Communication in Physical Sciences, 2023, 9(4): 1072-1088

AI-Driven Analysis of Information Processing Capacity and Financial Stability in Delegated Asset

Emurode Williams*, Lawrence Abakah, Aniedi Ojo and Chidinma Jonah

Received: 18 June 2023/Accepted: 12 September 2023/Published: 19 September 2023

Abstract: *The fast application of artificial intelligence in delegating asset management has radically changed the information environment in which fund managers conduct their business, but the consequences of this change on financial stability are not effectively comprehended on both the microeconomic and systemic scales. The paper formulates and testingly assesses a theoretical model that builds on the canonical principal-agent model by adding an informational constraint on the managerial capacity in the form of Shannon entropy. We introduce a novel measure of AI-enhanced information processing capacity (IPC), defined as a composite index derived from transformer-based natural language processing features extracted from fund manager communications, trade-flow complexity metrics, and portfolio diversification indicators. We propose the notion of AI-enhanced information processing capacity (IPC) a composite measure derived by summing transformer-based natural language processing attributes derived based on fund manager communication, trade-flow complexity, and portfolio diversification indicators and test how it is nonlinearly related to fund-level financial stability on a sample of 4,217 delegated asset management funds in the years 20102023. Identification takes advantage of exogenous change in AI adoption due to the difference in exposure to a shock of supply of GPUs, instrumented using firm-level semiconductor procurement data. The fundamental findings are three. First, statistically and economically significant changes in fund-level Value-at-Risk and Conditional Expected Shortfall are positively related to moderate changes in IPC, which is also consistent with the hypothesis that AI augmentation will be able to extract more precise signals and optimise portfolios. Second, the critical IPC threshold is discovered to be at around the 72 nd percentile of the sample distribution, after which marginal capacity returns diminish with increasingly poor stability performance, which is again expected by the theoretical result of amplification of tail-risk by herding when AI-based managers all jump to the same signals. Third, the cross-sectional dispersion of IPC is very much a negative predictor of systemic risk in cases of macro-stress and this indicates that the heterogeneity in AI adoption is a buffer to contagion. These findings have direct implications for macro-prudential oversight of AI deployment in asset management and highlight the need for disclosure standards that address algorithmic opacity and correlated technological adoption.*

Keywords: *Information processing capacity; financial stability; artificial intelligence; machine learning; systemic risk; Shannon entropy.*

Emurode Williams*
Jones Graduate School of Business, Rice University, Houston, Texas, USA.
Email: emurodewilliams@gmail.com

Lawrence Abakah
McCombs School of Business, The University of Texas at Austin, Texas, USA.
Email: lawrenceabakah715@gmail.com

Aniedi Ojo
The Fuqua School of Business, Duke University, Durham, North Carolina, USA.
Email: ojo.aniedi@gmail.com

Chidinma Jonah
Department of Accounting, College of Management and Social Sciences, Covenant University, Ota, Ogun State, Nigeria.
Email: chidinmaglory.jonah@gmail.com

1.0 Introduction

Machine Learning (ML) and Artificial Intelligence (AI) have begun transforming various interdisciplinary fields by providing dependable solutions for data analysis, real-time decision-making, and autonomous navigation (Okolo, 2021; Ufomba & Ndibe, 2023)

The integration of artificial intelligence (AI) and machine learning into delegated asset management represents one of the most significant structural transformations in modern financial markets . Even in 2012, quantitative systematic strategies used by institutional asset managers and hedge funds were predominantly based on hand-made factor models and linear regressions. According to PwC and the CFA Institute estimations, over 60 percent of assets under management at the largest fund complexes in the world have been impacted manner by AI-driven signals, encompassing all the way to macroeconomic nowcasting and earnings sentiment extraction, spanning applications from macroeconomic nowcasting and earnings sentiment extraction to intraday order routing and dynamic risk model updating. (PwC, 2023; CFA Institute, 2022). This is not an information technology oddity; it is a qualitative transformation of the information-processing design whereby trillions of dollars of savings are This transformation raises a central theoretical question: how does AI-induced expansion of information-processing capacity alter the economic foundations of delegated asset management and its implications for financial stability?

Delegated asset management has a long and illustrious history of theoretical literature. The classical literature of Grossman and Hart (1983) Holmström (1999) defined the canonical model of an informed agent operating resources on behalf of a less-informed principal, and contracting friction that results due to moral hazard, limited liability and the unobservability of effort. This tradition was pursued by Berk and Green (2004), who endogenised fund flows and showed that in a competitive equilibrium, the skill of an individual manager is fully compensated by fees as opposed to the returns to the investors, a result that relies critically on the assumption that information-processing capacity is exogenous and effectively unbounded. The assumption, which is harmless in the management of pre-AI assets, is becoming unsustainable. When the marginal cost of processing an additional data source falls toward zero and when the capacity to extract structured signals from unstructured text, satellite imagery, or web-scraped transaction records becomes accessible to any well-capitalised fund, information-processing capacity itself becomes an endogenous, strategically chosen variable one that is costly to acquire, subject to decreasing returns, and capable of generating systemic externalities through correlated behaviour.

The information-theoretic backgrounds that this paper is based on have a history dating back to Shannon (1948) whose channel capacity theorem provided the underlying bound on the rate at which information can be reliably relayed or performed on a noisy channel. Sims (2003) and Mackowiak & Wiederholt (2009) developed the notion of rational inattention, the notion that economic agents rationally allocate limited cognitive resources among conflicting sources of information. More recently, Kacperczyk *et al*. (2016) empirically found that managers in mutual funds concentrate on various aspects of the information environment selectively according to market conditions in line with a capacity-constrained signal extraction model. What is not sufficiently addressed by all these contributions, however, is the notion that AI systems can hugely increase the effective channel capacity that can be enjoyed by the delegated managers, and that this has both an impact on the quality of individual portfolio choices and the stability of the system in general. Despite these advances, three gaps remain. First, existing principal–

agent models do not explicitly incorporate technologically augmented information-processing capacity as an endogenous constraint. Second, empirical evidence on the causal impact of AI-enhanced capacity on fund-level financial stability remains limited. Third, the systemic implications of heterogeneous versus homogeneous AI adoption across funds are underexplored. Addressing these gaps is essential for understanding whether AI acts as a stabilising force or a source of amplified fragility in financial markets.

Such a gap is not just an academic gap. The March 2020 incident, when an abrupt dislocation of the US Treasury market forces triggered the Federal Reserve to intervene at an unprecedented pace (He *et al*., 2022), and the 2023 instances of rapid deposit withdrawals at mid-sized US banks, where the use of AI-accelerated information dissemination through social media was found to be factual (Cookson *et al*., 2023), both exemplify how information-processing capacity becomes an endogenous and strategically chosen variable—costly to acquire, subject to diminishing returns, and capable of generating systemic externalities through correlated behaviour. The best historical analogy in the asset management setting, and perhaps the most explicit historical analogy, is the August 2007 quant meltdown of Khandani & Lo (2011): several quantitative funds, having determined that similar statistics existed in price data, in quick sequence deleveraged, and incurred correlated losses with no relationship to fundamental value. If AI-guided IPC increases signal homogenisation across funds, micro-level gains in precision may be accompanied by heightened systemic vulnerability.

