Networks and Heterogeneous Media

NHM, 21(3): 1041–1068.
DOI: 10.3934/nhm.2026043
Received: 16 March 2026
Revised: 12 May 2026
Accepted: 13 May 2026
Published: 20 May 2026

https://www.aimspress.com/journal/NHM

Review

Agent-Based modeling in financial markets: Modeling frameworks, validation challenges, and emerging applications

Hanool Choi[1] **and Sunghee Choi**[2,*]

[1] Department of Management Information System, Keimyung University, 1095, Dalgubeol-daero, Daegu 42601, Republic of Korea
[2] Department of International Commerce, Keimyung University, 1095, Dalgubeol-daero, Daegu 42601, Republic of Korea

* **Correspondence**: Email: choisu@kmu.ac.kr; Tel: +82-535805434; Fax: +82-535805313.

Abstract: As financial markets have increasingly exhibited heterogeneity across participants, complex interactions among them, and out-of-equilibrium dynamics, have provided growing interest in agent-based modeling (ABM). Unlike traditional representative-agent models, ABMs generate macro-level market behavior from the bottom up through interactions among adaptive agents. In this paper, we review the development and applications of ABMs in financial markets and provide a structured perspective on the growing literature. Specifically, we reviewed three key design axes—agent heterogeneity, market mechanism fidelity, and interaction topology—and discussed how different modeling assumptions shape the ability of ABMs to reproduce stylized financial market phenomena. We also discussed two emerging domains in which ABMs are increasingly applied: Simulations using LLM-based agents and modeling approaches in green energy finance and climate-related financial systems. Here, we identified the current boundaries of financial ABM research and promising directions for future research.

Keywords: agent-based modeling; financial markets; heterogeneous agents; interaction networks; systemic risk; model validation; LLM-based agents; green finance; climate-related financial systems

1. Introduction

Modern financial theory is built on a coherent set of assumptions. Rational agents optimize, markets clear, and prices reflect available information. These premises form the shared foundation of the Efficient Markets Hypothesis, the Capital Asset Pricing Model, and the Black-Scholes framework alike. From this foundation, researchers derived an influential analytical toolkit, including stochastic calculus, fixed-point theorems, and linear factor models. For half a century, these tools have shaped academic inquiry and industry practice.

The same assumptions that make these models tractable, however, can limit their ability to represent several empirical regularities of financial markets, including fat-tailed return distributions, volatility clustering, excess trading volume, persistent deviations from fundamental value, and sudden crashes [1,2]. In representative-agent equilibrium frameworks, such phenomena are often handled only indirectly, such as through exogenous shocks or reduced-form frictions. The 2008 global financial crisis highlighted these limits, especially for models not designed to capture leverage spirals, liquidity freezes, and cascading defaults [3]. As Farmer and Foley [3] argued, equilibrium-based approaches are often not well suited to nonlinear feedback, path dependence, and network amplification. That critique does not mean such models are uninformative for all purposes; rather, it underscores that different modeling traditions are useful for different questions.

Unlike a representative-agent model, agent-based modeling (ABM) starts from a different set of premises. Rather than solving for equilibrium, ABM populates a system with heterogeneous autonomous agents, equips each with behavioral rules and interaction protocols, and lets macroscopic dynamics emerge through computation. The resulting framework draws on the mathematics of complex systems, such as nonlinear dynamical systems, stochastic processes, network theory, percolation, and evolutionary game theory. ABM is, in this sense, not simply a simulation method but a modeling paradigm suited to the inherent complexity of financial markets. Financial markets can be represented as networks of heterogeneous agents linked through trading relationships, information flows, and balance sheet exposures, with prices, liquidity, and systemic risk arising from local interactions. The percolation threshold $q \cdot \bar{k} > 1$ governing contagion cascades in interbank networks [4], the spectral radius condition $\lambda_{max}(M) > 1$determining instability in overlapping portfolio systems, and the Hawkes process structure capturing order-flow clustering in limit order books—each reflects tools from network science and stochastic process finding natural application in financial ABMs.

The intellectual roots of this approach run from Herbert Simon's bounded rationality and the Santa Fe Institute's complex economics program through to the policy-oriented ABMs used at some central banks and regulatory institutions. Many scholars interpret this development as a reconceptualization of the economy as a complex adaptive system rather than a mere technical extension [5–7]. The publication of Axtell and Farmer [8] in the Journal of Economic Literature also signals that ABM is part of a mainstream methodological discussion. Whether or not one adopts the stronger language of a paradigm shift, the practical case for ABM rests on its ability to represent heterogeneity, feedback, and interaction mechanisms that many other models treat only in reduced form. This is especially relevant to this special issue's focus on advanced mathematical methods in financial mathematics.

In this paper, we review agent-based modeling in financial markets. We focus on three areas: how ABMs are designed and what mathematics underlies those choices, the domains in which this approach has been especially informative, and the question of when simulation results are credible enough to inform policy. Surveys have provided valuable overviews [9–12], so our aim is not to replace them with a wholly new taxonomy. Rather, we use a three-axis design space—agent heterogeneity, market mechanism fidelity, and interaction topology—as an organizing device for comparing model choices and trade-offs. We also discuss two emerging but methodologically uneven areas: LLM-based agent simulation in economic contexts [13–15] and ABM applied to green finance and climate-related financial risk [16]. A further contribution concerns validation. The literature has long conflated calibration with validation, although these are distinct activities that require different standards of evidence. We therefore ask what counts as adequate evidence for models built for explanation, forecasting, or policy analysis, and we connect that discussion to estimation and inference practices in mathematical finance.

The paper is organized as follows. In Section 2, we establish the ABM paradigm, tracing its roots in bounded rationality, heterogeneous agent theory, and complexity economics, and introduce the three-axis design space. In Section 3, we develop the mathematical tools associated with each axis. In Section 4, we survey four major application domains, including market microstructure and systemic risk, regulatory policy evaluation, LLM-based agent simulation, and green finance. In Section 5, we address validation and connect it to formal calibration methods. In Section 6, we conclude with a critical assessment and research agenda.

2. ABM as a model framework in the financial markets

2.1. Core principles of the agent-based approach

2.1.1. Bottom-up architecture

Modern finance frameworks are built on assumptions such as rational optimization, market clearing, and informational efficiency. These assumptions have produced powerful tools for understanding pricing, portfolio choice, and comparative statics, but they do not capture every empirically relevant feature of financial markets. Excess volatility, unresolved asset-pricing puzzles, and heavy-tailed return distributions have therefore motivated complementary approaches [9]. In that context, agent-based modeling is useful as an alternative framework for studying anomalies and interaction-driven dynamics that remain difficult to address in more aggregated settings.

The central distinction of agent-based modeling from standard modern finance frameworks is architectural rather than normative. Traditional models typically begin with equilibrium conditions and derive aggregate implications from them. ABMs instead begin with a population of autonomous agents, specify behavioral rules and interaction protocols, and generate aggregate outcomes through computation. As Epstein [17] argued, “If you didn’t grow it, you didn’t explain its emergence.” This bottom-up logic does not make ABM universally superior, but it is especially helpful when the object of study involves adaptation, out-of-equilibrium adjustment, or emergent interaction effects. Arthur [6] likewise emphasized that the economy can be viewed as continually formed by evolving actions, strategies, and beliefs rather than as a fully specified static object.

2.1.2. Heterogeneity of agents

The treatment of agent diversity is a key criterion for distinguishing representative-agent models from agent-based models. As a representative-agent framework in macroeconomics, the Dynamic Stochastic General Equilibrium model aggregates market participants into a single optimizing entity. Kirman [18] argued that this type of aggregation misrepresents how heterogeneous populations behave. More broadly, Arrow emphasized that heterogeneity of expectations is one important reason why trade occurs in markets [10].

Agent-based models take this observation as a design principle. Agents differ in their information sets, forecasting rules, risk preferences, time horizons, wealth, and learning mechanisms, respectively. LeBaron [9] argued that dynamic heterogeneity is the definitive feature setting ABM apart from simpler models with fixed agent types. The distribution of agents and wealth across strategies changes continuously driven by learning, imitation, and evolutionary selection. This endogenous shift in the agent composition generates dynamics that equilibrium models cannot produce. Hommes [10] showed that even simple heterogeneous agent models can reproduce the full suite of stylized facts, which Cont [1] documented as universal properties of asset returns. They also include excess volatility, high trading volume, temporary bubbles, sudden crashes, clustered volatility, and fat-tailed return distributions.

2.1.3. Bounded rationality

Bounded rationality, introduced by Simon [19,20] as a challenge to the assumption of perfect rationality, argued that rational behavior must be compatible with agents' actual information and computational limits. Simon's contribution provided much of the philosophical justification for ABM. His central claim was that agents satisfice rather than optimize: They search for solutions that meet aspiration thresholds instead of maximizing a utility function.

Gigerenzer [21,22] provided the ecological argument for simple rules. This line of research argues that simple heuristics are not merely acceptable compromises but can outperform complex optimization strategies in uncertain environments. The idea of ecological rationality holds that a heuristic's value depends on its fit to the decision environment, not on computational sophistication alone. This "less-is-more" effect has direct implications for financial ABMs, where simple trend-following rules may at times outperform agents attempting full Bayesian updating.

Kahneman and Tversky's prospect theory [23] added another layer by showing that decisions depend on gains and losses relative to reference points rather than on absolute wealth alone. In financial contexts, loss-averse agents may hold losing positions too long, while overconfident agents may trade too frequently. Availability bias can likewise produce overreaction to salient news. These mechanisms reinforce the point that financial decision makers are not perfectly rational in the neoclassical sense.

2.1.4. Agent adaptive behavior: Learning, evolution, and strategy switching

If bounded rationality specifies what agents cannot do, adaptive behavior specifies what they can. Financial ABMs employ a spectrum of learning mechanisms. At the most complex end, genetic algorithms enable open-ended strategy innovation: In the Santa Fe Artificial Stock Market [24], each agent maintains a set of candidates' forecasting rules that a genetic algorithm periodically evaluates, recombines, and mutates based on realized accuracy. At an intermediate level, the evolutionary switching mechanism of Brock and Hommes [25] governs selection among a fixed set of strategy types such as fundamentalist, chartist, and contrarian through a discrete-choice model driven by past profitability. The intensity of choice parameter β controls how aggressively agents switch strategies. Kirman's [26] stochastic herding mechanism governs random switching between opinion states, driven by the tension between individual exploration and social conformity. Reinforcement learning, neural network forecasting, and deep reinforcement learning provide additional adaptive mechanisms.

