A[gobble]

Generative AI and Systemic Risk in Finance: A Conceptual Framework for Innovation, Stability, and Regulatory Design

Jörg Osterrieder[1*]

[1*]University of Twente, Faculty of Behavioural, Management and Social Sciences, Enschede, Netherlands.

Corresponding author(s). E-mail(s): j.r.osterrieder@utwente.nl;

Abstract

Generative artificial intelligence—encompassing large language models, generative adversarial networks, and diffusion models—is transforming financial services across modeling, trading, risk management, and compliance at a pace and breadth unprecedented in prior waves of financial technology adoption. Unlike earlier innovations that penetrated individual functional domains before diffusing, GenAI has entered multiple core functions of financial intermediation nearly simultaneously, raising systemic implications the stability literature has not addressed with adequate theoretical structure. Existing research catalogs applications or enumerates risks but provides no unified framework connecting firm-level adoption decisions to system-level stability outcomes through specified causal channels. This paper develops a Micro-Meso-Macro conceptual framework identifying sixteen channels—twelve risk and four benefit—through which GenAI affects systemic risk in finance. At the micro level, channels address model risk amplification, operational dependency on concentrated providers, competitive divergence, and human capital erosion through automation bias. At the meso level, channels capture algorithmic monoculture from shared foundation models, market microstructure transformation, information ecosystem distortion from synthetic content, and interconnectedness via shared technological infrastructure. At the macro level, channels encompass procyclicality amplification, regulatory arbitrage acceleration, too-connected-to-fail AI infrastructure, and cross-border regulatory fragmentation. Four benefit channels identify innovation pathways—enhanced risk detection, financial inclusion, improved regulatory surveillance, and market efficiency—each conditional on institutional safeguards and subject to identified tipping points. We derive sixteen testable propositions, each with identification strategies and falsification conditions. The framework provides a structured foundation for empirical research, evidence-based regulation, and strategic AI governance in financial institutions.

Keywords: Generative AI systemic risk financial stability large language models conceptual framework financial regulation model risk algorithmic monoculture

1 Introduction

The release of ChatGPT in November 2022 marked an inflection point in the financial industry's engagement with artificial intelligence. Within months, major banks, asset managers, and insurers began integrating large language models into functions spanning credit analysis, trading strategy development, regulatory compliance, and client advisory. The development of domain-specific foundation models—most notably BloombergGPT, a 50-billion-parameter model trained on a proprietary corpus of financial data Wu et al (2023)—signaled that generative AI (GenAI) was not merely a general-purpose technology being adapted to finance but a class of models increasingly purpose-built for financial applications. The pace of adoption has been rapid and broadly based: GenAI is now reshaping financial modeling and forecasting, risk management and stress testing, trading and market making, and compliance and regulatory reporting. This breadth of penetration distinguishes the current wave from prior episodes of financial technology adoption, which typically affected one or two functional domains before diffusing more widely. GenAI, by contrast, has entered multiple core functions of financial intermediation nearly simultaneously, raising questions about systemic implications that the financial stability literature has not yet addressed with adequate theoretical structure.

Generative AI refers to models that learn the underlying distribution of training data and generate novel samples from that distribution. The class encompasses several architectural families relevant to finance. Generative adversarial networks (GANs), introduced by Goodfellow et al (2014), learn to produce synthetic data by training a generator against a discriminator, with applications in synthetic financial data generation, privacy-preserving data sharing, and tail-risk scenario simulation Eckerli and Osterrieder (2021). Transformer-based architectures Vaswani et al (2017), scaled to large language models such as GPT-3 Brown et al (2020) and its successors, process and generate natural language text, enabling applications from automated analyst reports to real-time interpretation of earnings calls. Diffusion models, which generate samples through iterative denoising, represent a more recent entry with emerging applications in financial time-series generation. The defining characteristic unifying these architectures—and distinguishing them from discriminative or predictive machine learning—is creation rather than classification. Where a traditional credit scoring model assigns a probability to existing data, a generative model produces synthetic credit scenarios, fabricates plausible analyst commentary, or generates novel trading strategies. This generative capacity is the source of both the technology's transformative potential and its distinctive risks.

The dual nature of GenAI in finance constitutes a fundamental tension that existing literature has not resolved. On the innovation side, GenAI offers enhanced forecasting through non-linear pattern recognition, improved risk modeling through synthetic scenario generation, democratized analytics through accessible API-based tools, and more effective regulatory

---

[0] JEL Classification: G01, G18, G21, G28, O33.

surveillance through automated monitoring at scale. On the risk side, the same technology introduces hallucination in financial advice and risk assessment, algorithmic herding arising from shared model architectures, opacity that defeats existing validation frameworks, and provider concentration creating new forms of systemic interconnection. The International Monetary Fund has recognized that AI may amplify systemic risk through correlated behavior and operational concentration International Monetary Fund (2023), while Danielsson et al (2022) demonstrate formally that artificial intelligence can undermine financial stability through strategic complementarities that produce correlated actions across market participants. This tension is not merely a technical issue amenable to engineering solutions; it is a strategic governance challenge requiring financial institutions, regulators, and policymakers to manage innovation and stability as jointly determined outcomes rather than independent objectives.

Despite the significance of this challenge, the academic literature contains a notable gap. Existing work addresses either the applications of GenAI in finance or the risks of AI for financial stability, but no contribution provides an integrated framework connecting firm-level adoption decisions to system-level stability outcomes through specified causal channels. The gap can be characterized precisely. Eisfeldt and Schubert (2025), in their Annual Review of Financial Economics survey, provide a broad overview of generative AI and its implications for finance, with particular attention to occupational exposure and labor market effects; they do not develop a systemic risk framework or derive testable propositions linking GenAI adoption to financial instability. Ali et al (2025) conduct a PRISMA-based systematic review of 84 papers on AI and large language models in finance, identifying six thematic clusters; their contribution is bibliometric and taxonomic rather than theoretical, and it does not address systemic risk. Danielsson et al (2022) analyze AI and financial stability with formal rigor, but their analysis predates the LLM revolution and does not address mechanisms specific to generative models—hallucination, monoculture arising from shared foundation models, or the information ecosystem distortion that mass-produced synthetic content introduces. Practitioner and policy contributions from the IMF International Monetary Fund (2023) and the Financial Stability Board Financial Stability Board (2025) provide risk taxonomies grounded in regulatory experience, but these are not peer-reviewed academic frameworks and do not derive testable propositions with specified identification strategies and falsification conditions. Eckerli and Osterrieder (2021) survey GAN applications in finance, but their scope is limited to a single model family in a pre-LLM landscape. The result is a fragmented literature in which no single work provides a multi-layered conceptual framework connecting firm-level GenAI adoption to system-level financial stability through identified causal channels, testable propositions, and derived regulatory design implications.

This paper addresses that gap. The contribution is fourfold. First, it develops a three-layer conceptual framework—Micro, Meso, and Macro—grounded in the systemic risk theory of De Bandt and Hartmann (2000), the evolutionary economics architecture of Dopfer et al (2004), the multi-level perspective on technological transitions of Geels (2002), and the prudential regulation framework of Crockett (2000). The framework identifies sixteen channels—twelve risk channels and four benefit/innovation channels—through which GenAI adoption affects financial stability, organized by the level at which each channel operates. Second, it derives sixteen testable propositions, each linked to a specific channel and equipped with an identification strategy, primary data source, and explicit falsification condition, following theory-building methodology in management research Jaakkola (2020); MacInnis (2011). Third, it identifies five cross-layer feedback loops—three amplifying and two conditionally stabilizing—that determine whether GenAI's net systemic effect is destabilizing or beneficial, demonstrating that the outcome depends not on the technology itself but on the institutional conditions under which it is deployed. Fourth, it translates the framework's causal channels into regulatory design implications across model governance, systemic risk monitoring, operational resilience, information integrity, and proportionate regulation. The paper thus addresses GenAI in finance as a strategic technology governance challenge that spans firm strategy, market dynamics, and regulatory architecture.

The remainder of the paper is organized as follows. Section 2 surveys the empirical landscape of GenAI technologies in finance, establishing the application domains that motivate the framework. Section 3 develops the three-layer conceptual framework, defining the sixteen channels, four benefit channels, and five feedback loops. Section 4 derives the sixteen testable propositions with identification strategies and falsification conditions. Section 5 translates the framework into regulatory design implications. Section 6 discusses theoretical contributions, practical implications, and limitations. Section 7 concludes.

2 Generative AI Technologies in Finance

The application of generative artificial intelligence (GenAI) in finance has expanded rapidly since the public release of ChatGPT in November 2022, though foundational work on generative models in financial contexts predates this inflection point. This section surveys the principal domains in which GenAI is being deployed or actively researched, organized by financial application rather than by underlying technology. For comprehensive reviews of GenAI architectures, the reader is referred to Goodfellow et al (2014) on generative adversarial networks, Vaswani et al (2017) on transformer architectures, and Brown et al (2020) on large language models. The objective here is to establish the empirical landscape of GenAI adoption that the conceptual framework in Section 3 seeks to explain.

2.1 Financial Modeling and Forecasting

GenAI has entered financial modeling through multiple vectors. The most visible development has been the application of large language models (LLMs) to sentiment-driven financial prediction. Lopez-Lira and Tang (2023) demonstrated that ChatGPT can extract sentiment signals from news headlines with predictive power for next-day stock returns, outperforming established sentiment dictionaries. This finding is significant less for the predictive improvement itself than for what it reveals: a general-purpose language model, trained on no financial data specifically, can generate trading-relevant signals from unstructured text, suggesting that LLMs encode latent representations of market-relevant information. This has driven substantial interest in developing finance-specific foundation models.

Two domain-specific LLMs illustrate the current trajectory. BloombergGPT Wu et al (2023) is a 50-billion-parameter model trained on a proprietary corpus combining general-purpose text with Bloomberg's financial data archive, demonstrating superior performance on financial NLP benchmarks including sentiment analysis, named entity recognition, and financial question answering. FinGPT Yang et al (2023) pursues an alternative, open-source approach, fine-tuned on publicly available financial data. The contrast—proprietary versus open-source, closed data versus public data—foreshadows the provider concentration dynamics analyzed in Section 3 (channels M2 and S4), as the resources required to train competitive financial LLMs increasingly determine which organizations shape the analytical infrastructure of finance.

In parallel, generative adversarial networks (GANs) have established a substantial research presence in financial data generation. Eckerli and Osterrieder (2021) survey GAN applications in finance, documenting their use in synthetic market data generation, time-series simulation, and privacy-preserving data sharing. GANs are particularly valuable for generating realistic synthetic financial scenarios—including tail-risk events poorly represented in historical datasets—a capability directly relevant to the stress testing applications discussed in Section 2.2. More recently, diffusion models have emerged as a potential alternative to GANs for financial time-series generation, with Coletta et al (2025) demonstrating the application of denoising diffusion probabilistic models to synthetic market data generation, though this application remains at a relatively early stage compared to GANs.

2.2 Risk Management and Stress Testing

The application of GenAI to risk management addresses a fundamental limitation of traditional approaches: dependence on historical data for estimating risk distributions. Tail-risk scenarios—the low-probability, high-impact events that pose the greatest threat to financial stability—are inherently rare in historical records, and the non-stationarity of financial systems means that future crises need not resemble past ones. GANs trained on historical market data can generate synthetic scenarios that are statistically consistent with observed dynamics while exploring regions of the distribution that historical data do not adequately cover Eckerli and Osterrieder (2021), enabling stress testing against a richer set of adverse conditions than backward-looking methods permit.

