrm{E}^*(\delta_\ell) = (\ell_{\delta_\ell} - \ell_0)/\rho$. Under a charlatan account, $\mathbb{E}_\mathrm{C}L(n) = \ell_0 - \rho n$. Solving $\ell_0 - \rho n = -\ell_{\delta_\ell}$ gives $n_\mathrm{C}^*(\delta_\ell) = (\ell_{\delta_\ell} + \ell_0)/\rho$. If the prior is even, then $\ell_0 = 0$, so the two expressions coincide at $n^*(\delta_\ell) = \ell_{\delta_\ell}/\rho$.

For the sequential version, consider first an expert account and define the first-passage time $N_\mathrm{E} = \inf\{n \geq 0 : L(n) \geq \ell_{\delta_\ell}\}$. We first verify that $\mathbb{E}N_\mathrm{E} < \infty$. If $N_\mathrm{E} > n$, the process has not crossed the upper boundary by date n, so necessarily $L(n) < \ell_{\delta_\ell}$. Under an expert account, $L(n)$ is Gaussian with mean $\ell_0 + \rho n$ and variance $2\rho n$. Therefore $\Pr(N_\mathrm{E} > n) \leq \Phi([\ell_{\delta_\ell} - \ell_0 - \rho n]/\sqrt{2\rho n})$. The argument of Φ has an asymptotic of $-\sqrt{\rho n/2}$, so the Gaussian tail bound implies $\Pr(N_\mathrm{E} > n) \leq Ce^{-cn}$ for all sufficiently large n and some constants $C, c > 0$. Hence $\mathbb{E}N_\mathrm{E} = \sum_{n\geq 0}\Pr(N_\mathrm{E} > n) < \infty$.

Define the overshoot $\mathcal{O}_\mathrm{E} = L(N_\mathrm{E}) - \ell_{\delta_\ell} \geq 0$. It also has a finite expectation. Let $X_k = \rho + \sqrt{2\rho}\varepsilon_k$ denote the iid increments. If crossing occurs on date n, then $L(n-1) < \ell_{\delta_\ell}$ and $\mathcal{O}_\mathrm{E} = X_n - [\ell_{\delta_\ell} - L(n-1)] < X_n^+$. For $x > 0$, $\{N_\mathrm{E} = n, \mathcal{O}_\mathrm{E} > x\} \subseteq \{N_\mathrm{E} \geq n, X_n > x\}$. The event

$\{N_{\mathrm{E}} \geq n\}$ is measurable with respect to increments through date $n-1$ and is therefore independent of X_n. Summing over n gives $\Pr(\mathcal{O}_{\mathrm{E}} > x) \leq \sum_{n\geq 1} \Pr(N_{\mathrm{E}} \geq n) \Pr(X_1 > x) = \mathbb{E}N_{\mathrm{E}} \Pr(X_1 > x)$. Integrating this tail inequality over $x > 0$ yields $\mathbb{E}\mathcal{O}_{\mathrm{E}} \leq \mathbb{E}N_{\mathrm{E}}\, \mathbb{E}X_1^+ < \infty$.

The stopped sum representation is $L(N_{\mathrm{E}}) = \ell_0 + \sum_{k=1}^{N_{\mathrm{E}}} X_k$. The increments are integrable and $\mathbb{E}N_{\mathrm{E}} < \infty$, so Wald's identity applies and gives $\mathbb{E}L(N_{\mathrm{E}}) = \ell_0 + \rho\, \mathbb{E}N_{\mathrm{E}}$. By the definition of overshoot, $\mathbb{E}L(N_{\mathrm{E}}) = \ell_{\delta_\ell} + \mathbb{E}\mathcal{O}_{\mathrm{E}}$. Equating the two expressions yields $\mathbb{E}N_{\mathrm{E}} = [\ell_{\delta_\ell} - \ell_0 + \mathbb{E}\mathcal{O}_{\mathrm{E}}]/\rho$, and hence $\mathbb{E}N_{\mathrm{E}} \geq (\ell_{\delta_\ell} - \ell_0)/\rho = n_{\mathrm{E}}^*$.

For a charlatan account, apply the same argument to $-L(n)$. Its increments have mean ρ, the relevant upper boundary is ℓ_{δ_ℓ}, and the initial value is $-\ell_0$. If $N_{\mathrm{C}} = \inf\{n \geq 0 : L(n) \leq -\ell_{\delta_\ell}\}$ and $\mathcal{O}_{\mathrm{C}} = -L(N_{\mathrm{C}}) - \ell_{\delta_\ell} \geq 0$, then $\mathbb{E}N_{\mathrm{C}} = [\ell_{\delta_\ell} + \ell_0 + \mathbb{E}\mathcal{O}_{\mathrm{C}}]/\rho \geq n_{\mathrm{C}}^*$. Thus, the mean-evidence thresholds weakly understate the expected sequential time required to cross the corresponding posterior boundary.

Finally, every threshold above is proportional to $1/\rho = 2\sigma_r^2/\Delta^2$. Holding the confidence level and prior fixed, the threshold therefore increases linearly in σ_r^2 and decreases at the rate Δ^{-2} in the quality gap. This proves the lemma. ■

The visibility lock-in benchmark The mean-field benchmark has two accounts, one expert and one charlatan. Followers arrive at rate λ and audience evolves according to $\dot{F}_i(t) = \lambda q_i(t)$, where $q_i(t)$ is given by equation (5). Let $s_{\mathrm{C}}(t) = q_{\mathrm{C}}(t)$ be the charlatan account's share of the arriving flow.

Proposition A.1 (Visibility lock-in). *Suppose that at some date t_0,*

$$\eta_{\mathrm{C}} F_{\mathrm{C}}(t_0)^{\nu-1} > e^{\beta} \eta_{\mathrm{E}} F_{\mathrm{E}}(t_0)^{\nu-1}. \tag{A1}$$

Then equation (A1) continues to hold for every $t \geq t_0$ regardless of the subsequent path of beliefs, $F_{\mathrm{C}}(t)/F_{\mathrm{E}}(t) \to \infty$, and $s_{\mathrm{C}}(t) \to 1$.

Proof. Write $h_i(t) = e^{\beta p_i(t)}$. Since $p_i(t) \in [0,1]$, $1 \leq h_i(t) \leq e^{\beta}$ for both accounts and every date. Let $Q(t) = \kappa\eta_{\mathrm{C}} F_{\mathrm{C}}(t)^{\nu} h_{\mathrm{C}}(t) + \kappa\eta_{\mathrm{E}} F_{\mathrm{E}}(t)^{\nu} h_{\mathrm{E}}(t)$. The denominator is strictly positive because all efficiencies and follower bases are positive. Equation (5) implies $\dot{F}_i(t) = \lambda\kappa\eta_i F_i(t)^{\nu} h_i(t)/Q(t)$.

Step 1. Forward invariance of the lock-in region. Define $M(t) = \eta_{\mathrm{C}} F_{\mathrm{C}}(t)^{\nu-1} - e^{\beta}\eta_{\mathrm{E}} F_{\mathrm{E}}(t)^{\nu-1}$. Equation (A1) is exactly the condition $M(t_0) > 0$. Differentiating and substituting the audience laws gives $\dot{M}(t) = (\nu-1)\eta_{\mathrm{C}} F_{\mathrm{C}}^{\nu-2}\dot{F}_{\mathrm{C}} - (\nu-1)e^{\beta}\eta_{\mathrm{E}} F_{\mathrm{E}}^{\nu-2}\dot{F}_{\mathrm{E}}$, and hence $\dot{M}(t) = [(\nu-1)\lambda\kappa/Q(t)][\eta_{\mathrm{C}}^2 F_{\mathrm{C}}^{2\nu-2} h_{\mathrm{C}} - e^{\beta}\eta_{\mathrm{E}}^2 F_{\mathrm{E}}^{2\nu-2} h_{\mathrm{E}}]$. Using $h_{\mathrm{C}} \geq 1$ in the positive term and $h_{\mathrm{E}} \leq e^{\beta}$ in the negative term yields $\dot{M}(t) \geq [(\nu-1)\lambda\kappa/Q(t)][(\eta_{\mathrm{C}} F_{\mathrm{C}}^{\nu-1})^2 - (e^{\beta}\eta_{\mathrm{E}} F_{\mathrm{E}}^{\nu-1})^2]$. Factoring the difference of squares gives $\dot{M}(t) \geq a(t)M(t)$, where $a(t) = [(\nu-1)\lambda\kappa/Q(t)][\eta_{\mathrm{C}} F_{\mathrm{C}}^{\nu-1} + e^{\beta}\eta_{\mathrm{E}} F_{\mathrm{E}}^{\nu-1}] > 0$. Therefore, on any interval on which $M(t) \geq 0$, $\dot{M}(t) \geq 0$. If M were to hit zero for the first time at some $\tau > t_0$, continuity would imply $M(t) > 0$ on $[t_0, \tau)$ and hence M would be nondecreasing on that interval,

which is incompatible with $M(t_0) > 0$ and $M(\tau) = 0$. Thus $M(t) > 0$ for every $t \geq t_0$. In fact, $M(t) \geq M(t_0)$ for all such t.

Step 2. The audience ratio diverges. Let $u(t) = F_{\mathrm{C}}(t)/F_{\mathrm{E}}(t)$. Since both follower bases remain positive, $d\ln u(t)/dt = \dot{F}_{\mathrm{C}}/F_{\mathrm{C}} - \dot{F}_{\mathrm{E}}/F_{\mathrm{E}}$. Substituting the audience laws gives $d\ln u(t)/dt = [\lambda\kappa/Q(t)][\eta_{\mathrm{C}}F_{\mathrm{C}}^{\nu-1}h_{\mathrm{C}} - \eta_{\mathrm{E}}F_{\mathrm{E}}^{\nu-1}h_{\mathrm{E}}]$. Using $h_{\mathrm{C}} \geq 1$ and $h_{\mathrm{E}} \leq e^{\beta}$ gives $d\ln u(t)/dt \geq [\lambda\kappa/Q(t)]M(t) \geq [\lambda\kappa/Q(t)]M(t_0) > 0$. Hence $u(t)$ is strictly increasing.

Suppose for contradiction that $u(t)$ is bounded. Since it is increasing, it then converges to a finite limit $\bar{u} \geq u(t_0) > 0$. Total audience satisfies $d[F_{\mathrm{C}}(t) + F_{\mathrm{E}}(t)]/dt = \lambda[q_{\mathrm{C}}(t) + q_{\mathrm{E}}(t)] = \lambda$, so $F_{\mathrm{C}}(t) + F_{\mathrm{E}}(t) = F_{\mathrm{C}}(t_0) + F_{\mathrm{E}}(t_0) + \lambda(t - t_0)$. For sufficiently large t, this total is bounded below by $c_0 t$ for some $c_0 > 0$ and bounded above by $C_0 t$ for some finite C_0. Since $u(t) \leq \bar{u}$, $F_{\mathrm{E}}(t) = [F_{\mathrm{C}}(t)+F_{\mathrm{E}}(t)]/[1+u(t)] \geq c_0 t/(1+\bar{u})$. Since $u(t) \geq u(t_0) > 0$, $F_{\mathrm{C}}(t) = u(t)F_{\mathrm{E}}(t) \geq u(t_0)c_0 t/(1+\bar{u})$. Thus both follower bases grow at least linearly under the bounded-ratio hypothesis. They also grow at most linearly because neither exceeds total audience.

The denominator satisfies $Q(t) \leq \kappa e^{\beta}[\eta_{\mathrm{C}}F_{\mathrm{C}}(t)^{\nu} + \eta_{\mathrm{E}}F_{\mathrm{E}}(t)^{\nu}] \leq C_1 t^{\nu}$ for some finite C_1. To obtain a lower bound on $M(t)$ of the correct order, write $M(t) = \eta_{\mathrm{C}}F_{\mathrm{C}}(t)^{\nu-1}[1 - (e^{\beta}\eta_{\mathrm{E}}/\eta_{\mathrm{C}})u(t)^{-(\nu-1)}]$. The bracket is increasing in u, and $u(t) \geq u(t_0)$. Since $M(t_0) > 0$, $\varepsilon_0 = 1 - (e^{\beta}\eta_{\mathrm{E}}/\eta_{\mathrm{C}})u(t_0)^{-(\nu-1)} > 0$. Hence $M(t) \geq \eta_{\mathrm{C}}\varepsilon_0 F_{\mathrm{C}}(t)^{\nu-1} \geq C_2 t^{\nu-1}$ for some $C_2 > 0$ and all sufficiently large t.

Combining the two bounds yields $d\ln u(t)/dt \geq (\lambda\kappa C_2/C_1)t^{-1}$ for all sufficiently large t. Integrating from a sufficiently large deterministic date t_1 to t gives $\ln u(t) - \ln u(t_1) \geq (\lambda\kappa C_2/C_1)\ln(t/t_1)$, which diverges as $t \to \infty$. This contradicts $u(t) \leq \bar{u}$. Therefore $u(t) \to \infty$.

Step 3. The charlatan absorbs the arriving flow. The charlatan's choice share can be written as $s_{\mathrm{C}}(t) = [1 + \eta_{\mathrm{E}}F_{\mathrm{E}}(t)^{\nu}h_{\mathrm{E}}(t)/(\eta_{\mathrm{C}}F_{\mathrm{C}}(t)^{\nu}h_{\mathrm{C}}(t))]^{-1}$. Using $h_{\mathrm{E}} \leq e^{\beta}$, $h_{\mathrm{C}} \geq 1$, and $F_{\mathrm{E}}/F_{\mathrm{C}} = u^{-1}$ gives $s_{\mathrm{C}}(t) \geq [1 + (e^{\beta}\eta_{\mathrm{E}}/\eta_{\mathrm{C}})u(t)^{-\nu}]^{-1}$. Since $u(t) \to \infty$, the lower bound converges to one. Because $s_{\mathrm{C}}(t) \leq 1$, the squeeze theorem gives $s_{\mathrm{C}}(t) \to 1$. ■

Proof of Proposition 1 We use the arrival-indexed timing convention in Section 1.1. Conditional on the complete sequence of recommendation outcomes for every account, the belief path of account i can be written as the deterministic sequence $p_i(0), p_i(1), \ldots$ indexed by that account's own follower counter. The recommendation innovations are independent of the randomness that allocates follower arrivals. Hence, conditional on the performance histories, the only remaining randomness in the attention process is the arrival allocation itself.

Part (a). Strong monopoly for $\nu > 1$. Define $h_i(m) = e^{\beta p_i(m)}$. The posterior bound gives $1 \leq h_i(m) \leq e^{\beta}$ for every account and every counter. Attach to each pair (i, m) an independent unit exponential random variable $E_{i,m}$. When account i currently has $F_i(0) + m$ followers, assign it the waiting time $E_{i,m}/[\eta_i(F_i(0) + m)^{\nu}h_i(m)]$. If the current counters are $(m_1, \ldots, m_N)$, these waiting times are independent exponential random variables with rates $\eta_i(F_i(0) + m_i)^{\nu}h_i(m_i)$. For independent exponentials with rates $\lambda_1, \ldots, \lambda_N$, the probability that clock i rings first is $\lambda_i / \sum_j \lambda_j$.

Therefore, the probability that the next ring belongs to account i is exactly equation (5), with the common factor κ omitted because it cancels from the numerator and denominator. After a clock rings, the memoryless property allows the remaining clocks to be restarted without changing the law. The embedded process therefore reproduces the sequential attention process exactly.

The total embedded time required for account i to accumulate infinitely many additional followers is $T_i = \sum_{m=0}^{\infty} E_{i,m}/[\eta_i(F_i(0)+m)^{\nu}h_i(m)]$. Since $h_i(m) \geq 1$, $\mathbb{E}[T_i \mid \{p_i(m)\}_m] \leq \eta_i^{-1}\sum_{m=0}^{\infty}(F_i(0)+m)^{-\nu}$. The series is finite because $\nu > 1$. Hence the nonnegative random variable T_i has finite conditional expectation and is finite almost surely. Integrating over the performance histories preserves this conclusion.

We also need uniqueness of the smallest explosion time. Conditional on the belief path, write $T_i = a_{i,0}E_{i,0} + Y_i$, where $a_{i,0} = [\eta_i F_i(0)^{\nu}h_i(0)]^{-1} > 0$ and $Y_i = \sum_{m\geq 1} E_{i,m}/[\eta_i(F_i(0)+m)^{\nu}h_i(m)]$ is independent of $E_{i,0}$. Conditional on Y_i, the first term has a continuous density, so T_i has no atoms. Explosion clocks are independent across accounts conditional on the belief paths because they are functions of disjoint exponential sequences. Therefore $\Pr(T_i = T_j) = 0$ for every $i \neq j$. Since the number of accounts is finite, the minimizer $i^{\star} = \arg\min_i T_i$ is unique almost surely.

