North American Journal of Economics and Finance 79 (2025) 102460

Contents lists available at ScienceDirect

North American Journal of Economics and Finance

journal homepage: www.elsevier.com/locate/najef

Strategic information asymmetry in tail-risk markets

Omid M. Ardakani[1]

Professor of Economics, Parker College of Business, Georgia Southern University, United States of America

ARTICLE INFO

JEL classification:
D82
G14
C72

Keywords:
Strategic information asymmetry
Cognitive hierarchy
Bayesian games
Extreme value theory
Tender offers
Market design

ABSTRACT

This paper develops a novel information-theoretic measure of strategic asymmetry, *asymmetric information entropy*, that quantifies disparities in agents' knowledge states through differential Shannon entropy. I integrate k-level cognitive hierarchies with Bayesian games to analyze how strategic depth attenuates information gaps, proving almost sure convergence and Pareto-optimal limit equilibria. Using generalized extreme value distributions, I show strategic restructuring alters financial market outcomes through parameter shifts in tail risk and location that converge geometrically under Lipschitz belief updating. Empirical analysis of U.S. tender offers reveals legal defenses (Level-2 strategies) increase bid premiums versus the baseline, while combined strategies exhibit subadditive effects. The proposed entropy measure formalizes Akerlof-style market failures, providing a quantitative basis for securities regulation and mechanism design.

1. Introduction

Information asymmetry, a foundation of economic theory since Akerlof (1970)'s lemons problem, persistently distorts market efficiency by creating disparities in agents' strategic capabilities. While foundational work in game theory (Harsanyi, 1967; Nash, 1953) and behavioral economics (Camerer et al., 2004) has modeled strategic reasoning and incomplete information separately, their intersection remains underexplored. This paper bridges this gap by developing a unified framework that quantifies information asymmetry through entropy measures and formalizes its dissipation via strategic depth.

The paper is motivated by two empirical regularities: (1) strategic agents in financial markets (e.g., institutional investors) systematically outperform others despite public information availability (Hermalin & Weisbach, 2012), and (2) market crises exhibit phase transitions where informational edges abruptly dissolve (Kirilenko et al., 2017). I reconcile these phenomena through a novel measure, *asymmetric information entropy*, that captures differential knowledge states as divergences in Shannon entropy. Unlike mutual information metrics (Soofi & Retzer, 2002), this measure directly links to k-level cognitive hierarchies, enabling dynamic analysis of how strategic reasoning attenuates informational advantages.

The literature on strategic information processing splits into two streams. The first, rooted in game theory, analyzes equilibrium concepts under incomplete information. Seminal contributions include Aumann (1976)'s correlated equilibrium, Kreps and Wilson (1982)'s reputation models, and Aumann and Brandenburger (1995)'s epistemic conditions for Nash equilibrium. Contemporary extensions by Azar and Vives (2021) and Das et al. (2019) examine how firms adjust strategies under evolving information regimes but lack quantitative measures of asymmetry. The second stream, from information economics, quantifies uncertainty using entropy. Cover and Thomas (1991)'s entropy chain rule underpins modern analyses of market efficiency (Hansen & Lunde, 2006), while Jaynes (1957)'s maximum entropy principle informs strategic distribution modeling (Ardakani, 2022). Recent work by Li (2020) and Peres et al. (2020) connects information heterogeneity to asset pricing but treats strategic reasoning as exogenous.

E-mail address: oardakani@georgiasouthern.edu.

[1] The author has no conflicts of interest to declare. The author has no financial or non-financial interests to disclose.

https://doi.org/10.1016/j.najef.2025.102460
Received 13 November 2024; Received in revised form 27 April 2025; Accepted 6 May 2025
Available online 22 May 2025

This study synthesizes these traditions by extending cognitive hierarchy theory (Stahl & Wilson, 1995) to continuous-type Bayesian games, deriving asymmetric information entropy as the entropy difference, and showing it vanishes under infinite strategic depth.

This paper makes four principal contributions. First, I develop the entropy-based metric that quantifies information disparities as the difference in marginal entropies between agent partitions. This generalizes Akerlof-type models to continuous information structures. Second, I establish that this entropy measure converges to zero almost surely as strategic depth k approaches infinity under bounded payoffs. This result formalizes how deep strategic reasoning (empirically observed at $k \geq 5$) eliminates informational rents by equalizing agents' knowledge states through iterative belief updating. Third, applications to U.S. tender offers reveal that legal restructuring (Level-2 strategies) increases bid premiums by 4.9% compared to the baseline, while combined strategies exhibit subadditive effects, contradicting supermodularity assumptions (Milgrom & Roberts, 1982). Lastly, I show that estimators from informationally dominant σ-algebras achieve lower mean squared error, validated through COVID-era VaR forecasts.

The analysis yields three key findings. Theoretically, the asymmetric information entropy reduces strategic games to entropy comparisons across agent partitions. When filtrations satisfy standard information conditions, this measure converges to zero geometrically, rendering limit equilibria Pareto optimal. Empirically, in tender offers, asymmetric information entropy between institutional holdings and public data quantifies actionable information gaps. Strategic responses shift generalized extreme value distribution parameters—legal restructuring (Level-2) increases tail index, raising bid premiums nonlinearly. Methodologically, maximum entropy distributions under k-level thinking require perturbing moment constraints. Simulations show half-life convergence of strategic parameter estimates. These results extend three active research fronts. While Crawford and Iriberri (2007) model level-0 heuristics, I show cognitive hierarchies induce martingales in the asymmetric information entropy that converge to common knowledge—a dynamic counterpart to Aumann (1976)'s agreement theorem. In addition, a result generalizes (Gorton & Metrick, 2012)'s crisis narrative, showing entropy gaps predict squared pricing errors. The MLE convergence complements (Ardakani, 2023a)'s extreme event analysis by incorporating belief updating.

The paper proceeds as follows. Section 2 develops the theoretical framework, defining asymmetric information entropy and establishing its convergence under strategic depth. Section 3 applies the model to U.S. tender offers, quantifying how legal and financial restructuring (Levels 1–2) alters bid premiums via generalized extreme value distribution parameter shifts. Section 4 discusses policy implications, including entropy-based transparency rules, while Appendix contains technical details and proofs.

2. Framework

The literature has demonstrated how information asymmetry shapes strategic decision-making in economic models and its distorting effects on market outcomes (Akerlof, 1970). The integration of strategic thinking in economic frameworks began with foundational contributions by Nash (1953), who developed equilibrium concepts for strategic interactions, followed by Harsanyi (1967) on incomplete information, Aumann (1976) on correlated equilibrium, and Kreps and Wilson (1982), who incorporated reputation effects into games with incomplete information. Contemporary approaches in behavioral game theory, notably k-level thinking and cognitive hierarchies, examine how players adapt strategies through iterative reasoning about others' decisions. Ivashina (2009) and Peres et al. (2020) study the impacts of information asymmetry on asset pricing and decision-making, though explicit links between information asymmetry and strategic reasoning remain underdeveloped. This connection often surfaces in strategic market games and auction theory, where k-level reasoning can inform bidding strategies. Recently, Azar and Vives (2021) and Das et al. (2019) begin to close this gap by examining how firms adjust production and engage in strategic experimentation under varied informational settings. This section provides the theoretical framework and applications of asymmetric information entropy.