LLMs contribute through a complementary channel: the analysis of unstructured risk signals. Traditional risk models operate primarily on structured quantitative data—prices, returns, balance sheet figures. LLMs can process earnings call transcripts, regulatory filings, news articles, and social media commentary to identify early indicators of credit deterioration or operational disruption that precede their manifestation in quantitative metrics Li et al (2024).

Two caveats are essential. First, synthetic data generation does not eliminate model risk; it transforms it. A GAN generating unrealistic tail scenarios may induce overhedging or capital misallocation, substituting one form of estimation error for another. Second, LLM-based risk analysis introduces hallucination risk into a domain where fabricated information has severe consequences—precisely the model risk amplification analyzed in Section 3 (channel M1). These tools complement rather than replace traditional risk models, and their value is conditional on validation frameworks that current regulatory guidance does not yet adequately address.

2.3 Trading and Market Making

GenAI is reshaping trading and market making through mechanisms qualitatively distinct from prior generations of algorithmic trading. Traditional algorithmic strategies—high-frequency market making, statistical arbitrage, momentum—operate on predefined rules applied to structured data feeds. LLM-based trading agents, by contrast, interpret unstructured information in real time: parsing earnings call transcripts, evaluating regulatory announcements, and assessing news sentiment with a breadth and speed no human trading desk can match. Lopez-Lira and Tang (2023) provide early evidence that LLM-derived sentiment signals contain exploitable return predictability, confirming that the integration of natural language understanding into trading logic is not merely theoretical.

The concept of "LLM-as-agent" extends this further. Emerging research explores autonomous trading systems in which an LLM acts not merely as a signal generator but as a decision-making agent that interprets market conditions, formulates trading strategies, and executes orders with minimal human intervention Li et al (2024); Xiao et al (2024). The operational scope of such agents—capable of reasoning across asset classes, integrating qualitative and quantitative inputs, and adapting strategies in real time—represents an unprecedented expansion in the autonomy of automated trading systems.

Perhaps most consequentially, the prospect of agentic AI in financial markets is emerging: multiple autonomous AI agents interacting in market environments where no human participant is directly involved in individual transaction decisions. The dynamics of agent-to-agent markets—in which LLM-based systems trade with other LLM-based systems, each interpreting the other's behavior through learned representations—have no precedent in financial market history. The potential for emergent behaviors—flash crashes, liquidity spirals, coordination failures arising from interactions between agents trained on similar data—motivates the meso-level analysis of market microstructure transformation in Section 3 (channel S2). The speed, scale, and autonomy of GenAI-driven trading represent not merely a quantitative acceleration of existing trends but a qualitative transformation in how markets function.

2.4 Compliance and Regulatory Reporting

Financial compliance is among the most resource-intensive functions in regulated institutions, and GenAI offers substantial potential for both cost reduction and quality improvement. LLMs can parse, summarize, and cross-reference regulatory documents—a significant share of compliance workload given the volume and complexity of financial regulation across jurisdictions. More advanced applications involve automated compliance checking, in which GenAI systems monitor transactions and business activities against regulatory requirements in near real time, flagging potential violations before they result in enforcement actions.

The distinction between regulatory technology (RegTech) and supervisory technology (SupTech) is analytically important. RegTech refers to GenAI applications deployed by regulated institutions to manage their own compliance obligations: automated reporting, transaction monitoring, regulatory change management. SupTech refers to GenAI applications deployed by regulatory authorities to enhance their supervisory capabilities: market surveillance, anomaly detection in filings, and automated analysis of regulated entities' reports Arner et al (2017). The effectiveness of GenAI in improving financial system governance depends critically on the relative pace of adoption on both sides of this divide—a theme developed in Section 3 (benefit channel B3 and its tipping point).

The potential benefits are significant: reduced compliance costs, faster violation detection, more consistent regulatory interpretation. However, GenAI-powered compliance also introduces the risk of what Section 3 terms “compliance theater” (channel L2). An LLM can generate documentation that satisfies the formal requirements of regulatory review while the underlying activity remains non-compliant in substance. The very fluency that makes GenAI useful for legitimate compliance makes it equally useful for producing the appearance of compliance without the substance. Whether GenAI proves net positive for regulatory effectiveness depends on whether supervisory authorities develop commensurate GenAI capabilities or whether the technology asymmetry between industry and regulators widens further.

3 The Conceptual Framework

3.1 Framework Overview

The relationship between technological innovation and financial stability has long resisted simple characterization. Early treatments of systemic risk tended to operate at either the micro-prudential level, examining the soundness of individual institutions, or at the macro-prudential level, examining the resilience of the financial system as a whole Crockett (2000). This dichotomy, while analytically convenient, obscures a critical intermediate stratum: the emergent dynamics that arise when multiple institutions adopt the same technology simultaneously, producing collective behaviors that are not reducible to the sum of individual decisions. The present paper addresses this gap by proposing a three-layer conceptual framework—Micro, Meso, and Macro—through which the channels linking generative AI (GenAI) adoption to systemic risk in finance can be systematically identified and analyzed.

The framework draws on four established theoretical traditions. First, De Bandt and Hartmann (2000) provide the foundational distinction between micro-prudential risk (the failure of individual institutions), meso-level risk (contagion and correlated behavior within markets and sectors), and macro-prudential risk (the stability of the financial system as a whole). Second, Dopfer et al (2004) offer a formal micro-meso-macro architecture from evolutionary economics, in which the meso level constitutes the domain of “rules”—shared technologies, conventions, and behavioral patterns—that mediate between individual adoption decisions and system-level outcomes. Third, Crockett (2000) articulates the regulatory rationale for distinguishing micro-prudential from macro-prudential oversight, providing the institutional grounding for regulatory implications that follow from the framework. Fourth, Geels (2002) supplies the multi-level perspective on technological transitions, in which innovations emerge in niches (micro), challenge and reconfigure existing regimes (meso), and ultimately reshape the broader landscape (macro). Together, these traditions provide complementary lenses: systemic risk theory specifies the mechanisms of instability, evolutionary economics explains how technological adoption aggregates across levels, prudential regulation motivates the policy architecture, and the multi-level perspective captures the dynamics of technological displacement.

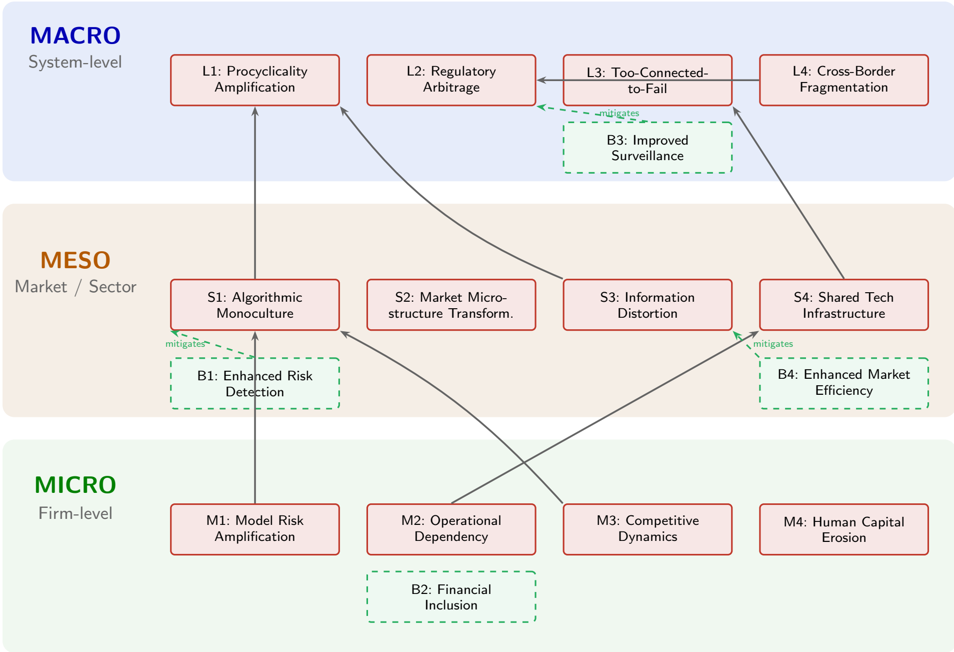

Fig. 1 Micro-Meso-Macro conceptual framework for GenAI and systemic risk. Solid red boxes denote risk channels (M1–M4, S1–S4, L1–L4); dashed green boxes denote benefit channels (B1–B4). Solid arrows indicate cross-layer causal pathways; dashed green arrows indicate mitigation relationships. The three shaded bands correspond to the Micro (firm-level), Meso (market/sector-level), and Macro (system-level) layers.

The choice of three layers is deliberate and consequential. A two-layer decomposition (micro-macro) would collapse meso-level phenomena—algorithmic herding, model monoculture, information ecosystem distortion—into either the firm level or the system level, forfeiting the ability to analyze emergent dynamics that constitute arguably the most distinctive risk contribution of GenAI. A decomposition into four or more layers would fragment the analysis without commensurate explanatory gain; the meso level already accommodates both market-level and sector-level dynamics. Three layers also map directly onto the existing regulatory architecture: micro-prudential supervision, market conduct regulation, and macro-prudential oversight, ensuring that the framework's risk channels translate naturally into regulatory design implications (developed in Section 5).

The boundaries between layers are defined as follows. The **Micro** layer encompasses the individual firm's GenAI adoption decision and its direct consequences for that firm's operations, risk profile, and competitive position; the layer's scope ends where interaction effects between firms begin. The **Meso** layer captures emergent dynamics arising from the interaction of multiple firms' GenAI adoption—market-level phenomena that cannot be reduced to individual behavior; its scope ends where dynamics become properties of the financial system as a whole. The **Macro** layer addresses system-wide stability properties that emerge from meso-level dynamics interacting with the institutional and regulatory environment. Figure 1 provides a schematic overview, mapping all twelve risk channels, four benefit channels, and five cross-layer feedback loops onto this three-layer architecture.

3.2 Micro Layer: Firm-Level Risk Channels

The micro layer identifies four channels through which an individual firm's GenAI adoption creates or amplifies risk. Each channel is distinguished by mechanisms specific to generative models as opposed to traditional machine learning or rule-based systems.

3.2.1 M1: Model Risk Amplification

Model risk—the potential for adverse consequences from decisions based on incorrect or misused model outputs—has been a recognized concern in financial regulation since at least the Federal Reserve's SR 11-7 guidance on model risk management Board of Governors of the Federal Reserve System (2011). GenAI introduces two qualitative escalations of this risk. The first is hallucination: generative models produce outputs that are fluent, internally coherent, and entirely fabricated Ji et al (2023). Unlike traditional statistical models, which may generate inaccurate point estimates, large language models (LLMs) fabricate plausible-sounding reasoning, citations, and analyses that are difficult to distinguish from legitimate outputs without independent verification. The second escalation is opacity at scale. While traditional machine learning models such as gradient-boosted trees or logistic regressions are already subject to interpretability concerns, GenAI models with billions of parameters exhibit emergent capabilities—behaviors that arise unpredictably at scale and were not explicitly programmed Wei et al (2022)—rendering existing model validation frameworks fundamentally insufficient.