Fix a realization with unique minimizer $i^{\star}$. For each rival $j \neq i^{\star}$, $T_j > T_{i^{\star}}$. The partial ring times of rival j increase to T_j, so only finitely many of its rings occur before time $T_{i^{\star}}$. In contrast, the partial ring times of account $i^{\star}$ increase to $T_{i^{\star}}$ and therefore contain infinitely many rings before or arbitrarily close to that time. The discrete arrival sequence is obtained by ordering all ring times. Consequently, after the last rival ring that occurs before $T_{i^{\star}}$, every remaining discrete arrival belongs to $i^{\star}$. Thus $i^{\star}$ receives all but finitely many followers. This proves part (a).

Part (b). Positive probability of charlatan monopoly and comparative statics. Fix a labeled charlatan account i. For readability, write this account as C. Define its actual explosion time T_{C} and the rival explosion times T_j as above. The event that this labeled account monopolizes is $\{T_{\mathrm{C}} < R_{\mathrm{C}}\}$, where $R_{\mathrm{C}} = \min_{j\neq \mathrm{C}} T_j$.

Construct a comparison system using the same exponential variables. Since $h_{\mathrm{C}}(m) \geq 1$, $T_{\mathrm{C}} \leq \bar{T}_{\mathrm{C}}$, where $\bar{T}_{\mathrm{C}} = \sum_{m\geq 0} E_{\mathrm{C},m}/[\eta_{\mathrm{C}}(F_{\mathrm{C}}(0)+m)^{\nu}]$. Since $h_j(m) \leq e^{\beta}$ for every rival, $T_j \geq \underline{T}_j$, where $\underline{T}_j = e^{-\beta}\sum_{m\geq 0} E_{j,m}/[\eta_j(F_j(0)+m)^{\nu}]$. Therefore $\{\bar{T}_{\mathrm{C}} < \min_{j\neq \mathrm{C}} \underline{T}_j\} \subseteq \{T_{\mathrm{C}} < R_{\mathrm{C}}\}$.

The comparison event has strictly positive probability. To see this directly, choose any $c > 0$. Each untilted explosion clock has support arbitrarily close to zero because all of a sufficiently long finite set of exponential increments can be small and the expected remaining tail can be made arbitrarily small. It also has unbounded support because its first exponential term has unbounded support. Hence $\Pr(\bar{T}_{\mathrm{C}} < c) > 0$ and $\Pr(\underline{T}_j > c) > 0$ for every rival j. These events are independent across accounts, so $\Pr(\bar{T}_{\mathrm{C}} < c,\ \underline{T}_j > c\ \forall j \neq \mathrm{C}) = \Pr(\bar{T}_{\mathrm{C}} < c)\prod_{j\neq \mathrm{C}} \Pr(\underline{T}_j > c) > 0$. This event is contained in the comparison winning event. Hence $\underline{q}_{\mathrm{C}}(\beta) = \Pr(\bar{T}_{\mathrm{C}} < \min_{j\neq \mathrm{C}} \underline{T}_j) > 0$. The bound uses only the universal inequalities $1 \leq h_i \leq e^{\beta}$ and therefore holds uniformly in the learning rate ρ.

Let $\bar{R}_{\mathrm{C}} = \min_{j\neq \mathrm{C}} \sum_{m\geq 0} E_{j,m}/[\eta_j(F_j(0)+m)^{\nu}]$. Then $\min_{j\neq \mathrm{C}} \underline{T}_j = e^{-\beta}\bar{R}_{\mathrm{C}}$ and $\underline{q}_{\mathrm{C}}(\beta) = \Pr(\bar{T}_{\mathrm{C}} <$

$e^{-\beta}\bar{R}_{\mathrm{C}}$). If $\beta_2 > \beta_1$, then $e^{-\beta_2}\bar{R}_{\mathrm{C}} < e^{-\beta_1}\bar{R}_{\mathrm{C}}$ almost surely, so the winning event at β_2 is contained in the winning event at β_1. Strict inclusion has positive probability. Indeed, conditional on any realization $r > 0$ of $\bar{R}_{\mathrm{C}}$, the interval $[e^{-\beta_2}r, e^{-\beta_1}r)$ has positive length, and $\bar{T}_{\mathrm{C}}$ has a density that is strictly positive on $(0, \infty)$. Hence $\Pr(e^{-\beta_2}\bar{R}_{\mathrm{C}} \leq \bar{T}_{\mathrm{C}} < e^{-\beta_1}\bar{R}_{\mathrm{C}}) > 0$. Therefore $\underline{q}_{\mathrm{C}}(\beta)$ is strictly decreasing in β.

As $\beta \to \infty$, $e^{-\beta}\bar{R}_{\mathrm{C}} \to 0$ almost surely. Because $\bar{T}_{\mathrm{C}} > 0$ almost surely, $\mathbb{1}\{\bar{T}_{\mathrm{C}} < e^{-\beta}\bar{R}_{\mathrm{C}}\} \to 0$ almost surely. The indicators are bounded by one, so dominated convergence gives $\underline{q}_{\mathrm{C}}(\beta) \to 0$.

The event that the labeled account C monopolizes is contained in the event that some charlatan monopolizes. Hence whenever the realized type profile contains at least one charlatan, $\Pr(\theta_{i^\star} = \mathrm{C}) \geq \Pr(i^\star = \mathrm{C}) \geq \underline{q}_{\mathrm{C}}(\beta) > 0$. Taking the maximum of these account-specific lower bounds over all realized charlatan accounts gives the stated lower bound for the union event.

We next establish the comparative statics in the labeled account's endowed efficiency and initial audience. Hold its entire belief path, all rival paths, and all exponential variables fixed. Increasing η_{C} strictly increases every denominator in the summands of T_{C} and therefore strictly decreases T_{C} for every realization. Increasing $F_{\mathrm{C}}(0)$ also strictly decreases every summand because $x \mapsto x^{-\nu}$ is strictly decreasing. Rival clocks are unchanged. Thus the winning event $\{T_{\mathrm{C}} < R_{\mathrm{C}}\}$ expands pathwise. Weak monotonicity of the winning probability follows immediately. Strict monotonicity follows from Lemma 5: the own clock is strictly lower with probability one, it is independent of the minimum rival clock under the exponential coupling, and the minimum rival clock has a strictly increasing distribution function on $(0, \infty)$.

For completeness, we derive the local slow-learning expansion used in the discussion after the proposition. Fix a counter m and let $L_i^\rho(m) = \ell_0 + s_i\rho m + \sqrt{2\rho m}\, Z_i$, where $s_i = 1$ for an expert, $s_i = -1$ for a charlatan, and Z_i is standard normal. Define $f(\ell) = \exp\{\beta p(\ell)\}$ with $p(\ell) = e^\ell/(1+e^\ell)$. At the prior log odds ℓ_0, $p(\ell_0) = P_{\mathrm{E}}$, $p'(\ell_0) = P_{\mathrm{E}}(1-P_{\mathrm{E}})$, and $f'(\ell_0) = e^{\beta P_{\mathrm{E}}}\beta P_{\mathrm{E}}(1-P_{\mathrm{E}})$. A second-order Taylor expansion around ℓ_0 gives $\mathbb{E}f(L_i^\rho(m)) = f(\ell_0) + f'(\ell_0)s_i\rho m + \frac{1}{2}f''(\ell_0)\mathbb{E}[(L_i^\rho(m) - \ell_0)^2] + o(\rho)$. The remainder is $o(\rho)$ because p and all of its derivatives are bounded, $f(\ell) = e^{\beta p(\ell)}$ has a bounded third derivative, and $\mathbb{E}|L_i^\rho(m) - \ell_0|^3 = O(\rho^{3/2})$ for fixed m. Now $\mathbb{E}[(L_i^\rho(m) - \ell_0)^2] = 2\rho m + \rho^2 m^2$, which is the same for the two types up to an $O(\rho^2)$ term. The second-order contribution is therefore common across types to first order in ρ and cancels in the difference. Consequently, $\mathbb{E}[h_{\mathrm{E}}(m)] - \mathbb{E}[h_{\mathrm{C}}(m)] = 2e^{\beta P_{\mathrm{E}}}\beta P_{\mathrm{E}}(1 - P_{\mathrm{E}})\rho m + o(\rho)$. Since both expectations converge to $e^{\beta P_{\mathrm{E}}}$, dividing the charlatan expectation by the expert expectation and using $(a+\varepsilon_C)/(a+\varepsilon_E) = 1 + (\varepsilon_C - \varepsilon_E)/a + o(\rho)$ gives $\mathbb{E}[h_{\mathrm{C}}(m)]/\mathbb{E}[h_{\mathrm{E}}(m)] = 1 - 2\beta P_{\mathrm{E}}(1 - P_{\mathrm{E}})\rho m + o(\rho)$. This is a local comparison of expected informational tilts at a fixed counter, not a monotonicity theorem for the nonlinear monopoly probability.

Part (c). The slow-learning limit. Consider any deterministic sequence $\rho_\ell \downarrow 0$ and construct the economies on a common probability space using the same standard-normal recommendation innovations and the same urn exponentials for every value of ρ_ℓ. For fixed account i and fixed counter m, $L_i^{\rho_\ell}(m) = \ell_0 + s_i\rho_\ell m + \sqrt{2\rho_\ell}\sum_{k=1}^m \varepsilon_{i,k}$. The deterministic term $s_i\rho_\ell m$ converges to zero.

The finite random sum $\sum_{k=1}^{m} \varepsilon_{i,k}$ is almost surely finite, while $\sqrt{2\rho_\ell} \to 0$. Hence $L_i^{\rho_\ell}(m) \to \ell_0$ almost surely. Since the set of pairs (i, m) is countable, intersecting the corresponding probability-one events gives simultaneous convergence for every account and every finite counter.

The logistic map is continuous, so $p_i^{\rho_\ell}(m) \to P_{\mathrm{E}}$ almost surely and $h_i^{\rho_\ell}(m) = e^{\beta p_i^{\rho_\ell}(m)} \to e^{\beta P_{\mathrm{E}}}$ almost surely. The embedded clock is $T_i^{\rho_\ell} = \sum_{m \geq 0} E_{i,m}/[\eta_i (F_i(0)+m)^\nu h_i^{\rho_\ell}(m)]$. Because $h_i^{\rho_\ell}(m) \geq 1$, the mth summand is bounded above by $E_{i,m}/[\eta_i (F_i(0) + m)^\nu]$. The sum of these dominating terms has finite expectation $\eta_i^{-1} \sum_m (F_i(0) + m)^{-\nu}$ and is therefore finite almost surely. Dominated convergence for series now gives $T_i^{\rho_\ell} \to e^{-\beta P_{\mathrm{E}}} \sum_{m \geq 0} E_{i,m}/[\eta_i (F_i(0) + m)^\nu]$ almost surely.

Let T_i^0 denote the limiting clock. The vector $(T_1^0, \ldots, T_N^0)$ has a unique minimizer almost surely. On the event of a unique minimizer I^0, define the positive gap $d = \min_{j \neq I^0} (T_j^0 - T_{I^0}^0) > 0$. Joint convergence of the finite clock vector implies that, for all sufficiently large ℓ, $\max_i |T_i^{\rho_\ell} - T_i^0| < d/3$. Then $T_{I^0}^{\rho_\ell} < T_j^{\rho_\ell}$ for every $j \neq I^0$, so the identity of the monopolist is eventually equal to I^0. Thus the winner converges almost surely under the coupling and hence in distribution.

The common factor $e^{-\beta P_{\mathrm{E}}}$ multiplies every limiting clock and does not affect the minimizer. Conditional on the realized initial pairs $(\eta_i, F_i(0))$, the winner therefore depends only on the independent exponential sequences. By Assumption 1, those exponential sequences and the initial pairs are independent of the type vector. Hence the limiting winner is independent of type conditional on the realized initial states. This proves part (c).

Part (d). The linear benchmark. Set $\nu = 1$. Conditional on the performance histories, account i is a pure-birth process whose rate when it has received m additional followers is $\lambda_i(m) = \eta_i (F_i(0) + m) h_i(m)$. The embedded explosion time satisfies $T_i \geq e^{-\beta} \eta_i^{-1} \sum_{m \geq 0} E_{i,m}/[F_i(0) + m]$. Decompose this series as $\sum_m 1/[F_i(0) + m] + \sum_m (E_{i,m} - 1)/[F_i(0) + m]$. The first series diverges because it is a shifted harmonic series. The terms in the second series are independent, have mean zero, and have variances $1/[F_i(0) + m]^2$. The sum of these variances is finite. Kolmogorov's convergence theorem therefore implies that the second series converges almost surely to a finite random variable. The sum of a divergent positive harmonic series and an almost surely finite correction diverges to $+\infty$. Hence $T_i = \infty$ almost surely for every account. No account can therefore accumulate infinitely many rings in finite embedded time, so strong monopoly by finite-time explosion is impossible.

Every account nevertheless receives infinitely many followers as continuous time tends to infinity. To see this, suppose account i received only $M < \infty$ additional followers. Its state would eventually remain fixed at $F_i(0)+M$, and its instantaneous rate would then remain bounded below by $\eta_i [F_i(0)+ M] > 0$ because $h_i \geq 1$. A counting process with a rate bounded below by a positive constant cannot have only finitely many future rings over an infinite time interval. This contradiction proves that every account's follower count diverges.

By Lemma 1, $p_i(m) \to \mathbb{1}\{\theta_i = \mathrm{E}\}$ almost surely as the account receives infinitely many followers. Hence $h_i(m) \to h_i^\infty = e^{\beta \mathbb{1}\{\theta_i = \mathrm{E}\}}$. Define $\lambda_i^\infty = \eta_i h_i^\infty$. Fix $\varepsilon \in (0, \lambda_i^\infty)$. For all sufficiently large counters m, $\lambda_i^\infty - \varepsilon \leq \eta_i h_i(m) \leq \lambda_i^\infty + \varepsilon$. At those counters, the birth rate is bounded between $(\lambda_i^\infty - \varepsilon)[F_i(0) + m]$ and $(\lambda_i^\infty + \varepsilon)[F_i(0) + m]$. Couple the process after that counter with two Yule

processes having the same current population and per-capita rates $\lambda_i^\infty - \varepsilon$ and $\lambda_i^\infty + \varepsilon$. Monotonicity of exponential waiting times gives the pathwise sandwich $Z_i^-(t) \le F_i(t) \le Z_i^+(t)$ after an irrelevant finite time shift. For a Yule process with per-capita rate $a > 0$, $e^{-at}Z(t)$ converges almost surely to a finite strictly positive random variable. Hence $t^{-1}\ln Z_i^-(t) \to \lambda_i^\infty - \varepsilon$ and $t^{-1}\ln Z_i^+(t) \to \lambda_i^\infty + \varepsilon$. The sandwich gives $\lambda_i^\infty - \varepsilon \le \liminf_{t\to\infty} t^{-1}\ln F_i(t) \le \limsup_{t\to\infty} t^{-1}\ln F_i(t) \le \lambda_i^\infty + \varepsilon$. Letting $\varepsilon \downarrow 0$ yields $t^{-1}\ln F_i(t) \to \lambda_i^\infty$ almost surely.

Suppose a unique account i° has the largest asymptotic growth rate. For every rival j choose $\varepsilon > 0$ small enough that $\lambda_{i^\circ}^\infty - \varepsilon > \lambda_j^\infty + \varepsilon$. The preceding exponential-rate limits imply $\ln[F_j(t)/F_{i^\circ}(t)]/t \to \lambda_j^\infty - \lambda_{i^\circ}^\infty < 0$, so $F_j(t)/F_{i^\circ}(t) \to 0$ almost surely. Since there are finitely many rivals, $F_{i^\circ}(t)/\sum_k F_k(t) = [1 + \sum_{j\neq i^\circ} F_j(t)/F_{i^\circ}(t)]^{-1} \to 1$. For an expert account $\lambda_i^\infty = \eta_i e^\beta$, while for a charlatan $\lambda_i^\infty = \eta_i$. This gives the stated condition for the identity of the dominant type. The limiting exponential rate does not depend on the finite initial audience $F_i(0)$, so a finite initial lead cannot overturn a strict asymptotic-rate ranking. If several accounts share the maximal rate, the argument only eliminates lower-rate accounts and makes no claim that the tied accounts have deterministic relative shares. This proves part (d). ■

Proof of Proposition 2 We verify the candidate separating profile $b_{\mathrm{E}}^* = 0$ and $b_{\mathrm{C}}^* = \bar{b}$, establish the informational limits used by the two attention blocks, prove incentive compatibility, and then prove each stated consequence.