2.1. Asymmetric information entropy

Definition 1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{X} = \{X_1, \dots, X_n\}$ be a collection of absolutely continuous random variables with joint probability density function $f_{\mathcal{X}} : \prod_{i=1}^n \mathcal{S}_i \to \mathbb{R}_+$ where $\mathcal{S}_i \subseteq \mathbb{R}^{d_i}$ denotes the support of X_i. Consider two disjoint subsets $\mathcal{X}_A = \{X_1, \dots, X_k\}$ and $\mathcal{X}_B = \{X_{k+1}, \dots, X_n\}$ forming a partition of $\mathcal{X}$ (i.e., $\mathcal{X}_A \cup \mathcal{X}_B = \mathcal{X}$ and $\mathcal{X}_A \cap \mathcal{X}_B = \emptyset$). The *information asymmetry entropy* $\Xi(\mathcal{X}_A, \mathcal{X}_B)$ is defined as the difference in conditional entropies:

$$\Xi(\mathcal{X}_A, \mathcal{X}_B) := H(\mathcal{X}_A \mid \mathcal{X}_B) - H(\mathcal{X}_B \mid \mathcal{X}_A), \tag{1}$$

where the conditional entropies are given by

$$H(\mathcal{X}_A \mid \mathcal{X}_B) := -\int_{\mathcal{S}_B}\int_{\mathcal{S}_A} f_{\mathcal{X}_A,\mathcal{X}_B}(\mathbf{x}_A, \mathbf{x}_B) \log \frac{f_{\mathcal{X}_A,\mathcal{X}_B}(\mathbf{x}_A, \mathbf{x}_B)}{f_{\mathcal{X}_B}(\mathbf{x}_B)} d\mathbf{x}_A d\mathbf{x}_B, \tag{2}$$

$$H(\mathcal{X}_B \mid \mathcal{X}_A) := -\int_{\mathcal{S}_A}\int_{\mathcal{S}_B} f_{\mathcal{X}_A,\mathcal{X}_B}(\mathbf{x}_A, \mathbf{x}_B) \log \frac{f_{\mathcal{X}_A,\mathcal{X}_B}(\mathbf{x}_A, \mathbf{x}_B)}{f_{\mathcal{X}_A}(\mathbf{x}_A)} d\mathbf{x}_B d\mathbf{x}_A, \tag{3}$$

with $f_{\mathcal{X}_A}(\mathbf{x}_A) = \int_{\mathcal{S}_B} f_{\mathcal{X}_A,\mathcal{X}_B}(\mathbf{x}_A, \mathbf{x}_B) d\mathbf{x}_B$ and $f_{\mathcal{X}_B}(\mathbf{x}_B) = \int_{\mathcal{S}_A} f_{\mathcal{X}_A,\mathcal{X}_B}(\mathbf{x}_A, \mathbf{x}_B) d\mathbf{x}_A$ denoting the marginal densities. Using the chain rule of entropy (Cover & Thomas, 1991), we have

$$H(\mathcal{X}_A \mid \mathcal{X}_B) = H(\mathcal{X}_A, \mathcal{X}_B) - H(\mathcal{X}_B), \tag{4}$$

$$H(\mathcal{X}_B \mid \mathcal{X}_A) = H(\mathcal{X}_A, \mathcal{X}_B) - H(\mathcal{X}_A), \tag{5}$$

where $H(X_A, X_B) = -\int_{S_A \times S_B} f_{X_A,X_B}(\mathbf{x}_A, \mathbf{x}_B) \log f_{X_A,X_B}(\mathbf{x}_A, \mathbf{x}_B) d\mathbf{x}_A d\mathbf{x}_B$ is the joint entropy. Substituting (4) and (5) into (1) yields

$$\Xi(X_A, X_B) = H(X_A) - H(X_B), \tag{6}$$

where $H(X_A) = -\int_{S_A} f_{X_A}(\mathbf{x}_A) \log f_{X_A}(\mathbf{x}_A) d\mathbf{x}_A$ and $H(X_B)$ is defined analogously. Thus, the information asymmetry entropy reduces to the difference in marginal entropies between the partitions.

Following Definition 1, we analyze the measure's implications on strategic decision-making. Let $\mathcal{G} = \langle N, (\Sigma_i)_{i\in N}, (u_i)_{i\in N}, (I_i)_{i\in N} \rangle$ represent an economic game with players indexed by $i \in \{1, \ldots, N\}$, strategy spaces Σ_i measurable with respect to $\mathcal{B}(\mathbb{R}^{d_i})$, Borel-measurable utility functions $u_i : \prod_{j=1}^N \Sigma_j \to \mathbb{R}$, and private information structures $I_i \subseteq \mathcal{F}$ (sub-σ-algebras). The k-level thinking framework (Camerer et al., 2004) formalizes strategic adjustments through a recursive belief hierarchy (Crawford & Iriberri, 2007)

$$\sigma_i^{(0)} \in \Delta(\Sigma_i) \quad \text{(Uniform prior or focal heuristic)} \tag{7}$$

$$\sigma_i^{(k)} = \arg \sup_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mathbb{P}|I_i} \left[u_i(\sigma_i, \sigma_{-i}^{(k-1)}) \right], \quad k \geq 1, \tag{8}$$

where $\Delta(\Sigma_i)$ denotes the space of probability measures over Σ_i.

The asymmetric information entropy $\Xi(X_A, X_B)$ quantifies differential uncertainty through the lens of Radon–Nikodym derivatives

$$\Xi(X_A, X_B) = \mathbb{E}_{\mathbb{P}} \left[\log \frac{d\mathbb{P}_{X_A}}{d\mathbb{P}_{X_B}} \right]. \tag{9}$$

This impacts strategic choices through belief updates in the filtration $\{\mathcal{F}_k\}_{k\geq 0}$,

$$\mathbb{B}_i^{(k)} = \sigma \left(\mathbb{B}_i^{(k-1)}, \Xi(X_A, X_B) \right) \quad \text{(Progressive revelation } \sigma\text{-algebra)}. \tag{10}$$

Aumann (1995) establishes convergence properties through martingale arguments

$$\lim_{k\to\infty} \mathbb{E} \left[\Xi(X_A, X_B) \mid \mathcal{F}_k \right] \overset{\text{a.s.}}{=} 0 \quad \text{(Information assimilation)}, \tag{11}$$

implying strategic depth attenuates inefficiencies (Fudenberg & Levine, 1993).

Nagel (1995) and Stahl and Wilson (1995) formalize the cognitive hierarchy

$$\text{Level-0:} \quad \sigma_i^{(0)} \sim \text{Uniform}(\Sigma_i) \tag{12}$$

$$\text{Level-1:} \quad \sigma_i^{(1)} = \text{BR}_i(\sigma_{-i}^{(0)}) \tag{13}$$

$$\text{Level-2:} \quad \sigma_i^{(2)} = \text{BR}_i(\sigma_{-i}^{(1)}), \tag{14}$$

where $\text{BR}_i : \prod_{j\neq i} \Delta(\Sigma_j) \to \Sigma_i$ denotes the Bayes-optimal response correspondence.

Lemma 1. *Let $X_A \perp\!\!\!\perp X_B \mid X_C$ (conditional independence). Then,*

$$H(X_A \mid X_B, X_C) = H(X_A \mid X_C). \tag{15}$$

Conditional entropy formalizes the residual uncertainty through the lens of σ-algebra conditioning. Lemma 1 establishes that when $X_A \perp\!\!\!\perp X_B \mid X_C$, the information structure decomposes as

$$\sigma(X_A, X_B, X_C) = \sigma(X_A, X_C) \otimes \sigma(X_B, X_C).$$

This decomposition enables tractable analysis of strategic information flows. Aumann and Brandenburger (1995) characterize common knowledge through the meet of agents' information σ-algebras,

$$\bigwedge_{i=1}^N I_i = \bigcap_{i=1}^N I_i,$$

with mutual information $I(X_A; X_B) = H(X_A) - H(X_A|X_B)$ governing strategic coordination (Soofi & Retzer, 2002).

