# (Deep) Learning to Trade: An Experimental Analysis of AI Trading and Market Outcomes

I. Gufler*, F. Sangiorgi†, E. Tarantino‡

March 19, 2025

**Abstract**

We investigate how AI-driven investors, modeled via deep reinforcement learning, operate in a calibrated financial market with realistic return predictability and price impact. We ask whether these AI agents can detect and exploit return predictability from public signals, decode prices to infer latent demand, and adjust for price impact, and we examine how these behaviors shape market efficiency and liquidity. In our simulations, AI agents' portfolio policies qualitatively mirror those of a rational benchmark. Quantitatively, however, they fall short when many interact: competition distorts the reward signals that AI agents learn from, making it harder for each to isolate its own price impact from that of others. This negative learning externality reduces trading profits, lowering market efficiency and liquidity relative to the rational benchmark.

**Keywords:** Reinforcement learning, Financial stability, Market efficiency, Algorithmic trading, Asset pricing

*Luiss University, ivangufler26@gmail.com

†Frankfurt School of Finance and Management, f.sangiorgi@fs.de

‡Luiss University, etaranti@gmail.com

# 1 Introduction

An open question in finance is whether reinforcement learning (RL) can help traders exploit return predictability and improve investment decisions. Financial markets exhibit complex price dynamics, where asset prices evolve in response to public information and shifts in investor demand. A key challenge for any trader is to detect and act upon patterns in returns while accounting for price formation mechanisms and market frictions. In theory, a rational agent with full knowledge of the market structure can infer latent demand from prices and optimally adjust trading strategies. However, whether artificial intelligence (AI) agents, without prior knowledge of the data-generating process, can achieve similar accuracy remains unclear. Can AI traders learn to interpret price signals effectively, infer shifts in demand, and adapt their trading behavior in ways that enhance or disrupt market efficiency and liquidity?

Addressing this question is central to understanding the role of AI-driven investment strategies in modern financial markets. However, testing it through direct stock market experiments is challenging due to the complexity and evolving structure of real-world trading environments. We conduct an analysis integrating a theoretically grounded financial market structure with an empirically calibrated, micro-founded asset demand framework. Within this setting, we allow deep reinforcement learning algorithms to make trading decisions, assessing their ability to learn from market dynamics without prior knowledge of the data-generating process.

Unlike traditional models that assume traders possess an exogenous structure for learning, our AI agents must identify and exploit return predictability solely through their trading experience, without prior knowledge of market structure. We compare their performance against a rational expectations benchmark, where theoretically informed traders can fully decode the information embedded in prices. This allows us to assess whether reinforcement learning can uncover and act upon predictable price patterns in a way that aligns with, or diverges from rational trading strategies. We contribute to the growing literature on algorithmic trading and market microstructure, offering new insights into the implications of AI-driven financial decision-making.

Financial markets have undergone profound changes due to advances in computing power and algorithmic sophistication. Early algorithmic trading, rooted in fixed, rule-based paradigms, has evolved into data-intensive machine learning systems capable of adjusting to changing market

conditions. Recent evidence suggests that machine learning can enhance the quality of price discovery by filtering information relevant to predicting future returns and by better analyzing corporate success (Bai et al., 2016; Dugast and Foucault, 2018; Farbood et al., forthcoming). At the same time, the disruptive effects of AI-powered trading on equity markets have prompted intensive research into how these new technologies alter market outcomes (Biais et al., 2015; Brogaard et al., 2015; Foucault et al., 2016).

Against this backdrop, we examine how RL agents learn to trade in a realistic market setting. Most existing work on RL employs the tabular Q-learning framework introduced by Watkins (1989) to model economic agents (e.g., Calvano et al., 2020; Colliard et al., 2023; Abada and Lambin, 2023; Johnson et al., 2023, among others). While these algorithms, designed for discrete action spaces, reduce computational complexity, they can lead to significant departures from optimal decision-making. This is particularly problematic when evaluating the accuracy of learning return predictability. To overcome this limitation, we employ deep reinforcement learning algorithms that support continuous action spaces and state representations. Leveraging neural networks, our AI traders can adapt and refine their strategies even in previously unseen market conditions. This is not possible when using standard tabular Q-learning.

Notably, we nest our sophisticated algorithms in an empirically plausible market environment, where the marginal investor's asset demand is calibrated from the data following the demand-based asset pricing approach of Koijen and Yogo (2019). We examine whether such AI traders can detect and exploit sources of predictability from observable fundamentals and partially hidden demand shocks, and how their trading strategies feed back into market efficiency and liquidity.[1]

In our demand-based asset pricing model prices respond both to publicly available signals and to shifts in latent investor demand. Large orders, often related to institutional rebalancing, sentiment, or private information, cause price adjustments that are not directly attributable to changes in observable fundamentals. If RL agents can infer these latent demand components from price dynamics, they may generate excess returns—yet also reshape the very patterns they aim to exploit. Related research shows that RL algorithms can exhibit non-myopic behavior in repeated economic

[1] We are aware of the recent developments in the literature on asset demand estimation (e.g., van der Beck, 2022; Fuchs et al., 2024). We will discuss how they affect the interpretation of our results. However, our emphasis is on the learning capabilities of the AI traders and the qualitative nature of their impact on stock markets.

interactions (Dou et al., 2023; Yang, 2024), occasionally coordinating on outcomes unexpected under classical equilibrium models. Moreover, the potential for behavioral biases in "model-free" algorithms has been emphasized in recent work on portfolio choice (Barberis and Jin, 2023), while the possibility of new forms of liquidity premia or mark-ups emerges when learning-based strategies confront adverse selection (Colliard et al., 2023).

Building on these insights, we construct a market simulator in which the agents do not take price paths as exogenous but modify them through their trades. This choice contrasts with certain computer science approaches that treat traders as atomistic, without impacting prices (Vyetrenko and Xu, 2019; Yang et al., 2020). By modeling endogenous price formation in a microfounded demand environment, we can capture how learning-based trading strategies jointly evolve with, and potentially amplify or dampen, prevailing return predictability.