Step 1. Coupling and posterior under the candidate profile. The triangular sequence is constructed on a common probability space. The type vector, endowed efficiencies, initial audiences, payoff coefficients, urn exponentials, and standardized return and behavioral innovations are fixed across economies. Only the return-noise scale, the two record lengths, the attention horizons, and the behavioral-signal precisions vary with the economy index m. This coupling is used only to establish convergence statements and does not change any economy's marginal distribution.

Under the candidate strategies, the equilibrium return means are $\mu_{\mathrm{E}}^* = a$ and $\mu_{\mathrm{C}}^* = (1-\bar{b})(-a) - \bar{b}h = -a - \bar{b}(h-a)$. Let $d^* = \mu_{\mathrm{E}}^* - \mu_{\mathrm{C}}^* = 2a + \bar{b}(h-a) > 0$. Then $\mathcal{I}^* = (d^*)^2/(2\sigma_r^2)$.

At either information horizon, let $\bar{r}$ denote the observed average return from n recommendations and let z denote the behavioral statistic used at that horizon. Under the candidate profile, Bayes' rule gives $p(\bar{r}, z) = P_{\mathrm{E}} f_{\mathrm{E}}(z)\phi_{\sigma_r/\sqrt{n}}(\bar{r} - \mu_{\mathrm{E}}^*)/[P_{\mathrm{E}} f_{\mathrm{E}}(z)\phi_{\sigma_r/\sqrt{n}}(\bar{r} - \mu_{\mathrm{E}}^*) + (1 - P_{\mathrm{E}}) f_{\mathrm{C}}(z)\phi_{\sigma_r/\sqrt{n}}(\bar{r} - \mu_{\mathrm{C}}^*)]$. Full support of both behavioral densities implies that the denominator is strictly positive for every realized $(\bar{r}, z)$. Thus posterior beliefs are defined by Bayes' rule after every signal realization, including realizations generated by a unilateral deviation. No off-path belief specification is needed.

The corresponding log posterior odds in favor of E equal $\ell_0 + \Lambda_r(\bar{r}) + \Lambda_z(z)$, where $\ell_0 = \ln[P_{\mathrm{E}}/(1-P_{\mathrm{E}})]$, $\Lambda_z(z) = \ln[f_{\mathrm{E}}(z)/f_{\mathrm{C}}(z)]$, and $\Lambda_r(\bar{r}) = n[(\bar{r} - \mu_{\mathrm{C}}^*)^2 - (\bar{r} - \mu_{\mathrm{E}}^*)^2]/(2\sigma_r^2)$.

Step 2. Early beliefs are asymptotically uninformative. Consider any of the four type-action combinations $(\mathrm{E}, 0)$, $(\mathrm{C}, \bar{b})$, $(\mathrm{C}, 0)$, and $(\mathrm{E}, \bar{b})$. Let its true mean be μ. Write $\bar{r} = \mu + \sigma_r Z/\sqrt{n_E}$ with Z standard normal. Expanding the record log-likelihood ratio around the true mean gives

$\Lambda_r(\bar{r}) = n_E[(\mu - \mu^*_{\mathrm{C}})^2 - (\mu - \mu^*_{\mathrm{E}})^2]/(2\sigma_r^2) + \sqrt{n_E}\, d^* Z/\sigma_r$. The first term equals $n_E \mathcal{I}^* \Gamma(\mu)$, where $\Gamma(\mu) = [(\mu - \mu^*_{\mathrm{C}})^2 - (\mu - \mu^*_{\mathrm{E}})^2]/(d^*)^2$. Because the structural means are fixed along the triangular sequence, $\Gamma(\mu)$ is a finite constant for each of the four cases. The second term has variance $n_E(d^*)^2/\sigma_r^2 = 2n_E\mathcal{I}^*$. Therefore, the deterministic term converges to zero, and the stochastic term converges to zero in probability whenever $n_E\mathcal{I}^* \to 0$. The early behavioral condition gives $\ln LR_{z,E} = o_p(1)$, equivalently $\Lambda_z(z_E) = o_p(1)$ up to the chosen orientation of the likelihood ratio. Hence the complete early log posterior odds converge in probability to ℓ_0. Continuity of the logistic map gives $p_i^E \to P_{\mathrm{E}}$ in probability under each type-action combination.

Step 3. Late records classify the equilibrium strategies and both deviations. At the late horizon, the same expansion with n_L gives $\Lambda_r(\bar{r}) = n_L\mathcal{I}^*\Gamma(\mu) + \sqrt{2n_L\mathcal{I}^*}\, Z$. Dividing by $n_L\mathcal{I}^*$ yields $\Lambda_r(\bar{r})/(n_L\mathcal{I}^*) = \Gamma(\mu) + \sqrt{2/(n_L\mathcal{I}^*)}\, Z$. Since $n_L\mathcal{I}^* \to \infty$, the stochastic term converges to zero in probability. Thus the normalized record log-likelihood ratio converges to $\Gamma(\mu)$.

For equilibrium expert play, $\mu = \mu^*_{\mathrm{E}}$, so $\Gamma(\mu^*_{\mathrm{E}}) = 1$. For equilibrium charlatan play, $\mu = \mu^*_{\mathrm{C}}$, so $\Gamma(\mu^*_{\mathrm{C}}) = -1$.

If a charlatan deviates to no slant, its true mean is $\mu^d_{\mathrm{C}} = -a$. Its distance from the charlatan equilibrium mean is $\mu^d_{\mathrm{C}} - \mu^*_{\mathrm{C}} = \bar{b}(h - a)$, while its distance from the expert equilibrium mean is $\mu^d_{\mathrm{C}} - \mu^*_{\mathrm{E}} = -2a$. Hence $\Gamma(\mu^d_{\mathrm{C}}) = [\bar{b}^2(h-a)^2 - 4a^2]/[2a + \bar{b}(h-a)]^2$. The denominator is strictly positive. The upper bound $\bar{b} < 2a/(h-a)$ in equation (11) implies $\bar{b}^2(h-a)^2 < 4a^2$, so $\Gamma(\mu^d_{\mathrm{C}}) < 0$.

If an expert deviates to maximal slant, its true mean is $\mu^d_{\mathrm{E}} = (1-\bar{b})a - \bar{b}h = a - \bar{b}(a+h)$. Its distance from the charlatan equilibrium mean is $\mu^d_{\mathrm{E}} - \mu^*_{\mathrm{C}} = 2a(1-\bar{b})$, while its distance from the expert equilibrium mean is $\mu^d_{\mathrm{E}} - \mu^*_{\mathrm{E}} = -\bar{b}(a+h)$. Therefore $\Gamma(\mu^d_{\mathrm{E}}) = [4a^2(1-\bar{b})^2 - \bar{b}^2(a+h)^2]/[2a + \bar{b}(h-a)]^2$. The numerator is negative exactly when $2a(1-\bar{b}) < \bar{b}(a+h)$, which is equivalent to $\bar{b} > 2a/(3a+h)$. This is the lower bound in equation (11). Hence $\Gamma(\mu^d_{\mathrm{E}}) < 0$.

The same lower bound also implies that the expert deviation has a negative raw mean. Indeed, $\mu^d_{\mathrm{E}} < 0$ is equivalent to $\bar{b} > a/(a+h)$. Since $h > a$, $2a/(3a+h) > a/(a+h)$ because $2a(a+h) > a(3a+h)$ is equivalent to $h > a$. Thus $\bar{b} > 2a/(3a+h)$ implies $\mu^d_{\mathrm{E}} < 0$.

The late behavioral likelihood ratio is $o_p(n_L\mathcal{I}^*)$. Hence it cannot change the sign of a record log-likelihood ratio that diverges at rate $n_L\mathcal{I}^*$ with a nonzero limiting coefficient. Under $(\mathrm{E}, 0)$ the log posterior odds therefore diverge to $+\infty$, while under $(\mathrm{C}, \bar{b})$, $(\mathrm{C}, 0)$, and $(\mathrm{E}, \bar{b})$ they diverge to $-\infty$. Consequently, $p_i^L \to 1$ under equilibrium expert play and $p_i^L \to 0$ under equilibrium charlatan play and under either unilateral deviation.

Step 4. Incentive compatibility for a charlatan. Fix the strategies of all rival accounts at the candidate profile and compare the charlatan's two actions. Under both $b_i = 0$ and $b_i = \bar{b}$, Step 3 gives $p_i^L \to 0$ in probability. Since Φ is continuous on the compact interval $[0, 1]$, it is bounded. Convergence in probability to zero and boundedness imply $\mathbb{E}_{\mathrm{C}}\Phi(p_i^L) \to \Phi(0)$ under both actions. Therefore, the difference in the two franchise terms converges to zero.

The effective multiplicative weight of the charlatan at the early information date is $\tilde{\eta}_i(b)e^{\beta p_i^E(b)}$. The ratio of the slant weight to the no-slant weight is $R_{E,m} = g(\bar{b})^{1/\chi} \exp\{\beta[p^E_{\mathrm{C},m}(\bar{b}) - p^E_{\mathrm{C},m}(0)]\}$.

By Step 2, both posteriors converge to P_{E}, so $R_{E,m} \to g(\bar{b})^{1/\chi} > 1$ in probability. At the late date the analogous ratio is $R_{L,m} = g(\bar{b})^{1/\chi} \exp\{\beta[p^L_{\mathrm{C},m}(\bar{b}) - p^L_{\mathrm{C},m}(0)]\}$. By Step 3, both late posteriors converge to zero, so $R_{L,m} \to g(\bar{b})^{1/\chi} > 1$ in probability. Choose any number r with $1 < r < g(\bar{b})^{1/\chi}$. Then the event $\mathcal{G}_m = \{R_{E,m} \geq r,\ R_{L,m} \geq r\}$ has probability converging to one.

On $\mathcal{G}_m$, the charlatan has a weakly larger own multiplicative weight throughout each block under slant than under no slant, with rival strategies fixed. We now construct an explicit finite-horizon monotone coupling. Pre-generate a rival-only urn path. After the rivals have collectively received r draws, let $B(r)$ be their total current weight and let the identity of the rival selected on the next rival draw be generated by a common uniform variable using probabilities proportional to the rival weights. Conditional on the next global draw not going to the charlatan, these rival-choice probabilities do not depend on the charlatan's own multiplicative weight. Because each rival weight is increasing in its own count, $B(r)$ is nondecreasing in r.

Couple the slant and no-slant economies with the same rival-only path and a common uniform variable U_n for the own-versus-rival decision at every global arrival n. Suppose inductively that after n arrivals the slant economy has received at least as many charlatan draws as the no-slant economy. Then it has received weakly fewer rival draws, say $r^S_n \leq r^0_n$. Its charlatan follower count is weakly larger, its own multiplicative coefficient is weakly larger on $\mathcal{G}_m$, and its rival total weight satisfies $B(r^S_n) \leq B(r^0_n)$. Therefore its probability of assigning arrival $n+1$ to the charlatan is weakly larger. Using the same U_{n+1}, whenever the no-slant economy assigns the arrival to the charlatan, the slant economy does so as well. The only asymmetric case is that the slant economy assigns the arrival to the charlatan while the no-slant economy assigns it to a rival, which preserves the induction inequality. If both assign the arrival to a rival, each consumes the next draw from the common rival-only path corresponding to its own rival-draw index. Thus the charlatan count under slant is weakly larger after every global arrival. Starting from equal initial counts proves the claim by induction for the entire block.

Applying this coupling first to the early block and then, conditional on the coupled early terminal states, to the later block gives $s^E_{i,m}(\bar{b}) \geq s^E_{i,m}(0)$ and $s^L_{i,m}(\bar{b}) \geq s^L_{i,m}(0)$ on $\mathcal{G}_m$. Outside $\mathcal{G}_m$, every share lies in $[0,1]$, so the negative part of either share difference is bounded by one. Since $\Pr(\mathcal{G}^c_m) \to 0$, the expected negative part of each share difference is $o(1)$.

Let $B_i(b) = \Delta_0(s^0_i)^\nu + \Delta_E[s^E_i(b)]^\nu + \Delta_L[s^L_i(b)]^\nu$. The difference in the attention component of the charlatan payoff is $A_i[g(\bar{b})^\nu B_i(\bar{b}) - B_i(0)]$. The initial share s^0_i is action-independent, so its contribution is exactly $\kappa_{\mathrm{C}i} = A_i[g(\bar{b})^\nu - 1]\Delta_0(s^0_i)^\nu > 0$. On $\mathcal{G}_m$, the latter share terms are nonnegative because $g(\bar{b})^\nu \geq 1$ and the slant shares are weakly larger. On the complement, their negative contribution is uniformly bounded by $A_i(\Delta_E + \Delta_L)$ times a constant independent of m. Hence their expected negative contribution is $o(1)$. Adding the $o(1)$ franchise difference gives $U_{\mathrm{C}i}(\bar{b}; b_{-i}) - U_{\mathrm{C}i}(0; b_{-i}) \geq \kappa_{\mathrm{C}i} + o(1)$. Since $\kappa_{\mathrm{C}i} > 0$, the difference is strictly positive for all sufficiently large m. Thus every charlatan strictly prefers maximal slant.

Step 5. Incentive compatibility for an expert. Under no slant, Step 3 gives $p^L_i(0) \to 1$ in probability.

Under maximal slant, Step 3 gives $p_i^L(\bar{b}) \to 0$ in probability. Continuity and boundedness of Φ therefore imply $\mathbb{E}_{\mathrm{E}}[\Phi(p_i^L(0)) - \Phi(p_i^L(\bar{b}))] \to \Phi(1) - \Phi(0)$. The limiting reputation gain from avoiding slant is $\Omega_i[\Phi(1) - \Phi(0)]$.

The largest possible attention-payoff gain from slant is bounded independently of the induced audience path. Since every audience share lies in $[0,1]$, $\Delta_0(s_i^0)^\nu + \Delta_E(s_i^E)^\nu + \Delta_L(s_i^L)^\nu \le \Delta_\Sigma$. The slant attention payoff is therefore at most $A_i g(\bar{b})^\nu \Delta_\Sigma$, while the no-slant attention payoff is nonnegative. Hence [attention payoff under slant] − [attention payoff under no slant] $\le A_i g(\bar{b})^\nu \Delta_\Sigma$. Condition (10) gives the strict margin $d_i = \Omega_i[\Phi(1) - \Phi(0)] - A_i g(\bar{b})^\nu \Delta_\Sigma > 0$. For sufficiently large m, the realized expected franchise advantage from no slant is within $d_i/2$ of its limit. It then exceeds the maximal possible attention advantage from slant by at least $d_i/2$. Thus $U_{\mathrm{E}i}(0; b_{-i}) > U_{\mathrm{E}i}(\bar{b}; b_{-i})$ for all sufficiently large m. Every expert strictly prefers no slant.

There are finitely many accounts. Taking the maximum of the finitely many economy thresholds from Steps 4 and 5 gives a common m_0 such that the candidate strategy profile is a Nash equilibrium for every $m \ge m_0$. This proves existence of the separating equilibrium. The argument does not rule out other equilibria and therefore makes no uniqueness claim.

Step 6. Proof of part (a). Fix a late return record $\bar{r}^L$. Under the candidate equilibrium, posterior odds of charlatan type relative to expert type are proportional to

$$[(1 - P_{\mathrm{E}})/P_{\mathrm{E}}][f_{\mathrm{C},L}(z^L)/f_{\mathrm{E},L}(z^L)][\phi_{\sigma_r/\sqrt{n_L}}(\bar{r}^L - \mu_{\mathrm{C}}^*)/\phi_{\sigma_r/\sqrt{n_L}}(\bar{r}^L - \mu_{\mathrm{E}}^*)].$$

For fixed $\bar{r}^L$, every factor except $f_{\mathrm{C},L}(z^L)/f_{\mathrm{E},L}(z^L)$ is constant in z^L. By the maintained MLRP, that likelihood ratio is strictly increasing in z^L. Therefore, the posterior odds of C are increasing in z^L, and so is the posterior probability of C. The return-chasing footprint follows from the definition of baited content as optimistic calls on recent winners.