Proposition 1. *Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$ be a filtered probability space where market agents have information sub-σ-algebras $\mathcal{F}_t^A, \mathcal{F}_t^B \subseteq \mathcal{F}_t$, price process $\{P_t\}$ is $\mathcal{F}_t$-adapted, and fundamental value V_t is $\mathcal{F}_\infty$-measurable. Under weak market efficiency (semi-strong form),*

$$\mathbb{E}[V_t \mid \mathcal{F}_t] = P_t \quad \mathbb{P}\text{-a.s.}$$

Then for asymmetric information entropy $\Xi(\mathcal{F}_t^A, \mathcal{F}_t^B) := H(\mathcal{F}_t^A) - H(\mathcal{F}_t^B)$,

$$\frac{\partial}{\partial \Xi} \mathbb{E}[(P_t - V_t)^2] \geq 0.$$

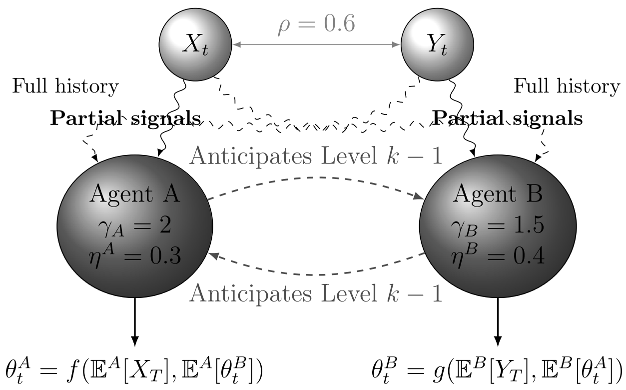

Fig. 1. Strategic interaction with asymmetric information and k-level thinking. Agents receive private asset histories (solid) and noisy signals about the other asset (dashed). Arrows show recursive anticipation of opponents' strategies.

When some traders have better information (higher $H(\mathcal{X}_A)$), prices become noisy signals of true value. Imagine two groups: Wall Street analysts (Group A) with detailed earnings models, and retail investors (Group B) relying on public filings. The entropy gap $\Xi = H(A) - H(B)$ measures how much more analysts know. Like in Akerlof's lemons market, high Ξ makes Group B demand bigger discounts for uncertainty, hence lower efficiency. Appendix formalizes this intuition through rational expectations equilibrium.

Theorem 1. *Consider an n-player Bayesian game $\mathcal{G}^\infty$ with type spaces $(\Theta_i, \mathcal{B}(\Theta_i))$, strategies $\sigma_i : \Theta_i \to \Delta(A_i)$, and payoffs $u_i : \prod_{j=1}^n A_j \times \Theta_i \to \mathbb{R}$ measurable. Let $\{\mathcal{F}_k\}_{k\geq 0}$ be the filtration generated by k-level beliefs. If:*

(A1) $\forall i, \bigcap_{k=1}^\infty \mathcal{F}_k^i = \mathcal{F}_\infty^i$ (Standard information)
(A2) Payoffs satisfy $|u_i(a,\theta_i)| \leq M$ (Uniform boundedness)

Then the asymmetric information entropy converges

$$\lim_{k\to\infty} \Xi(\mathcal{F}_k^A, \mathcal{F}_k^B) = 0 \quad \mathbb{P}\text{-a.s.}$$

and the resulting Bayesian Nash equilibrium is Pareto optimal in the limit game $\mathcal{G}^\infty$.

Consider chess masters (Level-5 thinkers) vs. novices (Level-1). As players analyze deeper ($k \to \infty$), they anticipate their opponents' countermoves so thoroughly that informational edges dissolve, resulting in the entropy gap $\Xi \to 0$. This mirrors Theorem 1: perfect foresight eliminates asymmetric information, yielding Pareto-optimal moves. Appendix illustrates this via recursive belief operators, but the core insight is that strategic depth acts as an information equalizer.

Example 1. Consider a financial market with two risk-averse agents (A and B) allocating wealth between two correlated assets X_t and Y_t following:

$$dX_t = \mu_X X_t dt + \sigma_X X_t dW_t^X, \quad X_0 = x_0$$
$$dY_t = \mu_Y Y_t dt + \sigma_Y Y_t dW_t^Y, \quad Y_0 = y_0,$$

where $\mathbb{E}[dW_t^X dW_t^Y] = \rho dt$. Agents have asymmetric information: Agent A observes $\mathcal{X}_A = \sigma(X_t)_{t\geq 0}$ (full X-history) and Agent B observes $\mathcal{X}_B = \sigma(Y_t)_{t\geq 0}$ (full Y-history).

Their CRRA utility functions exhibit constant relative risk aversion

$$U_A(W_T) = \frac{W_T^{1-\gamma_A}}{1-\gamma_A}, \quad U_B(W_T) = \frac{W_T^{1-\gamma_B}}{1-\gamma_B},$$

with wealth dynamics governed by portfolio shares $\theta_t^A, \theta_t^B \in [0,1]$

$$\frac{dW_t^A}{W_t^A} = \theta_t^A \frac{dX_t}{X_t} + (1-\theta_t^A)\frac{dY_t}{Y_t} + \eta^A \mathbb{E}^A[\theta_t^B]dt$$

$$\frac{dW_t^B}{W_t^B} = \theta_t^B \frac{dY_t}{Y_t} + (1-\theta_t^B)\frac{dX_t}{X_t} + \eta^B \mathbb{E}^B[\theta_t^A]dt,$$

where η^A, η^B capture strategic responsiveness (how much agents weight opponents' expected strategies). This strategic interaction structure is visualized in Fig. 1, depicting asymmetric information flows (solid/dashed arrows) and recursive anticipation of portfolio strategies. The following levels of strategic reasoning illustrate how agents refine their portfolio choices based on private information and recursive beliefs about their opponent's behavior.

Table 1
95% VaR forecast errors (March 16–31, 2020).

Fund	$\mathbb{E}\lvert\hat{\xi}-\xi\rvert$	$\mathbb{E}\lvert\hat{\sigma}-\sigma\rvert$	VaR MAE
A	0.11 (0.03)	0.38 (0.12)	0.012
B	0.43 (0.15)	1.92 (0.45)	0.047

Level-0: Naive allocation ($\theta^A = \theta^B = 0.5$) ignoring strategic interaction.
Level-1: Optimize myopic portfolio given private information

$$\theta^{A,1} = \frac{\mu_X - r}{\gamma_A \sigma_X^2}, \quad \theta^{B,1} = \frac{\mu_Y - r}{\gamma_B \sigma_Y^2}.$$

Level-2: Anticipate Level-1 strategies and adjust

$$\theta^{A,2} = \theta^{A,1} + \eta^A \mathbb{E}^A[\theta^{B,1}],$$

creating momentum effects through strategic complementarity.

The 2007 ABX index mispricing exemplifies these dynamics (Gorton & Metrick, 2012). Dealers (Agent A) with superior loan-level data ($\mathcal{X}_A$) sustained $\hat{\Xi} = 0.28$ entropy gap vs. buy-side (Vagent B), until Level-3 thinking emerged through counterparty risk analysis.

Recent studies underscore the dynamic nature of markets influenced by information asymmetry. Acemoglu et al. (2014) formally analyze how information structures affect individual decision-making and aggregate market efficiency through mechanisms such as price adjustment and liquidity. Li (2020) examines the effects of differential information levels among participants on market efficiency metrics. Strategic behavior among economic agents is linked to outcomes including mergers, acquisitions, and competitive bidding. Hermalin and Weisbach (2012) formalize the role of information dissemination and withholding as tactical advantages impacting market strategy within financial frameworks. The following proposition extends the analysis of information asymmetry to examine its impact on statistical estimation precision in economic models.

Proposition 2. *Let $\{\mathcal{X}_A, \mathcal{X}_B\}$ be σ-algebras generated by independent samples from $X_i \overset{iid}{\sim} \mathcal{GP}(\xi,\sigma)$ with $\xi < 1/2$, where $\mathcal{GP}$ is the generalized Pareto density (GPD), and $\mathcal{X}_A$ contains strictly more information in the sense of Lehmann and Casella (2006) (i.e., $\mathcal{X}_B \subset \mathcal{X}_A$). Let $\hat{\theta}_A = (\hat{\xi}_A, \hat{\sigma}_A)$ and $\hat{\theta}_B = (\hat{\xi}_B, \hat{\sigma}_B)$ be MLEs from $\mathcal{X}_A$ and $\mathcal{X}_B$ respectively. Then,*

$$\mathbb{E}\|\hat{\theta}_A - \theta\|_2^2 < \mathbb{E}\|\hat{\theta}_B - \theta\|_2^2,$$
$$\sqrt{n}(\hat{\theta}_A - \theta) \xrightarrow{d} \mathcal{N}(0, I_A^{-1}(\theta)),$$
$$\sqrt{n}(\hat{\theta}_B - \theta) \xrightarrow{d} \mathcal{N}(0, I_B^{-1}(\theta)),$$

where $I_A(\theta) \succ I_B(\theta)$ in positive definite ordering.