This leads to another central challenge for our analysis: how can we evaluate the extent to which these sophisticated AI traders are actually learning? A compelling question when using RL algorithms is whether their actions reflect an improvement in their understanding of the environment. To address this, we introduce a rational expectations benchmark that provides theoretical predictions against which we can compare the behavior of our AI-driven traders. A similar approach is taken by Colliard et al. (2023), Dou et al. (2023), and Yang (2024), though their focus is on algorithmic coordination. In our experimental analysis, our emphasis is on assessing the efficacy of learning in complex market settings. To this end, we design our algorithm's learning problem to ensure that they cannot tacitly collude.[2]

The rational benchmark we use is based on a standard asset pricing model with varying number of speculators taking a trading position under alternative initial volumes of wealth. To begin with, Proposition 1 states that an optimal portfolio weight should decline with adjusted market equity. Intuitively, higher market equity reflects a larger inferred latent demand, which—due to mean reversion—signals lower expected capital gains. Moreover, the optimal allocation of the agent to the risky asset is decreasing in the relative size of speculators' wealth. The rationale for this result is related to each agent's internationalization of their own price impact on optimal trading strategies, not the impact of the other agents. Proposition 2 then studies market efficiency. It

[2]Formally, we do this by including only exogenous variables in the state space, which do not depend on the past trading decisions of the AI traders. Specifically, the algorithms do not observe past equilibrium prices.

shows theoretically that market efficiency increases in the size of speculators and in their number. In an efficient benchmark, returns should not be predictable from publicly available signals. Thus, the extent of predictability is a measure market's inefficiency. Since more aggressive trading by speculators, which occurs when they increase in size or number, removes more predictable price patterns, we obtain the theoretical result. Finally, Proposition 3 proves that liquidity increases in speculators size and number. We measure liquidity by the price response to a short-lived shock in asset supply. Rational speculators anticipate that the temporary nature of the price drop leads to an expected capital gain. Thus, as their volume or number increases, by trading more aggressively, they absorb larger supply shocks, thereby increasing liquidity.

Finally, we test whether our AI traders perform as predicted by our theoretical model with rational speculators. Our calibrated simulations reveal that there is a significant qualitative and quantitative accordance between AI traders' behavior and that of rational agents when AI traders are small in size, or operate in isolation. However, as their number, or size increases the similarities are only of qualitative nature. The quantitative discrepancy is to be attributed to the presence of a negative learning externality. The externality is produced by the presence of other agents with price impact, whose identity is unknown by each algorithm, and whose actions introduce further noise in the environment. We identify and quantify this novel effect thanks to our controlled simulations. We show that such externality can amplify or dampen the discrepancy of the AI traders' behavior and our theoretical results.

Overall, our main contributions are twofold. First, we specify an empirically plausible environment where asset price fluctuations reflect both fundamentals and latent investor demand. Second, we show that AI traders employing a continuous-control RL algorithm adapt to this environment in ways that materially affect market outcomes. Their learning decisions can either stabilize or destabilize prices, depending on how effectively they incorporate new data and respond to shifts in demand. This endogeneity underscores the broader impact of AI-based trading on financial stability and efficiency, laying a foundation for further theoretical and empirical exploration.

## 2 The market environment

In this section, we develop a framework in which AI traders interact with a representative investor. We first derive equilibrium asset prices, taking the AI traders' portfolio choices as given. Next, we analyze learning from prices and return predictability in the data-generating process. We then outline a rational expectations benchmark, against which we compare the results of our AI-driven traders.

### 2.1 Equilibrium prices with representative investor

We consider a market for $N$ risky assets and a risk-less asset. Time is discrete and runs from 0 to $\infty$. For each risky asset $n = 1, \ldots, N$, we denote its price and dividend per share in period $t$ with $P_{n,t}$ and $D_{n,t}$, respectively, and its number of shares with $S_n$, which we assume constant over time. Risky asset prices are determined endogenously, as explained below. The risk-free asset is elastically supplied at an exogenous gross return $R_f$ constant over time.

The market is populated by a representative investor and $J$ AI traders. Let $S^j_{n,t}$ be the number of shares of asset $n$ held by the $j$-th AI trader at time $t$. The total holdings of asset $n$ across all AI traders are given by $S^{AI}_{n,t} = \sum_{j=1}^{J} S^j_{n,t}$ while the residual supply held by the representative investor is $\tilde{S}_{n,t} = S_n - S^{AI}_{n,t}$. Throughout the paper, we use lower-case letters to denote logarithms and $\Delta$ to indicate first differences.[3]

To model the representative agent's asset demand, we adapt the framework of Koijen and Yogo (2019) and specify the following log-exponential form:

$$\frac{w_{n,t}}{w_{0,t} + \gamma_t} = \delta_{n,t}; \quad \delta_{n,t} = \exp\left(\beta_0(p_{n,t} + s_n) + \sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t}\right), \tag{1}$$

where $w_{n,t}$ is the representative agent's portfolio weight in asset $n$, $w_{0,t}$ is the weight in the risk-free asset, and $\gamma_t$ represents the fraction of assets consumed by the representative agent, which we define below. The variables $\{x_{n,k,t}\}_{k=1}^{K-1}$ represent publicly observable characteristics for asset $n$ (e.g., book-to-value, profitability, etc.). The term $\epsilon_{n,t}$ represents the representative investor's latent demand for asset $n$, which captures non-fundamental influences on demand.

[3]For example, $p_{n,t} = \log(P_{n,t})$ represents the log price, and $\Delta S^{AI}_{n,t} = S^{AI}_{n,t} - S^{AI}_{n,t-1}$ denotes the change in AI traders' aggregate holdings.