Step 7. Proof of part (b). Equation (6) gives $\tilde{\eta}_i(\bar{b})/\tilde{\eta}_i(0) = g(\bar{b})^{1/\chi} > 1$. Holding the belief path fixed therefore multiplies the account's effective choice weight at every own counter by the same factor greater than one. Under the exponential embedding, every own waiting time is strictly reduced when the same exponential variables are used. Lemma 5 then gives a weakly larger monopoly probability, with strictness whenever the rival minimum clock has the stated full-support property. The finite-horizon monotone coupling in Step 4 gives the corresponding weak increase in the account's audience path at every arrival. Thus maximal slant raises effective engagement efficiency, the audience path, and the probability of monopoly when beliefs are held fixed.

Step 8. Proof of part (c). By Step 2, $p_{i,m}^E \to P_{\mathrm{E}}$ in probability for every account under the candidate equilibrium. Hence the early multiplicative coefficient $c_{i,m} = \tilde{\eta}_i(b_i^*)e^{\beta p_{i,m}^E}$ converges in probability to $c_i^0 = \tilde{\eta}_i(b_i^*)e^{\beta P_{\mathrm{E}}}$. Lemma 2 therefore implies that the identity of the strong monopolist in the infinite extension of the early block converges in probability to the winner of the pure superlinear urn with weights $\eta_i g(\bar{b})^{1/\chi} F_i^\nu$ for charlatans and $\eta_i F_i^\nu$ for experts. The common factor $e^{\beta P_{\mathrm{E}}}$ has no effect on the identity of the winner.

Consider the limiting pure urn and a fixed labeled account i. Let $\omega_\theta = \Pr(i \text{ wins the limiting early urn} \mid \theta_i = \theta)$. Use the same realization of the account's endowed efficiency, initial audience, own exponentials, and all rival clocks when comparing the two possible values of θ_i. Assumption 1 makes the distribution of the endowed variables the same under the two types. If the labeled account is an expert, its own clock is $T_i^{\mathrm{E}} = \sum_m E_{i,m}/[\eta_i(F_i(0)+m)^\nu]$ up to the common factor. If it is a charlatan, the slant strategy multiplies its rate by $g(\bar{b})^{1/\chi}$, so $T_i^{\mathrm{C}} = g(\bar{b})^{-1/\chi} T_i^{\mathrm{E}} < T_i^{\mathrm{E}}$ almost surely. The rival clocks do not change. By Lemma 5, $\omega_{\mathrm{C}} > \omega_{\mathrm{E}}$.

Exchangeability gives the unconditional probability that the labeled account wins as $1/N$. Conditioning on its type, $1/N = (1-P_{\mathrm{E}})\omega_{\mathrm{C}} + P_{\mathrm{E}}\omega_{\mathrm{E}}$. The probability that the winner is a charlatan is $\sum_{i=1}^N \Pr(\theta_i = \mathrm{C},\ i \text{ wins}) = N(1-P_{\mathrm{E}})\omega_{\mathrm{C}}$. Substituting the exchangeability identity gives $\pi_{\mathrm{C}}^E = (1-P_{\mathrm{E}})\omega_{\mathrm{C}}/[(1-P_{\mathrm{E}})\omega_{\mathrm{C}} + P_{\mathrm{E}}\omega_{\mathrm{E}}]$. Because $\omega_{\mathrm{C}} > \omega_{\mathrm{E}}$, $\pi_{\mathrm{C}}^E > 1 - P_{\mathrm{E}}$. Lemma 2 transfers this probability statement from the limiting pure urn to the infinite extension of the early urn in economy m.

It remains to connect the infinite early urn to the finite early block and then to the later continuation. This is a diagonal issue because both the urn law and the early horizon vary with m. Lemma 3 shows that at the actual terminal date $M_{E,m}$, the strong monopolist I_m of the infinite early urn enters the later block with an audience diverging in probability, while the vector of rival audiences is tight. Corollary 2 then shows that, under the bounded later posterior tilts, the probability that any rival overturns this lead in the later block converges to zero. Hence $\Pr_m(i_m^\star = I_m) \to 1$. Combining this with convergence of the early winner's type gives $\Pr_m(\theta_{i_m^\star} = \mathrm{C}) \to \pi_{\mathrm{C}}^E > 1 - P_{\mathrm{E}}$.

Step 9. Proof of part (d). Under no slant, $\alpha(\mathrm{C}, 0) = -a$. Under maximal slant, $\alpha(\mathrm{C}, \bar{b}) = (1-\bar{b})(-a) - \bar{b}h = -a - \bar{b}(h-a)$. Subtracting gives $\alpha(\mathrm{C}, \bar{b}) - \alpha(\mathrm{C}, 0) = -\bar{b}(h-a) < 0$ because $\bar{b} > 0$ and $h > a$. This proves part (d) and completes the proof. ■

Lemma 2 (Continuity of the early monopolist) *Conditional on the initial states and the equilibrium slant choices, define $c_{i,m} = \tilde{\eta}_i(b_i^*)e^{\beta p_{i,m}^E}$ and $c_i^0 = \tilde{\eta}_i(b_i^*)e^{\beta P_{\mathrm{E}}}$. Using independent unit exponentials, let $Z_i = \sum_{k\geq 0} E_{i,k}/(F_i(0)+k)^\nu$ and define $T_{i,m}^E = Z_i/c_{i,m}$ and $T_i^{E,0} = Z_i/c_i^0$. If $p_{i,m}^E \to P_{\mathrm{E}}$ in probability for every account, then the clock vector converges in probability to its limiting clock vector, and the identity of the early strong monopolist converges in probability to that of the pure urn with weights $\tilde{\eta}_i(b_i^*)F_i^\nu$.*

Proof. Since the early posterior is fixed within the early block, $c_{i,m}$ is constant over that block and the explosion clock factors as $T_{i,m}^E = Z_i/c_{i,m}$. Continuity of the exponential map gives $e^{\beta p_{i,m}^E} \to e^{\beta P_{\mathrm{E}}}$ in probability, so $c_{i,m} \to c_i^0$ in probability. Because N is finite, $\max_i |c_{i,m} - c_i^0| \to 0$ in probability.

Each Z_i is strictly positive almost surely because all summands are nonnegative and the first summand is strictly positive almost surely. It is finite almost surely because $\mathbb{E}Z_i = \sum_{k\geq 0}(F_i(0)+k)^{-\nu} < \infty$ for $\nu > 1$. Also $c_i^0 > 0$. Since inversion is continuous on $(0,\infty)$, $1/c_{i,m} \to 1/c_i^0$ in

probability. Multiplying by the almost surely finite Z_i and using Slutsky's theorem gives $T^E_{i,m} \to T^{E,0}_i$ in probability for every i. Finiteness of N then gives joint convergence of the clock vector.

The limiting clocks are independent across accounts. Each has an absolutely continuous distribution because $T^{E,0}_i = a_i E_{i,0} + Y_i$ with $a_i > 0$, $E_{i,0}$ exponential, and Y_i independent of $E_{i,0}$. Hence ties occur with probability zero. On the event of a unique limiting minimizer, the gap between the smallest and second-smallest clock is strictly positive. Any vector sufficiently close to the limiting vector has the same minimizer. The argmin map is therefore continuous almost surely at the limiting vector. The continuous mapping theorem gives convergence in probability of the winner. The common factor $e^{\beta P_{\mathrm{E}}}$ in every c^0_i cancels from the ranking, leaving the pure-urn weights $\tilde{\eta}_i(b^*_i)F^\nu_i$. ■

Lemma 3 (Finite-horizon inheritance of early monopoly) *Use the common exponential coupling of Lemma 2 and suppose $c_{i,m} \to c^0_i$ in probability for every account. Let I_m be the strong monopolist of the infinite early urn in economy m, and let H_m be the total number of arrivals received by all accounts other than I_m over the entire infinite early urn. Then $\{H_m\}_{m\geq 1}$ is tight. Consequently, whenever $M_{E,m} \to \infty$, $F^E_{I_m}(M_{E,m}) \to \infty$ in probability and $\max_{j\neq I_m} F^E_j(M_{E,m}) = O_p(1)$.*

Proof. We use the subsequence characterization of tightness. Take an arbitrary subsequence of economies. Since the finite vector $c_m = (c_{1,m}, \ldots, c_{N,m})$ converges in probability to the deterministic conditional-limit vector c^0, there exists a further subsequence along which $c_m \to c^0$ almost surely. Work on that further subsequence and on the common exponential probability space.

For account i, define the time of its rth ring by $S_{i,m}(r) = \sum_{k=0}^{r-1} E_{i,k}/[c_{i,m}(F_i(0)+k)^\nu]$ and its explosion time by $T_{i,m} = \sum_{k=0}^{\infty} E_{i,k}/[c_{i,m}(F_i(0)+k)^\nu]$. For every fixed r, $S_{i,m}(r) \to S_{i,0}(r)$ almost surely because it is a finite sum and $c_{i,m} \to c^0_i$. For the infinite sum, note that along almost every convergent realization, $c_{i,m}$ is eventually bounded below by $c^0_i/2$. The summands are then dominated by $2E_{i,k}/[c^0_i(F_i(0)+k)^\nu]$, whose series is finite almost surely. Dominated convergence for series therefore gives $T_{i,m} \to T_{i,0}$ almost surely.

With probability one, the limiting clocks have a unique minimizer I_0. Since the number of accounts is finite, the gap between the smallest and second-smallest limiting clock is positive. Therefore $I_m = I_0$ for all sufficiently large m along the selected subsequence.

Fix a losing account $j \neq I_0$. Since $T_{j,0} > T_{I_0,0}$ and $S_{j,0}(r) \uparrow T_{j,0}$, there exists a finite integer r_j such that $S_{j,0}(r_j) > T_{I_0,0}$. We may choose the realization outside the probability-zero event on which any finite losing partial ring time equals the winning explosion time. The almost sure convergence of $S_{j,m}(r_j)$ and $T_{I_m,m} = T_{I_0,m}$ then implies $S_{j,m}(r_j) > T_{I_m,m}$ for all sufficiently large m. Thus loser j experiences at most $r_j - 1$ rings before the winning account explodes. In the corresponding discrete urn, it therefore receives at most $r_j - 1$ additional followers over the entire infinite process.

Summing over the finitely many losing accounts gives the almost surely finite random bound $H_m \leq H^0 = \sum_{j\neq I_0}(r_j - 1)$ for all sufficiently large m along the selected subsequence. For any $\varepsilon > 0$, choose K with $\Pr(H^0 > K) < \varepsilon$. Then $\limsup_m \Pr(H_m > K) \leq \varepsilon$ along the further subsequence. If the original sequence were not tight, there would exist an $\varepsilon_0 > 0$ and a subsequence for which

no such uniform tail bound exists. The construction above would produce a further subsequence contradicting that property. Hence the original sequence $\{H_m\}$ is tight.

At the terminal date of the finite early block, the number of losing arrivals is no greater than the total number of losing arrivals in the infinite extension, namely H_m. Therefore every rival's terminal audience is at most its initial audience plus H_m, and the largest rival audience is $O_p(1)$ because initial audiences have bounded support and $H_m = O_p(1)$. The leader receives at least $M_{E,m} - H_m$ of the $M_{E,m}$ early arrivals. Since $M_{E,m} \to \infty$ deterministically and $H_m = O_p(1)$, $M_{E,m} - H_m \to \infty$ in probability. Adding the leader's positive initial audience gives $F^E_{I_m}(M_{E,m}) \to \infty$ in probability. ■

Lemma 4 (Stochastic persistence of a large early lead) *Suppose every account's later choice tilt satisfies $1 \leq e^{\beta p_i(m)} \leq e^{\beta}$. Let account i enter the later block with K followers while rival j enters with L_j, and define, using independent unit exponentials, $\overline{T}_i(K) = \sum_{m\geq 0} E_{i,m}/[\tilde{\eta}_i(K+m)^{\nu}]$ and $\underline{T}_j(L_j) = e^{-\beta}\sum_{m\geq 0} E_{j,m}/[\tilde{\eta}_j(L_j+m)^{\nu}]$. The actual explosion time of account i is no greater than $\overline{T}_i(K)$, and the actual explosion time of rival j is no less than $\underline{T}_j(L_j)$. If $K \to \infty$ while every L_j remains bounded, then the conditional probability that account i has the smallest actual explosion time converges to one.*

Proof. When the leader has received m additional later followers, its actual rate is $\tilde{\eta}_i e^{\beta p_i(m)}(K+m)^{\nu} \geq \tilde{\eta}_i(K+m)^{\nu}$. Using the same exponential variables term by term, its actual waiting time at every counter is therefore no larger than the corresponding waiting time in $\overline{T}_i(K)$. Summing gives $T_i \leq \overline{T}_i(K)$. For rival j, the actual rate is at most $e^{\beta}\tilde{\eta}_j(L_j+m)^{\nu}$, so the actual waiting time is at least $e^{-\beta}E_{j,m}/[\tilde{\eta}_j(L_j+m)^{\nu}]$. Hence $T_j \geq \underline{T}_j(L_j)$.

The leader's upper comparison clock converges to zero in probability. Indeed, $\mathbb{E}\overline{T}_i(K) = \tilde{\eta}_i^{-1}\sum_{m\geq 0}(K+m)^{-\nu}$. For $K \geq 2$, the integral comparison gives $\sum_{m\geq 0}(K+m)^{-\nu} \leq K^{-\nu} + \int_{K-1}^{\infty} x^{-\nu}dx = K^{-\nu} + (K-1)^{1-\nu}/(\nu-1)$. Thus $\mathbb{E}\overline{T}_i(K) = O(K^{1-\nu}) \to 0$. Markov's inequality implies $\overline{T}_i(K) \to 0$ in probability.

Now suppose the rival entry counts lie in a fixed bounded set. Each $\underline{T}_j(L_j)$ is strictly positive almost surely because its first summand is strictly positive almost surely. There are finitely many rivals, so $R = \min_{j\neq i}\underline{T}_j(L_j) > 0$ almost surely. Consequently, $\Pr(R \leq c) \downarrow \Pr(R = 0) = 0$ as $c \downarrow 0$. Given $\varepsilon > 0$, choose $c > 0$ such that $\Pr(R \leq c) < \varepsilon$. Then choose K large enough that $\Pr(\overline{T}_i(K) \geq c) < \varepsilon$. On the event $\{\overline{T}_i(K) < c < R\}$, $T_i \leq \overline{T}_i(K) < c < R \leq \min_{j\neq i} T_j$. Thus account i has the smallest actual explosion time. The reversal probability is at most 2ε, which proves the claim. ■

Corollary 2 (Random terminal states) *Suppose the leader enters the later block with $K_m \to \infty$ in probability and the vector of rival entry counts is tight. Under the bounded later tilts of Lemma 4, $\Pr_m(i^{\star}_m \neq I_m) \to 0$.*

Proof. Fix $\varepsilon > 0$. Tightness of the rival vector implies that there exists $A < \infty$ such that, for all sufficiently large m, the probability that any rival entry count exceeds A is below ε.

We make the comparison in Lemma 4 uniform over all states in which every rival count is at most A. Assumption 1 gives a common lower bound $\eta_\ell > 0$ and upper bound $\eta_h < \infty$ for endowed efficiency. Since the slant multiplier is either 1 or $g(\bar{b})^{1/\chi}$, effective efficiencies are bounded between $\underline{\eta} = \eta_\ell$ and $\overline{\eta} = \eta_h g(\bar{b})^{1/\chi}$. The leader's worst upper clock, for a given K, occurs at the smallest effective efficiency $\underline{\eta}$. Every rival's lower comparison clock is smallest when its entry count is A and its effective efficiency is $\overline{\eta}$. Therefore the probability of reversal is bounded above by the reversal probability in a fixed worst-case comparison system with leader clock $\overline{T}^{wc}(K) = \sum_m E_m/[\underline{\eta}(K+m)^\nu]$ and finitely many rival clocks $\underline{T}_j^{wc}(A) = e^{-\beta}\sum_m E_{j,m}/[\overline{\eta}(A+m)^\nu]$. By Lemma 4, this worst-case reversal probability converges to zero as $K \to \infty$. Choose K_0 so that it is below ε for every $K \geq K_0$.