The key tradeoff across these approaches is tractability versus behavioral richness. Models with a small number of strategy types and simple switching rules remain amenable to bifurcation analysis, while models with open-ended learning typically require pure simulation and can generate richer emergent dynamics. Whether such dynamics are more realistic, however, depends on the empirical context and on how the model is validated; for example, Di Domenico et al. [27] showed that reinforcement learning may converge toward relatively simple heuristics in a large-scale macroeconomic ABM.

2.2. A Three-dimensional approach to ABM

2.2.1. Navigating the choices that define an agent-based financial model

Building an agent-based model of a financial market requires a series of design choices. How heterogeneous should the agents be? How faithfully should the market mechanism be represented? What interaction structures connect agents to one another? We use three broad dimensions—agent heterogeneity, market mechanism fidelity, and interaction topology—as a parsimonious way to organize these choices. A model's position along each dimension affects the phenomena it can capture, the analytical tools that remain available, and the computational burden it imposes.

Our framework builds on earlier classification efforts in the literature rather than claiming to replace them. Chen et al [12] provided a multidimensional classification of financial ABMs organized by agent types, preference specifications, adaptive mechanisms, and stylized facts reproduced. LeBaron [9] organized his survey around design issues such as market mechanism, agent strategies, and learning parameters. Fagiolo and Roventini [11] identified ten ingredients of economic ABMs, including heterogeneity, bounded rationality, direct interactions, and network structures. The three-axis scheme used here is therefore intentionally incremental: It is meant as a compact map for comparing design trade-offs and situating models relative to existing work.

2.2.2. Three dimensions: Agent heterogeneity, market fidelity, and interaction topology

The first axis captures the degree and nature of agent differentiation. At one end are zero-intelligence agents [28], which generate random decisions subject to budget constraints. Moving along the axis, discrete heterogeneity introduces a small number of behavioral types such as fundamentalists, chartists, and noise traders with evolutionary switching mechanisms. At the far end, deep heterogeneity gives each agent a unique and continuously evolving cognitive apparatus, as in the Santa Fe Artificial Stock Market [24]. The substantive question is not simply whether more heterogeneity is better, but when additional heterogeneity changes explanatory power rather than merely adding complexity. Simpler heterogeneity preserves bifurcation analysis and statistical-mechanics treatments, whereas deeper heterogeneity can expand behavioral realism at higher computational and inferential cost. We develop the formal switching mechanisms and their mathematical properties in Section 3.1.

The second axis captures how faithfully the model represents the institutional mechanics of price formation. At the lowest end of fidelity, a Walrasian auctioneer clears markets by equating aggregate supply and demand. Intermediate models introduce price-impact functions and market-maker mechanisms. At the highest end of fidelity, full limit order book models replicate the actual mechanics of electronic exchanges, including limit orders, market orders, and price-time priority matching. The relevant question, again, is purpose: Simple clearing rules may be sufficient for studying broad price dynamics, whereas flash crashes, bid-ask spread dynamics, and high-frequency trading effects require explicitly microstructural mechanisms. In Section 3.2, we present mathematical tools for order-book dynamics, including Hawkes processes for order arrival and diffusion approximations for price formation.

The third axis captures the structure of interactions among agents. At the low-complexity end, mean-field interaction has agents influencing one another only indirectly through aggregate prices.

Moving along the axis, lattice and grid-based topologies import spatial structure from statistical physics. At the high-complexity end, empirically calibrated network structures such as random graphs, small-world networks, and scale-free networks capture the heterogeneous connectivity patterns observed in real financial systems. This dimension governs how information, opinions, and financial distress propagate through the population, and it has received greater research attention since the 2008 global financial crisis, especially because the crisis highlighted the systemic consequences of interconnection. We discuss percolation thresholds, contagion matrix analysis, and related results in Section 3.3.

2.2.3. Interpreting and navigating the design space

The three-axis design space is intended as a conceptual map for classifying models and guiding the construction of new ones. It is not exhaustive, nor is it a claim that every important ABM contribution can be reduced to three coordinates. One consistent pattern in the literature, however, is that most models concentrate near the edges of this space. They tend to be complex along one or two axes while remaining simple along the others. This reflects computational constraints and the scientific need to isolate causal mechanisms. A model that is simultaneously complex along all three axes would be difficult to calibrate, expensive to simulate, and difficult to interpret.

The framework also highlights how the field has evolved. Financial ABMs of the 1990s, such as the Santa Fe ASM, Lux-Marchesi, and Brock-Hommes models, emphasized agent heterogeneity while keeping market mechanisms and network structure relatively sparse. The 2000s brought greater attention to market-mechanism fidelity as limit order book models gained prominence, and the 2010 flash crash sharpened interest in microstructure. After 2008, the interaction-topology axis gained visibility as systemic-risk research expanded. Researchers increasingly combine heterogeneous learning, more realistic trading mechanisms, and empirically informed network structure, although not every application benefits equally from pushing all three axes at once.

Each axis poses a fundamental question. The heterogeneity axis asks how much cognitive and behavioral diversity is needed to generate the phenomenon of our interest. The market mechanism fidelity axis asks at what level of institutional details the relevant price formation operates. The topology axis asks through what channels and with what structure, agents influence one another. The answers to these three questions, taken together, define the model. In Section 3, we develop the formal tools corresponding to each axis in turn.

3. Core modeling approaches

In this section, we use formal elements as comparative tools rather than as ornaments of rigor. Some structures, such as low-dimensional switching rules, local stability conditions, or percolation thresholds, support analytical results and clearer falsifiability. Others, such as richly adaptive agents or high-fidelity market simulators, are primarily useful within simulation-based inquiry. Keeping this distinction explicit matters for inference and policy use: Added mathematical detail is valuable only when it clarifies mechanisms, improves empirical discipline, or changes the model's explanatory reach.

3.1. Agent heterogeneity

Agent-based models accommodate a spectrum of agent heterogeneity, ranging from fully homogeneous populations to richly differentiated ones [29]. Unlike standard mainstream economics, which often assumes homogeneous agents, ABMs incorporate diversity to reflect markets where participants follow different strategies [9,18]. Rather than excluding homogeneous agents, ABMs can use them as a useful baseline at one end of the spectrum, although laboratory evidence frequently rejects the null hypothesis of strictly uniform behavior [30]. At the other end, fully heterogeneous agents are endowed with individualized characteristics such as unique risk tolerance, distinct information sets, and varied decision rules, enabling the model to capture reaction speeds and decision patterns that more closely match real-world observations [10].

3.1.1. Fundamentalist–Chartist paradigm

A common approach introduces a small number of distinct behavioral types. In many financial market ABMs, agents are divided into fundamentalists, who trade based on intrinsic value signals, and chartists, who follow trends or technical analysis [25,31]. The fundamentalist–chartist split has been used to generate realistic price dynamics including bubbles and crashes. Moreover, where agents switch between strategies, the model can reproduce volatility clustering and other stylized facts [10,25]. When trend followers dominate, momentum builds, and prices deviate far from fundamentals; fundamentalists then provide a stabilizing counterforce [32].

The strategy switching mechanism is formally captured by the Brock–Hommes evolutionary selection rule. The fraction of agents adopting strategy h at time $t + 1$ is updated according to the following expression.

$$n_{h,t+1} = \exp\left[\beta\Pi_{h,t}\right] / \sum_{j} \exp\left[\beta\Pi_{j,t}\right]$$

where $\Pi_{h,t}$ denotes the recent profit of strategy h.

h and β are 'intensity of choice' [25]. The logit type formula means that strategies with higher past profits attract more agents, introducing nonlinearity and potential chaos into price dynamics. When β is large, even a simple two strategy model can exhibit bifurcations and chaotic price fluctuations as agents flip en masse between strategies, a phenomenon Brock and Hommes termed the 'rational route to randomness.' By grounding agent composition in past market performance rather than fixed assumptions, the formulation provides a rigorous basis for understanding why the mix of behavioral types shifts endogenously over time.

3.1.2. Time-varying trader types

ABM can incorporate multiple categories of agents drawn from empirical trader behavior. Liquidity providers, high frequency arbitrageurs, noise traders, and institutional versus retail investors all operate under distinct rules [3,9]. Using richer trader categories lets the model probe how these groups interact. Studies have shown that mixing several behavioral types improves the realism of simulated market patterns [33]. Each category can be assigned different goals, and the proportions of types can be calibrated to real market data.

A further dimension of heterogeneity is whether agents' decision rules are static or change over time. Static agents follow a predetermined strategy throughout the simulation [34]. To capture learning and adaptation, many ABMs incorporate agents whose rules evolve with experience or performance [24].

Evolutionary and genetic algorithms (GAs) update trading strategies by treating them like genes that can mutate and recombine. In the Santa Fe Artificial Stock Market, for example, agents maintain forecasting rules encoded as binary strings and periodically update them through selection, crossover, and mutation. This enables the distribution of strategies to change endogenously with market feedback [24].

In formal terms, a GA proceeds through iterative stages, population initialization, fitness evaluation, selection, crossover, and mutation, before advancing to the next generation. Rules are ranked by a fitness function such as profitability, and the strongest rules are recombined to produce new variants. Under broad conditions, expected average fitness tends to rise across generations [35]. GA-driven heterogeneity therefore keeps agent behavior fluid and enables the distribution of strategies to change over time, creating genuinely open-ended dynamics.