We assume that stock characteristics and latent demand follow the autoregressive processes:

$$x_{k,n,t+1} = c_{k,n} + \rho_{k,n} x_{k,n,t} + \eta_{k,n,t+1}, \tag{2}$$

$$\epsilon_{n,t+1} = c_{\epsilon,n} + \rho_{\epsilon,n} \epsilon_{n,t} + \xi_{n,t+1}, \tag{3}$$

where $\rho_{k,n}, \rho_{\epsilon,n} \in (0,1)$ are the autoregressive coefficients, and $\eta_{k,n,t+1}$ and $\xi_{n,t+1}$ are mean-zero normally distributed innovations, independent over time and across variables, with variances $Var(\eta_{k,n,t}) = \sigma^2_{\eta_{k,n}}, Var(\xi_{n,t}) = \sigma^2_{\xi_n}$.

To conduct our numerical experiments in Section 4, we estimate the representative investor demand coefficients in Eq. (1) and calibrate the parameters of the processes in Eqs. (2)–(3), as detailed in Section 3.

The representative investor's wealth $A$ evolves according to

$$A_t = \sum_{m=1}^{N} \tilde{S}_{m,t-1} \left(P_{m,t} + D_{m,t}\right) + S_{0,t-1} R_f + \Gamma_t, \tag{4}$$

where $S_{0,t-1}$ is the units of the risk-free asset held at time $t-1$ and $\Gamma_t$ is an exogenous inflow of resources which is specified below. Similarly, AI trader-$j$'s wealth $A^j$ evolves according to

$$A^j_t = \sum_{m=1}^{N} S^j_{m,t-1} \left(P_{m,t} + D_{m,t}\right) + S^j_{0,t-1} R_f. \tag{5}$$

The market clearing condition for each stock $n$ reads as follows:

$$A_t w_{n,t} + P_{n,t} S^{AI}_{n,t} = P_{n,t} S_n. \tag{6}$$

We will initially take AI traders' holdings in all securities $\{S^j_{n,t}\}$ as given and derive the equilibrium price function for each security as implied by (i) the representative investor's demand in Eq. (1), (ii) the market clearing condition in Eq. (6), and (iii) the representative agent and the AI investors asset dynamics in Eqs. (4)-(5) together with the respective budget constraints. In Appendix A.1,

we derive the equilibrium price function, which takes the form:

$$p_{n,t} = \frac{\beta_0 s_n - \tilde{s}_{n,t} + \sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t} + \log\left(D_{M,t} + S_{0,t-1}R_f + \Gamma_t + S_{0,t-1}^{AI}R_f - S_{0,t}^{AI}\right)}{1-\beta_0}, \tag{7}$$

where the aggregate dividend $D_{M,t}$ is given by

$$D_{M,t} = \sum_{m=1}^{N} S_m D_{m,t}.$$

We make the following additional assumptions to enhance analytical tractability:

**Assumption 1.** *The exogenous change in the representative investor's assets equals* $\Gamma_t = S_{0,t}^{AI} - S_{0,t-1}^{AI}R_f$.

**Assumption 2.** *The representative investor's consumption at time* $t$ *equals* $\gamma_t A_t = \lambda(D_{M,t} + S_{0,t-1}R_f)$*, where* $\lambda \in (0,1)$ *is a constant.*

**Assumption 3.** *The aggregate dividend grows at the constant rate* $g$*:* $D_{M,t} = (1+g)D_{M,t-1}$.

**Assumption 4.** *The parameters* $\lambda, g, R_f$ *satisfy the condition* $1 + g > R_f(1-\lambda)$.

Assumption 1 eliminates the indirect cross-asset effects of AI trading in the risk-free asset, as it simplifies the argument of the logarithm in the right-hand side of Eq. Eq. (7) to $D_{M,t} + S_{0,t-1}$. Assumption 2 implies that $D_{M,t} + S_{0,t-1}$ can be expressed as a weighted sum of past dividends, $D_{M,t-1}, \ldots, D_{M,0}$ (see Eq. (A5) in Appendix A.1). Assumption 3 removes aggregate uncertainty from the price function by ensuring that aggregate dividends grow at a deterministic rate. Assumption 4 guarantees the convergence of the sum in Eq. (A5) as $t \to \infty$.

In Appendix A.1, we prove that under Assumptions 1–4, the price function in Eq. (7) simplifies to

$$p_{n,t} = -s_n + \frac{-\log\left(1-\alpha_{n,t}^{AI}\right) + \sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t} + \log\left(D_{M,t}\right) + \phi}{1-\beta_0}, \tag{8}$$

where

$$\alpha_{n,t}^{AI} = \frac{S_{n,t}^{AI}}{S_n}; \quad \phi = \log\left(\frac{(1+g)}{1+g-R_f(1-\lambda)}\right).$$

The price function in Eq. (8) has an intuitive interpretation. The price of asset $n$ is inversely related to its supply $s_n$ and positively related to its observable characteristics, weighted by the demand coefficients, $\sum_{k=1}^{K-1} \beta_k x_{k,n,t}$. Additionally, it increases with the aggregate market dividend $D_{M,t}$. These variables are public information. The price is also influenced by the latent demand $\epsilon_{n,t}$, which, in contrast, is not publicly observable. In the next subsection, we examine how investors can infer this variable from market prices.

Finally, AI investors' demand also affects the price. Eq. (8) implies that the elasticity of price with respect to the residual supply held by the representative investor is[4]

$$-\frac{\partial p_{n,t}}{\partial \log(\tilde{S}_{n,t})} = \frac{1}{1-\beta_0}.$$

This means that if AI traders purchase an additional 1% of the asset supply, the price $P_{n,t}$ increases by $1/(1-\beta_0)\%$. Intuitively, the price impact is larger when $\beta_0$ is closer to one, that is, when the demand of the representative investor is less elastic.

## 2.2 Learning and predictability in the data-generating process

The representative investor's latent demand contributes to price variability by reflecting investor sentiment and subjective beliefs, capturing components of demand not directly explained by asset fundamentals. Koijen and Yogo (2019) document that this latent demand accounts for a significant portion of cross-sectional variation in stock returns.