In practice, a GenAI model deployed for credit risk assessment may hallucinate correlations between macroeconomic variables and default probabilities, producing risk estimates that embed fabricated causal relationships. An LLM generating investment memoranda may present confident analysis grounded in non-existent data. The consequences—mispriced assets, inadequate capital buffers, flawed lending decisions—are individually damaging and, through the meso-level channels discussed below, collectively destabilizing. The existing regulatory toolkit, designed for parametric models with identifiable assumptions Basel Committee on Banking Supervision (2009), lacks the conceptual vocabulary to address models whose failure mode is the generation of convincing falsehoods rather than measurable estimation error.

3.2.2 M2: Operational Dependency and Provider Concentration

The second micro-level channel concerns the operational dependencies created by GenAI adoption. The computational requirements for training and deploying foundation models—measured in hundreds of millions of dollars for frontier models—concentrate capability in a small number of providers, principally OpenAI, Anthropic, and Google Financial Stability Board (2025). Financial institutions that integrate these providers' APIs into core processes create single points of failure: a provider outage, security breach, policy change, or commercial dispute simultaneously disrupts all dependent institutions. This concentration is qualitatively different from traditional information technology outsourcing, where enterprise software markets typically support dozens of competing vendors offering substitutable products.

The regulatory significance of this dependency has been recognized through the Digital Operational Resilience Act (DORA, Regulation 2022/2554), which establishes requirements for ICT third-party risk management in financial services European Parliament and Council (2022). However, DORA was drafted before the full implications of foundation model concentration were apparent, and its provisions do not account for the absence of meaningful substitutability between GenAI providers whose models differ in architecture, training data, and output characteristics. From a strategic management perspective, this dependency represents a critical resource bottleneck Barney (1991) in which the most strategically important capability is controlled by entities outside the regulated financial sector.

3.2.3 M3: Competitive Dynamics and the AI Divide

GenAI adoption is not uniform across the financial sector. Large institutions with substantial technology budgets, proprietary data assets, and specialized talent pools are positioned to capture disproportionate benefits from GenAI deployment, while smaller institutions face qualitatively higher adoption barriers than with previous generations of financial technology Tushman and Anderson (1986). The scale advantages of foundation models—where training costs alone can exceed $100 million and effective deployment requires dedicated machine learning engineering teams—are orders of magnitude larger than for traditional analytics, creating what Rogers (2003) would characterize as a technology diffusion pattern heavily skewed toward well-resourced early adopters.

The systemic implications extend beyond competitive fairness. As GenAI-enabled institutions gain analytical advantages, market structure shifts toward greater concentration, potentially weakening the competitive discipline that constrains excessive risk-taking Vives (2016). From the perspective of technological transitions, this represents a competence-destroying discontinuity Tushman and Anderson (1986) in which a new paradigm renders existing capabilities obsolete, restructuring the competitive landscape with implications for market stability.

3.2.4 M4: Human Capital Erosion and Automation Bias

The fourth micro-level channel concerns the effect of GenAI adoption on human judgment within financial institutions. Parasuraman and Riley (1997) established that operators of automated systems exhibit systematic biases: over-reliance on automation outputs (automation bias), reduced vigilance in monitoring automated processes (complacency), and progressive loss of manual skills (deskilling). These phenomena have been documented across domains from aviation to medicine, and there is no reason to expect financial services to be exempt.

GenAI amplifies these risks through a mechanism unique to generative models: the production of outputs in natural language that mimic expert reasoning. A traditional credit scoring model produces a numerical score requiring human interpretation. An LLM produces a narrative risk assessment, complete with qualitative reasoning, caveats, and recommendations, that reads as though authored by a senior analyst. This natural language mimicry makes GenAI outputs uniquely persuasive and correspondingly more difficult to scrutinize critically Bansal et al (2021). Over time, financial professionals who routinely defer to GenAI-generated analyses may lose the independent analytical capacity that constitutes the human-in-the-loop safeguard assumed by virtually all model governance frameworks. The erosion is gradual and largely invisible until a stress event reveals the extent to which institutional judgment has been hollowed out.

3.3 Meso Layer: Market and Sector-Level Risk Channels

The meso layer captures dynamics that emerge from the interaction of multiple firms' GenAI adoption—phenomena that are not properties of any individual institution but arise from collective behavior. Following Dopfer et al (2004), the meso level is the domain of shared technological "rules" whose adoption by populations of agents generates emergent systemic properties.

3.3.1 S1: Algorithmic Monoculture and Herding

When multiple financial institutions deploy the same foundation model—or models trained on substantially overlapping data—their analytical outputs converge. Danielsson et al (2022) demonstrate that artificial intelligence can undermine market stability through strategic complementarities: when many agents optimize using similar models, their actions become correlated, amplifying market movements and reducing the diversity of opinion that stabilizes prices. GenAI introduces a deeper form of this monoculture. Traditional algorithmic herding arises from similarity in strategy type (e.g., momentum, mean-reversion); GenAI monoculture arises from shared training data, shared model architecture, and shared knowledge representations. The correlation operates not at the level of strategy but at the level of the underlying epistemic framework through which market conditions are interpreted.

In credit markets, convergent GenAI assessments may produce simultaneous credit tightening or expansion across institutions, amplifying credit cycles. In equity markets, LLM-based investment models interpreting the same earnings reports through the same trained weights generate correlated trading signals that amplify price movements. Under stress conditions, this convergence transforms what would otherwise be idiosyncratic portfolio adjustments into synchronized liquidation cascades. The critical distinction from traditional herding is that GenAI monoculture is infrastructural rather than behavioral—it is embedded in the shared technological substrate rather than arising from conscious strategic imitation.

3.3.2 S2: Market Microstructure Transformation

GenAI is rapidly reshaping market microstructure through three vectors: AI-powered market making, LLM-based order flow prediction, and autonomous trading agents. Each introduces novel dynamics distinct from prior generations of algorithmic trading. Traditional algorithmic market makers and high-frequency traders operate on predefined rules or narrow machine learning models optimized for specific signals. LLM-based trading agents, by contrast, can interpret unstructured information—earnings call transcripts, regulatory filings, news articles, social media sentiment—in real time, enabling fundamentally different trading logic that integrates qualitative judgment at algorithmic speed Brogaard et al (2014).

The risks are multifaceted. Flash crashes may become more frequent if correlated GenAI trading strategies produce simultaneous directional bets in response to ambiguous news events that LLMs interpret identically. Liquidity provision may exhibit an illusion of depth: apparent liquidity that evaporates under stress as GenAI market makers simultaneously widen spreads or withdraw from the market. Perhaps most consequentially, the emergence of autonomous agent-to-agent market dynamics—in which GenAI systems trade with other GenAI systems without direct human oversight—introduces interaction effects with no precedent in financial market history and no established theoretical framework for analysis.

3.3.3 S3: Information Ecosystem Distortion

The capacity of LLMs to generate fluent, structured financial text at scale introduces a qualitatively new risk to the information ecosystem on which price discovery depends. GenAI can mass-produce synthetic analyst reports, market commentary, earnings previews, and macroeconomic analyses that are difficult to distinguish from genuine expert-authored research. While empirical evidence on the market impact of AI-generated financial text remains limited, the IMF International Monetary Fund (2023) identifies synthetic content as an emerging risk vector, and rapid advances in LLM text generation capabilities Brown et al (2020) make this channel increasingly plausible. This capacity is unique to generative models; prior automation technologies could not fabricate convincing narrative analysis at scale.

The consequences for price discovery are significant. If prices reflect available information Fama (1970), flooding information channels with synthetic content containing no genuine private information degrades the signal-to-noise ratio upon which efficient pricing depends. Synthetic analyst reports may move prices while conveying no actual insight, introducing noise that misallocates capital. The problem is compounded by the difficulty of detection: as generative models improve, the arms race increasingly favors generation over identification. The cumulative effect is a market environment in which the informational foundations of price discovery are progressively undermined.

3.3.4 S4: Interconnectedness via Shared Technological Infrastructure

Traditional analyses of financial interconnectedness focus on balance sheet linkages: interbank lending, counterparty exposures, common asset holdings. GenAI introduces a new form of interconnection that is technological rather than financial. When multiple institutions rely on the same GenAI provider's API, they become linked through a shared infrastructure layer invisible to conventional network analysis Acemoglu et al (2015). An API outage, data poisoning attack, or model corruption at a shared provider simultaneously affects all dependent institutions, producing correlated failure patterns that propagate through operational channels rather than balance sheet channels.

This technological interconnection is unprecedented in degree. Financial institutions have long shared infrastructure—payment systems, clearinghouses—but these typically feature redundancy, standardized protocols, and regulatory oversight. GenAI provider APIs are proprietary, non-standardized, and largely unregulated. The resulting hidden network topology overlays the financial network: institutions appearing diversified in financial exposures may be critically correlated through technological dependencies. Mapping this network represents a first-order challenge for financial stability surveillance.

3.4 Macro Layer: System-Level Risk Channels

The macro layer addresses properties of the financial system as a whole that emerge from the interaction of meso-level dynamics with the institutional and regulatory environment. These are system-level phenomena in the sense intended by De Bandt and Hartmann (2000): they concern the capacity of the financial system to perform its core functions of intermediation, risk allocation, and payment under stress.

3.4.1 L1: Procyclicality Amplification

Procyclicality—the tendency of the financial system to amplify rather than dampen business cycle fluctuations—has been recognized as a central systemic risk concern since at least the debates surrounding Basel II Danielsson et al (2001). GenAI models trained on historical data inherently encode cyclical patterns: boom-period training data produces models that underestimate tail risk and overweight positive signals; bust-period training data produces the reverse. This is not merely a data quality issue amenable to correction through better sampling. GenAI's superior pattern-matching capability means it exploits cyclical patterns more aggressively than traditional models, amplifying the procyclical bias that simpler models already exhibit Danielsson et al (2022).

The macro-level implication is a financial system in which GenAI-driven risk assessments, credit decisions, and trading strategies collectively amplify cycles. During expansions, GenAI models across the system produce optimistic assessments that support excessive credit growth and asset price inflation. During contractions, the same models pivot to pessimistic assessments that tighten credit and accelerate price declines. The speed of GenAI-driven adjustment—operating at algorithmic timescales rather than human decision-making timescales—means that the procyclical amplification is faster and potentially sharper than in previous cycles, reducing the time available for policy intervention.

3.4.2 L2: Regulatory Arbitrage Acceleration

Regulatory arbitrage—the structuring of activities to minimize regulatory burden without genuinely reducing risk—is a persistent challenge in financial regulation Fleischer (2010). GenAI accelerates this challenge in two ways. First, LLMs can parse regulatory text across multiple jurisdictions, identify ambiguities and gaps, and suggest structuring approaches at a speed and scale impossible for human compliance teams. What previously required weeks of specialist legal analysis can be accomplished in minutes, tilting the balance of the perpetual "cat-and-mouse game" between regulators and regulated entities decisively in favor of the latter.

Second, GenAI enables what may be termed "compliance theater"—the appearance of regulatory compliance achieved through sophisticated automated reporting that satisfies the letter of regulatory requirements while the spirit is systematically undermined. An LLM can generate compliance documentation that passes regulatory review while the underlying activity exploits precisely the gaps the regulation was intended to close. The macro-level consequence is an erosion of regulatory effectiveness that is difficult to detect and even more difficult to address, as the surface indicators of compliance remain satisfactory while the substantive risk reduction that regulation is intended to achieve progressively diminishes.