3.2. Market mechanism fidelity

3.2.1. Stylized market clearing

Many early agent-based models use simplified price setting rules to clear the market at each time step [25]. Classic models like Kirman's herding model or Lux and Marchesi [31] use stylized price updates rather than explicit trade by trade simulation, preserving analytical tractability at the cost of abstracting away intraday trading mechanics. Such stylized price update rules can nonetheless be expressed with mathematical precision. The Lux–Marchesi (1999) formulation links the return to a weighted excess demand from fundamentalists and chartists, as shown below.

$$(P_{t+1} - P_t)/P_t = \alpha\left[N_f(V - P_t) + N_c(P_t - P_{t-1})\right] + \sigma\eta_t$$

where $N_f$and N_c are the numbers of fundamentalist and chartist traders, V is the fundamental value, $(P_t - P_{t-1})$ serves as a trend signal for chartists, and η_t is a noise term. Price rises when there is excess demand and falls under excess sell pressure. The difference equation lends itself to stability analysis. One can derive conditions under which too high an α or too many trend followers N_c destabilizes the fixed point $P_t = V$ and eigenvalues of the linearized dynamics reveal when heterogeneous beliefs lead to cycling or bifurcations to chaos.

3.2.2. Order book dynamics

Many modern ABMs incorporate a continuous double auction mechanism with a limit order book (LOB) to more closely replicate actual market conditions [36,37]. The LOB system yields emergent bid-ask spreads, order book depth, and other microstructural features. It also enables us to model order cancellation, liquidity droughts, and price impact.

Orders in the market are often modeled as Poisson or Hawkes processes to capture clustering. Denoting λ_buy(t) as the intensity of buy orders at time t, a Hawkes process takes the following form:

$$\lambda^{\text{buy}}(t) = \mu + \sum_{t_i<t} \phi(t - t_i)$$

where past orders increase future intensity via kernel φ [38]. This formulation helps capture the volatility clustering that ABMs often produce. When order-flow utilization approaches 1 (analogous to an M/M/1 queue at capacity), the order book becomes susceptible to instability, which is consistent with ABM-based studies of flash crashes.

Equilibrium price formation in a LOB is defined implicitly by the aggregate supply–demand condition $D^{\text{buy}}(P_t^*) = D^{\text{sell}}(P_t^*)$. Under certain simplifying assumptions, one can derive diffusion approximations that capture how the flow of individual orders aggregates into emergent price processes. These approximations take the form of an Ornstein–Uhlenbeck process for price deviations, or a random walk with drift driven by order imbalance. Both formulations connect agent-level actions directly to stochastic differential equations. The resulting price processes exhibit well-known statistical properties:

$$P(|r| > x) \sim Cx^{-\mu}$$

where μ is the tail exponent. ABMs reproduce $\mu \approx 3$, consistent with empirical asset returns [31]. Volatility clustering is captured by the autocorrelation of absolute returns $\mathrm{ACF}(|r|, \tau)$. Absolute returns remain positively correlated over long lags. This autocorrelation tends to decay as a power law rather than dying out quickly. The slow decay implies long memory in volatility, which is quantified by a Hurst exponent $H > 0.5$. Thus, these statistical measures link ABM output directly to GARCH and fractional Brownian motion models, which leads to quantitative comparison between simulated and real markets.

3.2.3. High-fidelity market microstructure

Some modern ABMs go beyond LOBs and simulate realistic order types, millisecond timestamps, and heterogeneous trader classes, including HFT agents [39]. One such model, designed to replicate the Flash Crash of May 6, 2010, operates at millisecond resolution and reproduces realistic stylized facts as well as liquidity withdrawal patterns. Such models are useful for testing market structure reforms, though adding these features makes calibration and computation considerably more demanding [3].

3.3. Interaction topology

Interaction topology influences emergent market outcomes in agent-based modeling [40]. Early models assumed that all agents interacted with one another equally, but more recent network-based models introduce more complex and realistic connectivity structures [4]. Whether the topology is random, small-world, or scale-free, the structural arrangement shapes how diffusion and systemic risk propagate [41–43]. For a concise overview of multi-layer network approaches to financial systemic risk, the importance of grounding network structure in real-world data, and open gaps in the literature, see Riccetti [44]. Early theoretical work on financial contagion focused on how network structure affects the propagation of shocks in financial systems. Allen and Gale [45] showed that interbank network topology is a significant factor in contagion among banks and hence in the propagation of financial distress.

We can represent an agent network as an adjacency matrix $A = (a_{ij})$, where agent states adapt by following the update rule:

$$x_i(t+1) = f\big(x_i(t)\big) + \sum_j a_{ij}\, g\left(x_j(t)\right)$$

Here, $f(\cdot)$ is agent i's own update rule and the second term aggregates neighbor influences. The difference equation on a graph can produce cascades that are mathematically similar to epidemic SIR dynamics. This structure lends itself to percolation theoretical analysis.

3.3.1. Information cascades and contagion

When agents are connected in a network, effects propagate through links and can trigger information cascades and contagion. Bornholdt [46] demonstrated that simple agent interactions in a spin-model framework can reproduce intermittent price dynamics and speculative bubbles similar to those observed in real markets. In financial contexts, a cascade may begin when a few investors sell and others follow after observing their neighbors' behavior, even in the absence of new fundamental information. Network topology can amplify or dampen such dynamics. Gai and Kapadia [4] described this as a robust-yet-fragile property: More interbank links can reduce the probability of isolated failures in normal times, yet a shock to a structurally important node can propagate system-wide.

Percolation theory offers the mathematical basis for this phenomenon. If we define q as the probability that a given link transmits a failure and $\bar{k}$ as the average network degree, a classic result shows that if $q \cdot \bar{k} > 1$, a giant cascade is likely to happen. If $q \cdot \bar{k} < 1$, contagion stays limited [4]. ABM experiments exhibit this threshold behavior. As leverage or connectivity surpasses a certain critical value, the number of defaulting entities can increase discontinuously from about 0% to about 100%. This jump reflects a phase transition in the fixed-point solution of the default fraction.

We can apply this framework to systemic risk by simulating overlapping portfolios or interbank networks. This yields a contagion matrix M, where M_{ij} represents the fractional capital loss imposed on institution i by the default of institution j. Whether a cascade unfolds depends on the spectral radius of M. If $\lambda_{max}(M) > 1$, one default can trigger more than one subsequent default, thereby leading to the entire systematic collapse. We can express this through the fixed-point equation for the default set X given initial shock H.

$$X = (I - M)^{-1}H = H + MH + M^2H + \cdots$$

The Neumann series diverges when $\lambda_{max}(M) \geq 1$, which corresponds to the phase transition from partial to full cascade observed in ABM simulations. We can extract M from simulated portfolio data and compute λ_{max} as a quick indicator of systemic risk. This measure complements what ABM simulations provide.

3.3.2. Spatial structure and local interactions

Agents interact with neighbors in a spatial environment such as a grid or a geographic map. Diffusion of epidemics or innovations and segregation dynamics can be studied using this topological structure. ABMs based on spatial network structures often exhibit local cluster effects [47]. Schelling's segregation model is a well-known illustration. In that model, agents of two types occupy a lattice and decide whether to move depending on local neighborhood composition; distinct clusters then emerge from local interactions without any global coordination.

Spatial structure has been used to study phenomena such as regional housing markets or the spread of information through local contacts. This can produce regional investment fads or bank runs restricted to certain areas. Such topological structure constrains interactions, leading to heterogeneous outcomes across the system. The policy implications can be significant. In a contagion model in the banking industry, adding network structure can reveal how a small cluster of tightly interlinked banks may serve as a systemic risk amplifier, even when the broader system appears stable. When we incorporate a topological structure, ABMs enable us to analyze cascading failures, diffusion patterns, and cluster formation that would remain invisible in a mixed model. We can then vary network topology to study how robustness and efficiency trade across market structures.

4. Application domains

4.1. Market structure and systemic risk

4.1.1. Background and motivation

Agent-based modeling provides one useful framework for studying financial market microstructure and its links to systemic risk because it represents markets as complex adaptive systems with heterogeneous traders and endogenous interactions [3]. This makes ABM well suited to examine how local trading behavior can scale into broader market instability, although not every systemic-risk question requires the same level of microstructural detail.

Within this setting, market microstructure refers to the rules and processes by which securities are traded, including continuous double auctions via a limit order book (LOB). ABMs explicitly model these mechanics, making it possible to study liquidity provision and withdrawal, order-flow imbalances, and failures of price discovery under stress. For questions in which the mechanism of instability is explicitly microstructural, such models function as an in silico "wind tunnel" for evaluating alternative trading environments and sources of fragility [36].

4.1.2. Mechanisms of instability

High-frequency trading (HFT) and algorithmic interactions can be an important source of microstructure instability. Kårvik et al. [48] developed a calibrated LOB model with heterogeneous trading frequencies and found that increasing the relative presence of HFT agents makes flash-instability events more frequent, as HFT algorithms amplify price declines by withdrawing liquidity procyclically. This aligns with empirical evidence from the May 6, 2010 Flash Crash [49] and with the documentation of thousands of ultrafast "mini" flash crashes in equity markets between 2006 and 2011 [50].

The role of HFT is nevertheless nuanced. While ABM research confirms that HFT proliferation can increase flash-crash frequency through procyclical order submission, Leal and Napoletano [51] found that HFT agents are among the first to resume trading after a crash, providing liquidity that supports mean reversion. This duality, HFT as a source of fragility and a contributor to rapid recovery, points to a stability–resilience tradeoff that is difficult to capture without simulation.

ABMs have also been used to reconstruct crash events. Gao et al. [39] developed a high-fidelity ABM calibrated to millisecond-level E-mini S&P 500 data and simulated the 2010 Flash Crash by inserting a single large institutional sell algorithm. Their analysis revealed nonlinear relationships between crash severity and market-maker inventory capacity, selling aggressiveness, and the reaction speed of fundamental traders, suggesting that liquidity evaporation, rather than order size alone, was the main driver of the simulated dislocation. More broadly, many ABM studies show that the most severe price breaks are preceded by gaps in the order book and that market microstructure can enter a qualitatively different regime under stress.