Although latent demand is unobservable, it influences prices in conjunction with publicly available, payoff-relevant information. Investors who understand the price formation process can use prices—together with other public signals—to infer latent demand. This process of information extraction from prices is central to the asset pricing literature on information frictions (e.g., Grossman and Stiglitz (1980)).

In this subsection, we first examine how investors would learn about latent demand and expected returns under the assumption that they are rational and fully understand the price formation process. This "fully revealing" rational expectations equilibrium serves as a benchmark against which we compare the outcomes of algorithmic learning.

[4]Notice that $\log\left(1-\alpha_{n,t}^{AI}\right) = \log\left(\frac{S_n - S_{n,t}^{AI}}{S_n}\right) = \log\left(\tilde{S}_{n,t}\right) - s_n$.

We define $P^*_{n,t}$ to be the "beginning of period price," or equivalently, the equilibrium price that would prevail if the AI traders did not change their portfolios from time $t-1$ to $t$ (i.e., $\Delta S^{AI}_{n,t} = 0$). By Eq. (8) we have:

$$p^*_{n,t} = f_n\left(\alpha^{AI}_{t-1}, \{x_{n,k,t}\}_{k=1}^{K-1}, D_{M,t}\right) + \frac{\epsilon_{n,t}}{1-\beta_0}, \tag{9}$$

where

$$f_n\left(\alpha^{AI}_{t-1}, \{x_{n,k,t}\}_{k=1}^{K-1}, D_{M,t}\right) = -s_n + \frac{-\log\left(1-\alpha^{AI}_{n,t-1}\right) + \sum_{k=1}^{K-1}\beta_k x_{k,n,t} + \beta_K + \log\left(D_{M,t}\right) + \phi}{1-\beta_0}. \tag{10}$$

Here, $p^*_{n,t}$ incorporates updated public information $\left\{x_{n,k,t}\right\}_{k=1}^{K-1}, D_{M,t}$ and latent demand $\epsilon_{n,t}$, but not the time-$t$ trading decisions by AI investors. Hence, $p^*_{n,t}$ is the "pre-trade" price level.

Consider a rational expectations benchmark in which,at the beginning of each period $t$, investors observe $p^*_{n,t}, \alpha^{AI}_{t-1}, \{x_{n,k,t}\}_{k=1}^{K-1}$, and $D_{M,t}$. Because investors understand the price function, grom Eq. (9) it follows that

$$E_t\left(\epsilon_{n,t}\right) = (1-\beta_0)\left[p^*_{n,t} - f_n\left(\alpha^{AI}_{t-1}, \{x_{n,k,t}\}_{k=1}^{K-1}, D_{M,t}\right)\right] = \epsilon_{n,t}, \tag{11}$$

where $E_t(\cdot)$ denotes the rational expectations operator. Thus, $p^*_{n,t}$ fully reveals the value of the latent demand to rational investors. In our numerical experiments, we assess the extent to which AI traders can replicate this learning process.

We next consider return predictability. For simplicity, we assume that the dividend yield $D_{n,t+1}/P_{n,t}$ is i.i.d. over time and focus on the capital gain.[5] From Eq. (8), the capital gain is

$$\frac{P_{n,t+1}}{P_{n,t}} = \exp\left(\frac{-\Delta\log\left(1-\alpha^{AI}_{n,t+1}\right) + \sum_{k=1}^{K-1}\beta_k \Delta x_{k,n,t+1} + \Delta\epsilon_{n,t+1} + \log\left(1+g\right)}{1-\beta_0}\right). \tag{12}$$

Consider the case where $\alpha^{AI}_{n,t+1} = 0$ and investors correctly anticipate the equilibrium value of $\alpha^{AI}_{n,t}$.

[5] See Koijen and Yogo (2019) for evidence that dividend yields explain only 0.4% of the cross-sectional variance of stock returns, suggesting that capital gains drive most price variation.

Using Eqs. (2)–(3), we obtain the expected capital gain:

$$E_t\left(\frac{P_{n,t+1}}{P_{n,t}}\bigg|_{\alpha^{AI}_{n,t+1}=0}\right) = \exp\left(-\frac{\sum_{k=1}^{K-1}\beta_k(1-\rho_{k,n})x_{k,n,t}+(1-\rho_{\epsilon_n})\epsilon_{n,t}}{1-\beta_0}\right)\left(1-\alpha^{AI}_{n,t}\right)^{\frac{1}{(1-\beta_0)}}\Phi, \quad (13)$$

where $\Phi$ is a constant value derived in Eq. (A15) in the Appendix. Eq. (13) shows that mean reversion in asset characteristics and latent demand induces return predictability in the data generating process.

In our numerical experiments, we explore how AI traders learn to trade on these predictable components and study the corresponding equilibrium outcomes.

### 2.3 The rational expectations equilibrium benchmark

In this subsection we consider a one-period trading game with strategic, rational speculators. This special case provides a benchmark against which we can compare the results of our numerical simulations with AI investors in Section 4.

Consider the setup from Section 2.1 where AI investors are replaced by rational speculators who enter the market at time $t$, trade, and then closer their positions and exit at time $t+1$. Each of the $J$ speculators allocates their wealth across risky asset $n$ and the riskless asset to maximize the expected next period wealth. Equivalently, each speculator maximizes the expectation next-period portfolio return

$$R^j_{p,t+1} = R_f + \theta^j_{n,t}\left(R_{n,t+1}-R_f\right); \quad R_{n,t+1} = \frac{P_{n,t+1}}{P_{n,t}} + DY_{n,t+1},$$

where $DY_{n,t+1} = \frac{D_{n,t+1}}{P_{n,t}}$ is the dividend yield which we assume is exogenous and unaffected by $P_{n,t}$. Prices are given by Eq. (8) where we denote $\alpha^S_{n,\tau} = \sum_{j=1}^{J}\alpha^j_{n,\tau}$ the fraction of the supply of shares held collectively by the speculators at time $\tau$. Since they enter at time $t$ and liquidate their holdings at time $t+1$, we have $\alpha^S_{t-1} = \alpha^S_{t+1} = 0$, whereas $\alpha^S_t$ is determined in equilibrium.