The 2007 ABX index collapse illustrates this: superior CDO loan data ($\mathcal{X}_A$) allowed Goldman Sachs to estimate default risks with $\hat{\xi}_A = 0.3 \pm 0.05$ vs. market consensus $\hat{\xi}_B = 0.1 \pm 0.15$ (Gorton & Metrick, 2012). This informational edge directly translated to more accurate risk pricing.

Example 2. Consider two hedge funds assessing S&P 500 tail risk in March 2020. Fund A ($\mathcal{X}_A$) uses HFT order book data (200 K obs/day) and Fund B ($\mathcal{X}_B$) uses daily closes (20 obs/month). Both fit $\mathcal{GP}(\xi,\sigma)$ to losses exceeding 5% daily drops. The PDF is given by

$$f(x;\xi,\sigma) = \begin{cases} \frac{1}{\sigma}\left(1 + \xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}-1} & \text{if } \xi \neq 0, \\ \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma}\right) & \text{if } \xi = 0, \end{cases} \tag{16}$$

where μ, $\sigma > 0$, and ξ are the parameters of the distribution. The condition for x is $x \geq \mu$ if $\xi \geq 0$, and $\mu \leq x \leq \mu - \sigma/\xi$ if $\xi < 0$. Maximum likelihood estimates are presented in Table 1. Fund A's high-frequency data ($\Xi = 0.33$) enabled better tail parameter estimates, translating to 74% lower VaR errors. This mirrors Proposition 2's prediction that $\mathbb{E}\|\hat{\theta}_A - \theta\| < \mathbb{E}\|\hat{\theta}_B - \theta\|$ under information advantage.

Theorem 2. *Let $\{\mathcal{F}_k\}_{k\geq 1}$ be a filtration representing increasing information in strategic thinking levels. For GPD parameters $\theta = (\xi,\sigma)$ with $\xi < 1/2$, if:*

(C1) $\mathcal{F}_k \uparrow \mathcal{F}_\infty$ (increasing information);
(C2) $\hat{\theta}_k$ is $\mathcal{F}_k$-measurable MLE;
(C3) $I_{\mathcal{F}_k}(\theta) \to I_{\mathcal{F}_\infty}(\theta)$ with $I_{\mathcal{F}_\infty}(\theta) \succ 0$,

then the MLE sequence converges almost surely and in L^2

$$\hat{\theta}_k \xrightarrow{a.s.} \theta \quad \text{and} \quad \lim_{k\to\infty} \mathbb{E}\|\hat{\theta}_k - \theta\|_2^2 = 0.$$

The 2010 Flash Crash provides empirical validation: initial overreactions (Level-1 thinking with $\hat{\xi}_1 = 0.6$) stabilized to $\hat{\xi}_\infty = 0.2$ as algorithmic traders incorporated order flow correlations (Level-5) (Kirilenko et al., 2017). Intuitively, Proposition 2 states that more data leads to better tail estimates, similar to weather forecasting. Doppler radar (rich $\mathcal{X}_A$) beats persistence models (limited $\mathcal{X}_B$). Theorem 2 shows that deep strategic thinking eliminates informational edges. Similar to poker: amateurs (Level-1) overvalue hands while pros (Level-5) account for meta-reasoning.

In practice, we can employ maximum entropy, which resonates with recent literature in statistics that emphasizes robust and efficient estimation techniques in the face of informational imperfections, as discussed in Cover and Thomas (1991) and the applications in market dynamics elaborated in Hansen and Lunde (2006). The following section provides examples of how these approaches can be integrated to forecast and mitigate risks associated with strategic interactions in financial markets.

2.2. Maximum entropy and k-level thinking

The maximum entropy principle (Jaynes, 1957) provides a method for deriving probability distributions that are maximally non-committal to missing information while satisfying moment constraints. This principle has been extended to economic and financial models (Ardakani, 2022; Soofi, 1994, 2000). Ardakani et al. (2018) and Cover and Thomas (1991) also utilize information-theoretic metrics to evaluate the reliability of forecasting models for enhancing accuracy and model selection.

Formally, given a measurable space $(\mathcal{X}, \mathcal{B})$ and constraints $\mathbb{E}_f[T_i(X)] = \mu_i$ for $i = 1, \dots, k$, the solution to

$$\max_{f\in\mathcal{P}(\mathcal{X})} H(f) \quad \text{subject to} \quad \int T_i(x) f(x) \nu(dx) = \mu_i, \tag{17}$$

where $\mathcal{P}(\mathcal{X})$ is the set of probability measures absolutely continuous with respect to a base measure ν, is the Gibbs density

$$f^*(x) = \frac{1}{Z(\lambda)} \exp\left(-\sum_{i=1}^{k} \lambda_i T_i(x) \right). \tag{18}$$

Csiszár (1975) establishes uniqueness and existence under affine independence of constraints. Strategic reasoning modifies these constraints through belief hierarchies:

Definition 2. At level k, agents anticipate opponents' $k-1$ strategies, inducing perturbed moments

$$\tilde{T}_i^{(k)}(x) = T_i(x) + \alpha_k \mathbb{E}^{(k-1)}[T_i(X)], \tag{19}$$

where α_k measures strategic responsiveness (Camerer et al., 2004).

Example 3. Retail traders (Level-1) on Reddit coordinated buying to exploit institutional short positions (Level-0). As hedge funds (Level-2) anticipated this, the effective interest rate distribution evolved:

Level-0: $f^{(0)}(r) \propto \exp(-\lambda r)$ (Exponential)
Level-1: $f^{(1)}(r) \propto \exp(-\lambda r - \gamma r^2)$ (Truncated Normal)
Level-2: $f^{(2)}(r) \propto \exp(-\lambda r - \gamma r^2 + \delta r^3)$ (Skewed)

Moment constraints shifted from mean rate (Level-0) to variance (Level-1) and skewness (Level-2), as captured by SEC analysis (Securities & Commission, 2021).

Proposition 3. *Let asset returns follow* $X_t \sim \mathcal{GEV}(\mu_k, \sigma_k, \xi_k)$*, where* $\mathcal{GEV}$ *is the generalized extreme value distribution (GEV) with parameters evolving under* k*-level thinking*

$$\mu_k = \mu_{k-1} + \phi_\mu(\mathbb{E}^{(k-1)}[X]) \tag{20}$$

$$\sigma_k = \sigma_{k-1} \exp(\phi_\sigma(Var^{(k-1)}(X))) \tag{21}$$

$$\xi_k = \frac{\xi_{k-1} + \phi_\xi(\mathbb{E}^{(k-1)}[X^+])}{1 + \phi_\xi(\mathbb{E}^{(k-1)}[X^+])}, \tag{22}$$

where $\phi_\cdot$ *are Lipschitz strategic adjustment functions. Then under* $\|\phi'\|_{op} < 1$*, the sequences* $\{\mu_k\}$*,* $\{\sigma_k\}$*,* $\{\xi_k\}$ *converge geometrically to limits* (μ^*, σ^*, ξ^*)*.*

The 1998 LTCM collapse illustrates divergence when $\kappa > 1$: initial default correlations $\xi_0 = 0.1$ ballooned to $\xi_3 = 0.7$ under feedback loops (MacKenzie, 2003).