Beyond single-market crashes, Paulin et al. [52] combined an intraday LOB simulation with a network model of institutions holding overlapping portfolios, showing how a flash crash in one asset can trigger margin calls and fire sales in others via a "margin spiral" mechanism [53]. Their hybrid ABM showed that systemic risk depends critically on portfolio crowding and algorithmic trading behavior, while suggesting that extremely high portfolio overlap can, in some cases, slow contagion by prompting coordinated regulatory responses.

Network-based analyses also suggest that greater financial connectivity does not necessarily improve stability. Battiston et al. [54] demonstrated that increased connectivity may initially enhance risk sharing but can ultimately amplify systemic risk, particularly when shocks propagate through tightly coupled financial networks.

4.2. AI agents for LLM-based economic simulation

4.2.1. From rule-based to LLM-based agents

Traditional ABM agents, whether rule-based heuristics, boundedly rational strategies, or simple reinforcement learners, follow predefined logic that limits their ability to handle complex and open-ended scenarios. They represent heterogeneity through manually specified parameters and usually lack flexible reasoning or natural-language interaction. Large language models (LLMs) offer a potentially different paradigm, but the evidence remains preliminary. Because they draw on large pre-trained corpora, LLM-based agents can engage in context-aware reasoning, maintain conversational state within a simulation, communicate through natural language, and adopt persona-specific behaviors through prompting, all without explicit programming of every behavior.

A growing body of work from 2022 to 2024 has explored whether LLMs can serve as stylized stand-ins for human subjects in economic experiments. Chen et al. [55] examined whether large language models exhibit economically interpretable decision patterns in standard choice tasks involving risk and social preferences. Brand et al. [56] studied whether LLM-generated responses can reproduce qualitative patterns observed in human survey data, suggesting that such models may be useful for exploratory market-research applications. Horton [13] found that ChatGPT, when prompted with classic behavioral-economics scenarios, reproduced patterns resembling fairness motives in dictator games and status quo bias. These findings are suggestive, but they should be interpreted as early evidence rather than as definitive validation.

LLM agents also exhibit recognizable strategic patterns. In Ultimatum Game experiments [57], higher offers lead to higher acceptance rates, broadly mirroring human behavior. The agent CICERO (Meta FAIR) combined an LLM for natural-language negotiation with strategic-reasoning modules and ranked in the top 10% of human players in online Diplomacy leagues, showing that LLM-driven agents can perform well in multi-agent environments requiring language and strategy. Even so, such results do not yet establish stable economic validity: LLM agents can behave inconsistently across runs and may fail to adapt optimally to opponent behavior.

4.2.2. LLM agents in financial market and macroeconomics simulations

Embedding LLM agents into full market simulations has begun to suggest that they can reproduce some competitive and systemic dynamics endogenously, although the evidence remains exploratory. Yang et al. [58] introduced TwinMarket, an LLM-based multi-agent framework designed to simulate socio-economic dynamics in financial markets. Agents interact through a simulated social-media environment, and the model generates cognitive biases, emotional responses, and herd behavior without predefined trading rules. Their results are best read as a proof of possibility rather than as settled evidence that LLM-based market agents are empirically robust. Several references in this subsection are therefore best interpreted as exploratory preprint or workshop evidence rather than as established empirical support.

Fish et al. [59] studied whether LLM-based agents can engage in collusive behavior in oligopolistic pricing environments. They found that agents often converge toward supra-competitive prices even without an explicit intention to collude, and they showed that small variations in prompt wording can materially affect the degree of collusion.

At the macro level, Li et al. [60] introduced EconAgent, a macroeconomic simulation framework in which LLM-based agents act as heterogeneous decision makers with respect to work and consumption. Agent interactions generate macroeconomic outcomes such as inflation and unemployment dynamics. Furthermore, in the domain of energy market, residential agents' interactions across regions determine their region-specific energy consumptions and expenditure. These results are promising, but they should be interpreted cautiously: at this stage, they show that LLM-based agents can generate a wider range of macro patterns, not that they outperform better-established ABM architectures across validation criteria.

4.2.3. Methodological challenges

Using LLM agents raises several methodological concerns. First, unlike static agents, LLM agents may produce different outputs across runs because of non-deterministic sampling, which complicates replication [61]. Second, they may introduce fictional events or references, that is, hallucinations, that are not present in the simulation state. Because hallucination undermines validity, researchers often constrain input-output formats or use verification prompts to reduce the problem [62]. Third, there are broader technical and conceptual issues: LLMs do not necessarily behave like self-interested market participants, prompt design can strongly affect outcomes, and substantial computational resources may limit population size and experimental scale [63].

A pragmatic response is to adopt hybrid architectures. In such settings, LLM agents may be reserved for key decision makers, while simpler rule-based agents populate the rest of the environment. Traditional ABM remains better suited to transparent, scalable, and tightly controlled experiments, whereas LLM-based agents may be more useful when human-like judgment, language, and adaptive communication matter [64]. For now, the most productive strategy is likely comparative: Use classical ABM as a baseline and LLM-based agents as an exploratory extension rather than as a direct replacement.

4.3. ABM and green finance

4.3.1. Conceptual scope and applications

Green energy finance broadly encompasses financial activities that support environmentally sustainable development, including carbon markets, ESG investing, climate-risk finance, and sustainable development financing. We focus on these four domains because they involve heterogeneous actors interacting under evolving regulatory frameworks and market incentives. Furthermore, the literature is uneven across subfields, so this section distinguishes relatively established applications from more prospective research directions. Because ABMs explicitly model decentralized decision making and nonlinear feedback, they can be useful for understanding these domains and for evaluating public policies and corporate strategies [65].

First, carbon markets are among the most established ABM applications in green finance. Emissions trading systems involve heterogeneous firms facing different abatement costs, regulatory constraints, and strategic trading decisions. Liu et al. [66] studied China's national carbon market with an ABM that incorporates thousands of heterogeneous power firms and uses multi-agent reinforcement learning to represent adaptive production, technology adoption, and allowance trading. Their results suggested that tighter emission caps raise carbon prices and stimulate greater adoption of low-carbon technologies among power producers.

Second, ESG-driven investment behavior is an increasingly important domain for ABM application. ESG markets involve strategic interactions among firms, institutional investors, and regulators, with disclosure rules and sustainability preferences shaping capital allocation. Zhao et al. [16] developed an ABM of ESG disclosure policies across the EU and China. Their results suggested that strict disclosure regimes can improve transparency but may discourage costly green upgrades when compliance burdens become excessive. A hybrid disclosure regime, initially lax and then progressively stricter, generated stronger long-run incentives for green transition.

Katsamakas and Sanchez-Cartas [67] examined competition among firms investing in ESG improvements. They found that early adoption can generate reputational or cost advantages leading to winner-takes-all outcomes, while universal ESG adoption may erode profits and incentivize strategic under-disclosure. ABM studies incorporating ESG-sensitive traders showed how sustainability preferences shape liquidity, volatility, and price discovery, where micro-level heterogeneity plays a key role.

Third, climate-risk research introduces systemic uncertainty through both physical risks, such as extreme weather events, and transition risks, such as policy-induced asset repricing. Dubbelboer et al. [68] developed an agent-based model incorporating flood risk and insurance markets. Their results showed that public-private insurance schemes can maintain affordability in the short run but become vulnerable under rising climate hazards and urban development pressures. More recent work integrates ABM into climate stress-testing frameworks, enabling regulators to simulate cascading defaults, insurance claims, and balance-sheet adjustments under alternative scenarios. These studies suggest that network effects and feedback loops matter materially for climate-related financial stability.

Finally, researchers have begun to apply ABM to green credit policies and sustainable supply-chain finance. Lamperti et al. [69] built an agent-based model to evaluate instruments such as green credit policies and green bonds. They showed that well-designed green financial policies can accelerate low-carbon investment while interacting in nontrivial ways with financial stability. Monasterolo and Raberto [70] used a stock-flow-consistent agent-based model to study green fiscal policy and green sovereign bonds, highlighting trade-offs between environmental objectives and financial stability. These contributions are promising, but the literature is smaller and less standardized than the more established ABM domains reviewed earlier. Here, we shed light on the importance of agent-based models to identify how heterogeneous the impact of green energy policies the agents, i.e., residential consumers as the heterogeneous regions where the agents stay.

4.3.2. Methodological challenges and future directions

Despite growing interest in this area, important challenges remain. Model calibration is difficult because historical data are limited in many emerging sustainability markets. Scaling ABMs to include large financial networks with multiple interacting sectors can also be computationally demanding. Researchers increasingly combine reinforcement learning, hybrid modeling, and policy-oriented stress testing [16]. As climate-related financial regulation intensifies, ABM appears to be a potentially useful framework for analyzing sustainable transitions, but its policy relevance will depend on stronger empirical discipline and clearer validation standards.

Advancing ABM in green finance will also require stronger interdisciplinary collaboration among economists, climate scientists, and financial engineers. Moreover, models often simplify either financial behavior or physical climate dynamics. Greater integration across these domains is therefore essential for building more credible ABM analyses of green financial systems.

5. The validation challenge

How do we know whether an ABM is scientifically credible? Validation refers to the degree to which model outcomes align with real-world data and empirical knowledge. This question is especially important for ABMs because their flexibility enables many behaviors to be generated by tuning parameters. Without convincing validation, an ABM may look like a sophisticated simulation exercise rather than a reliable scientific model. With proper validation, however, ABMs can provide evidence that complements other empirical and theoretical approaches [3].

5.1. Three validation goals

We identify three broad purposes that have guided ABM evaluation: replication of stylized facts, out-of-sample predictive accuracy, and credible policy counterfactuals. These purposes correspond to different uses of a model and therefore to different evidentiary standards. Table 1 summarizes their structural conflicts.

Table 1. Structural conflicts among the three ABM validation goals.

Validation goal	Primary purpose	Key criterion	Main risk
Stylized facts replication	Explanation/description	Qualitative and distributional match to known empirical regularities	Many causal structures can produce the same patterns, making it hard to identify the correct one
Out-of-Sample Prediction	Forecasting	Demonstrated accuracy on held-out data; benchmark comparison	Overfitting to historical data; adaptive markets erode predictability over time
Policy counterfactual	Policy analysis	Structural realism; retrodiction of known policy episodes; robustness	Agents may not adapt realistically to policy changes; stylized-fact fit alone may give false confidence

To make this more operational, it is useful to distinguish at least five evidence layers: (i) stylized-fact matching, (ii) micro-data calibration, (iii) out-of-sample predictive assessment, (iv) sensitivity and uncertainty analysis, and (v) structural robustness under policy counterfactuals. Not every ABM must satisfy all five layers, but a review centered on validation should make explicit which layer a given study reaches and what claims that level of evidence can support.