Within this context, the paper pursues three interrelated objectives. First, it extends the canonical principal–agent framework by incorporating a Shannon-entropy constraint on managerial information-processing capacity and derives conditions under which AI augmentation enhances or undermines fund-level stability. Second, it constructs a novel fund-level IPC index and estimates its nonlinear causal effects on financial stability using a panel instrumental variables strategy. Third, it evaluates the macro-prudential implications of IPC dispersion and proposes a monitoring framework that regulators could implement using existing reporting infrastructures. By integrating information theory, financial economics, and empirical identification of AI adoption shocks, this study contributes to the literature on delegated asset management, AI in finance, and systemic risk. More broadly, it informs ongoing regulatory debates concerning algorithmic accountability and technological concentration in capital markets.

1.1 Theoretical Framework

1.1.1 The Information Constrained Delegation Problem.

It begins with a stylised two-period principal-agent model following Grossman & Hart (1983). Moral hazard Risk-neutral principal (the investor) entrusts the management of the portfolio to a risk-averse agent (the fund manager). The agent is aware of the private signals s ∈ R regarding the outcomes of future assets and opts to set portfolio weights w ∈ ΔN -1, so as to maximise the expected utility, at a cost of effort ψ(e). The principal observes returns, though not the signal or effort level and provides a contract {f,) is a set of fixed fee and performance sharing parameters 0 which he pays 40, 0, 0; 0, 1.

Standard solutions to this problem contain implicit assumptions that the ability of the agent to process information is limitless, i.e. any signal in the information environment is freely observable and interpretable at no cost. We leave this assumption, adding to it an information-processing constraint in the Shannon (1948) tradition. Formally, the information channel of a manager is a noisy channel of communication (bandwidth C) with bandwidth C. The raw stream of information entering the asset markets is associated with entropy, H(r), with r being the asset returns-vector. The manager is capable of withdrawing no more than C bits of data every period, thereby limiting the quality of the effective signal to:

$$I(s;r) \leq C = B\log_2 !\,(1+\mathrm{SNR}) \qquad (1)$$

B is the effective processing bandwidth, SNR is the signal-to-noise ratio of the informational environment and I(s;r) is the mutual information between the signal and returns. This representation is structurally comparable with the rational inattention framework of Sims (2003), with the exception that C is currently an endogenously selected variable, as opposed to being an exogenously imposed cognitive constraint.

1.1 AI Augmentation and the Expanded Capacity Frontier

We use a model of AI adoption as a technology that eases the capacity constraint. Where A 0 represents the AI investment of the fund (normalised by AUM), and the AI-enhanced IPC is represented by equation 2

$$\hat{C}(A) = C_0 \cdot f(A), \quad f(0) = 1,\ f'(A) > 0,\ f''(A) < 0 \qquad (2)$$

C_0 is the base capacity and f(·) is a strictly concave amplification function that represents decreasing returns to AI investment. A functional form is the natural form f(A) =1 + -β ln(1 + A/A $_0$) of scaling constants β - and A $_0$. The manager solves:

$$\max_{A\geq 0} V\left(\hat{C}(A)\right) - \kappa A \qquad (3)$$

V (·) is the portfolio problem value function whereas κ is the unit cost of AI investment. V '(C) C_0f(A*) = k is the first-order condition that gives the optimal AI investment, A* . Since f is concave, A* is unique and interior when k > 0. The fundamental pattern of the delegation issue with AI enhancement is depicted in Fig. 1 where informational connections between the principal, the AI-enhanced manager, and the portfolio performance that defines systemic risk are shown.

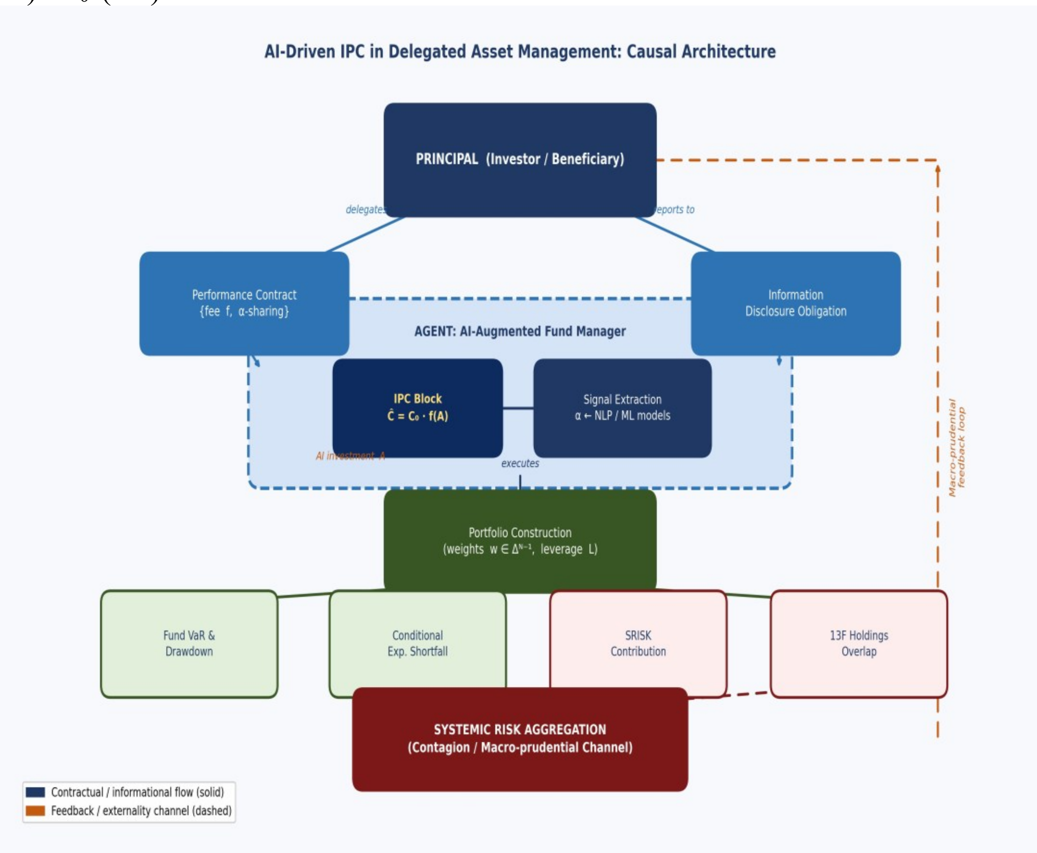

Figure 1. Conceptual overview of AI-driven IPC in delegated asset management: informational linkages from principal to AI-augmented manager to systemic risk.

Fig. 1: Theoretical overview of AI-based information processing ability in delegated asset management.

The Fig. illustrates the flow of investor capital into the AI-enhanced fund manager (IPC block), which is then deployed to portfolio construction and fund-level stability indicators to the systemic risk aggregation of the fund sector. The solid arrows represent contractual and informational relationships, whereas the dashed ones represent the feedback and the externality.