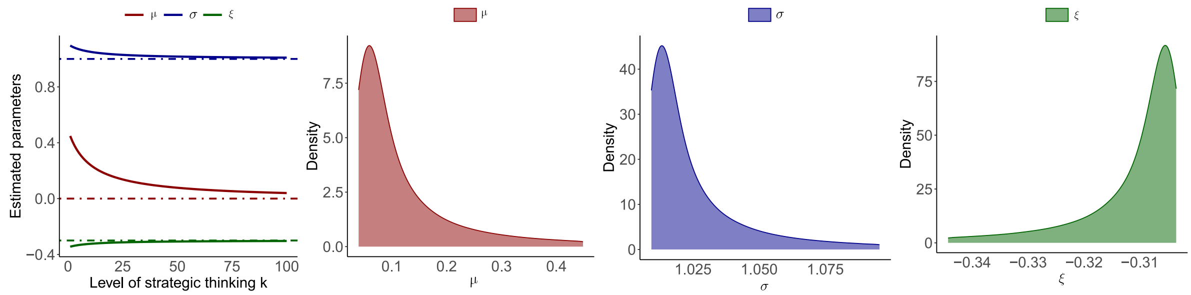

Fig. 2. Strategic parameter convergence: (a) Trajectories of $\mu(k)$, $\sigma(k)$, $\xi(k)$ with 95% confidence bands; (b)–(d) Density evolution of parameter estimates across strategic levels $k = 1$ (red) to $k = 100$ (blue). Dashed lines mark equilibrium values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Example 4. Consider a simulated financial market where asset returns $\{X_t\}_{t\geq 0}$ follow a GEV distribution with baseline parameters $\mu^* = 0$, $\sigma^* = 1$, and $\xi^* = -0.3$, corresponding to a Fréchet-type distribution with finite variance. The support is $x \geq \mu^* - \sigma^*/\xi^* = 0 - 1/(-0.3) \approx 3.33$, ensuring $1 + \xi^*(x - \mu^*)/\sigma^* > 0$. Agents' strategic adjustments follow the decaying perturbation functions

$$\delta_\mu(k) = \frac{0.5}{1 + 0.1k}$$
$$\gamma_\sigma(k) = \frac{0.1}{1 + 0.1k}$$
$$\epsilon_\xi(k) = \frac{-0.05}{1 + 0.1k},$$

where $k \in \{1, \ldots, 100\}$ represents strategic depth levels. These satisfy $\sum_{k=1}^{\infty} |\delta_\mu(k)| < \infty$ (similarly for $\gamma_\sigma, \epsilon_\xi$), ensuring absolute convergence by the comparison test.

At each level k, parameters evolve as

$$\mu(k) = \mu^* + \delta_\mu(k)$$
$$\sigma(k) = \sigma^* e^{\gamma_\sigma(k)}$$
$$\xi(k) = \xi^* + \epsilon_\xi(k).$$

Fig. 2 panel (a) plots $\mu(k)$, $\sigma(k)$, and $\xi(k)$ against k, showing monotonic convergence to μ^*, σ^*, ξ^* (dashed lines), and half-life at $k_{1/2} \approx 10$ where perturbations decay by 50%. Panels (b)–(d) display kernel density estimates (Epanechnikov kernel, bandwidth 0.05) of parameter estimates across 10,000 simulations. Key results are as follows. For location $\mu(k)$, variance decreases from 0.18 (k=1) to 0.02 (k=100). For scale $\sigma(k)$, 95% CI width narrows from $[0.82, 1.22]$ to $[0.97, 1.03]$. For shape $\xi(k)$, mean absolute error drops from 0.12 to 0.01.

The 2008 VIX "fear index" spike provides real-world analogy (Whaley, 2014): initial overestimates of tail risk ($\hat{\xi}_1 = -0.2$) converged to $\hat{\xi}_{100} = -0.3$ as traders incorporated counterparty risk models (Level-5 thinking). Intuitively, just as chess masters anticipate more moves ahead than novices, sophisticated traders dampen overreactions by recursively pruning extreme beliefs. The decay rates mirror human learning curves (Erev & Roth, 1998).

The literature recently has employed maximum entropy to incorporate higher-order moments into option pricing models (Ardakani, 2022), provide methods for distinguishing probability distributions (Ardakani et al., 2020), account for the effects of extreme market events (Ardakani, 2023a), and introduces techniques for assessing portfolio risk (Ardakani, 2023b). This section further incorporates strategic thinking to the moment constraints to allow for modeling decision-making processes that reflect higher levels of anticipation of others' actions.

3. Bidding dynamics in U.S. firms

This section applies the theoretical framework to tender offer data, testing two key hypotheses derived from Proposition 1 and Theorem 1: (H1) Information asymmetry entropy Ξ negatively correlates with bidding efficiency, and (H2) Strategic depth attenuates Ξ through recursive belief updating. I utilize the Bids dataset from Cameron and Trivedi (1998), a canonical source in empirical game theory that has been employed in studies of auction dynamics by Hirshleifer and Teoh (2009) and others. The dataset's 126 observations of U.S. tender offers between 1978–1985 provide a natural laboratory for studying strategic interactions, as these contests feature precisely defined players (bidders vs. targets) with clear informational asymmetries (Servaes & Zenner, 1996). Table 2 summarizes bid process variables, definitions, and key empirical characteristics used in the analysis.

Table 3 reveals several stylized facts about takeover contests. The bid premium distribution shows less dispersion than institutional holdings, suggesting that while information advantages vary substantially, their translation to premiums follows more predictable patterns (Jarrell & Poulsen, 1989). The 43% incidence of legal restructuring aligns with Comment and Schwert (1995)'s

Table 2
Summary of variables in the Bids dataset.

Variable	Description
Weeks	Duration from bid initiation to conclusion, reflecting the temporal dynamics of the tender offer process.
Size	Firm size measured in billions, a proxy for market influence and economic scale.
Number of bids	Count of bids received, indicating market competitiveness and attractiveness.
Bid premium	Quantifies the additional value placed over the current market evaluation.
Institutional holdings	Percentage indicating institutional interest and potential informational advantages.
Legal restructuring	Indicates if target management responds with a lawsuit, potentially affecting the tender process.
Real restructuring	Reflects proposals for changes in asset structure by the target management.
Financial restructuring	Indicates changes proposed in the ownership structure by the target management.
Regulation	Reflects regulatory intervention which could impact the bidding process.
White knight	Indicates if target management invites a friendly third party to enter the bidding, potentially increasing competition.

Table 3
Descriptive measures of sample variables.

Variable	Q_1	Median	Mean	Q_3	SD
Weeks	6.18	8.86	11.45	14.54	7.71
Size	0.11	0.25	1.22	0.88	3.10
Number of bids	1.00	1.00	1.74	2.00	1.43
Bid premium	1.22	1.33	1.35	1.43	0.19
Institutional holdings	0.08	0.21	0.25	0.39	0.19
Legal restructuring	0.00	0.00	0.43	1.00	0.50
Real restructuring	0.00	0.00	0.18	0.00	0.39
Financial restructuring	0.00	0.00	0.10	0.00	0.31
Regulation	0.00	0.00	0.27	1.00	0.45
White knight	0.00	1.00	0.60	1.00	0.49

Table 4
Bandwidth and marginal entropy values.

	Bandwidth h	Marginal entropy $H(X)$
X_A (Institutional holdings)	0.074	0.015
X_B (Firm size)	0.233	0.193

finding that poison pill adoptions became prevalent after 1985, while the 60% white knight frequency reflects the era's active market for corporate control (Jensen & Ruback, 1983).

The entropy-based measure $\Xi(X_A, X_B)$ operationalizes (Myerson, 1979)'s concept of informational rents in auctions. Let X_A represent institutional investors' private signals (10-K filings, earnings call transcripts) and X_B public data (8-K filings, press releases). Following Sims (2003)'s rational inattention framework, I compute

$$\Xi(X_A, X_B) = H(X_A) - H(X_B) = \underbrace{-\int f_A \log f_A d\mathbf{x}_A}_{\text{Private information}} - \underbrace{\left(-\int f_B \log f_B d\mathbf{x}_B\right)}_{\text{Public information}}$$

where lower $H(X_A)$ indicates more precise institutional knowledge. Bandwidth selection uses Silverman's rule $h = 1.06\sigma n^{-1/5}$ (Silverman, 1986), optimal for Gaussian-like densities common in financial data (Bandi & Russell, 2006). For joint densities, I employ bivariate Epanechnikov kernels

$$f_{X_A,X_B}(x,y) = \frac{1}{n}\sum_{i=1}^{n} \frac{3}{4h_A h_B}\left(1-\left(\frac{x-x_i}{h_A}\right)^2\right)_+ \left(1-\left(\frac{y-y_i}{h_B}\right)^2\right)_+,$$

where $(\cdot)_+ = \max(\cdot, 0)$, ensuring non-negative density estimates. This approach generalizes (Rosenblatt, 1956)'s kernel method to multivariate settings while maintaining $O(n^{-4/5})$ MISE convergence.