Since $K_m \to \infty$ in probability, $\Pr(K_m < K_0) < \varepsilon$ for all sufficiently large m. The event of reversal is contained in the union of three events: some rival count exceeds A, the leader count is below K_0, or reversal occurs in the worst-case comparison despite all counts satisfying the bounds. The union bound therefore gives $\Pr_m(i_m^\star \neq I_m) < 3\varepsilon$ for all sufficiently large m. Since ε is arbitrary, the probability converges to zero. ■

The deterministic benchmark in Appendix A and the stochastic continuation lemma express the same economic force in different environments. In the deterministic mean-field model, crossing the lock-in boundary makes reversal impossible. In the sequential urn, every finite lead can in principle be reversed, but Lemma 4 shows that the reversal probability converges to zero as the early lead diverges.

Lemma 5 (Pathwise ordering and strict monopoly comparison) *For account i, let the effective choice-weight path be $X_i(m) = \tilde{\eta}_i(b_i)e^{\beta p_i(m)}$ and define $T_i = \sum_{m\geq 0} E_{i,m}/[X_i(m)(F_i(0)+m)^\nu]$. If two coupled versions A and C satisfy $X_i^A(m) \geq X_i^C(m)$ for every m and $F_i^A(0) \geq F_i^C(0)$ almost surely, then $T_i^A \leq T_i^C$ almost surely and, holding the rival clocks fixed, the monopoly probability of version A is weakly larger. For strictness, let $R_i = \min_{j\neq i} T_j$. If R_i is independent of (T_i^A, T_i^C), its distribution function is strictly increasing on $(0,\infty)$, and $\Pr(T_i^A < T_i^C) > 0$, then $\Pr(T_i^A < R_i) > \Pr(T_i^C < R_i)$.*

Proof. Use the same own-account exponential sequence in versions A and C. For every counter m, $X_i^A(m) \geq X_i^C(m)$ and $F_i^A(0)+m \geq F_i^C(0)+m$. Since $\nu > 0$, the denominator of the mth summand in version A is at least as large as in version C. Thus every summand of T_i^A is weakly smaller, and summing gives $T_i^A \leq T_i^C$ almost surely.

Holding the rival clocks fixed, if version C wins then $T_i^C < R_i$. Since $T_i^A \leq T_i^C$, this implies $T_i^A < R_i$. Hence $\{T_i^C < R_i\} \subseteq \{T_i^A < R_i\}$, which proves weak monotonicity of the monopoly probability.

For strictness, let $F_R(t) = \Pr(R_i \le t)$. Independence of R_i from the two own clocks gives $\Pr(T_i^A < R_i) = \mathbb{E}[1 - F_R(T_i^A)]$ and $\Pr(T_i^C < R_i) = \mathbb{E}[1 - F_R(T_i^C)]$. Subtracting yields $\Pr(T_i^A < R_i) - \Pr(T_i^C < R_i) = \mathbb{E}[F_R(T_i^C) - F_R(T_i^A)]$. Because $T_i^A \le T_i^C$, the integrand is nonnegative. On the event $\{T_i^A < T_i^C\}$ it is strictly positive because F_R is strictly increasing. This event has positive probability by assumption, so the expectation is strictly positive.

We finally verify the full-support property used throughout the urn comparisons. In the urn applications, $X_j(m) = \tilde{\eta}_j e^{\beta p_j(m)} \ge \tilde{\eta}_j > 0$, so the expected coefficients of the tail clock are bounded by a constant multiple of $(F_j(0) + m)^{-\nu}$ and are summable for $\nu > 1$. Conditional on any effective-weight path, a rival clock can be written as $T_j = a_{j,0}E_{j,0} + Y_j$, where $a_{j,0} > 0$ and Y_j is independent of $E_{j,0}$. The tail Y_j has zero in its support. To see this, fix $\varepsilon > 0$. Choose K so large that the expected tail after K is below $\varepsilon/4$. Markov's inequality then gives positive probability that this infinite tail is below $\varepsilon/2$. The preceding finite collection of exponential terms can all be made small enough that their sum is below $\varepsilon/2$, an event with positive probability. Independence gives $\Pr(Y_j < \varepsilon) > 0$.

Conditional on $Y_j = y < t$, the random variable $a_{j,0}E_{j,0} + y$ has density $a_{j,0}^{-1} \exp[-(t-y)/a_{j,0}]$ at t. Integrating over Y_j gives a density for T_j that is strictly positive at every $t > 0$ because $\Pr(Y_j < t) > 0$. The first exponential term also gives unbounded support. Thus every rival explosion clock has an absolutely continuous distribution with full support on $(0, \infty)$. The minimum of finitely many independent rival clocks has distribution function $F_R(t) = 1 - \prod_{j \ne i}[1 - F_j(t)]$, which is strictly increasing on $(0, \infty)$. ■

Marginal first-order stochastic dominance of a one-date scalar weight is not enough to order explosion times when the entire belief path changes. The block timing of Proposition 2 avoids that difficulty. Early posterior tilts become asymptotically common while the early allocation is determined, and Lemma 4 handles the later block.

Pricing block details For a stock j with positive total visible coverage, the visibility share is $w_j = \sum_{i:\theta_i=\mathrm{C}} V_i \mathbb{1}\{i \text{ covers } j\} / \sum_i V_i \mathbb{1}\{i \text{ covers } j\}$. It is the charlatan share of visibility among the accounts covering the stock. Propositions 1 and 2 provide the mechanism that can raise this share above its population benchmark and leave it heterogeneous across stocks.

At date 1, retail investors observe the feed and submit demand according to equation (16). The parameter restriction $\delta^- < \delta$ captures asymmetric ability to trade on negative sentiment, and $\delta^- = 0$ is the binding case used in Proposition 4(c). For part (c) and Corollary 1, place the residual primitive vector $(R_{j,0}, f_j, u_j, e_j)$ on a common probability space independently of (w_j, BC_j). Its conditional distribution given $(w_j, \mathrm{BC}_j) = (w, x)$ is therefore invariant in (w, x). In particular, f_j is conditionally independent of S_j^{C} because S_j^{C} is a measurable function of $(R_{j,0}, e_j)$ plus a constant, while f_j is independent of both variables. The common-space construction permits pathwise comparisons across w and x and provides the domination needed for the endpoint limits as $w \uparrow 1$.

Risk-averse arbitrageurs absorb retail demand with finite capacity. The date-1 price displacement is $\Lambda_j D_j$, where $\Lambda_j = \Lambda(\mathrm{BC}_j)$ and $\Lambda' > 0$. The price returns to the fundamental at date

2, so the date-1 to date-2 abnormal return is $R_j = f_j - \Lambda_j D_j$. In the proofs below, we fix the borrowing-cost state x and abbreviate $\Lambda = \Lambda(x)$.

Proof of Proposition 3 *Part (a).* Let $v_i(t) = V_i(t)/\sum_k V_k(t)$ denote account i's visibility share. On the event that account $i^\star$ strongly monopolizes attention, every rival receives only finitely many additional followers while $F_{i^\star}(t) \to \infty$. Hence each rival follower base remains bounded and $F_j(t)/F_{i^\star}(t) \to 0$. Since $V_i(t) = \kappa\eta_i F_i(t)^\nu$, for every rival $V_j(t)/V_{i^\star}(t) = (\eta_j/\eta_{i^\star})[F_j(t)/F_{i^\star}(t)]^\nu \to 0$. With finitely many rivals, $v_{i^\star}(t) = [1 + \sum_{j\neq i^\star} V_j(t)/V_{i^\star}(t)]^{-1} \to 1$.

The visibility-weighted belief is $\bar\theta(t) = \sum_i v_i(t)p_i(t)$. For any winning account $i^\star$, $\bar\theta(t) - p_{i^\star}(t) = \sum_{j\neq i^\star} v_j(t)[p_j(t) - p_{i^\star}(t)]$. Because every posterior lies in $[0,1]$, $|\bar\theta(t) - p_{i^\star}(t)| \le \sum_{j\neq i^\star} v_j(t) = 1 - v_{i^\star}(t) \to 0$. If the winner is a charlatan, Lemma 1(a) gives $p_{i^\star}(t) \to 0$ in the counter-indexed benchmark, so $\bar\theta(t) \to 0$. If the winner is an expert, $p_{i^\star}(t) \to 1$, so $\bar\theta(t) \to 1$.

Let E_{E} and E_{C} denote the events that the unique winner is expert and charlatan, respectively. Since exactly one account wins, $\mathbb{1}_{E_{\mathrm{E}}} + \mathbb{1}_{E_{\mathrm{C}}} = 1$. The preceding pathwise limits imply $\bar\theta(t) \to \mathbb{1}_{E_{\mathrm{E}}}$ almost surely. Since $0 \le \bar\theta(t) \le 1$, dominated convergence gives, conditional on any realized type profile, $\lim_{t\to\infty} \mathbb{E}[\bar\theta(t) \mid \boldsymbol{\theta}] = \Pr(E_{\mathrm{E}} \mid \boldsymbol{\theta}) = 1 - \Pr(E_{\mathrm{C}} \mid \boldsymbol{\theta})$. If the profile contains at least one charlatan, Proposition 1(b) gives $\Pr(E_{\mathrm{C}} \mid \boldsymbol{\theta}) > 0$, so this limit is strictly below one. Integrating over the type vector gives $\lim_t \mathbb{E}\bar\theta(t) = \Pr(E_{\mathrm{E}}) < 1$ whenever charlatan profiles occur with positive probability.

For the slant block model, fix economy m. The later posterior $p^L_{i,m}$ is fixed during the later urn. If $i^\star_m$ is its strong monopolist, the same visibility-share argument gives $|\bar\theta_m(t) - p^L_{i^\star_m,m}| \le 1 - v_{i^\star_m,m}(t) \to 0$ almost surely. The difference is bounded by one, so convergence also holds in L^1. Therefore $\lim_{t\to\infty} \mathbb{E}_m\bar\theta_m(t) = \mathbb{E}_m[p^L_{i^\star_m,m}]$.

To compare this expectation with the type of the winning account, write $\mathbb{1}\{\theta_{i^\star_m} = \mathrm{E}\} = \sum_i \mathbb{1}\{i^\star_m = i\}\mathbb{1}\{\theta_i = \mathrm{E}\}$. Then $|p^L_{i^\star_m,m} - \mathbb{1}\{\theta_{i^\star_m} = \mathrm{E}\}| \le \sum_i \mathbb{1}\{i^\star_m = i\}|p^L_{i,m} - \mathbb{1}\{\theta_i = \mathrm{E}\}|$. Taking expectations and dropping the indicators gives $|\mathbb{E}_m[p^L_{i^\star_m,m}] - \Pr_m(\theta_{i^\star_m} = \mathrm{E})| \le \sum_i \mathbb{E}_m|p^L_{i,m} - \mathbb{1}\{\theta_i = \mathrm{E}\}|$. Step 3 in the proof of Proposition 2 gives convergence in probability of each bounded posterior to its type indicator. Boundedness converts this to L^1 convergence. Since N is finite, the sum on the right converges to zero. Proposition 2(c) gives $\Pr_m(\theta_{i^\star_m} = \mathrm{C}) \to \pi^E_{\mathrm{C}}$. Hence $\lim_{m\to\infty}\lim_{t\to\infty} \mathbb{E}_m\bar\theta_m(t) = 1 - \pi^E_{\mathrm{C}} < P_{\mathrm{E}}$.

Part (b). Equation (15) implies $S^V_j - S^{\mathrm{E}}_j = w_j(S^{\mathrm{C}}_j - S^{\mathrm{E}}_j)$. Conditional on (w_j, BC_j), the maintained primitive means give $\mathbb{E}S^{\mathrm{E}}_j = 0$ and $\mathbb{E}S^{\mathrm{C}}_j = \mu_{\mathrm{C}}(\bar b)$. Therefore $\mathbb{E}[S^V_j - S^{\mathrm{E}}_j \mid w_j, \mathrm{BC}_j] = w_j\mu_{\mathrm{C}}(\bar b)$. The right-hand side does not depend on the borrowing-cost state. Applying the law of iterated expectations over BC_j conditional on w_j yields $\mathbb{E}[S^V_j - S^{\mathrm{E}}_j \mid w_j] = w_j\mu_{\mathrm{C}}(\bar b)$. It is zero at $w_j = 0$, strictly positive for $w_j > 0$, strictly increasing in w_j because $\mu_{\mathrm{C}}(\bar b) > 0$, and, for $w_j > 0$, strictly increasing in $\mu_{\mathrm{C}}(\bar b)$. ■

Proof of Proposition 4 Fix a borrowing-cost state x, write $\Lambda = \Lambda(x)$, and condition throughout on $(w_j, \mathrm{BC}_j) = (w, x)$. Suppress the stock subscript. Under the pricing assumption, $S^{\mathrm{E}} = f + u$ has mean zero and variance $A = \tau_f^2 + \tau_u^2$, while $S^{\mathrm{C}} = \mu_{\mathrm{C}}(\bar{b}) + \zeta R_0 + e$ has mean $\mu_{\mathrm{C}}(\bar{b})$ and variance $B = \zeta^2 \tau_R^2 + \tau_e^2$. The two signals are independent because (f, u) is independent of (R_0, e). Their covariance is therefore zero. Visible sentiment satisfies $S^V = (1-w)S^{\mathrm{E}} + wS^{\mathrm{C}}$, so $\mathbb{E}S^V = w\mu_{\mathrm{C}}(\bar{b})$ and $\Sigma_V(w) = \mathrm{Var}(S^V) = A(1-w)^2 + Bw^2$. Also $\mathrm{Cov}(f, S^{\mathrm{E}}) = \tau_f^2$, $\mathrm{Cov}(f, S^{\mathrm{C}}) = 0$, and $\mathrm{Cov}(f, S^V) = (1-w)\tau_f^2$.

Part (a). Joint and univariate return projections. In parts (a) and (b), $\delta^- = \delta$ and $m = 0$, so $D = \delta S^V$ and $R = f - \Lambda\delta S^V$. All population projections include an intercept, so coefficients are computed from centered covariances.

Because S^{E} and S^{C} are uncorrelated, the covariance matrix of the two regressors is diagonal. The coefficient on S^{C} in the joint projection is therefore $\beta_{\mathrm{C}} = \mathrm{Cov}(R, S^{\mathrm{C}})/\mathrm{Var}(S^{\mathrm{C}})$. Now $\mathrm{Cov}(R, S^{\mathrm{C}}) = \mathrm{Cov}(f, S^{\mathrm{C}}) - \Lambda\delta\,\mathrm{Cov}(S^V, S^{\mathrm{C}})$. The first term is zero. Since $\mathrm{Cov}(S^V, S^{\mathrm{C}}) = (1-w)\,\mathrm{Cov}(S^{\mathrm{E}}, S^{\mathrm{C}}) + w\,\mathrm{Var}(S^{\mathrm{C}}) = wB$, we obtain $\beta_{\mathrm{C}} = -\Lambda\delta w$. It is nonpositive and is strictly negative when $w > 0$.

Similarly, $\beta_{\mathrm{E}} = \mathrm{Cov}(R, S^{\mathrm{E}})/\mathrm{Var}(S^{\mathrm{E}})$. Using $\mathrm{Cov}(f, S^{\mathrm{E}}) = \tau_f^2$ and $\mathrm{Cov}(S^V, S^{\mathrm{E}}) = (1-w)A$ gives $\mathrm{Cov}(R, S^{\mathrm{E}}) = \tau_f^2 - \Lambda\delta(1-w)A$. Dividing by A gives $\beta_{\mathrm{E}} = \tau_f^2/(\tau_f^2 + \tau_u^2) - \Lambda\delta(1-w)$. Differentiating with respect to x gives $\partial_x\beta_{\mathrm{C}} = -\Lambda'(x)\delta w \le 0$ and $\partial_x\beta_{\mathrm{E}} = -\Lambda'(x)\delta(1-w) \le 0$.

The univariate coefficient on visible sentiment is $\beta_V(w, x) = \mathrm{Cov}(R, S^V)/\Sigma_V(w)$. Since $\mathrm{Cov}(R, S^V) = (1-w)\tau_f^2 - \Lambda\delta\Sigma_V(w)$, we have $\beta_V(w, x) = (1-w)\tau_f^2/\Sigma_V(w) - \Lambda(x)\delta$. Define $G(w, x) = (1-w)\tau_f^2 - \Lambda(x)\delta\Sigma_V(w)$. Because $\Sigma_V(w) > 0$ for every $w \in [0, 1]$, β_V and G have the same sign.