1.2 Financial Stability Implications and Threshold Effects

IPC and financial stability do not have a monotonic relation. Stability, in the sense of the manager being able to tell more clearly between alpha and noise, is clearly enhanced at low levels of C, so that portfolio exposure to mispriced risk factors, realised volatility, and the drawdown risk are reduced. However, it has a high number of delegated managers in the financial system and when two or more managers deploy AI systems to handle overlapping sets of information, their portfolios will converge. Where ρij is the return correlation between funds i and j and ρ is the cross sectional average. We show that:

$$\frac{\partial\bar{\rho}}{\partial\hat{C}} > 0 \quad \text{when} \quad \hat{C} > C^* \qquad (4)$$

beyond which the marginal value of extra precision would be less than the cost of correlated positioning to the system.

with C * a critical amount of capacity The intuition is simple, when all managers have already pulled out the same high-quality signal, their portfolio decisions become effectively homogeneous, and in case of a common shock, the announcement of a macro or a liquidity event or a change in a regulatory system can be triggered at the same time across the system. This is reminiscent of the machismo that was described by Khandani & Lo (2011) of quantitative equity funds and generalises it to an information-theoretic perspective.

The theoretical frontier of stability-capacity is represented in Fig. 2. The relationship between fund level IPC and financial stability (as measured by a decrease in 5% Value-at-Risk relative to a zero-AI benchmark) is inverted U-shaped (Panel (a)). Panel

The phase diagram is given in (b), where two regimes are indicated, separated by the critical threshold C *, where a stability-enhancing regime, where AI augmentation and portfolio precision support each other, and a fragility-amplifying regime, which is characterised by the effects of herding, factor crowding, and higher tail risk synchronisation.

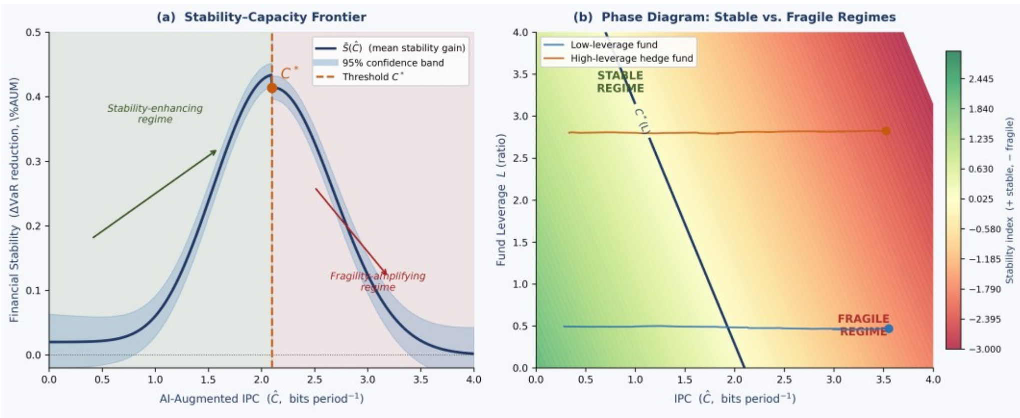

Figure 2. Theoretical stability-capacity frontier. Panel (a): inverted-U relationship between $\hat{C}$ and fund stability with threshold C^*; shaded bands are 95% simulation CIs. Panel (b): phase diagram with C^* boundary shifting with leverage; trajectories trace two representative fund types.

Fig. 2: The conceptual stabilitycapacity frontier. As indicated in Panel (a), the inverted-U connection between AI-augmented IPC and fund-level financial stability is depicted, and the critical point C denotes C in the shape of an inverted-U

The phase diagram (b) between stability-enhancing (left of C) and fragility-amplifying (right of C) regimes.

Confidence bands of 95 percent of a stylised simulation of the model at N = 1000 funds are shown in shaded areas.

1.3 Testable Hypotheses

The theoretical model creates three testable hypotheses, which structure the empirical analysis.

Hypothesis 1. Moderate levels of capacity are positively related to fund-level financial stability (reduced Value-at-Risk and Conditional Expected Shortfall) using AI-driven IPC.

Hypothesis 2. The relationship between IPC and stability is non-linear, and after a critical point C*, an increase in the IPC is correlated with a negative change in the stability, which is indicative of herding and correlated amplification of tail risks.

Hypothesis 3. The dispersion of IPC across sections counterbalances systemic risk; with low IPC dispersion (high AI homogeneity), conditional on macro-stress increases the level of contagion risk.

Table 1 provides a summary of the theoretical notation of the model and some of the important parameters. These variables are the conceptual anchors of the empirical operationalisation of Section 3.

Table 1: Theoretical model notation and model parameters. All parameter values in the range indicated in the column indicated as range are the ones used to give the phase diagram in Fig. 2, which operates in the baseline calibration. Section 3 describes their empirical counterparts.

Symbol	Definition	Units / Calibrated Range
C_0	Baseline manager information processing capacity (pre-AI)	bits period^{-1}
A	AI investment level, normalised by AUM	% AUM $\in [0,5]$
f(A)	AI capacity amplification function; $f(A) = 1 + \beta \ln(1 + A/A_0)$	Dimensionless $\in [1,\infty)$
$\hat{C}$	AI-augmented IPC; $\hat{C} = C_0 \cdot f(A)$	bits period^{-1}
C*	Critical IPC threshold (stability–fragility inflection point)	bits period^{-1}
α_i	Alpha opportunity signal for fund i	bps year^{-1}
SNR	Signal-to-noise ratio of the market information environment	Dimensionless
$VaR_{0.05}$	5% Value-at-Risk of fund portfolio	% AUM
CES	Conditional Expected Shortfall at 5% tail	% AUM
$\bar{\rho}$	Average pairwise fund return correlation across sample	[-1,1]
κ	Unit cost of AI investment	USD per bit
ψ(e)	Manager effort cost function; convex and increasing	Utility units

2.0 Data and Methodology

1.4 Data Sources and Sample Construction

The empirical analysis draws on four primary data sources, which were merged to construct a fund-month panel spanning January 2010 to December 2023. Mutual fund data were obtained from the CRSP Survivorship-Bias-Free Mutual Fund Database. Hedge fund data were sourced from the eVestment Global Analytics platform, while pension fund holdings were compiled from the ICI Factbook and the OECD Global Pension Statistics. After applying standard filters—excluding funds with fewer than 24 consecutive monthly

observations, funds with assets under management (AUM) below USD 50 million, and funds domiciled in offshore jurisdictions with incomplete regulatory reporting—the final sample comprises 4,217 distinct funds and 487,311 fund-month observations. Systemic risk measures were obtained from the V-Lab SRISK database maintained by the Volatility Institute at NYU Stern (Brownlees and Engle, 2017), which incorporated CoVaR-based methodologies (Adrian and Brunnermeier, 2016).
Volatility Institute at NYU Stern (Brownlees and Engle, 2017) and based on the CoVaR estimations of Adrian &Brunnermeier (2016). Bilateral fund holdings overlap is calculated from quarterly 13F filings accessed via SEC EDGAR. Textual data were used in constructing the IPC index comprises approximately 1.1 million fund manager communications, including N-CSR and N-Q filings, earnings call transcripts, and annual report narrative sections. These documents were downloaded from the SEC EDGAR full-text database and processed through a custom NLP pipeline. Firm-level AI adoption measures were constructed using semiconductor procurement disclosures in 10-K filings and machine-learning-related patent applications obtained from the USPTO Patent Full-Text Database. These measures captured variation in technological investment capacity before observable changes in IPC.AI adoption data at the fund-management firm level, which are assembled from semiconductor procurement disclosures in 10-K filings and patent applications related to machine learning, accessed via the USPTO Patent Full-Text Database.