The trading game unfolds as follows. Each speculator enters the market at $t$ with initial wealth $A^s_t = \omega P^*_{n,t}S_n$, where $\omega < 1/J$. That is, all speculators have the same initial wealth and collectively their wealth equals a fraction $J\omega$ of the beginning-of-period market capitalization of asset $n$. Each speculator observes the public information $\mathcal{I}_{n,t} = \{p^*_{n,t}, \{x_{n,k,t}\}_{k=1}^{K-1}, D_{M,t}\}$. Then,

conditional on this information, they simultaneously choose a risky asset portfolio share subject to no-short selling and no-borrowing constraints. That is, a speculator's strategy is a function $\theta : \mathcal{I} \to [0, 1]$. Given speculator choices of portfolio shares $\{\theta^j_{n,t}\}$, the asset price is determined as in Eq. (8), which, using the definition $S^j_{n,t} = A^s_t \theta^j_{n,t}/P_{n,t}$ to substitute for $S^j_{n,t}$ in Eq. (8) can be rewritten as an implicit function of $\{\theta^j_{n,t}\}$ as follows:

$$0 = \beta_0(p_{n,t} + s_n) - \log\left(S_n P_{n,t} - \sum_{j=1}^{J} \theta^j_{n,t} A^s_t\right) + \sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t} + \log(D_{M,t}) + \phi. \quad (14)$$

Since speculators close their positions in $t + 1$, $p_{n,t+1}$ equals

$$p_{n,t+1} = -s_n + \frac{\sum_{k=1}^{K-1} \beta_k x_{k,n,t+1} + \beta_K + \epsilon_{n,t+1} + \log\left(D_{M,t+1}\right) + \phi}{1 - \beta_0}. \quad (15)$$

When choosing their portfolio share, a rational speculator correctly anticipates the effect of their price impact on the expected portfolio return:

$$\frac{\partial E_t(R^j_{p,t})}{\partial \theta^j_{n,t}} = E_t\left[\frac{P_{n,t+1}}{P_{n,t}}\left(1 - \underbrace{\frac{\theta^j_{n,t}}{P_{n,t}}\frac{\partial P_{n,t}}{\partial \theta^j_{n,t}}}_{\text{price impact factor}}\right)\right] + E_t(DY_{n,t+1}) - R_f. \quad (16)$$

Taking other traders' strategies as given, we can calculate the price impact factor in Eq. (16) from implicit differentiation of Eq. (14) as follows:

$$\frac{\theta^j_{n,t}}{P_{n,t}}\frac{\partial P_{n,t}}{\partial \theta^j_{n,t}} = \frac{\theta^j_{n,t} A^s}{(1 - \beta_0) S_n P_{n,t} + \beta_0 \sum_{i=1}^{J} \theta^i_{n,t} A^s}. \quad (17)$$

For expositional simplicity we define "adjusted market equity" as

$$me^*_{n,t} = p^*_{n,t} + s_n - \frac{\log\left(D_{M,t+1}\right)}{1 - \beta_0}, \quad (18)$$

and we rewrite speculators' information set more compactly as $\mathcal{I}_{n,t} = \{me^*_{n,t}, \{x_{n,k,t}\}_{k=1}^{K-1}\}$. This is without loss of generality because (i) the joint observation of $\{me^*_{n,t}, \{x_{n,k,t}\}_{k=1}^{K-1}\}$ fully reveals the

latent demand,[6] and (ii) neither the capital gain in Eq. (12) nor the price impact factor in Eq. (17) depend on $D_{M,t}$.

**Equilibrium portfolio choice** An equilibrium of the trading game is a collection of portfolio shares $\{\theta^j_{n,t}\}$ that are optimal for each speculator given their information $\mathcal{I}_{n,t}$ and given the price functions in Eq. (14)-(15). The following proposition shows properties of the equilibrium.[7]

**Proposition 1.** *An equilibrium of the trading game exists, is unique, and is symmetric. The equilibrium portfolio weight in the risky asset is a function $\theta(\mathcal{I})$ such that*

*(i) $\theta$ is decreasing in adjusted market equity $me^*_{n,t}$;*

*(ii) $\theta$ is increasing in asset characteristic $x_{n,k,t}$ if and only if $\beta_k(\rho_{n,k} - \rho_{\epsilon,n}) > 0$;*

*(iii) $\theta$ is decreasing in each speculator's relative size $\omega$ (size effect) and, fixing speculators' aggregate relative size $\omega J$, $\theta$ is increasing in the number of speculators $J$ (competition effect).*

The comparative statics results in Proposition 1 are intuitive. First, for fixed values of $(x_{n,1,t}, \ldots, x_{n,K-1,t})$, a larger value of $me^*_{n,t}$ reveals a larger contemporaneous value of the latent demand (see Eq. (19) in Footnote 6), which, because of mean reversion, predicts a lower capital gain (see Eq. (13)). In response, speculators allocate a smaller fraction of their wealth to the risky asset.

Second, a larger value of the characteristic $x_{n,k,t}$ has two effects. First, it decreases the expectation of $\Delta x_{k,n,t+1}$ due to mean reversion, thereby affecting the expected capital gain depending on the sign of $\beta_k$ (see Eq. (19)). Second, for a fixed value of $me^*_{n,t}$, a larger value of the characteristic $x_{n,k,t}$ also affects the expectation of latent demand (see Eq. (19)). The overall effect depends on the relative mean reversion speeds of $\epsilon_n$ and $x_{n,k}$ times the sign of $\beta_k$ as shown in the proposition.