Suppose $\Lambda(x)\delta(\tau_f^2 + \tau_u^2) < \tau_f^2$. Then $G(0, x) = \tau_f^2 - \Lambda(x)\delta A > 0$, while $G(1, x) = -\Lambda(x)\delta B < 0$. Thus continuity gives at least one root in $(0, 1)$. To prove uniqueness, differentiate twice. Since $\Sigma_V(w) = A(1-w)^2 + Bw^2$, $\Sigma_V''(w) = 2(A+B)$. Therefore $\partial_{ww}G(w, x) = -2\Lambda(x)\delta(A+B) < 0$. Hence G is strictly concave. If $\partial_w G(0, x) \le 0$, strict concavity implies that $\partial_w G(w, x) < 0$ for every $w > 0$, so G is strictly decreasing and can cross zero only once. If $\partial_w G(0, x) > 0$, strict concavity implies that the derivative decreases strictly and crosses zero at most once. Thus G first increases to a unique maximum and then decreases strictly. Since $G(0, x) > 0$ and $G(1, x) < 0$, the zero must occur on the decreasing branch and is unique. Denote it by $\bar{w}(x)$. The equation $G(\bar{w}(x), x) = 0$ is exactly equation (18).

At the root, $\partial_w G(\bar{w}(x), x) < 0$. Indeed, if it were zero, strict concavity would make the root the unique maximizer, so the maximum value of G would be zero, contradicting $G(0, x) > 0$. Also $\partial_x G(w, x) = -\Lambda'(x)\delta\Sigma_V(w) < 0$. The implicit-function theorem therefore applies and gives $\bar{w}'(x) = -\partial_x G(\bar{w}(x), x)/\partial_w G(\bar{w}(x), x) < 0$. This establishes the threshold comparative static.

Finally, the expert coefficient remains positive at and above the threshold. At $w = \bar{w}$, $\Lambda\delta\Sigma_V(\bar{w}) = (1-\bar{w})\tau_f^2$. Multiplying both sides by $(1-\bar{w})/\Sigma_V(\bar{w})$ gives $\Lambda\delta(1-\bar{w}) = \tau_f^2(1-\bar{w})^2/\Sigma_V(\bar{w})$. Since $\bar{w} > 0$ and $B > 0$, $\Sigma_V(\bar{w}) = A(1-\bar{w})^2 + B\bar{w}^2 > A(1-\bar{w})^2$. Therefore $\Lambda\delta(1-\bar{w}) < \tau_f^2/A$, so $\beta_{\mathrm{E}}(\bar{w}) > 0$. Because $\beta_{\mathrm{E}}(w) = \tau_f^2/A - \Lambda\delta(1-w)$ is strictly increasing in w, it remains positive for every $w \ge \bar{w}$.

Part (b). Conditional return moments. The conditional mean is $\mathbb{E}R = \mathbb{E}f - \Lambda\delta\mathbb{E}S^V = -\Lambda\delta w\mu_{\mathrm{C}}(\bar{b})$. Its derivative with respect to visibility is $-\Lambda\delta\mu_{\mathrm{C}}(\bar{b}) < 0$. Its derivative with respect to the borrowing-cost state is $-\Lambda'(x)\delta w\mu_{\mathrm{C}}(\bar{b}) \leq 0$, with strict inequality when $w > 0$.

For the variance, $\mathrm{Var}(R) = \mathrm{Var}(f) + (\Lambda\delta)^2\,\mathrm{Var}(S^V) - 2\Lambda\delta\,\mathrm{Cov}(f, S^V)$. Substituting the primitive moments gives $\mathrm{Var}(R) = \tau_f^2 + (\Lambda\delta)^2\Sigma_V(w) - 2\Lambda\delta(1-w)\tau_f^2$. Differentiating with respect to w gives $\partial_w\,\mathrm{Var}(R) = (\Lambda\delta)^2\Sigma_V'(w) + 2\Lambda\delta\tau_f^2$. Now $\Sigma_V'(w) = -2A(1-w) + 2Bw \geq -2A$. Hence $\partial_w\,\mathrm{Var}(R) \geq 2\Lambda\delta[\tau_f^2 - \Lambda\delta A]$. Under the maintained condition $\Lambda\delta A < \tau_f^2$, the right-hand side is strictly positive, so conditional return variance is strictly increasing in w.

Part (c). Bullish and bearish charlatan signals. Now impose $\delta^- = 0$ and $m = 0$. Then $D = \delta\max\{S^V, 0\}$ and $R = f - \Lambda\delta\max\{S^V, 0\}$. Let $A_+(w) = \mathbb{E}[\max\{S^V(w), 0\} \mid S^{\mathrm{C}} \geq s]$ and $A_-(w) = \mathbb{E}[\max\{S^V(w), 0\} \mid S^{\mathrm{C}} \leq -s]$. Because f is independent of S^{C} and has zero mean, $\mathbb{E}[f \mid S^{\mathrm{C}} \geq s] = \mathbb{E}[f \mid S^{\mathrm{C}} \leq -s] = 0$. Therefore $m_+(w,x) = -\Lambda(x)\delta A_+(w)$ and $m_-(w,x) = -\Lambda(x)\delta A_-(w)$. By construction $A_-(w) \geq 0$, so $m_-(w,x) \leq 0$. Also $A_+(w) > 0$. If $w = 1$, then on the event $S^{\mathrm{C}} \geq s > 0$, $\max\{S^V, 0\} = S^{\mathrm{C}} \geq s$. If $0 \leq w < 1$, then conditional on any realization of S^{C}, the term $(1-w)S^{\mathrm{E}}$ contains the independent Gaussian fundamental component $(1-w)f$ with positive variance. Hence S^V has positive probability of being positive, so its positive part has strictly positive conditional expectation. Thus $m_+(w,x) < 0$.

To compare magnitudes, define $X = S^{\mathrm{E}}$, $Y = S^{\mathrm{C}}$, and $H_w(y) = \mathbb{E}[\max\{(1-w)X + wy, 0\}]$, where the expectation is over X. Independence gives $A_+(w) = \mathbb{E}[H_w(Y) \mid Y \geq s]$ and $A_-(w) = \mathbb{E}[H_w(Y) \mid Y \leq -s]$. For $w > 0$, H_w is strictly increasing. If $0 < w < 1$ and $y_2 > y_1$, then the integrand at y_2 is everywhere weakly larger than at y_1. On the event $\{(1-w)X + wy_1 \geq 0\}$, the difference equals $w(y_2 - y_1) > 0$. This event has positive probability because X has full support. Hence $H_w(y_2) > H_w(y_1)$. At $w = 1$, $H_1(y) = \max\{y, 0\}$, so the same strict ordering holds between s and $-s$.

On the bullish event $Y \geq s$, monotonicity gives $H_w(Y) \geq H_w(s)$. On the bearish event $Y \leq -s$, it gives $H_w(Y) \leq H_w(-s)$. Therefore $A_+(w) \geq H_w(s) > H_w(-s) \geq A_-(w)$ for every $w > 0$. Since $|m_\pm| = \Lambda\delta A_\pm$, we obtain $|m_+(w,x)| > |m_-(w,x)|$. At $w = 0$, $H_0(y) = \mathbb{E}[\max\{X, 0\}]$ does not depend on y, so $A_+(0) = A_-(0)$ and the two conditional return magnitudes are equal.

It remains to prove the endpoint limits. Use the common-space construction in the pricing assumption. For fixed primitive realization, $S^V(w) = (1-w)S^{\mathrm{E}} + wS^{\mathrm{C}} \to S^{\mathrm{C}}$ almost surely as $w \uparrow 1$. Moreover, $\max\{S^V(w), 0\} \leq |S^V(w)| \leq |S^{\mathrm{E}}| + |S^{\mathrm{C}}|$. The dominating variable has finite expectation. Because the conditioning events $\{S^{\mathrm{C}} \geq s\}$ and $\{S^{\mathrm{C}} \leq -s\}$ have positive probability and do not depend on w, dominated convergence applies to the corresponding conditional expectations.

On $\{S^{\mathrm{C}} \leq -s\}$, $\max\{S^V(w), 0\} \to 0$, so $A_-(w) \to 0$ and hence $m_-(w,x) \to 0$. On $\{S^{\mathrm{C}} \geq s\}$, $\max\{S^V(w), 0\} \to S^{\mathrm{C}}$, so $A_+(w) \to \mathbb{E}[S^{\mathrm{C}} \mid S^{\mathrm{C}} \geq s, w = 1, \mathrm{BC} = x]$. Therefore $m_+(w,x) \to -\Lambda(x)\delta\mathbb{E}[S^{\mathrm{C}} \mid S^{\mathrm{C}} \geq s, w = 1, \mathrm{BC} = x] < 0$. The fundamental term contributes zero to both conditional means because it is independent of S^{C} and centered.

Part (d). Retail-flow covariance and exclusion. Return to the linear-demand region, where $D =$

$\delta S^V + mR_0$. Using $S^V = (1-w)S^{\mathrm{E}} + wS^{\mathrm{C}}$, independence of the two signals, and $\mathrm{Cov}(R_0, S^{\mathrm{C}}) = \zeta\tau_R^2$, we obtain $\mathrm{Cov}(D, S^{\mathrm{C}}) = \delta w \,\mathrm{Var}(S^{\mathrm{C}}) + m\zeta\tau_R^2 \geq 0$. The first term is strictly positive whenever $w > 0$ because $\mathrm{Var}(S^{\mathrm{C}}) = B > 0$. The second is strictly positive whenever $m > 0$, $\zeta > 0$, and $\tau_R^2 > 0$.

Conditional on R_0, the centered charlatan signal is e. The centered visible signal is $(1-w)(f+u)+we$, and the direct term mR_0 becomes a constant. Hence $\mathrm{Cov}(D, S^{\mathrm{C}} \mid R_0) = \delta w \,\mathrm{Var}(e) = \delta w \tau_e^2$.

Finally, the exclusion coefficient follows from the identity $D = \delta S^V + mR_0$. If the three regressors $(R_0, S^V, S^{\mathrm{C}})$ have full rank after an intercept is included, the unique population projection of D on these regressors has coefficients $(m, \delta, 0)$. It remains only to verify full rank under the stated conditions. Center the three regressors and suppose $aR_0 + b[S^V - \mathbb{E}S^V] + c[S^{\mathrm{C}} - \mathbb{E}S^{\mathrm{C}}] = 0$ almost surely. Since $S^V - \mathbb{E}S^V = (1-w)(f+u) + w(\zeta R_0 + e)$ and $S^{\mathrm{C}} - \mathbb{E}S^{\mathrm{C}} = \zeta R_0 + e$, the coefficient on the independent nondegenerate variable $f+u$ is $b(1-w)$. When $w < 1$ and $\mathrm{Var}(S^{\mathrm{E}}) > 0$, this forces $b = 0$. The remaining relation is $(a + c\zeta)R_0 + ce = 0$ almost surely. Since e is independent of R_0 and has positive variance, $c = 0$. Since $\tau_R^2 > 0$, this then forces $a = 0$. The regressors are therefore linearly independent, so the population coefficient vector is unique and the coefficient on S^{C} is exactly zero. At $w = 1$, $S^V = S^{\mathrm{C}}$ up to the same mean, so the separate coefficient is not identified. If $\tau_e^2 = 0$, the charlatan signal is affine in R_0, and if $\tau_R^2 = 0$, the lagged-return regressor is degenerate. These are exactly the singular cases excluded in the statement. ■

The next lemma records the sign of the univariate visible-sentiment coefficient outside the predictive-at-zero region used in Proposition 4(a).

Lemma 6 (Sign of the visible-sentiment coefficient at high price impact) *Fix a borrowing-cost state x and write $T = \tau_f^2$, $A = \tau_f^2 + \tau_u^2$, $B = \zeta^2\tau_R^2 + \tau_e^2$, and $k = \Lambda(x)\delta$. The univariate visible-sentiment coefficient has the sign of $G(w, x) = T - kA + (2kA - T)w - k(A+B)w^2$. Define $k_c = \frac{T}{2A}(1 + \sqrt{1 + A/B})$. The following cases exhaust the parameter space.*

(a) If $k < T/A$, G has one root in $(0,1)$ and is positive below it and negative above it.

(b) If $k = T/A$, $w = 0$ is a root and a second root lies in $(0,1)$.

(c) If $T/A < k < k_c$, there are two interior roots $0 < w_L < w_H < 1$, and G is negative outside (w_L, w_H) and positive inside it.

(d) If $k = k_c$, $G \leq 0$ with equality only at its interior maximizer.

(e) If $k > k_c$, $G < 0$ on $[0,1]$.

Whenever the upper root exists, $w_H = [2kA - T + \sqrt{T^2 + 4kBT - 4k^2AB}]/[2k(A+B)]$ and is strictly decreasing in x. Whenever two interior roots exist, the lower root is strictly increasing in x.

Proof of Lemma 6 From the proof of Proposition 4(a), the sign of the univariate coefficient is the sign of $G(w,x) = (1-w)T - k[A(1-w)^2 + Bw^2]$. Expanding gives $G(w,x) = T - kA + (2kA - T)w - k(A+B)w^2$. This is a strictly concave quadratic in w because the coefficient on w^2 is $-k(A+B) < 0$.

Its discriminant as a quadratic in w is $\mathcal{D}(k) = (2kA-T)^2 + 4k(A+B)(T-kA) = T^2 + 4kBT - 4k^2AB$. Viewed as a quadratic in k, $\mathcal{D}(k)$ has leading coefficient $-4AB < 0$. Solving $\mathcal{D}(k) = 0$ gives $k = T[\sqrt{B} \pm \sqrt{A+B}]/(2A\sqrt{B})$. The root with the minus sign is negative, while the positive root is $k_c = T[1 + \sqrt{1 + A/B}]/(2A)$. Therefore $\mathcal{D}(k) > 0$ for $0 < k < k_c$, $\mathcal{D}(k) = 0$ at $k = k_c$, and $\mathcal{D}(k) < 0$ for $k > k_c$.

For case (a), $k < T/A$ implies $G(0) = T - kA > 0$, while $G(1) = -kB < 0$. Continuity gives an interior root. The product of the two algebraic roots is $(T - kA)/[-k(A+B)] < 0$, so one root is positive and one is negative. Hence the interior root is unique. Strict concavity and the endpoint signs imply that G is positive below the interior root and negative above it.

For case (b), $k = T/A$ gives $G(0) = 0$. The derivative at zero is $G'(0) = 2kA - T = T > 0$, so $G(w) > 0$ for all sufficiently small $w > 0$. Since $G(1) = -kB < 0$, continuity gives another root in $(0,1)$. The quadratic has at most two roots, so these are the only roots. Strict concavity gives the stated sign pattern.

For case (c), $k > T/A$ gives $G(0) < 0$, and still $G(1) < 0$. Since $k < k_c$, the discriminant is positive and the quadratic has two distinct real roots. Its maximizer is $w^* = (2kA - T)/[2k(A+B)]$. The numerator is positive because $k > T/(2A)$, and the denominator exceeds the numerator because $2k(A+B) - (2kA - T) = 2kB + T > 0$. Thus $w^* \in (0,1)$. Substituting the maximizer into the quadratic gives $G(w^*) = \mathcal{D}(k)/[4k(A+B)] > 0$. Since the endpoints are negative and the interior maximum is positive, continuity implies one root in $(0, w^*)$ and one in $(w^*, 1)$. Strict concavity implies there can be no others. The sign is positive between the roots and negative outside them.

For case (d), note that $k_c > T/A$ because $\sqrt{1 + A/B} > 1$. Thus the maximizer is interior. At $k = k_c$, the discriminant is zero, so the two roots coincide at the maximizer. A strictly concave quadratic with zero maximum is nonpositive everywhere and equals zero only at that point.

For case (e), $k > k_c$ gives a negative discriminant. Since the leading coefficient is negative, the quadratic has no real root and is negative everywhere. In particular it is negative on $[0,1]$.

The quadratic formula gives the two roots, when real, as $[2kA - T \pm \sqrt{\mathcal{D}(k)}]/[2k(A+B)]$. The upper root is therefore the expression stated in the lemma. At this root, $\partial_w G(w_H, x) = (2kA - T) - 2k(A+B)w_H = -\sqrt{\mathcal{D}(k)} < 0$. At the lower root, $\partial_w G(w_L, x) = \sqrt{\mathcal{D}(k)} > 0$. Because $k = \Lambda(x)\delta$ and $\Lambda'(x) > 0$, $\partial_x G(w,x) = -\Lambda'(x)\delta[A(1-w)^2 + Bw^2] < 0$. The implicit-function theorem gives $w_H'(x) = -\partial_x G/\partial_w G < 0$ and, when the lower root exists, $w_L'(x) = -\partial_x G/\partial_w G > 0$. Thus higher borrowing costs shrink the region of positive visible-sentiment predictability from both sides. ■

Differentiating the conditional return variance in Proposition 4(b) with respect to the borrowing-cost state gives $\partial_x \operatorname{Var}(R \mid w, x) = 2\Lambda'(x)\delta[\Lambda(x)\delta\Sigma_V(w) - (1-w)\tau_f^2] = -2\Lambda'(x)\delta G(w,x)$. Thus

borrowing costs raise conditional return variance exactly where visible sentiment is contrarian, lower it where visible sentiment remains positively informative, and have no first-order effect at a sign-change threshold.