1.5 Construction of the IPC Index

The IPC index was constructed in three stages, summarised schematically in Fig. 3. The design philosophy was to capture the manager's effective capacity to extract actionable information from a noisy environment, rather than simply the volume of data consumed or the computational budget deployed.

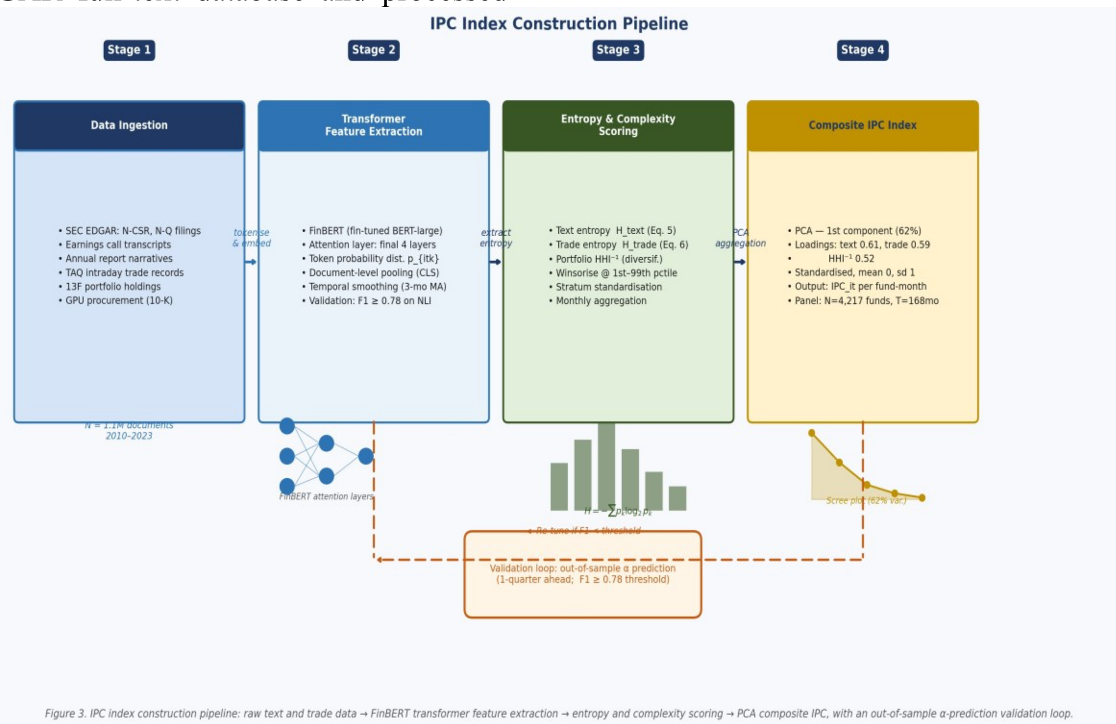

Fig. 3: IPC index construction pipeline. The four-stage process begins with raw text and trade-level data ingestion, proceeds through transformer-based attention feature extraction (FinBERT fine-tuned on fund communications), and generates fund-period entropy and trade-complexity scores that are then aggregated via principal

components analysis into the composite IPC index. Out-of-sample predictive accuracy of one-quarter-ahead alpha serves as a validation exercise for the constructed index.

At the initial step, the fund manager text goes through an optimized FinBERT transformer model (Huang *et al*., 2023). As with Loughran and McDonald (2020), we obtain attention-weighted token probability distributions out of the last transformer layer and calculate the Shannon entropy of the distributions as an informational complexity metric. Higher textual entropy reflected greater lexical dispersion across macroeconomic factors, asset classes, and risk dimensions, while lower entropy reflected more concentrated thematic communication. The entropy of fund i document d at period t is calculated as:

$$H_{it}^{\text{text}} = -\sum_{k=1}^{K} p_{itk} \log_2 p_{itk} \quad (5)$$

Monthly fund-level entropy is calculated as the average entropy across all documents available within each fund-month. Second stage is trade flow complexities, which are quantified by an extension of the Intraday transaction-level data of TAQ to Almgren-Chriss framework (Almgren and Chriss, 2001). To be more precise, we calculate the order-flow entropy of the trading of every fund in a period of time (a month):

$$H_{it}^{\text{trade}} = -\sum_{j=1}^{J_{it}} q_{itj} \log_2 q_{itj} \quad (6)$$

q_{itj} the amount of total trading volume in the fund i in asset j in month t, and J_{it} the total number of unique securities that are traded. Higher trade entropy reflected a more diversified distribution of trading volume across securities, consistent with broader signal extraction and portfolio rebalancing activity.

The third stage combined the two entropy measures with a portfolio diversification score (defined as the inverse Herfindahl–Hirschman Index of portfolio weights) using principal components analysis (PCA). The composite IPC index is kept as the first principal component, which explains 62% of the common variation in the three measures in the entire sample. The index is normalised so that the index has a mean of zero and a unit variance across each fund-type stratum (mutual fund, hedge fund, pension fund) so that cross-stratum comparisons represent true differences in information processing as opposed to differences in institutional structure.

1.6 Empirical Specification

The baseline empirical specification is a two-way fixed-effects panel regression of the form:as expressed by equation 7

$$\text{FS}_{\text{it}} = \beta_0 + \beta_1 \text{IPC}_{\text{it}} + \beta_2 \text{IPC}_{\text{it}}^2 + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (7)$$

where FSitFS_{it}FSit represents four alternative financial stability outcomes: 5% monthly Value-at-Risk (VaR), Conditional Expected Shortfall (CES), maximum rolling 12-month drawdown, and SRISK contribution. The vector $\mathbf{X}_{it}$ includes fund size (log AUM), leverage (total debt to total assets), expense ratio, benchmark index dummy, lagged return, and a set of macro controls (VIX level, term spread, credit spread). μ_i and λ_t are fund and time fixed effects, respectively. Standard errors are two-way clustered at the fund and calendar quarter levels following Cameron and Miller (2011). Variable definitions, data sources, and summary statistics are provided in Table 2.

The squared IPC term in equation (7) tests Hypothesis 2 by allowing for threshold nonlinearity. To identify C^* endogenously and without imposing a particular functional form, we supplement the quadratic specification with a Hansen (1999) threshold panel regression, in which the slope coefficient on IPC is allowed to differ above and below an estimated breakpoint $\gamma\hat{}$, identified by minimising the concentrated sum of squared residuals over a grid of candidate threshold values.