3.4.3 L3: Too-Connected-to-Fail AI Infrastructure

The concentration of foundation model capability in two to three providers creates a new category of systemically important entity that is not captured by existing regulatory frameworks. If a dominant GenAI provider—upon which hundreds of financial institutions depend for core risk management, trading, and compliance functions—suffers a catastrophic failure, the resulting simultaneous disruption across the financial system would constitute a systemic event. This extends the "too-big-to-fail" problem from financial institutions to technology providers, but with critical differences: these providers are not regulated as systemically important financial institutions, they do not hold capital buffers calibrated to their systemic footprint, and they operate outside the perimeter of financial supervisory authorities Financial Stability Board (2025).

The concentration is driven by natural monopoly dynamics: foundation model training requires compute investments measured in hundreds of millions of dollars, vast proprietary datasets, and specialized talent—barriers that foreclose competitive entry at the frontier. Unlike traditional enterprise software markets with dozens of substitutable vendors, the foundation model market exhibits extreme returns to scale that entrench incumbents. The resulting too-connected-to-fail dynamic is a structural vulnerability growing in proportion to financial sector GenAI adoption and for which no regulatory toolkit currently exists.

3.4.4 L4: Cross-Border Regulatory Fragmentation

GenAI models are inherently global: the same API serves institutions in New York, London, Singapore, and Lagos with identical outputs. Financial regulation, however, remains national, creating a fundamental mismatch between the jurisdictional scope of technology and the jurisdictional scope of oversight Bank for International Settlements (2021); Frost et al (2019). The EU AI Act imposes risk-based classification and compliance requirements; the United Kingdom pursues a principles-based, "pro-innovation" approach; the United States has adopted a sector-specific patchwork; and many emerging market jurisdictions have no GenAI-specific regulation at all.

This fragmentation creates regulatory arbitrage opportunities uniquely accessible through GenAI. Because GenAI APIs are globally uniform—unlike jurisdiction-specific financial products requiring restructuring for each market—institutions can route GenAI-dependent activities to the least restrictive jurisdiction with minimal adaptation cost. The result is a potential race to the bottom in AI governance. At the macro level, regulatory fragmentation produces coordination failures: each jurisdiction's perimeter covers only a fraction of the globally integrated GenAI infrastructure on which its financial system depends.

3.5 Innovation and Benefit Channels

A framework that addresses only risk channels would present an incomplete and potentially misleading account of GenAI's systemic implications. Four benefit channels, each with specified conditions and identified tipping points beyond which benefits invert into risks, complete the framework.

3.5.1 B1: Enhanced Risk Detection and Early Warning (Micro to Meso)

GenAI's capacity to identify complex, non-linear patterns in high-dimensional data offers genuine potential for improved risk detection. LLMs can analyze unstructured data—news, earnings call transcripts, social media, regulatory filings—to identify early warning signals that elude traditional quantitative models Gu et al (2020); Li et al (2024). Generative adversarial networks can produce synthetic stress scenarios that expand stress testing beyond historical worst-case scenarios.

This channel mitigates model risk (M1) when combined with robust validation, diverse model architectures, and human-in-the-loop oversight; at the meso level, improved early warning can reduce the correlated exposures driving algorithmic herding (S1). The *tipping point* is reached when institutions trust GenAI risk detection without validation—when the tool designed to detect risk becomes itself a source of false confidence. GenAI early warning systems that produce authoritative-sounding assessments going unchallenged replace one form of model risk with another that is harder to identify precisely because it mimics expert judgment.

3.5.2 B2: Financial Inclusion via Democratized Analytics (Micro to Macro)

GenAI lowers barriers to sophisticated financial analysis. Capabilities that previously required large quantitative teams become accessible through API-based tools with modest technical requirements Eisfeldt and Schubert (2025). For smaller institutions and emerging-market intermediaries, this democratization can partially offset the competitive disadvantage identified in channel M3. LLM-powered advisory tools extend financial guidance to underserved populations.

Conditions: broad, affordable access to GenAI tools and supportive regulatory frameworks. *Tipping point:* when GenAI advisory provides overconfident or hallucinated guidance to financially unsophisticated users. An LLM generating a confident but fabricated investment recommendation for a retail user who lacks the expertise to evaluate it does not democratize finance—it democratizes exposure to a novel form of mis-selling.

3.5.3 B3: Improved Regulatory Surveillance (Meso to Macro)

GenAI offers significant potential for regulatory technology (RegTech) and supervisory technology (SupTech). Automated reporting can improve compliance accuracy and timeliness; LLM-based surveillance tools can monitor market communications for manipulation signals and analyze filings for anomalies at scale Arner et al (2017).

This channel mitigates regulatory arbitrage (L2) under the critical condition that regulators adopt GenAI at a pace commensurate with industry. *Tipping point:* if regulators lag in GenAI capability, the technology asymmetry widens the arbitrage gap rather than closing it, as institutions use GenAI to evade oversight while regulators remain reliant on pre-GenAI methods.

3.5.4 B4: Enhanced Market Efficiency (Meso)

GenAI processes diverse information sources—textual, numerical, sentiment-based—faster and more comprehensively than human analysts, offering potential improvements in price discovery and allocative efficiency Gu et al (2020). Emerging evidence suggests GenAI may improve information processing speed, though the net effect on market efficiency depends on the balance between genuine analysis and synthetic content generation. This channel mitigates information distortion (S3) if synthetic content remains detectable.

Conditions: AI-generated content must be identifiable through labeling, watermarking, or detection classifiers. *Tipping point:* when synthetic financial content volume overwhelms genuine information, degrading the information ecosystem that efficient pricing requires. Whether GenAI is net positive or net negative for market efficiency is an empirical question that the propositions in Section 4 help resolve.

3.6 Cross-Layer Feedback Loops

The channels identified above do not operate in isolation. Cross-layer feedback loops create dynamic interactions in which micro-level risks aggregate into meso-level phenomena, meso-level dynamics generate macro-level consequences, and macro-level conditions reshape micro-level incentives. Three amplifying and two stabilizing loops are identified.

3.6.1 Amplifying Loops

Loop 1: Model Risk–Herding–Procyclicality Spiral (M1 to S1 to L1 to M1). Inadequate model governance leads to correlated errors across institutions sharing similar GenAI architectures. Correlated errors produce algorithmic herding as converging outputs drive convergent decisions. Herding amplifies business cycle dynamics through synchronized optimism in expansions and pessimism in contractions. The procyclical environment generates training data encoding the amplified cycle, reinforcing model errors in the next generation of GenAI models—a self-reinforcing spiral.

Loop 2: Operational Dependency–Interconnectedness–Regulatory Burden (M2 to S4 to L3 to M2). Provider concentration creates shared technological infrastructure, producing too-connected-to-fail dynamics that prompt regulatory intervention. Compliance costs disproportionately burden smaller institutions, further concentrating GenAI adoption among large providers and large financial institutions, reinforcing the original dependency. This loop illustrates how regulation can inadvertently amplify the concentration it seeks to address.

Loop 3: Information Distortion–Mispricing–Contagion (S3 to macro mispricing to micro losses to meso contagion to S3). Synthetic content flooding degrades the information environment, producing systematic mispricing. When mispriced assets revert toward fundamental value, unexpected losses propagate through interconnected balance sheets. The resulting uncertainty drives further demand for AI-generated "analysis," reinforcing the information distortion that initiated the cycle.

3.6.2 Stabilizing Loops

Loop 4: Risk Detection–Reduced Herding–Reduced Procyclicality (B1 to S1 mitigated to L1 mitigated). Diverse GenAI architectures deployed for risk detection improve identification of correlated exposures, mitigating monoculture and dampening procyclicality. This loop is *conditional on model diversity*: if all institutions deploy the same model for risk detection, the loop degrades from stabilizing to amplifying.

Loop 5: Surveillance–Reduced Arbitrage–Market Integrity (B3 to L2 mitigated to meso integrity to B3). Effective regulatory GenAI surveillance deters arbitrage, and the resulting market integrity generates higher-quality data that enhances detection capability. This loop is *conditional on regulatory pace*: if regulators lag industry in GenAI capability, the loop inverts as the widening capability gap becomes self-reinforcing.

3.6.3 Implications of the Feedback Structure

The feedback architecture yields a central insight: whether GenAI is net stabilizing or destabilizing is not determined by the technology itself but by the institutional conditions under which it is deployed. The amplifying loops operate unconditionally, emerging from the

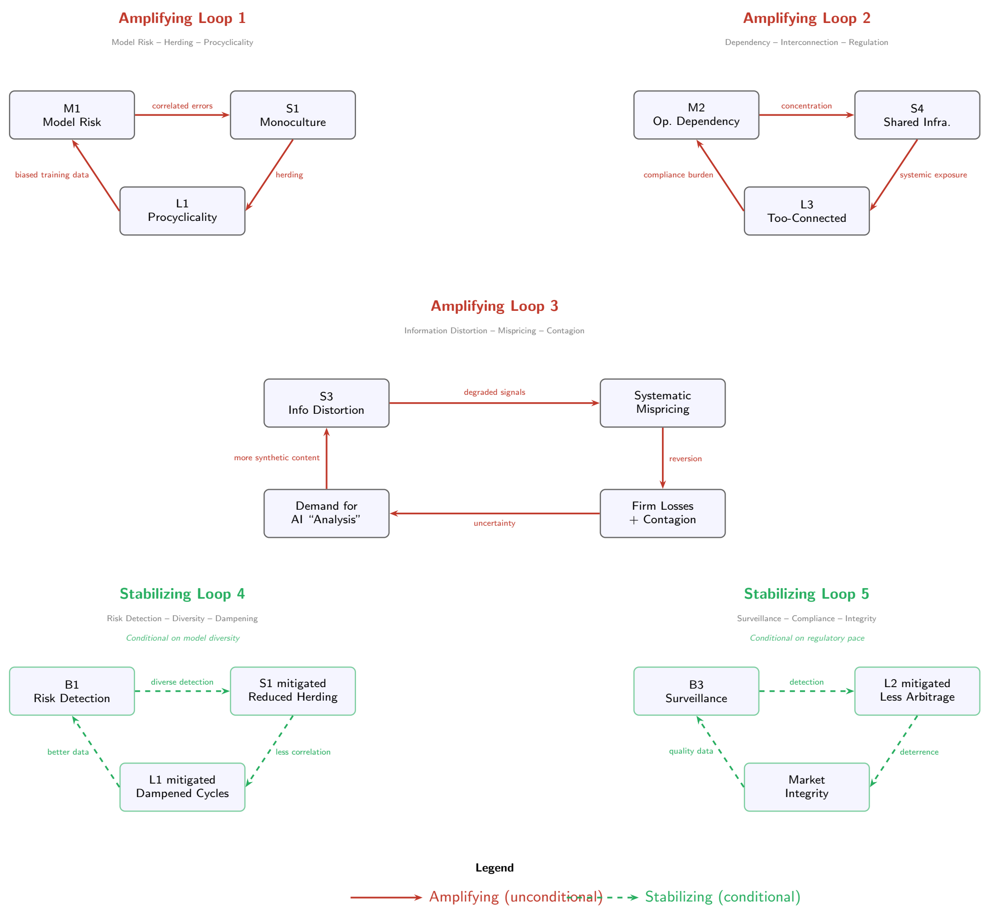

Fig. 2 Cross-layer feedback loops. Three amplifying loops (solid red arrows) operate unconditionally from structural properties of GenAI adoption. Two stabilizing loops (dashed green arrows) operate conditionally on model diversity and regulatory pace, respectively. Amplifying Loop 1 links model risk, algorithmic monoculture, and procyclicality in a self-reinforcing spiral. Amplifying Loop 2 links operational dependency, shared infrastructure, and too-connected-to-fail dynamics. Amplifying Loop 3 links information distortion, mispricing, and contagion. Stabilizing Loop 4 links risk detection to reduced herding and procyclicality. Stabilizing Loop 5 links regulatory surveillance to reduced arbitrage and market integrity.

structural properties of GenAI adoption and requiring active intervention to interrupt. The stabilizing loops are conditional, operating only when specific prerequisites—model diversity, regulatory capability, robust validation—are satisfied. This asymmetry implies that absent deliberate policy intervention, amplifying loops will tend to dominate, because the conditions for stabilizing loops demand coordinated action that markets will not produce spontaneously. The propositions in Section 4 translate this insight into testable hypotheses, and the regulatory implications in Section 5 address the institutional conditions under which stabilizing loops can be activated. Figure 2 provides a diagrammatic representation of the complete feedback structure across the three layers.