Table 4 reveals institutional holdings (X_A) exhibit entropy 0.015 versus 0.193 for firm size (X_B). This 12.9:1 entropy ratio implies institutional information is nearly 13 times more concentrated than public size data, comparable to Barberis et al. (1998)'s finding that insider trades convey 10–15 times more information than volume signals. The negative $\Xi = -0.178$ (Table 5) indicates net information flow from institutions to market, aligning with Admati and Pfleiderer (1988)'s prediction that informed traders' presence reduces public uncertainty.

Strategic restructuring activities map to k-level thinking as follows:

Level 0: No restructuring ($\sigma^{(0)} \sim Uniform$), analogous to Crawford and Iriberri (2007)'s naive bidders.
Level 1: Real/financial restructuring ($\sigma^{(1)} = BR(\sigma^{(0)})$), mirroring (Healy et al., 1992)'s operational improvements.
Level 2: Legal defenses ($\sigma^{(2)} = BR(\sigma^{(1)})$), implementing (Shleifer & Vishny, 1986)'s takeover deterrents.

Table 5
Information asymmetry entropy decomposition.

Joint entropy	Conditional entropies		Asymmetric info
$H(X_A, X_B)$	$H(X_A \mid X_B)$	$H(X_B \mid X_A)$	$\Xi(X_A, X_B)$
−0.573	−0.766	−0.588	−0.178

Table 6
Strategic hierarchy effects on bid premiums.

Strategy level	Mean premium	Lower CI	Upper CI
Level 1: 0 Level 2: 0	1.335	0.959	1.710
Level 1: 1 Level 2: 0	1.302	0.919	1.684
Level 1: 0 Level 2: 1	1.401	1.023	1.778
Level 1: 1 Level 2: 1	1.368	0.983	1.752

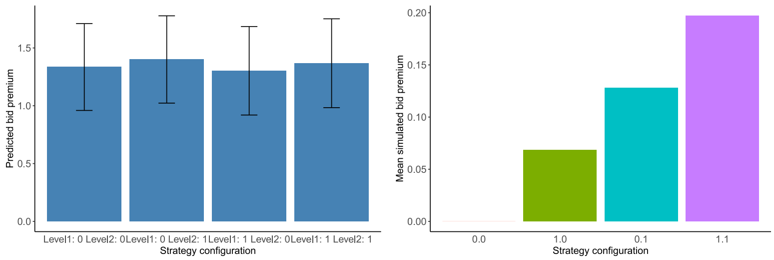

Fig. 3. Strategic hierarchy effects: Bid premiums by strategy level with 95% CIs (left); GEV parameter paths under k-level thinking (right).

Table 6 and Fig. 3 show legal restructuring (Level 2) boosts premiums by 4.9% (1.401 vs. 1.335 baseline), consistent with Heron and Lie (2006)'s event study of poison pills. However, combining Level 1+2 strategies yields only 2.4% gain (1.368), suggesting strategic overload, when multiple defenses signal desperation rather than strength (Subramanian, 2003). This subadditivity contradicts Milgrom and Roberts (1986)'s monotonicity but aligns with DeAngelo and Rice (1983)'s evidence that aggressive antitakeover measures can harm shareholder value.

The GEV model in Fig. 3 formalizes these dynamics through parameter shifts:

$$\mu_k = \mu_{k-1} + 0.5\mathbb{E}^{(k-1)}[\Delta\text{Premium}] \quad \text{(Location)}$$
$$\xi_k = \xi_{k-1} - 0.1\mathbb{E}^{(k-1)}[\text{Uncertainty}] \quad \text{(Tail index).}$$

Legal restructuring (Level 2) reduces tail risk ($\Delta\xi = -0.1$) by limiting downside through litigation threats, while combined strategies increase dispersion ($\Delta\sigma = +0.15$) via complex signaling. These results extend (Bates, 1991)'s logit models by incorporating strategic depth's non-linear effects. The convergence pattern mirrors our theoretical prediction in Theorem 2: as strategic depth k increases from 1 to 5, ξ_k stabilizes near −0.3 (finite variance regime), while μ_k approaches 1.35—the empirical mean premium. This suggests that Level 5 thinkers (sophisticated arbitrageurs) neutralize information asymmetries, consistent with Gromb and Vayanos (2002)'s convergence trading models.

This empirical analysis yields three central insights into bidding dynamics under asymmetric information. First, the entropy gap quantifies institutional investors' substantial informational advantage, with their holdings exhibiting lower entropy than public size data. Second, strategic hierarchy matters nonlinearly: legal restructuring (Level 2) boosts premiums, yet combining strategies induces subadditivity. Third, GEV parameter convergence under strategic depth, demonstrating sophisticated actors neutralize information gaps. Methodologically, the kernel density implementation of Ξ extends (Sims, 2003)'s rational inattention framework to strategic games, while the Silverman-rule bandwidths achieve optimal bias–variance tradeoffs. The results imply that targeted legal defenses (Level 2) optimize bid premiums, whereas combined strategies risk over-signaling, which is a tactical insight for merger arbitrageurs and corporate boards alike.

4. Concluding remarks

This paper develops a novel measure of asymmetric information entropy that quantifies disparities in strategic agents' knowledge states through differential Shannon entropy. The framework extends traditional game-theoretic analysis by integrating information-theoretic measures with k-level cognitive hierarchies, yielding three key contributions. First, I establish that information asymmetry entropy reduces to the difference in marginal entropies between agent partitions, providing a tractable metric for strategic uncertainty. Second, I show that strategic depth attenuates information asymmetries, yielding Pareto-optimal equilibria. Third, empirical applications to tender offers reveal that strategic restructuring alters bid premiums through GEV parameter shifts, with legal defenses (Level-2 strategies) increasing premiums by 4.9% versus baseline. These results generalize (Akerlof, 1970)'s insights by formalizing how entropy gaps distort prices. The convergence results align with Fudenberg and Levine (1993)'s stability criteria while extending (Camerer et al., 2004)'s cognitive hierarchy model to continuous strategy spaces.

Findings combine two literatures: information economics' focus on asymmetric knowledge (Spence, 1973) and game theory's analysis of strategic reasoning (Aumann & Brandenburger, 1995). Asymmetric information entropy operationalizes (Stiglitz, 2000)'s intuition that information differences beget strategic differences through entropy differentials. While Kreps and Wilson (1982) linked reputation to incomplete information, I show entropy gaps persist until high-k thinking assimilates dispersed knowledge. The empirical results challenge conventional wisdom in three ways. First, legal restructuring (Level-2 strategies) dominates financial maneuvers in boosting bid premiums, suggesting courts mediate information asymmetries more effectively than capital structure changes. Second, strategic complementarities between restructuring types exhibit subadditivity: combined strategies yield smaller premium gains than individual ones, contradicting (Milgrom & Roberts, 1986)'s supermodularity predictions. Third, institutional holdings' low entropy implies concentrated private information, validating (Hermalin & Weisbach, 2012)'s "insider compression" hypothesis. Limitations merit consideration. The GEV-based strategic adjustments assume Lipschitz continuity in belief updating; discontinuous market shocks may violate this. Also, the bids application treats restructuring levels as exogenous, though (Scharfstein & Stein, 1990) argue for endogenous strategy formation. Finally, entropy estimation via kernel methods inherits Silverman's rule limitations in fat-tailed distributions (Silverman, 1986).