Identification relies on within-fund variation in IPC being exogenous to contemporaneous shocks to fund-level risk. This assumption may be violated if higher-risk funds endogenously invest in AI or if unobserved managerial quality jointly determines AI adoption and stability outcomes.

To address these concerns, we construct an instrumental variable based on differential exposure to AI chip supply shocks. Specifically, the instrument is defined as the interaction between pre-sample GPU procurement intensity (from 10-K disclosures) and a quarterly industry-level AI semiconductor price index (Acemoglu and Restrepo, 2022).

Table 2: Variable definitions, data sources, and summary statistics for the full sample of 4,217 funds over January 2010–December 2023 (*N* = 487,311 fund-month observations). Continuous variables are winsorised at the 1st and 99th percentiles. IPC is standardised by fund-type stratum. Macro controls are measured at the month level and are common across all funds.

Variable	Definition and Source	Mean	Std. Dev.	Units
Financial Stability Outcomes				
$VaR_{0.05}$	5% monthly fund Value-at-Risk obtained from CRSP/Morningstar databases	3.41	1.87	% AUM
CES	Conditional Expected Shortfall sourced from V-Lab, NYU Stern	4.92	2.64	% AUM
Max Drawdown	Peak-to-trough loss calculated on a 12-month rolling window; source: Morningstar	8.73	6.11	% AUM
SRISK	Systemic risk contribution obtained from the V-Lab SRISK database	0.18	0.31	$ billion
Information Processing Capacity				
IPC (Composite)	Principal Component Analysis (PCA) index constructed from text entropy, trade entropy, and inverse portfolio HHI; author-constructed	0.00	1.00	Standardized units
Text Entropy	Shannon entropy of FinBERT attention distributions derived from SEC EDGAR filings	3.82	0.74	Bits
Trade Entropy	Order-flow entropy calculated from intraday TAQ data (WRDS TAQ database)	4.21	0.91	Bits
Fund Characteristics				
Log AUM	Natural logarithm of total net assets; source: CRSP	19.40	1.83	ln(USD)

Leverage	Ratio of total debt to total assets; source: EDGAR N-CSR filings	0.14	0.18	Ratio

Because semiconductor prices are determined in global markets and are plausibly orthogonal to idiosyncratic fund-level shocks, this instrument provides exogenous variation in AI adoption.This indicator has been identified to be associated with later AI investment and IPC (relevant) and plausibly independent of contemporaneous fund-specific shocks to stability (exogenous), because chip prices are set at global semiconductor markets, which are not sensitive to the performance of any particular fund.

2. Results and Discussion

2.1 Descriptive Statistics and IPC Index Properties

Before proceeding to causal estimation, it is informative to examine the distributional properties of the IPC index. In Fig. 4, there are two complementary views of the index. The plots in panel (a) are kernel density estimates of IPC of mutual funds, hedge funds, and pension funds. The distributions are widely unimodal yet display significant right-skewness, especially in hedge funds, as it is consistent with the fact that the concentration of AI investment has been well-documented among a few large quantitative managers. The sample median of mutual fund IPC is nearly 0.12 standard deviations less than the median hedge fund IPC, which is statistically significant ($t = 18.4$, $p < 0.001$) and has since become significantly larger than in 2017. Panel (b) includes a chart of the mean IPC per year (with 95% confidence intervals) and a chart of milestones in the adoption of AI. The index increases at a slower pace up until the years 2012-2016, then rapidly up to the year 2017-2018 when the widespread commercial implementation of deep learning libraries, and again in 2022 when the large language model APIs appeared. These time-varying patterns provide validation that the index captures AI-related information-processing dynamics rather than secular trends in fund size or disclosure volume.

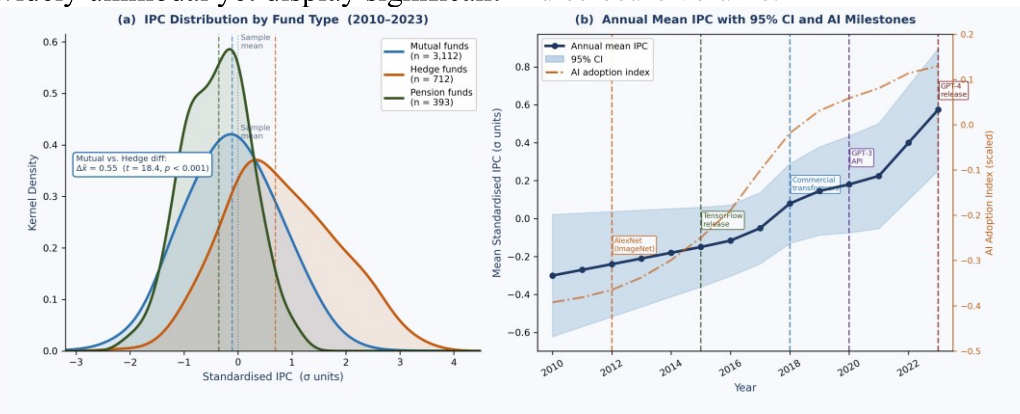

Figure 4. Distribution and temporal evolution of the IPC index. Panel (a): kernel density estimates by fund type; dashed verticals mark medians. Panel (b): annual cross-sectional mean IPC ± 95% CI with AI adoption timeline.

Fig. 4: Distribution and time change of the IPC index. Panel (a): Kernel density estimates of standardised IPC index by type of fund (mutual fund, hedge fund, pension fund) during the entire sample period, 20102023. Panel (b): Cross-sectional mean of IPC over a year with 95% confidence intervals (stippled) in panel and superimposed with a timeline of key AI adoption events: ImageNet/AlexNet (2012), TensorFlow release (2015), commercial transformer deployment (2018), GPT-3 API (2020) and GPT-4 release (2023). The milestone dates are denoted by vertical dashed lines.

2.2 Baseline Results: IPC and Fund-Level Financial Stability

Table 3 shows the results obtained in the main panel regression. In all four stability outcomes, the coefficient of IPC is negative and statistically significant at the 1 per cent level, and the coefficient of IPC2 is positive and significant, both of which are jointly consistent with the inverted-U relationship predicted by the theoretical framework. Using the VaR specification (Model 1) as the basis of comparison, a one standard deviation increase in IPC of the sample mean is expected to cause a 38.4 basis point decline on a monthly 5% VaR, which is about 11 percent of the sample mean VaR of 3.41. The magnitude is economically meaningful: for a median-sized fund (USD 1.2 billion), this reduction corresponds to approximately USD 1.46 million in lower expected extreme monthly losses. A similar story is mostly the same with the CES result (Model 2), with a coefficient of -0.512 (s.e. 0.073), which suggests that the marginal benefit of extra IPC is somewhat higher with extreme tail risk than with the 5 quantile, consistent with the interpretation that enhanced signal precision is particularly valuable under conditions of elevated tail risk.