4 Testable Propositions

The conceptual framework developed in Section 3 identifies twelve risk channels and four benefit channels through which GenAI adoption may affect financial stability. To move from conceptual identification to empirical tractability, this section derives sixteen testable propositions—twelve addressing risk and four addressing benefit channels. Each proposition follows logically from a specific causal channel, states a directional hypothesis, and specifies an identification strategy, primary data source, and falsification condition. This approach follows established theory-building methodology in management research Jaakkola (2020); Eisenhardt and Graebner (2007), in which propositions serve as the bridge between conceptual frameworks and subsequent empirical investigation.

4.1 Risk Propositions

Proposition 1 (M1—Model Risk Amplification). *Financial institutions that adopt GenAI-based credit scoring models exhibit higher unexpected loss rates during economic downturns than matched institutions using traditional scoring methods.* Derived from the hallucination and opacity mechanisms in channel M1. Identification: DID comparing GenAI-adopting and non-adopting banks around a recessionary shock. Data: FFIEC Call Report charge-offs and loan loss provisions. Falsified if GenAI-adopting institutions show equivalent or lower unexpected loss rates during downturns.

Proposition 2 (M2—Operational Dependency and Provider Concentration). *Higher concentration of GenAI provider usage among financial institutions is associated with larger aggregate operational losses during provider outage events.* Derived from the single-point-of-failure mechanism in channel M2. Identification: event study centered on documented API outages. Data: OCC and FDIC operational loss data cross-referenced with provider dependency disclosures. Falsified if operational losses show no systematic relationship with provider concentration.

Proposition 3 (M3—Competitive Dynamics and the AI Divide). *GenAI adoption widens the Gini coefficient of banking profitability, with adopting institutions capturing disproportionate profit share relative to non-adopters.* Derived from the competence-destroying discontinuity mechanism in channel M3. Identification: panel DID estimator following Callaway and Sant'Anna (2021) for staggered adoption. Data: FFIEC performance data. Falsified if the profitability distribution narrows or remains unchanged following GenAI adoption.

Proposition 4 (M4—Human Capital Erosion and Automation Bias). *Analyst teams using GenAI advisory tools exhibit decreased forecast dispersion without corresponding improvement in forecast accuracy.* Derived from the automation bias and deskilling mechanisms in channel M4. Identification: staggered DID comparing teams before and after GenAI tool adoption. Data: I/B/E/S analyst forecasts. Falsified if dispersion declines are accompanied by proportional accuracy gains, indicating genuine improvement rather than homogenization.

Proposition 5 (S1—Algorithmic Monoculture and Herding). *Institutional investors relying on the same foundation model exhibit higher pairwise return correlations during market stress episodes than those using diverse models or no GenAI.* Derived from the shared epistemic framework mechanism in channel S1. Identification: DID instrumented by exogenous variation in foundation model adoption (e.g., staggered API access). Data: 13-F holdings and SEC technology disclosures. Falsified if return correlations during stress are unrelated to foundation model commonality.

Proposition 6 (S2—Market Microstructure Transformation). *Higher GenAI-based trading volume is associated with increased intraday volatility and reduced market-maker diversity.* Derived from the correlated directional trading and liquidity illusion mechanisms in channel S2. Identification: cross-market comparison exploiting differential GenAI adoption rates across asset classes. Data: TAQ and LOBSTER order-level data. Falsified if markets with higher GenAI trading volume exhibit equivalent or lower volatility and stable market-maker participation.

Proposition 7 (S3—Information Ecosystem Distortion). *A higher proportion of AI-generated analyst reports in a stock's information environment is associated with increased return volatility and decreased post-earnings announcement drift.* Derived from the signal-to-noise degradation mechanism in channel S3. Identification: RDD exploiting the sharp increase in LLM capability around November 2022. Data: I/B/E/S report data and CRSP returns. Falsified if AI-generated content prevalence shows no relationship with volatility or if post-earnings announcement drift increases.

Proposition 8 (S4—Interconnectedness via Shared Technological Infrastructure). *Higher correlation in GenAI API usage across financial institutions predicts elevated tail-risk co-dependence among those institutions.* Derived from the hidden network topology mechanism in channel S4. Identification: time-series DCC models with API usage correlation proxied by technology spending disclosures. Data: CRSP returns. Falsified if tail-risk co-dependence is orthogonal to API usage correlation after controlling for balance-sheet interconnections.

Proposition 9 (L1—Procyclicality Amplification). *GenAI-based credit risk models produce larger forecast error variance across business cycle phases than traditional models applied to the same portfolios.* Derived from the cyclical pattern exploitation mechanism in channel L1. Identification: comparative backtesting across complete credit cycles. Data: Moody's and S&P rating transition and default data. Falsified if GenAI models exhibit equivalent or lower cross-cycle forecast error variance.

Proposition 10 (L2—Regulatory Arbitrage Acceleration). *Financial institutions deploying GenAI compliance tools receive fewer enforcement actions per unit of risk exposure than matched institutions without such tools, controlling for actual risk levels.* Derived from the compliance theater mechanism in channel L2. Identification: PSM of GenAI-adopting and non-adopting institutions. Data: FDIC and OCC enforcement action databases. Falsified if GenAI compliance tool adoption shows no association with enforcement frequency after controlling for underlying risk.

Proposition 11 (L3—Too-Connected-to-Fail AI Infrastructure). *Higher Herfindahl–Hirschman Index (HHI) of AI infrastructure provision to the financial sector Granger-causes higher systemic risk measures (CoVaR, SRISK) among dependent institutions.* Derived from the too-connected-to-fail mechanism in channel L3. Identification: Granger causality framework. Data: industry surveys on provider market share and NYU V-Lab systemic risk measures. Falsified if provider concentration has no predictive power for systemic risk after controlling for traditional concentration metrics.

Proposition 12 (L4—Cross-Border Regulatory Fragmentation). *Greater divergence in GenAI governance stringency across jurisdictions is associated with increased cross-border* *regulatory arbitrage activity in AI-dependent financial services.* Derived from the jurisdictional mismatch mechanism in channel L4. Identification: panel analysis exploiting variation in regulatory adoption timing. Data: constructed GenAI regulatory stringency index and BIS locational banking statistics. Falsified if cross-border activity patterns are unrelated to regulatory stringency differentials.

4.2 Benefit Propositions

Proposition 13 (B1—Enhanced Risk Detection and Early Warning). *Financial institutions deploying GenAI-based risk monitoring detect credit deterioration signals earlier than matched institutions relying on traditional early warning systems.* Derived from the non-linear pattern recognition mechanism in channel B1. Identification: matched-pair design comparing detection lead times. Data: rating agency credit transition data. Falsified if GenAI-equipped institutions show no improvement in early detection relative to matched controls.

Proposition 14 (B2—Financial Inclusion via Democratized Analytics). *The introduction of GenAI-powered advisory tools is associated with increased retail market participation among historically underrepresented demographic groups.* Derived from the barrier-lowering mechanism in channel B2. Identification: DID exploiting staggered platform-level rollouts of GenAI advisory features. Data: Survey of Consumer Finances (SCF) and brokerage account microdata. Falsified if participation rates among underrepresented groups are unchanged or decline following introduction.

Proposition 15 (B3—Improved Regulatory Surveillance). *Regulatory agencies deploying GenAI-based market surveillance achieve higher manipulation detection rates with shorter detection lags than prior non-GenAI regimes.* Derived from the surveillance capacity mechanism in channel B3. Identification: before-and-after comparison at agencies adopting GenAI tools. Data: SEC and FCA enforcement data. Falsified if detection rates and lag times are unchanged or worsen following adoption.

Proposition 16 (B4—Enhanced Market Efficiency). *Higher GenAI trading participation is associated with faster price convergence toward fundamental value, as measured by reduced variance ratios at intermediate horizons.* Derived from the information processing capacity mechanism in channel B4. Identification: variance ratio tests following Lo and MacKinlay (1988) across securities with varying GenAI trading participation. Data: TAQ. Falsified if variance ratios show no improvement or deteriorate with increased GenAI trading presence.

4.3 Summary and Limitations

The sixteen propositions advanced above are theoretical constructs derived from the causal channels of the framework; they await empirical adjudication. Several limitations warrant acknowledgment. First, the identification strategies assume observability of GenAI adoption decisions, which remains imperfect given limited mandatory disclosure requirements. Second, the staggered nature of GenAI adoption creates challenges for clean counterfactual construction, though recent econometric advances in heterogeneous treatment effects Callaway and Sant'Anna (2021) mitigate this concern. Third, the propositions do not address interaction effects between channels, which the cross-layer feedback loops of Section 3.6 suggest may be substantial. The balance between the twelve risk propositions and four benefit propositions should not be interpreted as a prior judgment that risks outweigh benefits; this balance is itself an empirical question, and a central one. The regulatory design implications developed in Section 5 depend critically on which propositions find empirical support and under what institutional conditions the benefit channels dominate the risk channels.

5 Regulatory Design Implications

The framework in Section 3 identifies twelve risk channels, four benefit channels, and five feedback loops; Section 4 specifies the conditions under which they activate. This section translates those findings into regulatory design implications across five domains. Each recommendation is derived from specific framework channels.

5.1 Model Governance and Validation

Linked channels: M1 (Model Risk Amplification), B1 (Enhanced Risk Detection)

Current model governance frameworks—the Federal Reserve's SR 11-7 guidance Board of Governors of the Federal Reserve System (2011), the Basel Committee's principles for effective risk data aggregation Basel Committee on Banking Supervision (2009), and the Prudential Regulation Authority's SS1/23 on model risk management Prudential Regulation Authority (2023)—were designed for parametric and traditional machine learning models whose failure modes are quantifiable estimation errors. Channel M1 establishes that generative models introduce a qualitatively distinct failure mode: hallucination, the production of fluent, internally coherent, and entirely fabricated outputs. This failure mode is not captured by conventional validation metrics such as backtesting accuracy, discriminatory power, or calibration error, because a model that fabricates plausible causal relationships will not be detected by frameworks testing only whether outputs fall within expected statistical ranges.