[6] Notice that Eq. (9) for $\alpha^{AI}_{n,t-1} = 0$ and Eq. (18) imply

$$E_t(\epsilon_{n,t}) = (1-\beta_0)me^*_{n,t} - \left(\sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \phi\right) = \epsilon_{n,t}, \tag{19}$$

[7] In the text of the proposition, "increasing" and "decreasing" should be understood as "weakly increasing" and "weakly decreasing" because of the possibility of binding no-short sale or borrowing constraint, in which case the portfolio choice is at the corner $\theta = 0$ or $\theta = 1$.

Third, the optimal allocation to the risky asset is decreasing in the relative size of speculators' wealth $\omega$ because of their price impact. Finally, the optimal allocation to the risky asset is increasing in the number of speculators $J$ fixing the overall wealth of the speculators. This is because each speculator only internalizes their individual price impact, not the price impact of others. As a result, when a fixed amount of wealth is divided by a larger number of speculators, they individually trade more aggressively on return predictability, approaching the competitive case when the number of speculators becomes infinite.

Next, we analyze how speculators' trading affects two key market characteristics: market efficiency and liquidity.

**Market Efficiency.** In the benchmark model with only a representative investor, price changes are predictable from public information. When risk-neutral speculators enter, they compete for profits based on this predictability. To capture this effect formally, we consider a measure of semi-strong market efficiency, whereby returns should not be predictable from publicly available information. Hence, the extent to which returns can be forecast from public signals directly measures market (in)efficiency.

In equilibrium, the (log) capital gain can be written as

$$p_{n,t+1} - p_{n,t} \;=\; g_n(\mathcal{I}_t) + e_{n,t+1},$$

where $\mathcal{I}_t$ is the public information set, $g_n(\mathcal{I}_t)$ is the predictable component of the capital gain as implied by Eqs. (14)-(15), and $e_{n,t+1}$ is the unpredictable component. $g_n(\cdot)$ and $e_{n,t+1}$ are derived in Eqs. (A19)-(A20) in the Appendix. We measure market efficiency as the fraction of variance that is not predictable by public information:

$$\mathcal{ME} = \frac{\mathrm{Var}\big(e_{n,t+1}\big)}{\mathrm{Var}\big(p_{n,t+1} - p_{n,t}\big)}. \tag{20}$$

If speculators completely eliminate predictable components, we have $\mathcal{ME} = 1$. In equilibrium, increasing either the speculators' size or their number makes them trade collectively more aggressively (Proposition 1), thereby raising $\mathcal{ME}$.

An alternative measure of return predictability is the following direct measure of "pre-

dictable strength"

$$\mathcal{P} \;=\; E\big[\,|E_t(R^e_{n,t+1})|\big],$$

which captures the average absolute deviation of expected excess returns from zero. In the model, speculators' trading on the predictive signals drives $|E_t(R^e_{n,t+1})|$ toward zero, more so when they have more capital or when they compete more aggressively.

**Proposition 2.** *Market efficiency increases in speculators' relative size $\omega$ and, fixing speculators' aggregate relative size $\omega J$, in the number of speculators $J$.*

**Market Liquidity.** Next, we measure liquidity by the price response to an exogenous "shock" in asset supply. Suppose (log) supply increases from $s_n$ to $s_n + \sigma$ at time $t$ before reverting back to $s_n$ at $t+1$. In the case where only the representative investor is present (i.e., assuming $\alpha_t^{AI} = 0$), Eq. (8) implies that the price impact of this supply shock is given by $\frac{\partial p_{n,t}}{\partial \sigma} = -1$, so the price drops one-for-one with supply. In contrast, when speculators are active in the market, they recognize that the price drop is temporary and anticipate an expected capital gain when the supply reverts at $t+1$. As a result, they increase their demand for the asset, partially absorbing the supply shock and reducing its price impact.

We define liquidity as

$$\mathcal{L} = 1 + E\left(\left|\frac{\partial p_{n,t}}{\partial \sigma}\right|_{\sigma=0}\right). \tag{21}$$

This measure captures the inverse of price impact: higher values of $\mathcal{L}$ indicate that the market can absorb larger supply shocks with smaller price distortions, reflecting greater liquidity.

In the absence of speculators, the full price impact of a supply shock leads to $\mathcal{L} = 0$. Conversely, if speculators fully absorb the shock, there is no price impact, implying $\mathcal{L} = 1$. As $\omega$ (the relative size of speculators) or $J$ (the number of speculators) increase, speculators trade more aggressively on mispricing, allowing the market to absorb larger shocks, thereby reducing price impact and increasing liquidity.

**Proposition 3.** *Liquidity increases with the relative size of speculators $\omega$ and, for a fixed total speculator presence $\omega J$, liquidity increases with the number of speculators $J$.*

### 2.3.1 Assessing algorithmic learning

In contrast to the rational benchmark, to correctly use the information in $\mathcal{I}_t$ for predicting returns and choosing portfolio weights, AI investors must simultaneously learn (i) the equilibrium mapping that links $\mathcal{I}_t$ to future returns, and (ii) the mapping from actions (i.e., portfolio weights) to rewards (i.e., portfolio returns). The algorithm's decision takes the form of portfolio weights, so we can only test the joint hypothesis that the algorithm learns both to make predictions and use them optimally in its portfolio decisions.

# 3 Empirical implementation

In this section, we outline the calibration and estimation of our model to reflect empirically observed market conditions. First, we outline the demand estimation procedure for the representative investor, based on Koijen and Yogo (2019). Second, we detail how we simulate stock characteristics, dividends, and latent demand processes. Third, we describe our deep reinforcement learning (DRL) approach and how it is integrated into the environment. Finally, we outline how we run and evaluate our simulation experiments.