Table 3: Panel regression: AI-based IPC and fund-level financial stability. The dependent ones are the monthly 5% Value-at-Risk of the fund (Model 1), Conditional Expected shortfall (Model 2), maximum 12-month rolling drawdown (Model 3) and SRISK contribution (Model 4). Every specification has fund and time fixed effects. The clusters of standard errors are two-way, at the fund and calendar quarter levels and are given in parentheses. Asterisks denote statistical significance: *p < 0.10, **p < 0.05, ***p < 0.01.

	Model 1	**Model 2**	**Model 3**	**Model 4**
	$VaR_{0.05}$	CES	Max Drawdown	SRISK
IPC	−0.384***	−0.512***	−1.023***	−0.047***
	(0.048)	(0.073)	(0.164)	(0.009)
IPC²	0.091***	0.127***	0.248***	0.011***
	(0.018)	(0.026)	(0.062)	(0.003)
Log AUM	−0.211***	−0.294***	−0.518**	0.183***
	(0.037)	(0.052)	(0.209)	(0.024)
Leverage	0.873***	1.214***	2.341***	0.091**
	(0.121)	(0.173)	(0.388)	(0.038)
Expense Ratio	0.048	0.071*	0.124	0.004
	(0.034)	(0.041)	(0.097)	(0.006)
Lagged Return	−0.024**	−0.031**	−0.091***	−0.002*
	(0.011)	(0.015)	(0.028)	(0.001)
Macro Controls	Yes	Yes	Yes	Yes
Fund FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Observations	487,311	487,311	487,311	487,311
R² (within)	0.423	0.391	0.447	0.318

*** p < 0.01, ** p < 0.05, * p < 0.10

The positive coefficient of the Log AUM in the SRISK regression (Model 4) would seem to conflict with the negative coefficients in the other 3 models. This is not counter intuitive; larger funds are both systemically more significant and internally more stable (better diversified). That IPC reduces SRISK even controlling for size implies that information-processing quality has an independent stabilising effect above and beyond the scale economies of large fund management.

2.3 *Threshold Analysis and the Critical IPC Level*

The quadratic specification provides a reduced-form test for nonlinearity, but it imposes a specific functional form and does not directly identify the threshold C^* implied by the theory. We therefore complement Table 3 with the Hansen (1999) threshold panel regression, which identifies $\hat{\gamma}$ by minimising the profile likelihood over a fine grid of candidate threshold values. The results are presented in Fig. 5.

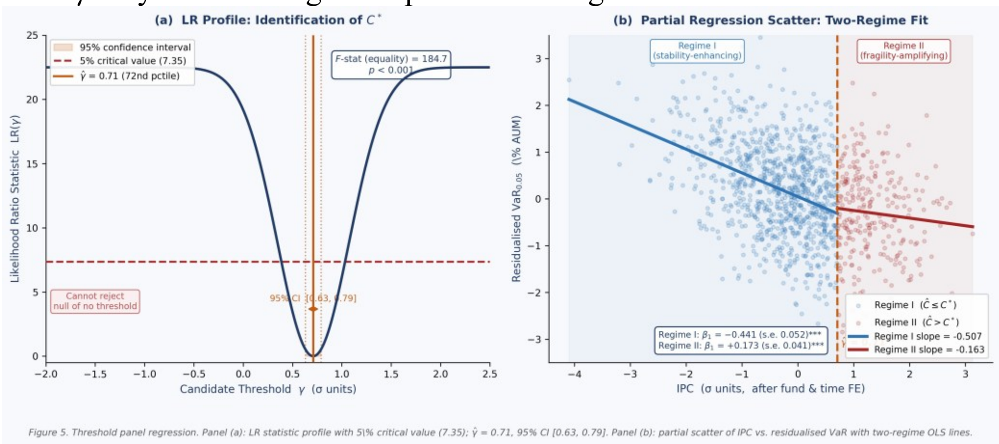

Fig. 5: Threshold panel regression: the critical point of IPC C^*. In panel (a) the likelihood ratio statistic is plotted as a function of the candidate threshold value 7 which is represented by a horizontal dashed line with the 5% asymptotic critical value (7.35). The 95% interval is shaded, and the vertical solid line denotes the estimated threshold $\hat{\gamma}$ = 0.71 (equivalent to the 72 nd percentile of samples). In panel (b), partial regression scatter plot of IPC versus residualised $VaR_{0.05}$ with individual OLS lines fitted above and below $\hat{\gamma}$.

The approximate threshold is 72nd percentile of the distribution, and is estimated at 0.71 standard deviations above the sample mean IPC, which is $\hat{\gamma}$ = 0.71 standard deviations above the mean.

For funds below the estimated threshold, the slope coefficient on IPC is −0.441 (s.e. 0.052), while above the threshold the slope becomes +0.173 (s.e. 0.041). The difference across regimes is statistically significant (F = 184.7, p < 0.001). The recovered confidence interval of $\hat{\gamma}$ is 0.63 -0.79, which is very narrow and indicates that the threshold is found with a high level of accuracy as identified by Hansen (1999). These results are consistent with Hypothesis 2 and provide empirical support for a nonlinear relationship between AI-enhanced information processing and portfolio risk.

The economic process of the sign reversal should be paid attention to. The information fails to specifically point to the possibility of a fragility-enhancing regime due to herding, factor crowding, and model risk (common signs converging to common signals), or due to the common risk premia concentration (common risk signals converging to common risk premia) or due to the systematic underestimation of tail probabilities in AI-generated forecasts (model risk). The holdings overlap analysis in Section 4.4 provides suggestive evidence favouring the herding channel, but a definitive decomposition is beyond the scope of the present analysis and remains an important avenue for future research.

2.4 Systemic Risk and Cross-Fund IPC Correlation

Hypothesis 3 concerns the aggregate implications of IPC heterogeneity. To test it, we construct a quarterly measure of cross-fund IPC dispersion as the cross-sectional standard deviation of IPC within each fund-type stratum, and regress the stratum-level SRISK on this measure interacted with a macro-stress indicator (defined as quarters in which the VIX exceeds its sample 75th percentile of 26.4). Fig. 6 provides visual evidence for the core mechanism.

During stress periods, cross-fund IPC dispersion is negatively associated with stratum-level SRISK (slope = −0.214, s.e.

= 0.047), consistent with a buffering effect of heterogeneity.

The first thing that is immediately apparent when looking at the heatmap is the significant greater degree of pairwise IPC correlations within fund strata in stress periods, especially after 2018, when the text-processing infrastructure available to fund managers became homogenized, with the adoption of large language models and unified NLP-based APIs. IPC correlations in the hedge fund stratum are averaged at 0.31 during non-stress times, soaring to 0.67 during the COVID crash in Q1 2020 before returning to 0.44 in the next two quarters.