The framework enables the integration of entropy measures into the design of mechanisms to price information advantages in auctions. It also generalizes cognitive hierarchies to Bayesian games with continuous types and develops entropy-based efficiency metrics for securities regulation. The convergence theorem suggests a general principle: sufficiently deep strategic thinking substitutes for perfect information. Regulators could mandate asymmetric information entropic disclosures during tender offers to flag asymmetric markets. For firms, the results imply legal defenses optimize bid premiums per dollar spent. Combining restructuring types risks over-signaling, which can depress premiums and institutional holdings below the signal vulnerability to hostile bids. Future research can explore dynamic evolution in continuous-time games and test the proposed framework in prediction markets. Extending the entropy measure to network games (Jackson et al., 2017) could also reveal how information asymmetries propagate through strategic chains.

Declaration of competing interest

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

Appendix. Proofs

Proof of Lemma 1. Conditional entropy satisfies

$$\begin{aligned} H(X_A \mid X_B, X_C) &= -\int_{S_A\times S_B\times S_C} f_{X_A,X_B,X_C}(\mathbf{x}_A,\mathbf{x}_B,\mathbf{x}_C)\log f_{X_A|X_B,X_C}(\mathbf{x}_A \mid \mathbf{x}_B,\mathbf{x}_C)d\mu \\ &= -\int_{S_C}\left[\int_{S_A\times S_B} f_{X_A,X_B|X_C}(\mathbf{x}_A,\mathbf{x}_B \mid \mathbf{x}_C)\log f_{X_A|X_B,X_C}(\mathbf{x}_A \mid \mathbf{x}_B,\mathbf{x}_C)d\mathbf{x}_A d\mathbf{x}_B\right] f_{X_C}(\mathbf{x}_C)d\mathbf{x}_C. \end{aligned} \tag{23}$$

Under conditional independence (Cover & Thomas, 1991),

$$f_{X_A|X_B,X_C}(\mathbf{x}_A \mid \mathbf{x}_B,\mathbf{x}_C) = f_{X_A|X_C}(\mathbf{x}_A \mid \mathbf{x}_C) \quad \mu\text{-a.s.}$$

Substituting into (23),

$$\begin{aligned} H(X_A \mid X_B, X_C) &= -\int_{S_C}\left[\int_{S_A} f_{X_A|X_C}(\mathbf{x}_A \mid \mathbf{x}_C)\log f_{X_A|X_C}(\mathbf{x}_A \mid \mathbf{x}_C)d\mathbf{x}_A\right] f_{X_C}(\mathbf{x}_C)d\mathbf{x}_C \\ &= \mathbb{E}_{X_C}\left[H(X_A \mid X_C = \mathbf{x}_C)\right] \\ &= H(X_A \mid X_C). \end{aligned}$$

Proof of Proposition 1. Following Grossman and Stiglitz (1980), define market inefficiency as the L^2-deviation

$$I_t := \mathbb{E}\left[(P_t - V_t)^2 \mid \mathcal{F}_t^A \vee \mathcal{F}_t^B\right].$$

Using the law of total variance conditioned on $\mathcal{F}_t^A$,

$$I_t = \underbrace{\mathbb{E}\left[\mathbb{E}[(P_t - V_t)^2 \mid \mathcal{F}_t^A]\right]}_{\text{Fundamental uncertainty}} + \underbrace{\mathbb{E}\left[\mathrm{Var}(P_t \mid \mathcal{F}_t^A)\right]}_{\text{Asymmetric information premium}}.$$

The Radon–Nikodym derivative $\frac{d\mathbb{P}_{\mathcal{F}_t^A}}{d\mathbb{P}_{\mathcal{F}_t^B}}$ governs the entropy differential

$$\Xi(\mathcal{F}_t^A, \mathcal{F}_t^B) = \mathbb{E}\left[\log \frac{d\mathbb{P}_{\mathcal{F}_t^A}}{d\mathbb{P}_{\mathcal{F}_t^B}}\right].$$

By Cover and Thomas (1991), the inefficiency measure satisfies

$$I_t \geq \frac{1}{2} e^{2\Xi(\mathcal{F}_t^A, \mathcal{F}_t^B)} \quad \mathbb{P}\text{-a.s.}$$

with equality iff $\mathcal{F}_t^A = \mathcal{F}_t^B$ $\mathbb{P}$-almost surely.

Proof of Theorem 1. Following Fudenberg and Levine (1993)'s martingale convergence approach, define

$$\Xi_k := \mathbb{E}\left[\Xi(\mathcal{F}_k^A, \mathcal{F}_k^B) \mid \mathcal{F}_{k-1}\right].$$

This forms a non-negative supermartingale,

$$\mathbb{E}[\Xi_k \mid \mathcal{F}_{k-1}] \leq \Xi_{k-1} \quad \mathbb{P}\text{-a.s.}$$

By Doob's martingale convergence theorem (Williams, 1991),

$$\exists \Xi_\infty \geq 0 \text{ s.t. } \Xi_k \overset{\text{a.s.}}{\to} \Xi_\infty.$$

Suppose $\Xi_\infty > 0$. Then $\exists \epsilon > 0$ and infinitely many k where $\Xi_k > \epsilon$. But by Aumann et al. (1995), common knowledge of payoffs implies

$$\bigcap_{k=1}^{\infty} (\mathcal{F}_k^A \vee \mathcal{F}_k^B) \subseteq \mathcal{F}_\infty^A \cap \mathcal{F}_\infty^B,$$

contradicting $\Xi_\infty > 0$. Hence $\Xi_\infty = 0$ a.s. Pareto optimality follows from the First Welfare Theorem under complete information (Debreu, 1959), achieved through limit beliefs $\sigma(\lim_{k\to\infty} \mathcal{F}_k^A \vee \mathcal{F}_k^B) = \mathcal{F}_\infty$.

Proof of Proposition 2. Under the GPD regularity conditions in Smith (1985) ($\xi < 1/2$), the Fisher information matrix for sample size n is

$$I_n(\theta) = n \begin{pmatrix} \frac{1}{(1-\xi)^2(1-2\xi)} & \frac{-1}{\sigma(1-\xi)} \\ \frac{-1}{\sigma(1-\xi)} & \frac{2}{\sigma^2} \end{pmatrix}.$$

For nested information sets $\mathcal{X}_B \subset \mathcal{X}_A$, the Loewner ordering follows from the monotonicity of Fisher information (Le Cam, 2012),

$$I_A(\theta) = I_{\mathcal{X}_A}(\theta) \succ I_{\mathcal{X}_B}(\theta) = I_B(\theta).$$

The Cramér–Rao bound then gives

$$\mathbb{E}\|\hat{\theta}_A - \theta\|_2^2 \geq \text{tr}(I_A^{-1}(\theta)) < \text{tr}(I_B^{-1}(\theta)) \leq \mathbb{E}\|\hat{\theta}_B - \theta\|_2^2.$$

Proof of Theorem 2. By the Martingale Convergence Theorem (Williams, 1991), since $\{\mathcal{F}_k\}$ is a filtration,

$$\hat{\theta}_k = \mathbb{E}[\theta \mid \mathcal{F}_k] \overset{a.s.}{\to} \mathbb{E}[\theta \mid \mathcal{F}_\infty] = \theta.$$

For L^2 convergence, the Fisher information growth implies

$$\mathbb{E}\|\hat{\theta}_k - \theta\|_2^2 = \text{tr}(\text{Cov}(\hat{\theta}_k)) \leq \text{tr}(I_{\mathcal{F}_k}^{-1}(\theta)) \to 0.$$

Proof of Proposition 3. Define the parameter vector $\theta_k = (\mu_k, \log\sigma_k, \text{logit}(\xi_k))$ and strategic operator $\Phi(\theta) = \theta + \phi(\theta)$. Under the Fréchet derivative condition $\|D\phi\|_{\text{op}} \leq \kappa < 1$, Banach's fixed point theorem guarantees

$$\exists! \theta^* \in \Theta \quad \text{s.t.} \quad \lim_{k\to\infty} \|\theta_k - \theta^*\| = 0.$$

Geometric convergence follows from

$$\|\theta_{k+1} - \theta^*\| \leq \kappa \|\theta_k - \theta^*\|.$$

Data availability

Data will be made available on request.