Most ABM studies begin with stylized-fact matching. Models like Lux and Marchesi [31] showed that a simple ABM with heterogeneous agents can generate fat-tailed returns and volatility clustering, patterns that orthodox equilibrium models often capture only indirectly. Parameters are then adjusted until simulated data match empirical regularities such as return kurtosis, volatility autocorrelation, or price–volume relationships. However, as Guerini and Moneta [71] argued, stylized-fact matching is a relatively weak form of evidence because many causal structures can produce the same statistical patterns. It is therefore a useful starting point, but not a sufficient validation strategy on its own.

Table 2. Illustrative classification of selected ABM studies by validation standard and evidentiary strength.

Study/model family	Domain/purpose	Calibration target	Evidence reached	OOS?	Main limitation
Lux–Marchesi [31]	Asset pricing/stylized facts	Return distribution and volatility moments	Stylized-fact matching	No	Weak identification; many mechanisms fit the same moments
Kårvik et al. [48]	LOB-HFT instability	LOB liquidity and HFT frequencies	Microstructure calibration + stress scenarios	No	Venue-specific assumptions; limited transportability
Paulin et al. [52]	Flash-crash contagion/systemic risk	Portfolio overlap and margin shocks	Structural scenario analysis	No	Preprint; stylized network assumptions
Poledna et al. [72]	Macro forecasting/policy	Macro time series and sectoral aggregates	Out-of-sample predictive assessment	Yes	Structural uncertainty across countries and regimes
Liu et al. [66]	Carbon market/policy	Firm emissions and allowance trading	Calibration + policy counterfactuals	No	Limited historical data for external validation
Zhao et al. [16]	ESG disclosure/transition	Disclosure regime and upgrade costs	Policy-counterfactual simulation	No	Evidence remains model-dependent and institution-specific

Out-of-sample prediction sets a stricter standard. Financial ABMs were initially designed mainly as explanatory tools rather than forecasting engines. A notable exception is Poledna et al. [72], whose calibrated ABM forecast key macroeconomic aggregates as accurately as standard VAR and DSGE models and was used to assess the economic impact of COVID-19 lockdowns. Moreover, Catalano et al. [73] applied the same large-scale ABM architecture to green-transition scenarios. More generally, the framework introduced by Poledna et al. [72] is a general macroeconomic ABM rather than a model designed for pandemic applications; researchers have adapted it to Canada and Italy [74,75]. Even so, few ABMs have achieved this standard, partly because rich models risk overfitting historical data while simpler ones lack the structure needed for broader use.

Policy counterfactuals are the most demanding goal. The model must generate credible conditional outcomes when policy rules change, and agents must adjust their behavior accordingly. Direct validation is impossible because the counterfactual world cannot be observed. Researchers therefore rely on retrodiction of known policy episodes, micro-level evidence on behavior, sensitivity and uncertainty analysis, and robustness checks across parameter ranges and institutional assumptions. These goals are not always compatible, which is why validation should be tied explicitly to model purpose. Also, Table 2 provides an illustrative, non-exhaustive mapping of several major studies cited in this review to validation standards and evidentiary strength.

5.2. Prediction difficulties in financial ABMs

Predictive validation has received far less attention than stylized-fact replication in the ABM literature for a few reasons. First, the field historically prioritized explanation over forecasting. The Santa Fe Institute's classic stock market ABM was intended to show how boundedly rational traders could produce realistic dynamics, not to predict prices. Without the discipline of forecast evaluation, models proliferated that fit the past well but were never vetted for future accuracy.

Second, abundant data does not necessarily make calibration easier. Using rich micro-level financial data is not straightforward, as it requires estimating agent decision rules and aggregate dynamics at the same time. ABMs' nonlinear and path-dependent nature makes this difficult, since standard econometric methods do not apply well [76]. Moreover, adjusting parameters manually to fit historical data risks overfitting. When a model has many free parameters, it can match past patterns while also picking up random noise that does not appear in new data.

Third, financial markets are adaptive. If an ABM identifies a pattern that traders can profit from, actual market participants may exploit it until it weakens or disappears. This problem is not unique to ABMs; it is a general challenge for predictive models in finance. It does, however, mean that even when an ABM shows some forecasting ability, that ability may not persist across different market conditions.

Finally, structural breaks, such as new regulations, the rise of HFT, and the emergence of crypto markets, can leave any model poorly suited to new conditions if it was calibrated to an earlier era. In this respect, ABMs may be better positioned than many alternative models because heterogeneous and adaptive agents can, in principle, be designed to respond to changing environments. Richer data, machine-learning-based agent training [77], and regular model updates may help address this limitation. Even so, ABMs may not outperform simpler statistical models for short-horizon financial forecasting.

5.3. The difference between calibration and validation

The ABM literature often conflates two distinct concepts. Calibration refers to selecting parameters to fit known data or stylized facts, whether at the level of aggregate moments or micro-level behavior. Validation refers to evaluating performance on criteria not already fitted. A common misuse labels a model "validated" because it reproduces empirical patterns that were implicitly used to develop it, creating circular reasoning. Windrum et al. [78] termed this indirect calibration and noted that while it demonstrates consistency with data, it does not confirm underlying model assumptions or predictive power.

5.3.1. Calibration, validation, and verification

Clarifying this distinction requires transparency about what information was used to build or tune the model versus what independent tests were applied afterward. Best practice involves explicit sequential steps. We first calibrate parameters to match a defined set of stylized facts or moments, then validate by testing separate criteria, such as out-of-sample data, facts not targeted in calibration, or causal-structure comparisons as in Guerini and Moneta [71]. A layered framework helps organize these steps. Level 1 refers to qualitative reproduction of core stylized facts. When those facts are not used to tune the model, Level 1 can count as a weak form of validation; when they are used as calibration targets, it should be understood as calibration-linked assessment rather than independent validation. Level 2 requires quantitative distributional fit after calibration. Level 3 requires out-of-sample predictive performance or the reproduction of novel facts not targeted during calibration. Many published ABMs satisfy Level 1 or 2 but have not been subjected to Level 3 tests, a gap the field is gradually closing.

A related confusion is between verification and validation. Verification ensures the model is correctly implemented, while validation ensures it is empirically realistic. Verification is an internal consistency check and calling it validation conflates model correctness with model truth. Reserving "validation" for the model's relationship with external reality is essential for scientific credibility.

5.3.2. Mathematical tools for calibration

Calibrating an ABM empirically means finding parameter values that make the model's output match patterns observed in real data. This is a complex optimization problem because ABMs typically have many parameters and no simple formula linking them to model output. Simulated method of moments (SMM) and Bayesian MCMC minimize the distance between simulated and empirical moment vectors, as shown below.

$$\min_{\theta} |m_{\text{data}} - m_{\text{sim}}(\theta)|$$

Here, m denotes moment vectors such as return kurtosis, volatility autocorrelation, and price volume correlations. Because ABMs are nonlinear and path-dependent, the mapping $\theta \rightarrow m_{\text{sim}}(\theta)$ has no closed form and must be estimated by running the simulation repeatedly. This is computationally expensive. A single calibration exercise may require thousands of simulation runs.

Machine-learning surrogates address this bottleneck by approximating $m_{\text{sim}}(\theta)$ with a learned regression $\hat{m}(\theta)$, which can be evaluated at negligible cost once trained. The surrogate guides the optimization toward promising regions of parameter space, reducing the total number of full simulations required by orders of magnitude. Bayesian approaches additionally yield a posterior distribution over θ rather than a point estimate, quantifying parameter uncertainty and providing a natural measure of how tightly the model is constrained by data. This is directly relevant to the falsifiability concerns discussed in Section 5.1.

Beyond moment matching, policy optimization within ABMs can be framed as stochastic control. We define a cost function $J(\theta)$ over policy parameter θ, for example a transaction tax rate or circuit-breaker threshold, and seek the following.

$$\min_{\theta} \mathbb{E}[L(\theta)]$$

The expectation is estimated via repeated ABM runs. Reinforcement learning treats the regulator as an agent adjusting θ across simulation episodes, providing a flexible search strategy when $J(\theta)$ is non-convex and high-dimensional. Response surfaces are often well-approximated by low-dimensional surrogates, enabling efficient policy optimization layered on top of the ABM. These mathematical tools connect ABM output to the rigorous estimation and optimization frameworks of classical econometrics, strengthening the method's credibility for empirical and policy applications.

5.4. Proper validation for model objectives

Validation standards should match the goal the model is designed to achieve. For a descriptive or explanatory ABM, validation emphasizes qualitative alignment with known patterns, plausible mechanisms, and evidence that the stylized facts are not produced only by ad hoc tuning. For a predictive ABM, the bar is stronger: Forecasting performance should be evaluated on held-out data relative to relevant benchmarks, and uncertainty should be reported explicitly. For models calibrated to micro data, the mapping from empirical behavior to agent rules should also be documented transparently.

Table 3. Checklist for financial ABM practitioners.

Step	Key questions	Design implication
Define purpose	Explain stylized facts? Evaluate policy? Assess systemic risk? Test microstructure?	Do not mix objectives without priority. A single primary purpose drives all subsequent choices.
Align complexity	Is each agent rule/state variable traceable to a real mechanism critical for the stated purpose?	Include only necessary heterogeneity. Trim features that cannot be justified by the research question.
Plan validation	Which stylized facts or historical episodes serve as calibration targets? Which are reserved for out-of-sample tests?	Instrument outputs to enable direct comparison. Reserve held-out data before calibration begins.
Document Trade-offs	Which design trade-off (realism, cognitive richness, flexibility, fidelity) was accepted, and why?	Make limitations explicit so readers can critique the design relative to the stated purpose.