## 3.1 Demand estimation

We calibrate the demand of the representative investor using data on US investors' holdings from SEC 13F filings, combined with asset characteristics from Compustat and CRSP. The sample spans the period from 1982:Q2 to 2021:Q4. We consider log market equity (me) and five additional exogenous stock characteristics: log book equity (be), investment growth (inv), dividend-to-book equity (div), profitability (prof), and market beta (mkt).

We begin by estimating individual investors' demand functions using the log-exponential specification of Koijen and Yogo (2019), and then aggregate these estimates to construct a representative demand curve. Following Koijen and Yogo (2019), for each investor $i = 1, \ldots, I$ and each quarter $t = 1, \ldots, T$, we estimate the following equation, where the elasticity of investor $i$ is identified using cross-sectional variations in stock characteristics:

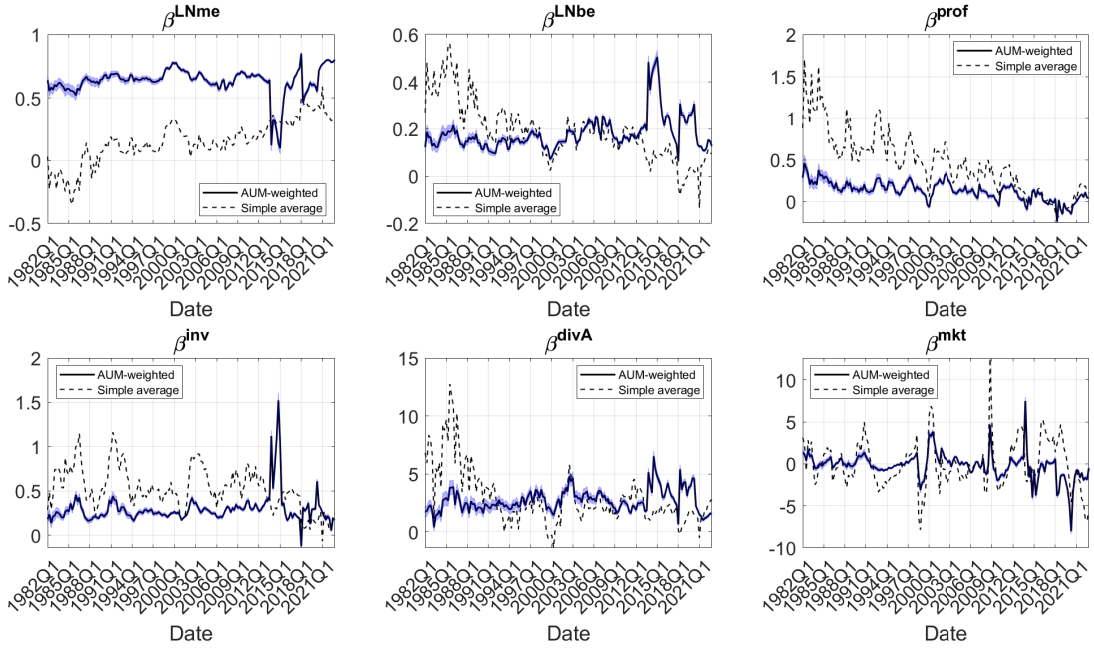
$$\frac{w_{i,n,t}}{w_{i,0,t}} = \exp\{\beta^{me}_{i,t}\,\mathrm{me}_{t,n} + \beta^{be}_{i,t}\,\mathrm{be}_{t,n} + \beta^{prof}_{i,t}\,\mathrm{prof}_{t,n} + \beta^{inv}_{i,t}\,\mathrm{inv}_{t,n} + \beta^{div}_{i,t}\,\mathrm{div}_{t,n} + \beta^{mkt}_{i,t}\,\mathrm{mkt}_{t,n} + \beta^{0}_{i,t} + \epsilon_{i,n,t}\}. \tag{22}$$

The dependent variable represents the holdings (in US dollars) of risky asset $n$ by investor $i$ at time $t$, relative to the outside asset (i.e., cash and other non-equity holdings). The error term $\epsilon_{i,n,t}$ captures latent demand, reflecting investor sentiment, private information, and beliefs.

To address price endogeneity, we follow Koijen and Yogo (2019) and instrument market equity using its counterfactual value, assuming that all other investors hold an equal-weighted portfolio within their investment universe.

Figure 1 illustrates the estimated demand coefficients, showing how average demand sensitivities (across investors) vary by stock characteristic and over time.

Figure 1: Estimated demand coefficients for individual investors.




Next, we use these estimated coefficients to compute the demand of a representative investor.

## 3.2 Calibration of the representative investor's demand

The demand coefficients of the representative investor are computed as an assets-under-management (AUM) weighted average of the individual investors' demand coefficients. To ensure

time-invariant parameters, we take the AUM-weighted average across the full sample period:

$$\bar{\beta}^k = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{I}\frac{AUM_{i,t}}{\sum_{i=1}^{I}AUM_{i,t}}\beta_{i,t}^k. \tag{23}$$

Table 1 reports the estimated values of these representative demand coefficients.

Table 1: Demand coefficients of the representative investor

| $\bar{\beta}^{me}$ | $\bar{\beta}^{be}$ | $\bar{\beta}^{prof}$ | $\bar{\beta}^{inv}$ | $\bar{\beta}^{div}$ | $\bar{\beta}^{mkt}$ |
|---|---|---|---|---|---|
| 0.6327 | 0.1810 | 0.1274 | 0.2947 | 2.6960 | -0.1984 |

*Notes:* This table reports the demand coefficients used to calibrate the representative investor. Coefficients are an assets-under-management weighted average of individual investors' demand coefficients estimated according to Eq. 22.

Similarly, the latent demand of the representative investor is computed as the AUM-weighted average of individual investors' latent demand:

$$\epsilon_{n,t} = \sum_{i=1}^{I}\frac{AUM_{i,t}}{\sum_{i=1}^{I}AUM_{i,t}}\epsilon_{i,n,t}. \tag{24}$$

This calibrated demand function provides the foundation for the equilibrium price computation in Section 2.1.