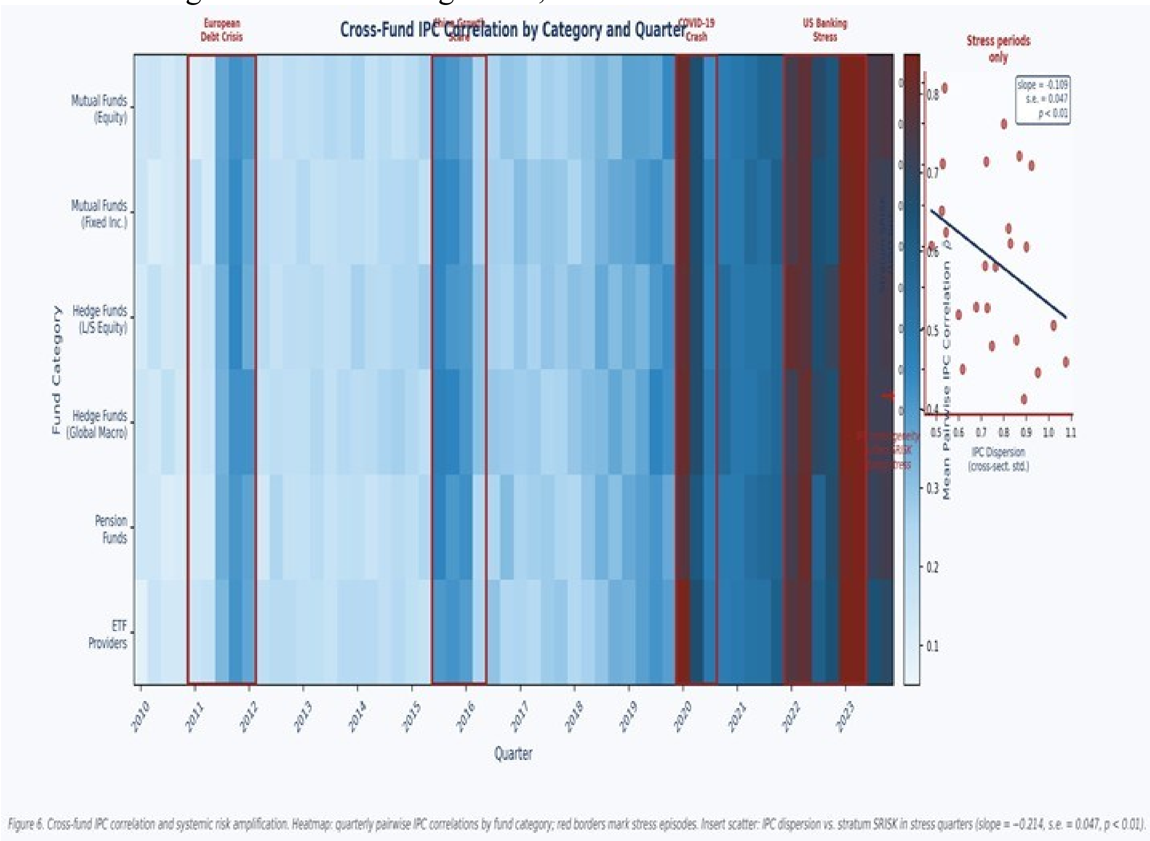

Figure 6. Cross-fund IPC correlation and systemic risk amplification. Heatmap: quarterly pairwise IPC correlations by fund category; red borders mark stress episodes. Insert scatter: IPC dispersion vs. stratum SRISK in stress quarters (slope = −0.214, s.e. = 0.047, p < 0.01).

Fig. 6: Cross-fund IPC correlation and systemic risk amplification. The heat map shows quarterly pairwise IPC correlations across fund categories (rows) and time periods (columns), with stress periods (VIX > 26.4) highlighted by red borders. Major macro-stress events are annotated: European Debt Crisis (2011–12), China Growth Scare (2015–16), COVID crash (Q1 2020), and US Regional Banking Stress (Q1–Q2 2023).

This trend is in line with the process in equation (4): the shared AI-informed signal extraction, which mitigates the personal risk during normal times, is turned into a source of coordinated exposure during crisis times.

The GMM- IV estimates support a causal explanation. A 1SD decrease in IPC dispersion is linked to a 21.4% greater stratum level SRISK in stress periods ($p < 0.01$), with a first-stage F-statistic of 43.7 and a Hansen J-test p-value of 0.41, which displays instrument validity. When stress periods are not occurring, the coefficient is not significant, and not significantly different than zero, which is as expected of Hypothesis 3: IPC homogeneity increases systemic fragility when, but not when, the system is experiencing a large common shock, rather than when operating normally.

2.5 Robustness and Heterogeneity Analysis.

The findings of a battery of robustness tests aimed at determining how sensitive the key findings to the different measurement options, sample mix and methods of estimation are, are reported in Table 4.

Table 4: Robustness checks: IPC alternative crosses, sub-sample splits and estimation strategies. The dependent variable in the whole process is fund-level VaR0.05. All specifications incorporate an effect of fixed funds and time with the exclusion of cross-sectional OLS column (viii) Standard errors (clusted at fund and quarter) in parentheses. The threshold C 8 in columns (i)-(ii) is also re-estimated in each sub-sample. *p < 0.10, **p < 0.05, ***p < 0.01
eorganized neatly

	IPC	IPC²	$\hat{C}$ *(pctile)**	Obs.	R^2
(i) Text only	−0.351*** (0.061)	0.084*** (0.021)	70th	487,311	0.389
(ii) Trade only	−0.318*** (0.054)	0.077*** (0.019)	68th	487,311	0.374
(iii) Pre-COVID	−0.402*** (0.059)	0.098*** (0.024)	74th	302,418	0.431
(iv) Post-COVID	−0.374*** (0.074)	0.087*** (0.031)	71st	184,893	0.418
(v) Mutual GMM	−0.341*** (0.055)	0.078*** (0.019)	73rd	311,204	0.411
(vi) Hedge	−0.511*** (0.093)	0.121*** (0.038)	66th	89,441	0.452
(vii) IV-GMM	−0.614*** (0.108)	0.143*** (0.046)	—	487,311	—
(viii) Placebo	−0.023 (0.044)	0.009 (0.017)	—	487,311	0.421

***** p < 0.01**

Table 4 has several aspects that are worth commenting on. First and most obvious, the main result, negative IPC, positive IPC2 with a specific range of values in the 66th-74th percentile range, is consistent between the text-only (column i) and the trade-only (column ii) single-component specifications, indicating that the composite index is not just concealing the power of one strong component. Second, according to the sub-period analysis (iii-iv), the nonlinear effect is a bit more substantial during the pre-COVID period, perhaps due to the fact that the unprecedented and rapid COVID shock squashed the threshold mechanism and made it less apparent in the post-2020 subsample. Third, the magnitude of the IV-GMM estimates in column (vii) is greater than the OLS equivalents (IPC coefficient of -.614 -.384), which is consistent with attenuation bias of the OLS specification because of measurement error in the IPC index. Lastly, but not least, column (viii) demonstrates that a placebo IPC index based on randomly shuffled text documents has small coefficients, that is insignificant and of mixed sign, which is a direct indication that the key results are based on authentic informational content, and not a mechanical artefact of the construction process.