An ABM designed for policy analysis becomes credible when it satisfies several additional conditions. The model should first be consistent with empirical data at micro and macro levels before any policy experiment is run. It should represent the institutional and behavioral structures most relevant to the policy question, and agents should be allowed to adjust their behavior when policy changes rather than mechanically holding rules fixed. Policy conclusions should remain reasonably stable across sensitivity tests, uncertainty analysis, and alternative behavioral assumptions. If similar policies have historical precedents, the model should reproduce the broad direction of observed reactions. Finally, the channels through which the policy affects outcomes should be clear enough that domain experts can evaluate the underlying mechanism rather than only the simulated result.

Each model should therefore be assessed for criteria appropriate to its purpose. An explanatory ABM should not be rejected solely because it is not a forecasting tool, and a predictive ABM should not be used for policy advice without structural grounding. In practice, different tasks may call for different models: Simple statistical models often remain useful for short-horizon forecasting, while ABMs are better suited to scenario analysis, stress testing, and mechanism exploration. Hybrid strategies can also be productive when they combine empirical discipline with structural richness [3].

Building on the validation principles discussed above, we propose a practical checklist in Table 3 to help practitioners align model design with intended use. Each step should be completed before coding begins and documented transparently in publications.

6. Conclusions

In the preceding sections, we examined what financial ABMs can do, where their limits lie, how they should be validated, and which newer domains remain exploratory. We now assess their comparative advantages and fundamental limitations, offer practical guidance for practitioners, and discuss directions for the field.

6.1. Comparative advantages of financial ABMs

Financial ABMs have several important strengths, but their value depends on matching the model to the question and on applying appropriate validation standards. Four advantages deserve particular emphasis.

Simple heterogeneous-agent models with trading heuristics can generate the statistical regularities of asset returns, including heavy tails, volatility clustering, and excess kurtosis, as emergent outcomes of agent interactions rather than exogenous assumptions. These patterns, which orthodox equilibrium models must impose explicitly, arise naturally from the fundamentalist–chartist dynamics and herding mechanisms described in Section 3.

Because agent decision rules are specified structurally, ABMs can impose hypothetical policies and observe resulting dynamics in ways that closed-form models cannot. Tightening a mortgage cap in a housing ABM, for example, shows how each household adjusts leverage, how credit allocation shifts, and whether unintended distributional consequences emerge. This capacity to generate structured counterfactual scenarios in novel settings, including macroprudential rules, climate-transition shocks, and new market mechanisms, is one of ABM's most useful contributions, although its policy value depends on empirical discipline, sensitivity analysis, and institutional realism.

ABMs also capture the mechanics of trading, including order books, intraday liquidity, and flash crashes, at levels of detail that equilibrium models cannot reach. High-fidelity ABMs have replicated flash crash episodes and shown that such events arise from internal feedback loops rather than exogenous news, offering insights relevant to market design and regulatory reform.

Finally, ABMs are well suited for studying how shocks propagate through networks of heterogeneous institutions. Traditional representative-agent models failed to anticipate 2008-style cascading defaults and liquidity spirals. By enabling multiple agent types, explicit network connections, and adaptive behavior under stress, ABMs can replicate crisis dynamics after the fact and explore systemic vulnerabilities in advance, serving as a testing ground for the financial system that linear models cannot provide.

6.2. Fundamental limitations and design trade-offs

Financial ABMs face fundamental limitations that manifest as design trade-offs. Improving one dimension of a model inevitably sacrifices another, which is why no single ABM performs well across all validation criteria.

Richly detailed ABMs gain descriptive realism at the cost of analytical tractability and computational manageability. As Farmer and Foley [3] observed, rational expectation models stripped away most of the structure of a real economy precisely because nonlinearity and complexity are incompatible with equilibrium methods. ABMs reintroduce that complexity but must be explored via simulation, yielding outputs that are understood statistically rather than analytically. The KISS vs. KIDS debate captures this tension. "Keep It Simple, Stupid" favors minimal models for clarity, while "Keep It Descriptive, Stupid" favors realistic detail for policy relevance. The right balance depends on the model's purpose.

Increasing agent sophistication, from fixed heuristics to reinforcement learning to LLM-based agents, can improve behavioral realism but at steep cost to interpretability. When agents operate via complex internal states, it becomes difficult to isolate which assumptions drive macro-outcomes, undermining the model's explanatory power. A simpler companion model is often needed to interpret the output of a cognitively rich simulation. The LLM-agent frontier discussed in Section 4.2 poses this trade-off most sharply: it offers unprecedented behavioral realism, but at considerable cost to transparency and falsifiability.

ABMs typically have many free parameters that can be tuned to fit a wide range of empirical patterns. This flexibility enables good fit to historical data but risks low falsifiability: a model that can explain any outcome after the fact makes few sharp predictions and is difficult to reject. As Richiardi [79] noted, the freedom afforded by ABMs often leads to models with too many degrees of freedom that are difficult to falsify. The remedy is to limit flexibility deliberately, by fixing parameters from prior literature, reserving held-out data, and conducting systematic sensitivity analysis, so that the model makes testable predictions rather than merely rationalizing known facts.

High-fidelity ABMs that simulate sub-second order-book mechanics or full bilateral contract networks are computationally intensive and limited in scope. They may cover only a few hours of trading or a small institution network. More abstract ABMs scale to larger populations and longer horizons but sacrifice micro-realism. No single model can encompass every level of financial system detail while remaining usable for broad analysis or extensive sensitivity testing. Segmentation by purpose is therefore unavoidable. Microstructure ABMs work well for market design questions, while macro-scale ABMs are better suited for systemic risk analysis. Linking models across scales remains an area of active development. We summarize these principles as a checklist in Table 3 in Section 5.

6.3. Core findings and proper scope

In this review, we highlight several core findings about financial ABMs. Financial markets can be fruitfully modeled as complex adaptive systems in which interacting heterogeneous agents generate macro-level regularities and contagion patterns that are difficult to recover in representative-agent settings. Feedback loops and network structures are first-order drivers of financial stability. Leverage dynamics, interbank exposures, and portfolio overlaps can generate instabilities that are emergent properties of the system rather than simple sums of individual decisions. ABMs therefore fill genuine gaps left by more aggregated models, especially when the question concerns crises, interaction effects, or unprecedented scenarios for which historical analogies are weak.

The proper scope of financial ABM follows directly from these strengths and limitations. ABMs are best suited for questions in which heterogeneity, network interactions, or out-of-equilibrium dynamics are first-order drivers, including understanding crises, wealth distributions, macroprudential policies, and climate transitions. In such domains, the ability to generate qualitative outcomes, such as the conditions under which a liquidity spiral or cascade becomes possible, is more valuable than any single point forecast. What ABMs cannot yet reliably do is point prediction or policy optimization under a single objective function. Forecasting next quarter's GDP or the exact timing of the next market crash lies outside their credible scope. The appropriate use is scenario planning, stress testing, and mechanism exploration rather than precise forecasting.

ABMs should complement, not replace, other approaches. In systemic-risk analysis, network metrics, empirical methods, and ABM simulations are often most useful combined. When multiple models agree, confidence grows; when they diverge, the reasons for disagreement become scientifically informative. Financial ABMs can therefore serve as virtual experimental laboratories rather than as wholesale substitutes for equilibrium or econometric methods. This complementary positioning is where they add durable value.

This complementary view is also reinforced by adjacent literature on systemic fragility, nonlinear dynamics, and empirical robustness. Work on modern financial engineering and systemic risk [80], nonlinear economic dynamics [81], bank-default prediction [82], extreme-event losses [83], market sentiment and nonlinear price dynamics [84], and return-distribution robustness [85] does not substitute for core ABM studies, but it provides useful comparative benchmarks for thinking about contagion, instability, and evidentiary standards.

6.4. Outlook

The outlook for financial ABM is cautiously optimistic, provided the field moves toward greater standardization and closer integration with mainstream economic science. Several converging trends suggest the next decade can be one of consolidation and steady progress.

Just as econometrics matured through standard tests and diagnostics, ABM is developing protocols for validation, documentation, and comparison. We anticipate the emergence of benchmark models, analogous to benchmark DSGEs, that become standard tools for specific topics: A well-calibrated stock market ABM against which new policy designs can be tested, or a baseline climate-finance ABM used by multiple research groups. Documentation standards such as the ODD protocol, adopted in social simulation, will likely become expected in journal submissions, making results across studies easier to compare and build upon.

Future ABMs will likely be more data driven. Machine learning will help estimate agent behavior rules from micro-data, including large datasets on investor trades, firm balance sheets, and loan records, and ABMs will then simulate scenarios using those learned behaviors. Running ABMs in large ensembles to produce outcome distributions that can be statistically analyzed will raise the scientific standard of the field. When ABM structures can be formally tested and refined based on data, the method will have achieved genuine scientific cumulativeness.

Two emerging domains, LLM-based agents and climate-finance ABMs, pose validation challenges that methods only partly address. LLM agents may broaden behavioral repertoire, but they remain difficult to falsify and highly sensitive to prompting and system configuration. Climate-focused ABMs address counterfactual futures with limited historical precedent, which makes structural validation, uncertainty analysis, and cross-model comparison more important than simple back-testing. Developing robust standards for these domains remains an open and important research task.

Financial ABM is also drawing on computer science, physics, and ecology, and this cross-disciplinary exchange will intensify as publication in leading economics journals demands greater rigor. We anticipate ABM becoming part of the standard curriculum for financial economics, with shared computational infrastructure lowering entry barriers and open-source base models enabling researchers to build upon work rather than starting from scratch.

Agent-based modeling in financial markets has moved from a marginal research program to an established, though methodologically diverse, part of economic inquiry. Phenomena such as fat tails, flash crashes, and contagion cascades can now be generated, studied, and partially explained within ABM frameworks. The next task is not to declare victory, but to turn these models into more reliable research and policy tools through rigorous validation, clearer comparative standards, and stronger cumulative practice. If that effort succeeds, ABM will continue to broaden the range of questions that financial economics can address realistically.