Extended validation frameworks for GenAI must therefore incorporate at least four additional dimensions. First, output monitoring must move beyond statistical performance to include semantic verification—automated cross-referencing of model-generated claims against authoritative data sources to detect hallucinated content. Second, adversarial testing must probe GenAI models' vulnerability to prompt injection, data poisoning, and adversarial inputs designed to elicit harmful outputs in financial contexts. Third, benchmark suites specific to financial GenAI applications must be developed, testing not only accuracy but also the rate and severity of hallucination under domain-relevant conditions. Fourth, validation must address emergent capabilities Wei et al (2022)—behaviors that arise unpredictably at scale—through ongoing monitoring rather than one-time pre-deployment assessment.

Channel B1 conditions this implication: when combined with robust validation and diverse architectures, GenAI genuinely enhances risk detection. The regulatory objective is not to suppress GenAI deployment but to ensure that validation infrastructure evolves commensurately with the technology. Governance calibrated to parametric assumptions will progressively lose its capacity to constrain M1 while failing to activate B1.

5.2 Systemic Risk Monitoring

Linked channels: S1 (Algorithmic Monoculture), S4 (Shared Technological Infrastructure), L1 (Procyclicality Amplification), L3 (Too-Connected-to-Fail)

Established systemic risk measures—CoVaR Adrian and Brunnermeier (2016) and SRISK Brownlees and Engle (2017)—quantify systemic risk through balance sheet exposures, market-based correlations, and capital shortfall estimates. These measures do not capture the concentration forms that channels S1, S4, and L3 identify: correlation arising from shared model architectures, overlapping training data, and dependence on common providers rather than financial linkages. An institution well-diversified in financial exposures may be critically correlated with peers through a shared GenAI infrastructure layer invisible to conventional network analysis.

Two supplementary indicators are proposed. The first, an *AI Concentration Index*, would measure financial sector dependence on individual GenAI providers by aggregating data on API usage, model deployment, and provider revenue concentration across regulated institutions. This index operationalizes channels S4 and L3: as concentration rises, vulnerability to correlated operational failure and too-connected-to-fail dynamics increases proportionally. The second, a *Model Diversity Index*, would measure heterogeneity of GenAI approaches across the sector—encompassing architecture, training data provenance, and fine-tuning methodology. This index operationalizes channel S1: declining diversity signals rising vulnerability to algorithmic herding and the procyclical amplification (L1) that the framework's first amplifying feedback loop describes.

Both indices require mandatory disclosure of GenAI provider relationships, model architectures, and training data characteristics—information not currently collected in any supervisory reporting framework. Their construction extends macro-prudential surveillance into the technological infrastructure domain, reflecting the framework's finding that financial interconnectedness now operates through technological as well as balance sheet channels.

5.3 Operational Resilience

Linked channels: M2 (Operational Dependency), S4 (Shared Technological Infrastructure), L3 (Too-Connected-to-Fail)

The Digital Operational Resilience Act (DORA, Regulation 2022/2554) establishes ICT third-party risk management requirements for financial services European Parliament and Council (2022). However, as channel M2 identifies, DORA was drafted before foundation model concentration was apparent. Its provisions assume meaningful substitutability between vendors—an assumption that does not hold for frontier GenAI providers, where training costs exceeding hundreds of millions of dollars create natural monopoly dynamics Financial Stability Board (2025).

Three GenAI-specific resilience requirements follow. First, multi-provider strategies must be mandated for systemically important GenAI applications, ensuring no single provider failure simultaneously disrupts critical functions. This addresses L3 directly: too-connected-to-fail risk diminishes with provider redundancy. Second, model portability requirements—technical standards enabling migration between providers—must reduce switching costs and mitigate the lock-in that channel M2 identifies. Without portability, multi-provider mandates become economically impractical. Third, mandatory failure scenario testing must encompass GenAI-specific contingencies: provider API outages, model degradation, data poisoning events, and simultaneous multi-institution disruption from a shared provider failure. Just as banks must demonstrate capital adequacy under adverse macroeconomic scenarios, institutions dependent on GenAI infrastructure must demonstrate operational continuity under adverse technological scenarios—ensuring the system can absorb shocks that individual institutions cannot diversify away independently.

5.4 Information Integrity

Linked channels: S3 (Information Ecosystem Distortion), B4 (Enhanced Market Efficiency)

Channel S3 identifies an unprecedented risk: generative models can produce synthetic financial content—analyst reports, market commentary, earnings analyses—indistinguishable from expert-authored research at a scale that can overwhelm the information ecosystem on which price discovery depends. Channel B4 establishes that GenAI enhances market efficiency only when AI-generated content remains identifiable. When that condition fails, synthetic content degrades rather than improves the signal-to-noise ratio.

The EU AI Act (Regulation 2024/1689) establishes transparency requirements for AI-generated content, including obligations to mark synthetic outputs in machine-readable formats European Parliament and Council (2024). These provisions require adaptation for financial markets: labeling requirements for AI-generated financial content must be mandated for regulated entities, ensuring that GenAI-produced research and recommendations are clearly identified. This is not merely a consumer protection measure; it is a market integrity requirement grounded in the informational foundations of efficient pricing Fama (1970).

Market surveillance must also adapt. Synthetic analyst reports designed to move asset prices constitute a novel manipulation form that existing tools—calibrated for insider trading, spoofing, and layering—cannot identify. Supervisory authorities must develop detection capabilities matching the generative capacity of the models being surveilled, operationalizing the arms race dynamic that channel S3 describes. Failure to adapt creates the conditions under which Loop 3—linking information distortion to mispricing and contagion—operates unchecked.

5.5 Proportionate and Adaptive Regulation

Linked channels: All channels, especially L4 (Cross-Border Regulatory Fragmentation), L2 (Regulatory Arbitrage Acceleration), B2 (Financial Inclusion), B3 (Improved Regulatory Surveillance)

The framework implies that regulatory design must be proportionate to the systemic significance of specific GenAI applications. GenAI deployed in core risk management, credit decisioning, and algorithmic trading warrants the highest scrutiny, while applications in customer service or internal knowledge management pose lower systemic risk and merit lighter requirements. This proportionality follows from the observation that channels M1, S1, and L1 activate most strongly in decision-critical applications where outputs directly influence capital allocation and risk pricing.

Channel L4 identifies the jurisdictional mismatch between globally uniform GenAI APIs and nationally bounded regulatory perimeters. International coordination is essential to prevent the arbitrage that channels L2 and L4 jointly enable. Existing multilateral frameworks—the Financial Stability Board's work on AI in financial services Financial Stability Board (2025), the Bank for International Settlements Innovation Hub, and bilateral supervisory cooperation arrangements—provide institutional foundations but lack the enforcement capacity and technical specificity that GenAI governance requires. Harmonized minimum standards for GenAI model governance in financial services would reduce the scope for arbitrage while preserving jurisdictional flexibility.

Regulatory sandboxes—controlled environments in which innovative applications are tested under supervisory observation—offer a mechanism for balancing the innovation benefits of channels B2 and B3 against the risks the framework identifies. The Financial Conduct Authority's regulatory sandbox and the Monetary Authority of Singapore's AI governance framework demonstrate the viability of this approach Cornelli et al (2024). Sandboxes are well-suited to GenAI because they allow observation of emergent behaviors that static pre-deployment assessment cannot predict. By establishing conditions under which stabilizing feedback loops (Loops 4 and 5) can be activated while amplifying loops are contained, sandboxes operationalize the framework's central insight: the systemic consequences of GenAI depend not on the technology itself but on the institutional conditions governing its deployment.

6 Discussion and Future Research

6.1 Limitations

Several limitations warrant candid acknowledgment. First and most fundamentally, the framework is conceptual rather than empirical. Its sixteen channels, five feedback loops, and sixteen propositions are theoretically derived; none has been subjected to empirical testing. The framework's value lies in organizing a fragmented literature, identifying causal mechanisms, and generating falsifiable hypotheses—contributions that Jaakkola (2020) identifies as the core functions of conceptual work in management research. Until the propositions in Section 4 are adjudicated against data, the framework remains a structured conjecture, not a demonstration.

Second, the technology is evolving at a pace that challenges any static analysis. New foundation models, application paradigms, and regulatory instruments will emerge between submission and publication. The framework mitigates this limitation by operating at the level of structural properties—opacity, monoculture, concentration, procyclicality—rather than specific models or providers. The channels linking shared architectures to algorithmic herding (S1) depend on architectural convergence as a structural fact, not on whether the converged model is GPT-4, Claude, or a successor. This property-level framing provides durability, though it cannot anticipate qualitatively novel risk mechanisms that future paradigms may introduce.

Third, the literature and examples reflect a predominantly Western, developed-market perspective. China's Interim Measures for the Management of Generative AI Services (the AIGC regulations), the Monetary Authority of Singapore's FEAT principles and Veritas framework, India's expanding fintech ecosystem, and mobile-first financial AI in sub-Saharan Africa represent contexts where channel weights may differ: provider concentration (M2, L3) may be less acute under state-backed AI programs; the AI divide (M3) may be sharper where technology infrastructure is less developed. Cross-country validation is essential before the framework can claim generality.

Fourth, the sixteen channels are those with the strongest current theoretical or empirical support. Other channels may emerge—environmental costs of model training, intellectual property disputes, or labor market disruptions feeding back into credit risk. The framework claims systematic coverage of currently identifiable mechanisms, not exhaustiveness. Fifth, the framework is static: it does not model the temporal dynamics through which channels evolve as adoption diffuses and regulation responds. Operational dependency (M2) may diminish if open-source models achieve competitive performance; information distortion (S3) may intensify as generative quality improves.

6.2 Future Research Directions

The framework defines a structured research agenda along several methodological axes.

Empirical testing of propositions. The most direct extension is systematic testing using the identification strategies specified in Section 4. SEC 10-K and 10-Q filings increasingly contain AI adoption disclosures that can be mined to construct firm-level GenAI adoption measures. CRSP and Compustat provide financial performance and market data for Propositions 1, 3, 5, 7, and 8. FFIEC Call Reports supply banking-specific variables—charge-offs, loan loss provisions, operational losses—for Propositions 1, 2, and 3. USPTO and EPO patent data (IPC/CPC classifications for machine learning and natural language processing) offer innovation proxies. I/B/E/S analyst forecast data enables testing of Propositions 4 and 7. The staggered nature of GenAI adoption creates variation for difference-in-differences designs, though researchers must attend to heterogeneous treatment effects under staggered adoption Callaway and Sant'Anna (2021).

Agent-based modeling. The feedback loops are amenable to computational simulation. Agent-based models can populate a synthetic financial system with heterogeneous institutions, vary GenAI monoculture levels, and simulate market dynamics under different adoption and stress scenarios—addressing questions difficult to resolve empirically: How much model diversity prevents herding-driven instability? At what concentration threshold does a provider outage become systemic? Agent-based approaches have established precedent in financial stability research and suit the emergent dynamics the framework describes.

Natural experiments. Credible causal inference requires exogenous variation. Documented API outages at major providers (OpenAI has experienced multiple significant disruptions since 2023) provide event-study opportunities for Proposition 2. The phased EU AI Act implementation creates regulatory shocks for Propositions 10 and 12. Major model releases serve as capability shocks traceable through high-frequency trading data.

Cross-country comparative studies. The EU's prescriptive approach, the United Kingdom's principles-based framework, Singapore's sandbox model, and the United States' sector-specific patchwork represent distinct regulatory treatments whose effects on adoption, stability, and innovation can be compared, testing Proposition 12 directly.