Proof of Corollary 1 Under the conditions of Proposition 4(c), $m_+(w,x) = -\Lambda(x)\delta A_+(w)$ and $m_-(w,x) = -\Lambda(x)\delta A_-(w)$. The invariant conditional law makes $A_+(w)$ and $A_-(w)$ independent of the borrowing-cost state x. Also $A_+(w) > 0$ and $A_-(w) \geq 0$. Therefore $\partial_x m_+(w,x) = -\Lambda'(x)\delta A_+(w) < 0$, while $\partial_x m_-(w,x) = -\Lambda'(x)\delta A_-(w) \leq 0$. Since both conditional means are nonpositive, $|m_+| = \Lambda\delta A_+$ and $|m_-| = \Lambda\delta A_-$. For every $w > 0$, Proposition 4(c) gives $A_+(w) > A_-(w)$. Hence $\partial_x[|m_+(w,x)| - |m_-(w,x)|] = \Lambda'(x)\delta[A_+(w) - A_-(w)] > 0$. At $w = 0$, $A_+(0) = A_-(0)$, so the magnitude gap is zero. The endpoint limits as $w \uparrow 1$ follow directly from the endpoint limits proved in Proposition 4(c). ■

B Measurement and Classification Details

This appendix collects the descriptive and validation evidence that supports the measurement inputs used in the model and real-time tests. The main text uses these estimates to discipline learnability, construct posterior coverage weights, and form out-of-sample portfolios.

B.1 Sample construction and classification robustness

Appendix Table D.6 traces each sample from the raw data. The Bloomberg feed contains 112.5 million records, including 18,348,209 platform messages used for text classification. Matching tweets to CRSP returns at the user-stock-day level yields 6,649,510 observations, of which 2,478,688 have non-neutral sentiment and enter the alpha estimation. We collect 139,401 platform accounts through the API, match 105,535 to Bloomberg usernames, and obtain 52,607 distinct accounts after pooling users with fewer than 20 lifetime posts. Among them, 29,391 have at least two return-matched directional observations and a measured alpha, 29,387 enter the mixture estimation after dropping four users with degenerate standard errors, 22,017 have follower counts, and 22,931 have complete tweeting-strategy covariates.

We assess robustness to both posting activity and track-record length. The baseline filter requires at least 20 tweets and 5 followers, leaving 12,337 users. A stricter filter requires at least 50 tweets and 10 followers and excludes users averaging 100 or more tweets per day, leaving 7,286 users. A third filter retains the top activity decile, leaving 2,938 users. The estimated skill distribution is stable across these samples, as shown in Section B.3. Requiring at least 3, 5, 10, 20, or 50 scored recommendations produces charlatan shares of 40.5%, 44.0%, 47.7%, 53.3%, and 52.8%, respectively, compared with 41.7% in the baseline. Adverse composition therefore does not arise from short histories and becomes stronger among users with longer records. The extreme charlatan mean moderates from -10.0 to -3.1, nevertheless, showing that short histories affect the tails more than the estimated composition.

Flooring degenerate standard errors yields the same ranking. At the 10th-percentile floor, the charlatan and expert shares are 33.9% and 22.5%, and at the 25th-percentile floor they are 22.5% and 13.2%. Allowing Student-t measurement error with 3 to 30 degrees of freedom produces charlatan shares between 42.0% and 54.5%. Finally, a month-block bootstrap of the full classification procedure gives a standard deviation of 2.3 percentage points for the charlatan share. Its mean is 49.8% because duplicated months reduce measured standard errors, so we interpret the dispersion as a conservative measure of classification uncertainty rather than as a centered confidence interval.

B.2 Independent text validation

In addition to the social sentiment scores assigned by Bloomberg's machine-learning algorithm, our data contain the headline text of each tweet. Bloomberg's social sentiment scores are based on a

Table B.1: Bloomberg Sentiment Validation

Panel A: Confusion matrix (N tweets)				
	Own text-based sentiment			
Bloomberg sentiment	Negative	Neutral	Positive	Total
Negative	630,250	716,470	173,337	1,520,057
Neutral	980,171	8,593,079	2,251,059	11,824,309
Positive	48,785	832,545	1,296,747	2,178,077
Total	1,659,206	10,142,094	3,721,143	15,522,443

Panel B: Validation statistics	
Correlation (Bloomberg vs. own)	0.374
Overall agreement rate	67.8%
Agreement | Bloomberg positive	59.5%
Agreement | Bloomberg negative	41.5%
Agreement | Bloomberg neutral	72.7%

Notes: This table validates Bloomberg's social sentiment scores against an independent text-based sentiment classifier applied to the tweet headline text. Panel A reports the confusion matrix (rows show Bloomberg sentiment, columns show own text-based sentiment). Panel B reports summary statistics. The text-based classifier uses keyword rules to identify positive, negative, and neutral sentiment from the raw tweet text. The sample includes 15.5 million tweets with non-missing Bloomberg sentiment.

proprietary supervised learning algorithm (Bloomberg, 2016). To validate the scores, we construct an independent text-based sentiment classifier using the headline field. Our rule-based classifier identifies positive sentiment from words such as buy, long, bullish, breakout, undervalued, and negative sentiment from words such as sell, short, bearish, crash, overvalued. Tweets with no sentiment words or equal positive and negative counts are classified as neutral.

Table B.1 reports the validation results. The correlation between Bloomberg sentiment and our text-based measure is 0.37, which is economically meaningful for discrete $\{-1, 0, 1\}$ variables. The overall agreement rate is 67.8%. Conditional on Bloomberg classifying a tweet as positive, our classifier agrees 59.5% of the time. Conditional on Bloomberg classifying a tweet as negative, agreement is 41.5%. Disagreements are concentrated between adjacent categories (neutral versus positive, or neutral versus negative) rather than between positive and negative. The two measures rarely assign opposite signs. The distribution of own-classified sentiment is somewhat more dispersed (24.0% positive versus Bloomberg's 14.0%), reflecting the broader keyword dictionary. The independent classifier is positively associated with Bloomberg sentiment and rarely assigns the opposite directional sign, providing useful external validation of the Bloomberg measure. It does not by itself establish that the results are independent of the proprietary algorithm; that would require re-estimating the pipeline under the independent classification, which we view as a valuable extension.

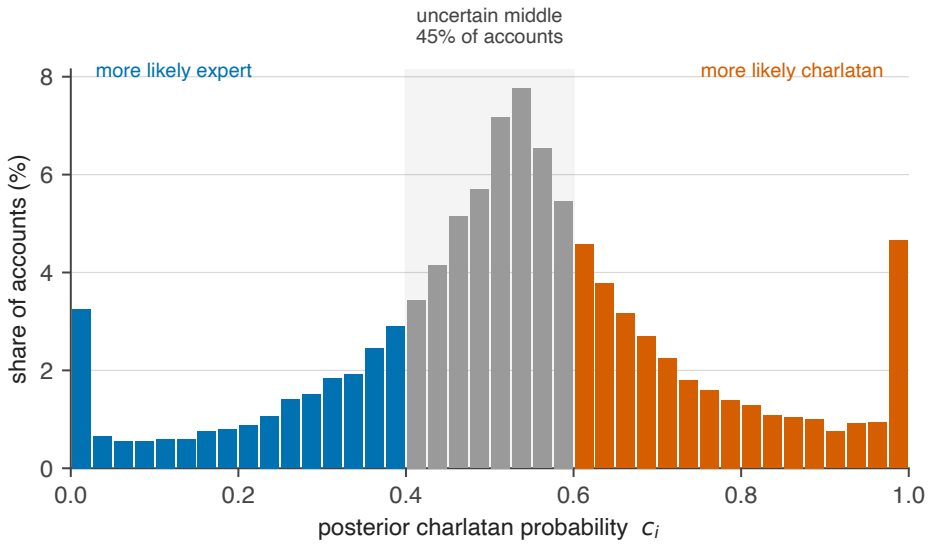


Figure B.1: Distribution of the Posterior Charlatan Probability

Notes: The figure shows the distribution of the posterior charlatan probability c_i across the 29,387 accounts in the mixture estimation, under the baseline two-sided sieve with $K = 3$ components per side. The shaded band marks posteriors between 0.4 and 0.6, the uncertain middle, which contains 45% of accounts under the $K = 3$ fit.

B.3 Advice-quality composition and directional-content validation

Table 2 assigns the largest population share to negative true alpha. This appendix reports the detailed posterior distribution, verifies that negative-alpha classifications reflect directional advice rather than social chatter, and examines return chasing and sample composition.

Under the previous specification retained here for comparability, the average posterior probabilities of the positive, zero, and negative categories are 29%, 29%, and 42%. Classifying users by their largest posterior probability produces shares of 25%, 16%, and 59%, respectively. Under either rule, negative-alpha accounts are the largest group. Posterior mean monthly alpha averages -0.48%, with a standard deviation of 2.22%. The bottom decile earns at most -2.01% per month, while the top decile earns at least 1.07%. Most users nevertheless lie much closer to zero. Figure B.1 shows the distribution of the posterior charlatan probability under the baseline two-sided $K = 3$ sieve: 45% of accounts have posteriors in the uncertain middle between 0.4 and 0.6 (35% under the $K = 2$ fit), so for a large share of the population the record does not reveal the sign of latent quality.

The composition is stable across the filters in Section 2.1. Re-estimating the model on the unfiltered sample (22,014 users; three accounts with follower counts lack activity information), the baseline filter ($\geq$20 tweets, $\geq$5 followers; 12,337 users), the strict filter ($\geq$50 tweets, $\geq$10 followers; 7,286 users), and the top activity decile (2,938 users) gives estimated charlatan shares of 41.7%, 41.6%, 41.3%, and 40.7%, expert shares of 29.0% to 29.6%, zero-mass shares of 29.3% to 29.7%, and posterior mean alphas of -0.48 to -0.39, so each share varies by at most one percentage point across the four filter samples. These estimates are computed under the previous specification with a

point mass at zero and are retained for comparability; recomputation under the two-sided baseline is queued. The result is therefore not driven by inactive accounts or bot-like posting, and the broader descriptive taxonomy of platform users to which these composition statistics belong is developed in Kakhbod, Kazempour, Livdan, and Schürhoff (2025).[1]

Negative advice value reflects advice rather than chatter. We classify all 18.3 million tweets as directional, commentary, or informational using keyword and regular-expression rules. Directional posts contain forward-looking buy or sell views, informational posts report news or data without taking a position, and the remaining posts are classified as commentary. A hand-labeled sample of 500 tweets is used to evaluate the procedure.

Table B.2 shows that 22.9% of tweets are directional, 66.0% are commentary, and 11.1% are informational. Directional content is nearly as common among expert users as among charlatan users, accounting for 20.5% and 24.1% of their tweets. Charlatan users are therefore not simply social posters whose commentary is mistakenly treated as advice.

The return evidence is sharper. Charlatan users earn −2.68% per month on directional posts, compared with 4.13% for expert users. The corresponding alphas also differ within commentary and informational posts, and the correlation between all-tweet and directional-tweet alpha is 0.64. Poor directional recommendations therefore account for an economically important part of measured negative advice value. Because the baseline classification uses all non-neutral posts, this exercise shows that the negative-alpha group contains genuine recommendations rather than only commentary. It is not an independent validation of the user classification. That validation comes from the out-of-sample tests in Section 5.1, which classify users on prior data and evaluate their later directional calls.

The classifier attains 57.6% accuracy in the hand-labeled sample. Its precision for directional tweets is 60.6%, while recall is only 37.7%, largely because implicit calls and position disclosures often contain no directional keyword. The procedure therefore understates directional content. Unless this measurement error differs systematically by skill, it attenuates differences across user types.

Commentary also contains information that is not captured by explicit directional language. Among commentary posts that our word-list procedure labels neutral, expert users earn 5.14% per month and charlatan users earn −3.08%. These estimates are close to the corresponding results for all commentary, 5.81% and -3.75%. The skill difference therefore extends beyond explicit recommendations.

[1]The distribution differs sharply from those estimated for professional analysts and curated platforms. Crane and Crotty (2020) classify 94% of professional analysts as expert, while Dim (2022) estimates an expert share of 56% among Seeking Alpha contributors. The weaker selection on an open platform is consistent with the negative return predictability of the platform's aggregate sentiment documented by Giannini, Irvine, and Shu (2018).

Table B.2: Tweet Content Classification

	Expert	Zero mass	Charlatan
Directional share of tweets (%)	20.5	23.6	24.1
Mean alpha (% per month)			
All tweets	+5.93	+0.41	−3.65
Directional only	+4.13	+0.40	−2.68
Commentary only	+5.81	+0.38	−3.75

Notes: We classify all 18,348,209 tweets as directional, commentary, or informational using keyword and regular-expression rules; 22.9% are directional, 66.0% commentary, and 11.1% informational. The table reports, by user type, the directional share of tweets and mean monthly alpha computed separately for each content category. Users are assigned to the category with the highest posterior probability under the previous three-type classification, retained here for comparability.

Return-chasing controls. Charlatan users may earn negative returns because they recommend recent winners before those stocks revert. We therefore estimate, separately for each user, $SocSent_{i,j,t}AbnRet_{j,t+1,t+20} = \tilde{\alpha}_i^{RC} + b_{1,i}Ret_{j,t-5,t} + b_{2,i}Ret_{j,t-20,t} + \epsilon_{i,j,t}$, and re-estimate the mixture using the return-chasing-adjusted intercepts.

Under the previous specification, the negative-alpha share falls from 41.7% to 32.2%, and the rank correlation between baseline and adjusted alphas is 0.61. Return chasing therefore explains part of measured negative advice value, consistent with Proposition 2(d), but the adjusted charlatan share remains larger than the expert share of 24.8%. Negative skill is not generated solely by extrapolation followed by short-run reversal.

B.4 Expanding-window classification stability

Classification stability. Table B.3 compares classifications at the midpoint estimation window with those from the final window for 8,577 users observed at both dates. The indices $t = 8$ and $t = 16$ count nested estimation vintages rather than calendar quarters. The diagonal dominates. Among users classified at the midpoint as expert, zero-mass, or charlatan under the previous specification, 65.8%, 86.6%, and 79.9% retain their classification. Direct reversals are rare, with 4.2% of charlatan users becoming expert and 8.2% of expert users becoming charlatan. Most movement is into the near-zero region as additional data shrink marginal alphas toward zero.

Rank correlations tell the same story. The midpoint and full-sample correlations are 0.849 for the expert posterior and 0.867 for posterior mean alpha. Table B.3 also shows correlations above 0.88 between adjacent expanding windows and 0.53 even between quarters 3 and 8. These results support the real-time classifications used in Section 5.1.

C Slant-Mechanism Validation

This appendix retains only the evidence the endogenous-slant mechanism needs: the association of content strategies with advice quality and with audience size. The broader descriptive landscape

Table B.3: Classification Stability Across Estimation Vintages

Retention and reversals, midpoint ($t = 8$) to full sample ($t = 16$), % of row group			
Midpoint classification	Expert	Zero mass	Charlatan
Retained in full sample	65.8	86.6	79.9
Direct reversal to opposite group	8.2		4.2
Spearman rank correlations across expanding-window pairs			
Period pair	$\rho(\Pr(\text{exp}))$	$\rho(\mathbb{E}[\alpha])$	% Same type
$t = 3 \to t = 8$	0.530	0.558	67.7
$t = 5 \to t = 10$	0.902	0.913	87.5
$t = 8 \to t = 12$	0.896	0.908	91.6
$t = 8 \to t = 16$	0.849	0.867	88.8
$t = 12 \to t = 16$	0.884	0.891	91.4

Notes: The table merges the nested-window transition evidence with the expanding-window stability statistics. The estimation windows are nested; t indexes the number of quarters of data used, the midpoint estimate uses the first 8 quarters (through approximately mid-2015), and the full sample uses all 16 quarters. In the top block, each user is classified by the type with the highest posterior probability; retention is the share of users whose midpoint classification is unchanged in the full sample, and direct reversals are moves between the expert and charlatan groups (the remainder moves into the near-zero region). The sample contains 8,577 users with skill estimates at both horizons. In the bottom block, Spearman rank correlations are computed for Pr(expert) and $\mathbb{E}[\alpha_i]$ across period pairs, and % Same type is the fraction of users whose binary classification, expert if Pr(expert) > Pr(charlatan) and charlatan otherwise, is unchanged; the pairs use the users available at both dates (3,202 for $t = 3 \to t = 8$, 8,336 for $t = 5 \to t = 10$, 8,577 for the $t = 8$ pairs, and 18,728 for $t = 12 \to t = 16$). All statistics are computed under the previous specification with a point mass at zero and are retained for comparability; recomputation under the two-sided baseline is queued.

of posting roles and audience gradients on this platform is documented in Kakhbod, Kazempour, Livdan, and Schürhoff (2025) and is not reproduced here.