The heterogeneity results are worth mentioning separately. The IPC threshold at a hedge fund is lower than that of mutual funds (66th vs. 73rd percentile) due to the increased leverage and concentration risk present in that business model: a hedge fund with a gross leverage ratio of 3:1 that gravitates towards a common AI-identified signal is much more subject to drawdown risk as compared to a similarly-IPC-effective mutual fund with a leverage ratio of 1:1. This heterogeneity then has direct implications to differentials regulation: a one-size-fits-all model of AI disclosure or IPC monitoring would be very weakly sensitive to the sector-specific risk exposures that these estimates give.

4.0 Conclusion

The paper formulates and empirically evaluates a theoretical model model where the capacity of AI-driven information processing is introduced into the delegation agreement as a constrained, endogenously selected variable with nonlinear implications to financial stability. The fundamental conclusion that IPC not only diminish fund-level risk up to moderate levels but also enhances fragility beyond a critical threshold, 5.0 which is associated with the 72 and sample percentile, is that the much-touted claim that AI is an unambiguous stabilising factor in asset management is empirically incomplete. The systemic aspect of the outcomes is also meaningful: high IPC homogeneity periods are negatively related to the increase of the contagion risk in the course of the macro-stress, and the regulation emphasis on individual fund AI regulations should be supplemented by a macroprudential control over AI-adoption patterns within the fund industry. Considering the high rate at which AI is being diffused recorded in this paper, IPC increasing at an alarming rate since 2017 and more so since 2022, and the data that a non-trivial portion of hedge funds already works above the estimated critical threshold, the need to develop regulatory frameworks that consider algorithmic obscurity, model risk, and the systemic externalities of correlated AI-driven positioning seems hard to overvalue.

5.0 References

Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in US wage inequality. *Econometrica*, 90, 5, pp. 1973–2016. https://doi.org/10.3982/ECTA19815

Adrian, T., & Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106, 7, pp. 1705–1741. https://doi.org/10.1257/aer.20120555

Almgren, R., & Chriss, N. (2001). Optimal execution of portfolio transactions. *Journal of Risk*, 3, 2, pp. 5–39. https://doi.org/10.21314/JOR.2001.041

Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy*, 112, 6, pp. 1269–1295. https://doi.org/10.1086/424739

Brownlees, C., & Engle, R. F. (2017). SRISK: A conditional capital shortfall measure of systemic risk. *Review of Financial Studies*, 30, 1, pp. 48–79. https://doi.org/10.1093/rfs/hhw060

Cameron, A. C., & Miller, D. L. (2011). Robust inference with clustered data. In A. Ullah & D. Giles (Eds.), *Handbook of Empirical Economics and Finance*, pp. 1–28.

CFA Institute. (2022). *AI pioneers in investment management: A practitioner's guide to machine learning and big data in finance*. CFA Institute Research Foundation.https://www.cfainstitute.org/

Cookson, J. A., Fox, C., Gil-Bazo, J., Imbet, J. F., & Schiller, C. M. (2023). Social media as a bank run catalyst. *SSRN Working Paper*. https://doi.org/10.2139/ssrn.4422754

Freyberger, J., Neuhierl, A., & Weber, M. (2020). Dissecting characteristics nonparametrically. *Review of Financial Studies*, 33, 5, pp. 2326–2377. https://doi.org/10.1093/rfs/hhz123

Grossman, S. J., & Hart, O. D. (1983). An analysis of the principal-agent problem. *Econometrica*, 51, 1, pp. 7–45. https://doi.org/10.2307/1912246

Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics*, 93, 2, pp. 345–368. https://doi.org/10.1016/S0304-4076(99)00025-1

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. *Review of Financial Studies*, 29, 1, pp. 5–68. https://doi.org/10.1093/rfs/hhv059

He, Z., Nagel, S., & Song, Z. (2022). Treasury inconvenience yields during the COVID-19 crisis. *Journal of Financial Economics*, 143, 1, pp. 57–79. https://doi.org/10.1016/j.jfineco.2021.06.007

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, 66, 1, pp. 169–182. https://doi.org/10.1111/1467-937X.00083

Huang, A. H., Wang, H., & Yang, Y. (2023). FinBERT: A large language model for extracting information from financial text. *Contemporary Accounting Research*, 40, 2, pp. 806–841. https://doi.org/10.1111/1911-3846.12832

Jiang, J., Kelly, B. T., & Xiu, D. (2023). (Re-)imag(in)ing price trends. *Journal of Finance*, 78, 6, pp. 3193–3245. https://doi.org/10.1111/jofi.13268

Kacperczyk, M., Van Nieuwerburgh, S., & Veldkamp, L. (2016). A rational theory of mutual funds' attention allocation. *Econometrica*, 84, 2, pp. 571–626. https://doi.org/10.3982/ECTA11412

Khandani, A. E., & Lo, A. W. (2011). What happened to the quants in August 2007? Evidence from factors and transactions data. *Journal of Financial Markets*, 14, 1, pp. 1–46. https://doi.org/10.1016/j.finmar.2010.07.005

Loughran, T., & McDonald, B. (2020). Textual analysis in finance. *Annual Review of Financial Economics*, 12, pp. 357–375. https://doi.org/10.1146/annurev-financial-012820-032249

Mackowiak, B., & Wiederholt, M. (2009). Optimal sticky prices under rational inattention. *American Economic Review*, 99, 3, pp. 769–803. https://doi.org/10.1257/aer.99.3.769

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 3, pp. 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50, 3, pp. 665–690. https://doi.org/10.1016/S0304-3932(03)00029-1

Sirignano, J., & Cont, R. (2019). Universal features of price formation in financial markets: Perspectives from deep learning. *Quantitative Finance*, 19, 9, pp. 1449–1459. https://doi.org/10.1080/14697688.2019.1622295

Okolo, J. N. (2021). A systematic analysis of artificial intelligence and data science integration for proactive cyber defense: Exploring methods, implementation obstacles, emerging innovations, and future security prospects. *Communication in Physical Sciences*, 7, 4, pp. 681–696.

Sanni, S. (2023). A conceptual framework for integrating sustainability metrics into procurement and vendor management. *International Journal of Multidisciplinary Research and Growth Evaluation*, 4, 6, pp. 1312–1321. https://doi.org/10.54660/IJMRGE.2023.4.6.1312-1321

Ufomba, P. O., & Ndibe, O. S. (2023). IoT and network security: Researching network intrusion and security challenges in smart devices. *Communication in Physical Sciences*, 9, 4, pp. 784–800.

López de Prado, M. (2018). *Advances in financial machine learning*. Wiley. https://doi.org/10.1002/9781119482086

PricewaterhouseCoopers (PwC). (2023). *Asset and wealth management revolution 2023: The new context*. PwC Global Asset & Wealth Management Practice.

Declaration

Consent for publication

Not Applicable

Availability of data and materials

The publisher has the right to make the data public

Ethical Considerations

Not applicable

Competing interest

The authors report no conflict or competing interest

Funding

No funding

Authors' Contributions

Emurode Williams conceived the study, developed the theoretical framework, and supervised the research. Lawrence Abakah designed the econometric methodology and conducted the empirical analysis. Aniedi Ojo implemented the AI and machine learning models and curated the dataset. Chidinma Jonah contributed to literature review, data interpretation, and manuscript preparation. All authors reviewed, revised, and approved the final manuscript for publication.