Emerging frontiers. Agentic AI—autonomous systems executing multi-step financial tasks without human intervention—introduces agent-to-agent market dynamics not fully captured by the current framework. Multi-modal GenAI may alter information channels through which risk signals propagate. Open-source and decentralized AI models may reshape the concentration dynamics of channels M2, S4, and L3, potentially substituting provider concentration risk for unvalidated model proliferation risk.

6.3 Practical Implications

The framework carries direct implications for practitioners and policymakers, framed for the management audience of this journal.

For chief risk officers and risk committees, the twelve risk channels provide a due diligence checklist for GenAI vendor selection and deployment governance: Does this deployment create provider concentration (M2)? Does it share a foundation model with competitors, increasing monoculture risk (S1)? Does the governance framework address hallucination rather than only estimation error (M1)?

For chief technology officers and boards, the three layers map onto organizational decision levels. Micro-level channels inform adoption strategy—which functions to automate, which providers to select, what validation to build. Meso-level channels inform industry positioning and collaborative approaches to shared infrastructure risk. Macro-level channels inform regulatory engagement and proactive disclosure strategy.

For regulators, the proposed AI Concentration Index and Model Diversity Index (Section 5.2) provide actionable monitoring tools extending macro-prudential surveillance to technological infrastructure.

For researchers, the sixteen propositions with specified identification strategies, data sources, and falsification conditions provide a structured agenda for a decade of empirical investigation.

7 Conclusion

This paper has developed a three-layer conceptual framework—Micro, Meso, and Macro—for analyzing the relationship between generative AI adoption and systemic risk in finance. The framework identifies twelve risk channels distributed across firm-level dynamics (model risk amplification, operational dependency, competitive dynamics, human capital erosion), market-level emergent phenomena (algorithmic monoculture, market microstructure transformation, information ecosystem distortion, technological interconnectedness), and system-level properties (procyclicality amplification, regulatory arbitrage acceleration, too-connected-to-fail infrastructure, cross-border regulatory fragmentation), together with four benefit channels (enhanced risk detection, financial inclusion, improved regulatory surveillance, enhanced market efficiency). For each channel, the paper derives a testable proposition with an identification strategy, data source, and falsification condition.

The central finding is that GenAI is neither purely beneficial nor inherently destabilizing. The net systemic effect is determined by institutional conditions: the degree of model diversity across the financial sector, the rigor of validation frameworks, and the pace at which regulatory capability evolves relative to industry adoption. Five cross-layer feedback loops constitute the mechanism through which these conditions determine outcomes. Three amplifying loops operate unconditionally from structural properties of GenAI adoption—the model risk–herding–procyclicality spiral, the dependency–interconnectedness–regulatory burden cycle, and the information distortion–mispricing–contagion chain. Two stabilizing loops operate conditionally: risk detection reduces herding and procyclicality only when model diversity is maintained; regulatory surveillance reduces arbitrage only when supervisory capability keeps pace with industry. This asymmetry implies that absent deliberate intervention, amplifying dynamics will dominate—not because the technology is inherently destabilizing, but because the conditions for stability require coordinated action that markets will not produce spontaneously.

The sixteen propositions constitute a structured, falsifiable research agenda whose feasibility is demonstrated by the data sources specified for each—SEC filings, CRSP/Compustat, FFIEC Call Reports, I/B/E/S forecasts, TAQ data, and enforcement databases. The regulatory implications—extended model validation, AI Concentration and Model Diversity Indices, operational resilience requirements, information integrity safeguards—provide a policy architecture grounded in identified causal mechanisms rather than precautionary intuition, calling for evidence-based design informed by progressive empirical testing.

For the management community, the implication is direct: strategic AI governance is not an optional enhancement to technology management but a determinant of whether GenAI strengthens or undermines financial stability. Firm-level adoption decisions aggregate through meso-level dynamics into macro-level consequences that reshape the competitive and regulatory environment. Institutions that treat GenAI governance as a compliance obligation rather than a strategic imperative will find themselves exposed to risks they helped create and poorly positioned to capture the benefits the technology offers.

The institutional choices made in the next several years—by firms governing GenAI deployments, by regulators designing oversight frameworks, and by researchers generating the evidence to guide both—will determine whether generative AI becomes a source of enhanced resilience or an amplifier of the systemic fragilities it was deployed to address.

Data Availability Statement

No datasets were generated or analyzed during this study. All sources are cited in the reference list.

Declarations

Conflict of Interest

Author Jörg Osterrieder serves as lead guest editor of the Springer collection "Generative AI in Finance: Innovation, Risk, and Systemic Stability." This manuscript was handled independently and underwent standard double-blind peer review without involvement of the guest editorial team.

Use of AI-Assisted Tools

The authors used AI-assisted tools during the preparation of this manuscript. The authors reviewed and edited all content and take full responsibility for the publication.

References

Acemoglu D, Ozdaglar A, Tahbaz-Salehi A (2015) Systemic risk and stability in financial networks. American Economic Review 105(2):564–608. https://doi.org/10.1257/aer.20130456

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

Ali H, Zafar MB, Aysan AF (2025) Generative AI in finance: Replicability, methodological contingencies, and future research directions. Finance Research Letters

Arner DW, Barberis J, Buckley RP (2017) FinTech, RegTech, and the reconceptualization of financial regulation. Northwestern Journal of International Law & Business 37:371–414

Bank for International Settlements (2021) Regulating big techs in finance. BIS Bulletin 45, Bank for International Settlements

Bansal G, Nushi B, Kamar E, et al (2021) Does the whole exceed its parts? the effect of AI explanations on complementary team performance. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems https://doi.org/10.1145/3411764.3445717

Barney J (1991) Firm resources and sustained competitive advantage. Journal of Management 17(1):99–120. https://doi.org/10.1177/014920639101700108

Basel Committee on Banking Supervision (2009) Revisions to the Basel II market risk framework. Tech. rep., Bank for International Settlements, Basel

Board of Governors of the Federal Reserve System (2011) SR 11-7: Guidance on model risk management. Tech. rep., Board of Governors of the Federal Reserve System, Washington, DC

Brogaard J, Hendershott T, Riordan R (2014) High-frequency trading and price discovery. Review of Financial Studies 27(8):2267–2306. https://doi.org/10.1093/rfs/hhu032

Brown TB, Mann B, Ryder N, et al (2020) Language models are few-shot learners. In: Advances in Neural Information Processing Systems, pp 1877–1901

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

Callaway B, Sant'Anna PHC (2021) Difference-in-differences with multiple time periods. Journal of Econometrics 225(2):200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Coletta A, Prata M, Conti M, et al (2025) Generation of synthetic financial time series by diffusion models. Quantitative Finance https://doi.org/10.1080/14697688.2025.2528697

Cornelli G, Doerr S, Gambacorta L, et al (2024) Regulatory sandboxes and fintech funding: Evidence from the UK. Review of Finance 28(1):203–233

Crockett A (2000) Marrying the micro- and macro-prudential dimensions of financial stability. BIS Speeches Remarks before the Eleventh International Conference of Banking Supervisors, Basel, 20–21 September 2000

Danielsson J, Embrechts P, Goodhart C, et al (2001) An academic response to Basel II. Special Paper 130, LSE Financial Markets Group

Danielsson J, Macrae R, Uthemann A (2022) Artificial intelligence and systemic risk. Journal of Banking & Finance 140:106290. https://doi.org/10.1016/j.jbankfin.2021.106290

De Bandt O, Hartmann P (2000) Systemic risk: A survey. In: Working Paper No. 35. European Central Bank

Dopfer K, Foster J, Potts J (2004) Micro-meso-macro. Journal of Evolutionary Economics 14(3):263–279. https://doi.org/10.1007/s00191-004-0193-0

Eckerli F, Osterrieder J (2021) Generative adversarial networks in finance: An overview. arXiv preprint arXiv:210606364

Eisenhardt KM, Graebner ME (2007) Theory building from cases: Opportunities and challenges. Academy of Management Journal 50(1):25–32. https://doi.org/10.5465/amj.2007.24160888

Eisfeldt AL, Schubert G (2025) Generative AI and finance. Annual Review of Financial Economics 17:363–393. https://doi.org/10.1146/annurev-financial-112923-020503

European Parliament and Council (2022) Regulation (EU) 2022/2554 on digital operational resilience for the financial sector (DORA). Official Journal of the European Union

European Parliament and Council (2024) Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (AI Act). Official Journal of the European Union

Fama EF (1970) Efficient capital markets: A review of theory and empirical work. Journal of Finance 25(2):383–417. https://doi.org/10.2307/2325486

Financial Stability Board (2025) Artificial intelligence in finance: Systemic implications and regulatory approaches. Tech. rep., Financial Stability Board, Basel

Fleischer V (2010) Regulatory arbitrage. Texas Law Review 89:227–290

Frost J, Gambacorta L, Huang Y, et al (2019) BigTech and the changing structure of financial intermediation. BIS Working Papers 779, Bank for International Settlements

Geels FW (2002) Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. Research Policy 31(8–9):1257–1274. https://doi.

org/10.1016/S0048-7333(02)00062-8

Goodfellow IJ, Pouget-Abadie J, Mirza M, et al (2014) Generative adversarial nets. In: Advances in Neural Information Processing Systems, pp 2672–2680

Gu S, Kelly B, Xiu D (2020) Empirical asset pricing via machine learning. Review of Financial Studies 33(5):2223–2269. https://doi.org/10.1093/rfs/hhaa009

International Monetary Fund (2023) Global financial stability report: AI and financial stability. Tech. rep., International Monetary Fund, Washington, DC

Jaakkola E (2020) Designing conceptual articles: Four approaches. AMS Review 10:18–26. https://doi.org/10.1007/s13162-020-00161-0

Ji Z, Lee N, Frieske R, et al (2023) Survey of hallucination in natural language generation. ACM Computing Surveys 55(12):1–38. https://doi.org/10.1145/3571730

Li H, Guo S, Che H, et al (2024) Large language model agent in financial trading: A survey. arXiv preprint arXiv:240806361

Lo AW, MacKinlay AC (1988) Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1(1):41–66. https://doi.org/10.1093/rfs/1.1.41

Lopez-Lira A, Tang Y (2023) Can ChatGPT forecast stock price movements? return predictability and large language models. arXiv preprint arXiv:230407619

MacInnis DJ (2011) A framework for conceptual contributions in marketing. Journal of Marketing 75(4):136–154. https://doi.org/10.1509/jmkg.75.4.136

Parasuraman R, Riley V (1997) Humans and automation: Use, misuse, disuse, abuse. Human Factors 39(2):230–253. https://doi.org/10.1518/001872097778543886

Prudential Regulation Authority (2023) SS1/23: Model risk management principles for banks. Tech. rep., Bank of England, London

Rogers EM (2003) Diffusion of Innovations, 5th edn. Free Press, New York

Tushman ML, Anderson P (1986) Technological discontinuities and organizational environments. Administrative Science Quarterly 31(3):439–465. https://doi.org/10.2307/2392832

Vaswani A, Shazeer N, Parmar N, et al (2017) Attention is all you need. In: Advances in Neural Information Processing Systems, pp 5998–6008

Vives X (2016) Competition and Stability in Banking: The Role of Regulation and Competition Policy. Princeton University Press, Princeton, NJ

Wei J, Tay Y, Bommasani R, et al (2022) Emergent abilities of large language models. In: Transactions on Machine Learning Research

Wu S, Irsoy O, Lu S, et al (2023) BloombergGPT: A large language model for finance. arXiv preprint arXiv:230317564

Xiao Y, Li MM, Deng Z, et al (2024) TradingAgents: Multi-agents LLM financial trading framework. arXiv preprint arXiv:241220138

Yang H, Liu XY, Wang CD (2023) FinGPT: Open-source financial large language models. arXiv preprint arXiv:230606031

Table 1 Summary of all sixteen channels through which GenAI adoption affects systemic risk in finance.