Use of AI tools declaration

The authors declare that AI tools were used in the creation of this article, primarily for writing assistance, grammar editing, and citation checking. The authors have reviewed all AI-generated content and take full responsibility for the accuracy and integrity of this article.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2024S1A5A2A01029162).

Conflict of interest

The authors declare there is no conflict of interest.

Author contributions

Both authors contributed equally to the literature review, discussion, writing, and revision of the paper.

References

1. R. Cont, Empirical properties of asset returns: Stylized facts and statistical issues, *Quant. Finance*, **1** (2001), 223–236. https://doi.org/10.1080/713665670
2. B. Mandelbrot, The variation of certain speculative prices, *J. Bus.*, **36** (1963), 394–419. https://doi.org/10.1086/294632
3. J. D. Farmer, D. Foley, The economy needs agent-based modelling, *Nature*, **460** (2009), 685–686. https://doi.org/10.1038/460685a
4. P. Gai, S. Kapadia, Contagion in financial networks, *Proc. R. Soc. A*, **466** (2010), 2401–2423. https://doi.org/10.1098/rspa.2009.0410
5. A. P. Kirman, Complexity and economic policy: A paradigm shift or a change in perspective? A review essay on David Colander and Roland Kupers's complexity and the art of public policy, *J. Econ. Lit.*, **54** (2016), 534–572. https://doi.org/10.1257/jel.54.2.534
6. W. B. Arthur, Foundations of complexity economics, *Nat. Rev. Phys.*, **3** (2021), 136–145. https://doi.org/10.1038/s42254-020-00273-3
7. R. Holt, J. B. Rosser, D. Colander, The complexity era in economics, *Rev. Polit. Econ.*, **23** (2011), 357–369. https://doi.org/10.1080/09538259.2011.583820
8. R. L. Axtell, J. D. Farmer, Agent-based modeling in economics and finance: Past, present, and future, *J. Econ. Lit.*, **63** (2025), 197–287. Available from: http://www.aeaweb.org/articles?id=10.1257/jel.20221319.
9. B. LeBaron, Agent-based computational finance, in *Handbook of Computational Economics* (eds. L. Tesfatsion and K. L. Judd), Elsevier, (2006), 1187–1233. https://doi.org/10.1016/S1574-0021(05)02024-1
10. C. H. Hommes, Heterogeneous agent models in economics and finance, in *Handbook of Computational Economics* (eds. L. Tesfatsion and K. L. Judd), Elsevier, (2006), 1109–1186. https://doi.org/10.1016/S1574-0021(05)02023-X

11. G. Fagiolo, A. Roventini, Macroeconomic policy in DSGE and agent-based models redux: New developments and challenges ahead, *J. Artif. Soc. Soc. Simul.*, **20** (2017), 1. https://doi.org/10.18564/jasss.3280
12. S. H. Chen, C. L. Chang, Y. R. Du, Agent-based economic models and econometrics, *Knowl. Eng. Rev.*, **27** (2012), 187–219. https://doi.org/10.1017/S0269888912000136
13. J. J. Horton, Large language models as simulated economic agents: What can we learn from Homo Silicus, *NBER Work. Pap.*, No. 31122, 2023. https://doi.org/10.3386/w31122.
14. Q. Zhao, J. Wang, Y. Zhang, Y. Jin, K. Zhu, H. Chen, et al., CompeteAI: Understanding the competition dynamics in large language model-based agents, arXiv:2310.17512, 2023. https://doi.org/10.48550/arXiv.2310.17512
15. C. Ziems, W. Held, O. Shaikh, J. Chen, Z. Zhang, D. Yang, Can large language models transform computational social science, *Comput. Linguist.*, **50** (2024), 237–291. https://doi.org/10.48550/arXiv.2305.03514
16. J. Zhao, M. Polukarov, C. Ventre, Green Disclosure policies and market dynamics: Evidence from agent-based ESG models, *Auton. Agents Multi-Agent Syst.*, **40** (2026), 6. Available from: https://link.springer.com/article/10.1007/s10458-026-09734-y.
17. J. M. Epstein, Agent-based computational models and generative social science, *Complexity*, **4** (1999), 41–60. https://doi.org/10.1002/(SICI)1099-0526(199905/06)4:5<41::AID-CPLX9>3.0.CO;2-F
18. A. P. Kirman, Whom or what does the representative individual represent? *J. Econ. Perspect.*, **6** (1992), 117–136. https://doi.org/10.1257/jep.6.2.117
19. H. A. Simon, A behavioral model of rational choice, *Q. J. Econ.*, **69** (1955), 99–118. https://doi.org/10.2307/1884852
20. H. A. Simon, The sciences of the artificial, MIT Press, 1969. Available from: https://mitpress.mit.edu/9780262691918/the-sciences-of-the-artificial/.
21. G. Gigerenzer, D. G. Goldstein, Reasoning the fast and frugal way: Models of bounded rationality, *Psychol. Rev.*, **103** (1996), 650–669. https://doi.org/10.1037/0033-295X.103.4.650
22. G. Gigerenzer, P. M. Todd, and the ABC Research Group, *Simple Heuristics That Make Us Smart*, Oxford University Press, 1999. Available from: https://global.oup.com/academic/product/simple-heuristics-that-make-us-smart-9780195143812.
23. D. Kahneman, A. Tversky, Prospect theory: An analysis of decision under risk, *Econometrica*, **47** (1979), 263–292. https://doi.org/10.2307/1914185
24. W. B. Arthur, J. H. Holland, B. LeBaron, R. Palmer, P. Tayler, Asset pricing under endogenous expectations in an artificial stock market, in *The Economy as an Evolving Complex System II* (eds. W. B. Arthur, S. Durlauf, and D. Lane), Addison-Wesley, 1997, 15–44. Available from: https://www.pearson.com/en-us/subject-catalog/p/the-economy-as-an-evolving-complex-system-ii/P200000004233/9780201323665.
25. W. A. Brock, C. H. Hommes, Heterogeneous beliefs and routes to chaos in a simple asset pricing model, *J. Econ. Dyn. Control*, **22** (1998), 1235–1274. https://doi.org/10.1016/S0165-1889(98)00011-6
26. A. P. Kirman, Ants, rationality, and recruitment, *Q. J. Econ.*, **108** (1993), 137–156. https://doi.org/10.2307/2118498

27. J. Di Domenico, P. Ragazzini, M. Catalano, A. Glielmo, L. Romeo, L. Riccetti, Reinforcement learning for firms in a large-scale macroeconomic ABM: Emerging pricing strategies and forecasting performance, SSRN, 2025, 5533858. Available from: https://ssrn.com/abstract=5533858.
28. D. K. Gode, S. Sunder, Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality, *J. Polit. Econ.*, **101** (1993), 119–137. https://doi.org/10.1086/261868
29. L. Tesfatsion, K. L. Judd, Handbook of Computational Economics, in *Agent-Based Computational Economics*, Elsevier, 2006. Available from: https://www.sciencedirect.com/handbook/handbook-of-computational-economics/vol/2.
30. C. M. Anderson, C. A. Holt, C. R. Plott, Information cascades in the laboratory, *Am. Econ. Rev.*, **87** (1997), 847–862. https://www.jstor.org/stable/2951328
31. T. Lux, M. Marchesi, Scaling and criticality in a stochastic multi-agent model of a financial market, *Nature*, **397** (1999), 498–500. https://doi.org/10.1038/17290
32. T. Lux, Stochastic behavioral asset-pricing models and the stylized facts, in *Handbook of Financial Markets: Dynamics and Evolution* (eds. T. Hens and K. R. Schenk-Hoppé), Elsevier, 2009, 161–215. Available from: https://www.sciencedirect.com/book/9780123742582/handbook-of-financial-markets.
33. C. Chiarella, G. Iori, J. Perelló, The impact of heterogeneous trading rules on the limit order book and order flows, *J. Econ. Dyn. Control*, **33** (2009), 525–537. https://doi.org/10.1016/j.jedc.2008.08.001
34. B. LeBaron, Evolution and time horizons in an agent-based stock market, *Macroecon. Dyn.*, **5** (2001), 225–254. https://doi.org/10.1017/S1365100501019058
35. J. H. Holland, Adaptation in natural and artificial systems, University of Michigan Press, 1975. Available from: https://press.umich.edu/Books/A/Adaptation-in-Natural-and-Artificial-Systems2.
36. C. Chiarella, G. Iori, A simulation analysis of the microstructure of double auction markets, *Quant. Finance*, **2** (2002), 346–353. https://doi.org/10.1088/1469-7688/2/5/303
37. T. Preis, S. Golke, W. Paul, J. J. Schneider, Multi-agent-based order book model of financial markets, *Europhys. Lett.*, **75** (2006), 510–516. https://doi.org/10.1209/epl/i2006-10139-0
38. A. G. Hawkes, Spectra of some self-exciting and mutually exciting point processes, *Biometrika*, **58** (1971), 83–90. https://doi.org/10.1093/biomet/58.1.83
39. K. Gao, P. Vytelingum, S. Weston, W. Luk, C. Guo, High-frequency financial market simulation and flash crash scenarios analysis: An agent-based modelling approach, arXiv:2208.13654, 2024. https://doi.org/10.48550/arXiv.2208.13654
40. M. O. Jackson, *Social and Economic Networks*, Princeton University Press, 2008.
41. R. Albert, A. L. Barabási, Statistical mechanics of complex networks, *Rev. Mod. Phys.*, **74** (2002), 47–97. https://doi.org/10.1103/RevModPhys.74.47
42. H. Choi, S. H. Kim, J. Lee, Role of network structure and network effects in diffusion of innovations, *Ind. Mark. Manage.*, **39** (2010), 170–177. https://doi.org/10.1016/j.indmarman.2008.08.006
43. E. Lee, J. Lee, J. Lee, Reconsideration of the winner-take-all hypothesis: Complex networks and local bias, *Manage. Sci.*, **52** (2006), 1838–1848. Available from: https://www.jstor.org/stable/20110658.