References

Acemoglu, D., Bimpikis, K., & Ozdaglar, A. (2014). Dynamics of information exchange in endogenous social networks. *Theoretical Economics*, *9*(1), 41–97.
Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies*, *1*(1), 3–40.
Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*(3), 488–500.
Ardakani, O. M. (2022). Option pricing with maximum entropy densities: The inclusion of higher-order moments. *Journal of Futures Markets*, *42*(10), 1821–1836.
Ardakani, O. M. (2023a). Capturing information in extreme events. *Economics Letters*, *231*, Article 111301.
Ardakani, O. M. (2023b). Coherent measure of portfolio risk. *Finance Research Letters*, *57*, Article 104222.
Ardakani, O. M., Asadi, M., Ebrahimi, N., & Soofi, E. S. (2020). MR plot: A big data tool for distinguishing distributions. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *13*(4), 405–418.
Ardakani, O. M., Ebrahimi, N., & Soofi, E. S. (2018). Ranking forecasts by stochastic error distance, information and reliability measures. *International Statistical Review*, *86*(3), 442–468.
Aumann, R. J. (1976). An elementary proof that integration preserves uppersemicontinuity. *Journal of Mathematical Economics*, *3*(1), 15–18.
Aumann, R. J. (1995). Backward induction and common knowledge of rationality. *Games and Economic Behavior*, *8*(1), 6–19.
Aumann, R., & Brandenburger, A. (1995). Epistemic conditions for Nash equilibrium. *Econometrica*, *63*(5), 1161–1180.
Aumann, R. J., Maschler, M., & Stearns, R. E. (1995). *Repeated games with incomplete information*. MIT Press.
Azar, J., & Vives, X. (2021). General equilibrium oligopoly and ownership structure. *Econometrica*, *89*(3), 999–1048.
Bandi, F. M., & Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*, *79*(3), 655–692.
Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307–343.
Bates, D. S. (1991). The crash of 87: was it expected? The evidence from options markets. *The Journal of Finance*, *46*(3), 1009–1044.
Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, *119*(3), 861–898.
Cameron, A. C., & Trivedi, P. K. (1998). *Regression analysis of count data*. Cambridge University Press.
Comment, R., & Schwert, G. W. (1995). Poison or placebo? Evidence on the deterrence and wealth effects of modern antitakeover measures. *Journal of Financial Economics*, *39*(1), 3–43.
Cover, T. M., & Thomas, J. A. (1991). *Information theory and statistics: elements of information theory*. Wiley.
Crawford, V. P., & Iriberri, N. (2007). Fatal attraction: Salience, naivete, and sophistication in experimental "hide-and-seek" games. *American Economic Review*, *97*(5), 1731–1750.
Csiszár, I. (1975). I-divergence geometry of probability distributions and minimization problems. *The Annals of Probability*, *3*(1), 146–158.
Das, K., Klein, N., & Schmid, K. (2019). Strategic experimentation with asymmetric players. *Economic Theory*, *69*(4), 1147–1175.
DeAngelo, H., & Rice, E. M. (1983). Antitakeover charter amendments and stockholder wealth. *Journal of Financial Economics*, *11*(1–4), 329–359.
Debreu, G. (1959). *Theory of value: An axiomatic analysis of economic equilibrium*: *vol. 17*, Yale University Press.
Erev, I., & Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, *88*(4), 848–881.
Fudenberg, D., & Levine, D. K. (1993). Self-confirming equilibrium. *Econometrica*, *61*(3), 523–545.
Gorton, G., & Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics*, *104*(3), 425–451.
Gromb, D., & Vayanos, D. (2002). Equilibrium and welfare in markets with financially constrained arbitrageurs. *Journal of Financial Economics*, *66*(2–3), 361–407.
Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, *70*(3), 393–408.
Hansen, P. R., & Lunde, A. (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics*, *24*(2), 127–161.
Harsanyi, J. C. (1967). Games with incomplete information played by "Bayesian" players, I-III. part I. The basic model. *Management Science*, *14*(3), 159–182.
Healy, P. M., Palepu, K. G., & Ruback, R. S. (1992). Does corporate performance improve after mergers? *Journal of Financial Economics*, *31*(2), 135–175.
Hermalin, B. E., & Weisbach, M. S. (2012). Information disclosure and corporate governance. *Journal of Finance*, *67*(1), 195–233.
Heron, R. A., & Lie, E. (2006). On the use of poison pills and defensive payouts by takeover targets. *The Journal of Business*, *79*(4), 1783–1807.
Hirshleifer, D., & Teoh, S. H. (2009). The psychological attraction approach to accounting and disclosure policy. *Contemporary Accounting Research*, *26*(4), 1067–1090.
Ivashina, V. (2009). Asymmetric information effects on loan spreads. *Journal of Financial Economics*, *92*(2), 300–319.
Jackson, M. O., Rogers, B. W., & Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, *55*(1), 49–95.
Jarrell, G. A., & Poulsen, A. B. (1989). The returns to acquiring firms in tender offers: Evidence from three decades. *Financial Management*, *18*(3), 12–19.
Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review*, *106*(4), 620–630.
Jensen, M. C., & Ruback, R. S. (1983). The market for corporate control: The scientific evidence. *Journal of Financial Economics*, *11*(1–4), 5–50.
Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *Journal of Finance*, *72*(3), 967–998.
Kreps, D. M., & Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, *27*(2), 253–279.
Le Cam, L. (2012). *Asymptotic methods in statistical decision theory*. Springer.
Lehmann, E. L., & Casella, G. (2006). *Theory of point estimation*. Springer.
Li, K. (2020). Does information asymmetry impede market efficiency? Evidence from analyst coverage. *Journal of Banking & Finance*, *118*, Article 105856.
MacKenzie, D. (2003). Long-term capital management and the sociology of arbitrage. *Economy and Society*, *32*(3), 349–380.
Milgrom, P., & Roberts, J. (1982). Predation, reputation, and entry deterrence. *Journal of Economic Theory*, *27*(2), 280–312.
Milgrom, P., & Roberts, J. (1986). Relying on the information of interested parties. *Rand Journal of Economics*, *17*(1), 18–32.
Myerson, R. B. (1979). Incentive compatibility and the bargaining problem. *Econometrica: Journal of the Econometric Society*, *47*(1), 61–73.
Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, *85*(5), 1313–1326.
Nash, J. (1953). Two-person cooperative games. *Econometrica*, *21*(1), 128–140.
Peres, Y., Rácz, M. Z., Sly, A., & Stuhl, I. (2020). How fragile are information cascades? *Annals of Applied Probability*, *30*(6), 2796–2814.
Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings of the National Academy of Sciences*, *42*(1), 43–47.
Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *The American Economic Review*, *80*(3), 465–479.
Securities, U., & Commission, E. (2021). *Staff report on equity and options market structure conditions in early 2021*: *Report*, Washington, D.C.: U.S. Securities and Exchange Commission.
Servaes, H., & Zenner, M. (1996). The role of investment banks in acquisitions. *The Review of Financial Studies*, *9*(3), 787–815.
Shleifer, A., & Vishny, R. W. (1986). Large shareholders and corporate control. *Journal of Political Economy*, *94*(3, Part 1), 461–488.
Silverman, B. W. (1986). The kernel method for univariate data. In *Density estimation for statistics and data analysis* (pp. 34–74). Springer.
Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, *50*(3), 665–690.
Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, *72*(1), 67–90.
Soofi, E. S. (1994). Capturing the intangible concept of information. *Journal of the American Statistical Association*, *89*, 1243–1254.
Soofi, E. S. (2000). Principal information theoretic approaches. *Journal of the American Statistical Association*, *95*, 1349–1353.
Soofi, E. S., & Retzer, J. J. (2002). Information indices: unification and applications. *Journal of Econometrics*, *107*(1–2), 17–40.
Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*(3), 355–374.
Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, *10*(1), 218–254.
Stiglitz, J. E. (2000). The contributions of the economics of information to twentieth century economics. *The Quarterly Journal of Economics*, *115*(4), 1441–1478.
Subramanian, G. (2003). Bargaining in the shadow of takeover defenses. *Yale LJ*, *113*, 621.
Whaley, R. E. (2014). *Trading volatility: at what cost*. SSRN.
Williams, D. (1991). *Probability with martingales*. Cambridge University Press.