Under Assumption 1, observable engagement need not reveal quality. The endogenous-slant mechanism of Proposition 2 gives a sharper prediction. Optimistic content can both attract followers and signal negative advice value, while winner chasing leaves a negative performance footprint. We estimate the marginal association between individual strategies and posterior quality using

$$Quality_i = \alpha + \beta\, TweetingActivity_i + \delta\, FractionPositive_i + \gamma^\top TweetingStrategy_i + \varepsilon_i, \tag{C2}$$

where $TweetingActivity_i$ is the logarithm of one plus the number of tweets, and $TweetingStrategy_i$ includes return chasing, contrarian tweeting, herding, and exposure to short-sale-constrained stocks.

Table C.4 shows that optimism is the clearest marker of negative advice value. Moving from entirely negative to entirely positive content raises the probability of being a charlatan by 7.6 percentage points and lowers the probability of being an expert by 5.2 percentage points. Return chasing is associated with negative advice value but not with a larger audience; decomposing it shows that the relation comes from bullish recommendations of recent winners, while bearish recommendations of recent losers contain no comparable quality signal. Contrarian tweeting instead predicts expert status, as does tweeting negatively about stocks with tighter short-sale constraints. Some observable behaviors therefore reveal quality, but the signals that attract attention need not

Table C.4: Content Strategies, Advice Quality, and Audience

	Advice quality (pp)		Audience
	Pr(expert)	Pr(charlatan)	$\log(1+F_i)$
$FractionPositive_i$	-5.16^{***}	7.63^{***}	0.48^{***}
	(0.73)	(0.76)	(0.07)
$ReturnChasing_i$	-2.16	4.31^{**}	-0.68^{***}
	(1.74)	(1.83)	(0.14)
$ContrarianTweet_i$	7.76^{***}	-1.24	-1.00^{***}
	(1.97)	(1.99)	(0.15)
$TweetingActivity_i$	-1.08^{***}	-0.52^{***}	0.76^{***}
	(0.12)	(0.13)	(0.01)
N	22,931	22,931	17,320

Notes: The table merges the strategy-quality and audience regressions and reports only the strategies used by the slant mechanism. The quality block reports regressions of posterior skill probabilities (in percentage points) on user activity and tweeting strategies, from the specification that also includes herding and short-sale-constraint exposure controls; the audience block reports the regression of the log of one plus the user's follower count as of February 2018 on the same regressors (equation (C3)). Activity is the logarithm of one plus the number of non-neutral tweets. Return chasing and contrarian tweeting are based on recommendations involving stocks in the extreme deciles of past-week returns. Standard errors in parentheses are robust to heteroskedasticity. *, **, and *** denote significance at the 10%, 5%, and 1% levels. The quality regressions retain the previous three-type posteriors as dependent variables and are retained for comparability; the two-sided rebuild is queued.

be the ones associated with quality. The regressions retain the previous three-type posteriors as dependent variables; the two-sided rebuild collapses these posteriors into the expert posterior and its complement.

We examine audience formation using

$$\log(1+F_i) = \alpha + \beta\, TweetingActivity_i + \delta\, FractionPositive_i + \gamma^{\top} TweetingStrategy_i + \varepsilon_i. \qquad \text{(C3)}$$

The audience column of Table C.4 reports the estimates. Activity is strongly associated with audience size, with an elasticity of 0.76 conditional on the other strategies. Optimism also attracts followers while predicting negative advice value. It is therefore the observable behavior that most closely matches the slant mechanism of Proposition 2. Winner chasing predicts poor performance but carries a negative follower coefficient, so it is better interpreted as a performance footprint of baited calls than as the source of their engagement advantage.

These regressions describe equilibrium associations rather than causal effects. Effort responds to audience through the first-order condition underlying equation (4), while audience responds to effort through exposure. The follower regression evaluates the model's prediction about which behaviors accompany a larger audience. Section 4.1 uses the reverse projection of activity on followers to discipline the amplification exponent ν.

Taken together, the evidence shows that attention rewards activity and optimism rather than demonstrated skill. Broad behavioral roles offer little information about quality, and the content choice that attracts followers most clearly is also the strongest observable marker of negative advice

value.

D Additional Figures and Tables

E Estimation of the Mixture Model

Within the expert and charlatan populations, we allow for K^+ and K^- one-sided sieve components, respectively, each with distinct mean quality. Let π_k^+ be the weight of expert-side component k and π_k^- the weight of charlatan-side component k; the components approximate the quality distributions F_{E} and F_{C} and are not economic types. The comparison specification of the earlier draft additionally carries a point mass π^0 at zero; it is retained below only so that both variants are covered by one likelihood. Latent quality α is then distributed according to

$$f(\alpha) = \mathbb{1}\{\alpha > 0\} \sum_{k=1}^{K^+} \pi_k^+ g(\alpha; \mu_k^+) + \pi^0 \, \mathbb{1}\{\alpha = 0\} + \mathbb{1}\{\alpha < 0\} \sum_{k=1}^{K^-} \pi_k^- g(\alpha; \mu_k^-), \tag{E4}$$

subject to

$$\begin{aligned} &\sum_{k=1}^{K^+} \pi_k^+ + \pi^0 + \sum_{k=1}^{K^-} \pi_k^- = 1, \\ &\mu_k^+ > 0 \quad \forall\, k \le K^+, \qquad \mu_k^- < 0 \quad \forall\, k \le K^-. \end{aligned} \tag{E5}$$

Here μ_k^+ and μ_k^- denote the expected abnormal returns of the k-th expert and charlatan subtype, respectively, and $g(\alpha; \mu)$ is the density for each component.

We use the maximum-entropy distribution consistent with the sign constraints. This is the exponential distribution for expert types, $g(\alpha; \mu) \equiv \frac{1}{\mu} \exp(-\frac{\alpha}{\mu})$, $\mu > 0$, and its sign-reversed counterpart for charlatan types ($\mu < 0$). Given that $\tilde{\alpha}_i = \alpha_i + \epsilon_i$, the distribution of measured alphas can be calculated as the convolution of f and a mean-zero Normal distribution with standard deviation $\tilde{\sigma}_i$, given by

$$\mathcal{G}(\tilde{\alpha}_i; \tilde{\sigma}_i, \Theta) = (f * \phi_{\tilde{\sigma}_i})(\tilde{\alpha}_i), \tag{E6}$$

where $*$ is the convolution operator, $\phi_{\tilde{\sigma}_i}$ denotes the Normal density with mean zero and standard deviation $\tilde{\sigma}_i$, and $\Theta = (\mu_1^+, \ldots, \mu_{K^+}^+, \mu_1^-, \ldots, \mu_{K^-}^-, \pi_1^+, \ldots, \pi_{K^+}^+, \pi_1^-, \ldots, \pi_{K^-}^-)$ is the vector of parameters. We maximize the likelihood function $\mathcal{L}(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_I; \tilde{\sigma}_1, \ldots, \tilde{\sigma}_I, \Theta) = \prod_{i=1}^{I} \mathcal{G}(\tilde{\alpha}_i; \tilde{\sigma}_i, \Theta)$ to estimate Θ.[2]

[2]Let X be an exponential variable with mean μ and Y be a mean-zero Normal variable with standard deviation σ. Their sum $Z = X + Y$ has the closed-form density $h(x; \mu, \sigma) = \frac{1}{2\mu} \exp\big(\frac{\sigma^2}{2\mu^2} - \frac{x}{\mu}\big) \times \big(1 - \mathrm{erf}\big(\frac{\sigma}{\sqrt{2}\mu} - \frac{x}{\sqrt{2}\sigma}\big)\big)$, where erf is the error function.

F Alternative Specifications for the Distribution of True Alphas

Table F.8 reports parameter estimates for alternative specifications, including the earlier variants that carry a point mass at zero; these are retained for comparison with the two-sided baseline of Table 2. Panel A varies the number of sieve components per side within the zero-mass family. The likelihood and both information criteria improve considerably from one to two components. At three components per side, the estimated zero mass collapses to zero, so the specification coincides with the two-sided baseline up to the one redundant parameter; this is the information-criterion evidence that the data do not require a point mass at zero once the one-sided distributions are estimated flexibly. Kolmogorov-Smirnov goodness-of-fit tests reject the one-component specification essentially always in simulation, while the richer specifications are rarely rejected, consistent with the information criteria. Panel B reports the results of fitting the zero-mass specification over different horizons H with two exponentials on the $\alpha > 0$ side, two exponentials on the $\alpha < 0$ side, and a mass at $\alpha = 0$. We calculate excess returns over the next 1, 2, 5, 10, and 20 trading days using the Fama-French five-factor model.

Table D.5: Full-Sample Versus Real-Time Classifications

	Full-sample	Real-time (embargoed)
Charlatan-exposure sort $C_{j,m}$, Q5−Q1, 20 days (31 months)		−0.90 (−4.23)
Out-of-sample charlatan-bullish portfolio, 20 days (monthly)		−0.588 (−2.91)

Notes: The table collects the paper's headline return coefficients under full-sample posteriors and under real-time posteriors, in which tweets in calendar quarter Q are weighted by embargoed posteriors estimated on data through quarter $Q-1$; each vintage uses only recommendations whose full 20-trading-day return window closed by its classification date. Exposure entries are Fama-MacBeth Q5−Q1 spreads in 20-day FF5-adjusted returns under the month-m formation, month-$m+1$ return design (Table 6). The out-of-sample portfolio return is from Table 8 and exists only in real-time form. t-statistics in parentheses.

Table D.6: Sample Construction

Sample	N	Used in
Panel A: Records		
Bloomberg feed, 7/13/2013–1/1/2017	112,543,629	
of which news stories	38,203,500	news-sentiment controls
of which X (Twitter) tweets	55,991,920	not used (no user identities)
of which platform messages	18,348,209	Tables B.2, B.1
User-stock-days matched to CRSP returns	6,649,510	Section 2.1; Table 9
of which directional (non-neutral)	2,478,688	alpha estimation (24)
Panel B: Users		
Platform accounts collected via API (2/2/2018)	139,401	
Matched to Bloomberg usernames	105,535	
Distinct accounts after pooling <20 lifetime posts	52,607	
of which with positive activity and follower counts	39,183	Regression (31); Figure 4
With measured alpha (≥2 directional obs.)	29,391	Table 1
In mixture estimation (nondegenerate s.e.)	29,387	Table 2; Appendix B.3
With follower counts	22,017	App. B.3 (filter composition)
With complete strategy covariates	22,931	Table C.4; audience block: 17,320
Baseline filter (≥20 tweets, ≥5 followers)	12,337	App. B.3 (filter composition)
Strict filter (≥50 tweets, ≥10 followers, <100/day)	7,286	App. B.3 (filter composition)
Top-decile activity	2,938	App. B.3 (filter composition)

Notes: This table traces every analysis sample from the raw data. Panel A reports record counts. Panel B reports user counts. The last column lists the tables and figures that use each sample.

Table D.7: Skill Consistency Across Return Horizons

Horizon pair (days)	1, 2	2, 5	5, 10	10, 20	1, 20
$\rho\big(\Pr(\text{expert})\big)$	0.668	0.613	0.672	0.657	0.271
$\rho\big(\mathbb{E}[\alpha_i \mid \tilde{\alpha}_i]\big)$	0.686	0.631	0.691	0.681	0.281

Notes: This table reports the consistency of skill classification across return horizons $H \in \{1, 2, 5, 10, 20\}$ trading days. Rows report Spearman rank correlations of Pr(user i expert) and of $\mathbb{E}[\alpha_i \mid \tilde{\alpha}_i]$ for adjacent horizon pairs and for the extreme 1-day versus 20-day pair. The sample includes 28,520 users with skill estimates at all five horizons.

Table F.8: Robustness: Alternative Specifications of the Mixture Model

Panel A: Model estimates for different number of sub-groups K

	(1) $K^+ = K^- = 1$		(2) $K^+ = K^- = 2$		(3) $K^+ = K^- = 3$	
	μ_k (%)	π_k (%)	μ_k (%)	π_k (%)	μ_k (%)	π_k (%)
Expert component 3					8.52	4.7
Expert component 2			8.31	5.0	1.58	22.5
Expert component 1	4.47	15.4	1.46	24.2	0.12	19.5
Zero mass	0.00	64.2	0.00	29.2	0.00	0.0
Charlatan component 1	-5.73	20.3	-1.33	33.6	-0.43	25.3
Charlatan component 2			-9.95	8.1	-1.96	21.3
Charlatan component 3					-10.75	6.8
N	29,387		29,387		29,387	
Log Likelihood	-88,209		-87,610		-87,589	
AIC	176,426		175,236		175,202	
BIC	176,459		175,302		175,302	

Panel B: Model estimates for different forecast horizon H

	(1) $H = 1$		(2) $H = 2$		(3) $H = 5$		(4) $H = 10$		(5) $H = 20$	
	μ_k	π_k	μ_k	π_k	μ_k	π_k	μ_k	π_k	μ_k	π_k
Expert component 2	4.16	1.3	3.77	1.8	4.87	2.6	7.10	3.0	8.31	5.0
Expert component 1	0.41	18.8	0.57	19.3	0.84	19.5	1.30	20.6	1.46	24.2
Zero mass	0.00	54.5	0.00	49.4	0.00	39.2	0.00	38.1	0.00	29.2
Charlatan component 1	-0.35	22.8	-0.49	26.3	-0.64	33.6	-0.94	30.9	-1.33	33.6
Charlatan component 2	-3.04	2.7	-3.72	3.1	-5.45	5.2	-6.79	7.3	-9.95	8.1
N	30,676		30,282		30,078		29,961		29,387	
Log Likelihood	-46,792		-56,092		-68,632		-79,041		-87,610	
AIC	93,600		112,201		137,281		158,097		175,236	
BIC	93,666		112,268		137,347		158,164		175,302	

Panel C: Overlap-robust standard-error inputs ($H = 20$, $K^+ = K^- = 2$)

	(1) Newey-West		(2) Non-overlapping		(3) Day-clustered	
	μ_k (%)	π_k (%)	μ_k (%)	π_k (%)	μ_k (%)	π_k (%)
Expert component 2	7.07	17.5	7.00	5.1	7.87	7.0
Expert component 1	1.81	27.6	1.55	40.3	1.56	26.7
Zero mass	0.00	2.6	0.00	0.0	0.00	21.3
Charlatan component 1	-1.95	32.0	-1.60	48.1	-1.46	34.8
Charlatan component 2	-8.47	20.3	-8.01	6.4	-9.49	10.3
N	29,391		23,008		29,081	

Notes: Panel A reports the results of fitting mixture models with one, two, and three components for expert and charlatan users. Means and probabilities are reported in percentage points. Panel B reports the results of fitting mixture models over different horizons H with two exponentials on $\alpha > 0$, two exponentials on $\alpha < 0$, and a mass at $\alpha = 0$. We calculate excess returns over the next 1, 2, 5, 10, and 20 trading days using the Fama-French five-factor model. The first number at the top of the columns shows the horizon of future returns. The measured alpha ($\tilde{\alpha}$) for each user is the average of signed adjusted returns after her tweets. For each horizon, the first column shows the mean of each component (μ's), and the second column shows the weight of the component in the mixture (π's). Means and probabilities are in percentage points. Panel C re-fits the baseline specification using standard-error inputs that are robust to the overlap of 20-day returns within a user. These inputs are Newey-West standard errors with $H - 1$ lags on each user's time-ordered signed-return series, alphas computed on non-overlapping 20-day windows, and standard errors clustered by calendar day within user. All estimates are computed from a single pipeline and log file.