Layer	Code	Name	Type	Core Mechanism	GenAI-Specific Feature	Prop.
Micro	M1	Model Risk Amplification	Risk	Adverse consequences from decisions based on incorrect or misused model outputs	Hallucination produces fluent, fabricated outputs indistinguishable from legitimate analysis; emergent capabilities defy existing validation	P1
Micro	M2	Operational Dependency & Provider Concentration	Risk	Single points of failure from integrating concentrated GenAI provider APIs into core processes	Two to three frontier providers with no meaningful substitutability, unlike traditional IT outsourcing with dozens of vendors	P2
Micro	M3	Competitive Dynamics & the AI Divide	Risk	Scale advantages in GenAI adoption create competence-destroying discontinuity favoring large institutions	Training costs exceeding $100M and specialized talent requirements produce adoption barriers orders of magnitude larger than prior analytics	P3
Micro	M4	Human Capital Erosion & Automation Bias	Risk	Over-reliance on automation, reduced vigilance, and progressive deskilling among financial professionals	Natural language mimicry of expert reasoning makes GenAI outputs uniquely persuasive and difficult to scrutinize critically	P4
Meso	S1	Algorithmic Monoculture & Herding	Risk	Correlated analytical outputs from institutions deploying the same foundation models or overlapping training data	Monoculture is infrastructural rather than behavioral—embedded in shared model architecture and knowledge representations	P5
Meso	S2	Market Microstructure Transformation	Risk	AI-powered market making, LLM-based order flow prediction, and autonomous trading agents reshape trading dynamics	Agents interpret unstructured information at algorithmic speed, enabling agent-to-agent market dynamics without human oversight	P6
Meso	S3	Information Ecosystem Distortion	Risk	Mass-produced synthetic analyst reports and commentary degrade the signal-to-noise ratio for price discovery	Generative models fabricate convincing narrative financial analysis at scale; detection arms race favors generation over identification	P7
Meso	S4	Interconnectedness via Shared Tech. Infrastructure	Risk	Shared GenAI provider APIs create a technological network overlay invisible to conventional financial network analysis	Proprietary, non-standardized, largely unregulated API dependencies produce correlated failure through operational channels	P8
Macro	L1	Procyclicality Amplification	Risk	GenAI models trained on historical data encode and exploit cyclical patterns more aggressively than traditional models	Superior pattern matching at algorithmic timescales amplifies procyclical bias faster than human decision-making cycles allow	P9
Macro	L2	Regulatory Arbitrage Acceleration	Risk	LLMs parse regulatory text across jurisdictions to identify gaps and suggest structuring approaches at unprecedented speed	Enables compliance theater—automated reporting satisfying the letter of requirements while systematically undermining the spirit	P10
Macro	L3	Too-Connected-to-Fail AI Infrastructure	Risk	Catastrophic failure of a dominant GenAI provider simultaneously disrupts hundreds of dependent financial institutions	Natural monopoly dynamics in foundation model training foreclose competitive entry; providers operate outside financial supervisory perimeters	P11
Macro	L4	Cross-Border Regulatory Fragmentation	Risk	Jurisdictional mismatch between globally uniform GenAI APIs and nationally bounded regulatory perimeters	Globally identical API outputs enable routing AI-dependent activities to least restrictive jurisdictions with minimal adaptation cost	P12
Benefit	B1	Enhanced Risk Detection & Early Warning	Benefit	Non-linear pattern recognition in high-dimensional data and synthetic stress scenario generation improve risk identification	LLMs analyze unstructured data for early warning signals that elude traditional quantitative models	P13
Benefit	B2	Financial Inclusion via Democratized Analytics	Benefit	API-based GenAI tools lower barriers to sophisticated financial analysis for smaller institutions and underserved populations	Capabilities previously requiring large quantitative teams become accessible with modest technical requirements	P14
Benefit	B3	Improved Regulatory Surveillance	Benefit	GenAI-based RegTech and SupTech improve compliance accuracy, detection rates, and monitoring scale	LLM surveillance tools analyze market communications and filings for manipulation signals at previously infeasible scale	P15
Benefit	B4	Enhanced Market Efficiency	Benefit	Faster, more comprehensive processing of diverse information sources improves price discovery and allocative efficiency	Simultaneous integration of textual, numerical, and sentiment-based information at speeds exceeding human analyst capacity	P16

Table 2 Summary of sixteen testable propositions with identification strategies, data sources, and falsification conditions.

#	Type	Statement	Ch.	ID Strategy	Data Source	Falsification
P1	Risk	Institutions adopting GenAI credit scoring exhibit higher unexpected loss rates during downturns than matched non-adopters	M1	DID comparing GenAI-adopting and non-adopting banks around recessionary shock	FFIEC Call Report charge-offs and loan loss provisions	GenAI adopters show equivalent or lower unexpected loss rates during downturns
P2	Risk	Higher GenAI provider concentration among financial institutions is associated with larger aggregate operational losses during outage events	M2	Event study centered on documented API outages	OCC and FDIC operational loss data; provider dependency disclosures	Operational losses show no systematic relationship with provider concentration
P3	Risk	GenAI adoption widens the Gini coefficient of banking profitability, with adopters capturing disproportionate profit share	M3	Panel DID following Callaway and Sant'Anna (2021) for staggered adoption	FFIEC performance data	Profitability distribution narrows or remains unchanged following adoption
P4	Risk	Analyst teams using GenAI tools exhibit decreased forecast dispersion without corresponding improvement in accuracy	M4	Staggered DID comparing teams before and after GenAI tool adoption	I/B/E/S analyst forecasts	Dispersion declines are accompanied by proportional accuracy gains
P5	Risk	Investors relying on the same foundation model exhibit higher pairwise return correlations during stress than those using diverse models	S1	DID instrumented by exogenous variation in foundation model adoption	13-F holdings and SEC technology disclosures	Return correlations during stress are unrelated to foundation model commonality
P6	Risk	Higher GenAI-based trading volume is associated with increased intraday volatility and reduced market-maker diversity	S2	Cross-market comparison exploiting differential GenAI adoption rates	TAQ and LOBSTER order-level data	Markets with higher GenAI trading show equivalent or lower volatility
P7	Risk	Higher proportion of AI-generated analyst reports is associated with increased return volatility and decreased post-earnings announcement drift	S3	RDD exploiting sharp increase in LLM capability around Nov. 2022	I/B/E/S report data and CRSP returns	AI-generated content prevalence shows no relationship with volatility
P8	Risk	Higher correlation in GenAI API usage across institutions predicts elevated tail-risk co-dependence	S4	Time-series DCC models; API usage proxied by technology spending	CRSP returns	Tail-risk co-dependence is orthogonal to API usage correlation
P9	Risk	GenAI credit risk models produce larger forecast error variance across business cycle phases than traditional models	L1	Comparative backtesting across complete credit cycles	Moody's and S&P rating transition and default data	GenAI models exhibit equivalent or lower cross-cycle forecast error variance
P10	Risk	Institutions deploying GenAI compliance tools receive fewer enforcement actions per unit of risk exposure	L2	PSM of GenAI-adopting and non-adopting institutions	FDIC and OCC enforcement action databases	No association with enforcement frequency after controlling for risk
P11	Risk	Higher HHI of AI infrastructure provision Granger-causes higher systemic risk measures (CoVaR, SRISK)	L3	Granger causality framework	Industry surveys; NYU V-Lab systemic risk measures	Provider concentration has no predictive power for systemic risk
P12	Risk	Greater divergence in GenAI governance stringency is associated with increased cross-border regulatory arbitrage	L4	Panel analysis exploiting variation in regulatory adoption timing	GenAI regulatory stringency index; BIS locational banking statistics	Cross-border activity unrelated to regulatory stringency differentials
P13	Benefit	Institutions deploying GenAI risk monitoring detect credit deterioration signals earlier than matched institutions	B1	Matched-pair design comparing detection lead times	Rating agency credit transition data	No improvement in early detection relative to matched controls
P14	Benefit	Introduction of GenAI advisory tools is associated with increased retail market participation among underrepresented groups	B2	DID exploiting staggered platform-level rollouts	Survey of Consumer Finances; brokerage account microdata	Participation rates unchanged or decline following introduction
P15	Benefit	Regulatory agencies deploying GenAI surveillance achieve higher manipulation detection rates with shorter lags	B3	Before-and-after comparison at agencies adopting GenAI tools	SEC and FCA enforcement data	Detection rates and lag times unchanged or worsen
P16	Benefit	Higher GenAI trading participation is associated with faster price convergence toward fundamental value	B4	Variance ratio tests following Lo and MacKinlay (1988)	TAQ data	Variance ratios show no improvement or deteriorate

Table 3 Regulatory design implications mapped to framework channels.

Domain	Linked Channels	Current Gap	Recommendation
Model Governance & Validation	M1, B1	SR 11-7 and Basel model governance frameworks address parametric estimation errors but lack vocabulary for hallucination—the generation of fluent, fabricated outputs undetectable by conventional backtesting	Extend validation to include semantic verification against authoritative sources, adversarial testing for prompt injection and data poisoning, financial GenAI benchmark suites measuring hallucination rates, and ongoing monitoring of emergent capabilities
Systemic Risk Monitoring	S1, S4, L1, L3	CoVaR and SRISK capture balance sheet and market-based correlations but miss technological concentration—correlation from shared model architectures, overlapping training data, and common provider dependence	Develop an AI Concentration Index measuring provider dependence and a Model Diversity Index measuring architecture heterogeneity; mandate disclosure of GenAI provider relationships, model architectures, and training data characteristics
Operational Resilience	M2, S4, L3	DORA assumes meaningful vendor substitutability that does not hold for frontier GenAI providers with natural monopoly dynamics and training costs exceeding hundreds of millions of dollars	Mandate multi-provider strategies for systemically important GenAI applications, establish model portability standards to reduce switching costs, and require failure scenario testing for GenAI-specific contingencies including simultaneous multi-institution disruption
Information Integrity	S3, B4	Existing surveillance tools are calibrated for insider trading, spoofing, and layering but cannot detect synthetic analyst reports designed to move asset prices—a novel manipulation form enabled by generative models	Mandate labeling of AI-generated financial content by regulated entities, develop supervisory detection capabilities matching generative model capacity, and adapt the EU AI Act transparency requirements specifically for financial market contexts
Proportionate & Adaptive Regulation	L4, L2, B2, B3	Globally uniform GenAI APIs face nationally bounded regulatory perimeters, enabling cross-border arbitrage with minimal adaptation cost; no harmonized minimum standards exist for GenAI in financial services	Calibrate scrutiny to systemic significance of specific GenAI applications, pursue international harmonization of minimum GenAI governance standards through FSB and BIS frameworks, and deploy regulatory sandboxes to observe emergent behaviors while balancing innovation and stability