44. L. Riccetti, Agent-based multi-layer network simulations for financial systemic risk measurement: A proposal for future developments, *Int. J. Microsimul.*, **15** (2022), 44–61. https://doi.org/10.34196/ijm.00262
45. F. Allen, D. Gale, Financial contagion, *J. Polit. Econ.*, **108** (2001), 1–33. Available from: https://www.jstor.org/stable/10.1086/262109.
46. S. Bornholdt, Expectation bubbles in a spin model of markets: Intermittency from frustration across scales, *Int. J. Mod. Phys. C*, **12** (2001), 667–674. https://doi.org/10.1142/S0129183101001845
47. H. Choi, B. Lee, Examining network externalities and network structure for new product introduction, *Inf. Technol. Manage.*, **13** (2012), 183–199. https://doi.org/10.1007/s10799-012-0125-x
48. G. A. Kårvik, J. Noss, J. Worlidge, D. Beale, The deeds of speed: An agent-based model of market liquidity and flash episodes, *Bank Engl. Staff Work. Pap.*, **743** (2018). Available from: https://www.bankofengland.co.uk/working-paper/2018/the-deeds-of-speed-an-agent-based-model-of-market-liquidity-and-flash-episodes.
49. A. Kirilenko, A. S. Kyle, M. Samadi, T. Tuzun, The flash crash: High-frequency trading in an electronic market, *J. Finance*, **72** (2017), 967–998. https://doi.org/10.1111/jofi.12498
50. N. Johnson, G. Zhao, E. Hunsader, H. Qi, N. Johnson, J. Meng, et al., Abrupt rise of new machine ecology beyond human response time, *Sci. Rep.*, **3** (2013), 2627. https://doi.org/10.1038/srep02627
51. S. J. Jacob Leal, M. Napoletano, Market stability vs. market resilience: Regulatory policies experiments in an agent-based model with low- and high-frequency trading, *J. Econ. Behav. Organ.*, **157** (2019), 15–41. https://doi.org/10.1016/j.jebo.2017.04.013
52. J. Paulin, A. Calinescu, M. Wooldridge, Understanding flash crash contagion and systemic risk: A micro-macro agent-based approach, arXiv:1805.08454, 2018. https://doi.org/10.48550/arXiv.1805.08454
53. M. K. Brunnermeier, L. H. Pedersen, Market liquidity and funding liquidity, *Rev. Financ. Stud.*, **22** (2009), 2201–2238. https://doi.org/10.1093/rfs/hhn098
54. S. Battiston, D. Delli Gatti, M. Gallegati, B. Greenwald, J. E. Stiglitz, Liaisons dangereuses: Increasing connectivity, risk sharing, and systemic risk, *J. Econ. Dyn. Control*, **36** (2012), 1121–1141. https://doi.org/10.1016/j.jedc.2012.04.001
55. X. Chen, Y. Liu, X. Li, The emergence of economic rationality of GPT, *Proc. Natl. Acad. Sci.*, **120** (2023), e2316205120. https://doi.org/10.1073/pnas.2316205120
56. J. Brand, A. Israeli, D. Ngwe, Using GPT for market research, SSRN, 2023, 4395751. Available from: https://ssrn.com/abstract=4395751.
57. F. Guo, GPT agents in game theory experiments, arXiv:2305.05516, 2023. https://doi.org/10.48550/arXiv.2305.05516
58. Y. Yang, Y. Zhang, M. Wu, K. Zhang, Y. Zhang, H. Yu, et al., TwinMarket: A scalable behavioral and social simulation for financial markets, arXiv:2502.01506, 2025. https://doi.org/10.48550/arXiv.2502.01506
59. S. Fish, Y. A. Gonczarowski, R. I. Shorrer, Algorithmic collusion by large language models, arXiv:2404.00806, 2024. https://doi.org/10.48550/arXiv.2404.00806

60. N. Li, C. Gao, M. Li, Y. Li, Q. Liao, EconAgent: Large language model-empowered agents for simulating macroeconomic activities, arXiv:2310.10436, 2024. https://doi.org/10.48550/arXiv.2310.10436
61. B. Atil, S. Aykent, A. Chittams, L. Fu, R. J. Passonneau, E. Radcliffe, et al., Non-determinism of "deterministic" LLM system settings in hosted environments, *Proc. 5th Workshop Eval. Compar. NLP Syst.*, 2025. Available from: https://aclanthology.org/events/eval4nlp-2025/.
62. X. X. Lin, LLM-based agents suffer from hallucinations: A survey of taxonomy, methods, and directions, arXiv:2509.18970, 2025. https://doi.org/10.48550/arXiv.2509.18970
63. C. Gao, X. Lan, Z. Lu, J. Mao, J. Piao, H. Wang, et al., Large language models empowered agent-based modeling and simulation: A survey and perspectives, *Humanit. Soc. Sci. Commun.*, 2024. https://doi.org/10.48550/arXiv.2312.11970
64. P. Taillandier, J. D. Zucker, A. Grignard, B. Gaudou, N. Q. Huynh, A. Drogoul, Integrating LLM in agent-based social simulation: Opportunities and challenges, arXiv:2507.19364, 2025. https://doi.org/10.48550/arXiv.2507.19364
65. C. Fang, T. Ma, Stylized agent-based modeling on linking emission trading systems and its implications for China's practice, *Energy Econ.*, **92** (2020), 104916. https://doi.org/10.1016/j.eneco.2020.104916
66. S. Liu, P. Zhou, M. Wang, A. Xu, An agent-based approach to modeling power firms' emission reduction strategies and market dynamics, *Appl. Energy*, **400** (2025), 126590. https://doi.org/10.1016/j.apenergy.2025.126590
67. E. Katsamakas, J. M. Sanchez-Cartas, A computational model of the competitive effects of ESG, *PLOS ONE*, **18** (2023), e0284237. https://doi.org/10.1371/journal.pone.0284237
68. J. Dubbelboer, I. Nikolic, K. Jenkins, J. Hall, An agent-based model of flood risk and insurance, *J. Artif. Soc. Soc. Simul.*, **20** (2017), Article 6. https://doi.org/10.18564/jasss.3135
69. F. Lamperti, V. Bosetti, A. Roventini, M. Tavoni, T. Treibich, Three green financial policies to address climate risks, *J. Financ. Stab.*, **34** (2018), 208–221. https://doi.org/10.1016/j.jfs.2021.100875
70. I. Monasterolo, M. Raberto, The EIRIN flow-of-funds behavioural model of green fiscal policies and green sovereign bonds, *Ecol. Econ.*, **144** (2018), 228–243. https://doi.org/10.1016/j.ecolecon.2017.07.029
71. M. Guerini, A. Moneta, A method for agent-based models validation, *J. Econ. Dyn. Control*, **82** (2017), 125–141. https://doi.org/10.1016/j.jedc.2017.06.001
72. S. Poledna, M. G. Miess, C. Hommes, K. Rabitsch, Economic forecasting with an agent-based model, *Eur. Econ. Rev.*, **151** (2023), 104306. https://doi.org/10.1016/j.euroecorev.2022.104306
73. M. Catalano, J. Di Domenico, L. Riccetti, Testing climate NGFS scenarios through the lens of a large-scale ABM for the Italian economy, SSRN, 2025, 5277018. Available from: https://ssrn.com/abstract=5277018.
74. C. Hommes, M. He, S. Poledna, M. Siqueira, Y. Zhang, CANVAS: A Canadian behavioral agent-based model for monetary policy, *J. Econ. Dyn. Control*, **172** (2025), 104986. https://doi.org/10.1016/j.jedc.2024.104986
75. J. Di Domenico, M. Catalano, L. Riccetti, Scaling and forecasting in a data-driven agent-based model: Applications to the Italian macroeconomy, *Econ. Model.*, **147** (2025), 107046. https://doi.org/10.1016/j.econmod.2025.107046

76. T. Lux, R. C. J. Zwinkels, Empirical validation of agent-based models, in *Handbook of Computational Economics* (eds. C. Hommes and B. LeBaron), Elsevier, 2018, 437–488. Available from: https://www.sciencedirect.com/handbook/handbook-of-computational-economics/vol/4.
77. A. Turrell, Agent-based models: Understanding the economy from the bottom up, *Bank Engl. Q. Bull.*, (2016), 173–188. Available from: https://www.bankofengland.co.uk/quarterly-bulletin/2016/q4/agent-based-models-understanding-the-economy-from-the-bottom-up.
78. P. Windrum, G. Fagiolo, A. Moneta, Empirical validation of agent-based models: Alternatives and prospects, *J. Artif. Soc. Soc. Simul.*, **10** (2007), Article 8. Available from: https://www.jasss.org/.
79. M. G. Richiardi, The future of agent-based modeling, *East. Econ. J.*, **43** (2017), 271–287. Available from: https://jasss.soc.surrey.ac.uk/10/2/8.html.
80. G. Orlando, M. Bufalo, H. Penikas, C. Zurlo, *Modern Financial Engineering: Counterparty, Credit, Portfolio and Systemic Risks*, World Scientific, 2022. https://doi.org/10.1142/12725
81. G. Orlando, A. N. Pisarchik, R. Stoop, *Nonlinearities in Economics: An Interdisciplinary Approach to Economic Dynamics, Growth and Cycles*, Springer, 2021. Available from: https://link.springer.com/book/10.1007/978-3-030-70982-2.
82. G. Orlando, G. Chironna, Predicting bank defaults in Italy: A comparative analysis of conventional and machine learning approaches, *Econ. Anal. Policy*, **89** (2026), 788–833. https://doi.org/10.1016/j.eap.2025.12.002
83. G. Ascione, M. Bufalo, G. Orlando, Cost and severity of natural catastrophes in extreme events: Implications for society and insurances, *Ann. Oper. Res.*, 2025. https://doi.org/10.1007/s10479-025-06708-3
84. M. Lampart, A. Lampartová, G. Orlando, On risk and market sentiments driving financial share price dynamics, *Nonlinear Dyn.*, **111** (2023), 16585–16604. https://doi.org/10.1007/s11071-023-08702-5
85. G. Orlando, M. Bufalo, Empirical evidences on the interconnectedness between sampling and asset returns' distributions, *Risks*, **9** (2021), 88. https://doi.org/10.3390/risks9050088