### 3.3 Simulating stock characteristics and latent demand

We simulate stock characteristics, dividends, and latent demand for each stock. Specifically, we select ten stocks that remained active throughout the entire sample period (1982:Q2 to 2021:Q4). The list of stocks used to assess the AI investor's trading ability is presented in Table 2.

For each of the five stock characteristics—log book equity, profitability, investment, dividend-to-book equity, and market beta—as well as latent demand, we estimate an AR(1) process using the simulated method of moments (SMM), based on Eqs. (2)-(3). Using data from CRSP and Compustat, we estimate stock-specific parameters, including the constant, autoregressive coefficient, and variance of innovations, and then generate simulated series for each variable. Details on the AR(1) process estimation and summary statistics of the mean and standard deviation of stock characteristics from the data and the simulation are provided in Appendix A.3.

For dividends, we compute the dividend yield as the ratio of the current period dividend, $D_{n,t}$, to the previous period's price, $P_{n,t-1}$. In our simulations, we sample the dividend yields from

Table 2: Company Information and Business Sectors

| Ticker | Company Name | Business Sector |
|---|---|---|
| IBM | International Business Machines Corporation | Information Technology |
| AXP | American Express Company | Financials |
| CPB | Campbell Soup Company | Consumer Staples |
| GE | General Electric Company | Industrials |
| XOM | Exxon Mobil Corporation | Energy |
| INTC | Intel Corporation | Information Technology |
| KO | The Coca-Cola Company | Consumer Staples |
| PFE | Pfizer Inc. | Health Care |
| SYY | Sysco Corporation | Consumer Staples |
| XRX | Xerox Holdings Corporation | Information Technology |

*Notes:* This table reports the ticker, company name and business sector for each stock in our dataset.

the data series with replacement. Appendix A.3 reports the mean and standard deviation of the simulated exogenous variables.

### 3.4 Calibration of other parameters

**Consumption rate ($\lambda$).** The consumption rate $\lambda$ is set to 0.99. Since this parameter does not affect stock returns, it is chosen to ensure that equilibrium price levels are on the same scale as those observed in the data.

**Aggregate dividend growth ($g$).** The aggregate dividend growth rate, $g$, is calibrated by minimizing the mean absolute error (MAE) between empirical stock returns and those obtained from simulations. This process incorporates simulated stock characteristics, dividends, and latent demand. The optimal value is found to be 0.4%.

**Variance of innovations of latent demand.** To align realized and simulated return volatility, we calibrate a parameter ($\alpha$) to scale the variance of innovations in latent demand. The optimal value of $\alpha$ is determined by minimizing the MAE between the standard deviation of stock log returns from empirical data and those obtained using simulated exogenous variables. The estimated value of $\alpha$ is 0.4916.

**Risk-free rate ($R_f$).** We calibrate the risk-free rate using the 3-month Treasury bill rate from 1982:Q2 to 2021:Q4, retrieved from FRED Economic Data. The annualized average over this period

is 3.54%.

### 3.5 Reinforcement Learning Model

In this study, we employ the Deep Deterministic Policy Gradient (DDPG) algorithm (Lillicrap et al., 2015) to model AI traders. The choice of DDPG over traditional tabular Q-learning methods is motivated by its ability to handle high-dimensional state spaces and continuous action spaces, which are crucial in our portfolio choice application where portfolio weights cannot be discretized without loss of generality. Unlike standard Q-learning, which is limited to discrete action sets and suffers from the curse of dimensionality, DDPG leverages deep function approximation and policy-based learning, making it more suited to complex financial decision-making (Yang et al., 2020).

The DDPG algorithm is based on an actor-critic architecture. The critic network, parameterized by $\Theta^Q$, approximates the state-action value function $Q(\mathcal{S}, a)$ and evaluates the expected future rewards of a given state-action pair. The actor network, parameterized by $\Theta^\mu$, determines the optimal action given the current state and updates the policy in a continuous action space. To ensure stable training, the algorithm employs target networks, denoted as $\Theta^{Q'}$ and $\Theta^{\mu'}$, which are updated gradually to mitigate instability during learning. Additionally, a replay buffer stores past transitions $\{\mathcal{S}, a, \pi, \mathcal{S}'\}$, breaking temporal correlations and improving learning efficiency.

At each time step, the AI trader observes the current market state $\mathcal{S}$, selects an action $a$, and receives a reward $\pi$. The transition tuple $\{\mathcal{S}, a, \pi, \mathcal{S}'\}$ is stored in the replay buffer $\mathcal{B}$. Training updates occur by sampling minibatches of size $|B|$ from $\mathcal{B}$ and computing target Q-values:

$$y\left(\pi, \mathcal{S}'\right) = \pi + \gamma \Theta^{Q'}\left(\mathcal{S}', \Theta^{\mu'}\left(\mathcal{S}'\right)\right), \tag{25}$$

where $\gamma$ is the discount factor. The critic network minimizes the loss:

$$L\left(\Theta^Q\right) = \mathbb{E}\left[\left(y - Q\left(\mathcal{S}, a|\Theta^Q\right)\right)^2\right]. \tag{26}$$

The actor network updates its policy by maximizing the expected return:

$$\mathbb{J}(\Theta^\mu) = \mathbb{E}[Q(\mathcal{S}, \mu(\mathcal{S}))], \tag{27}$$

with its gradient computed as:

$$\nabla_{\Theta^\mu}\mathbb{J}(\Theta^\mu) \approx \mathbb{E}\left[\nabla_a Q(\mathcal{S}, a)\nabla_{\Theta^\mu}\mu(\mathcal{S}|\Theta^\mu)\right]. \tag{28}$$

Target networks are softly updated according to:

$$\Theta^{\mu'} \leftarrow \tau\Theta^\mu + (1-\tau)\Theta^{\mu'},$$
$$\Theta^{Q'} \leftarrow \tau\Theta^Q + (1-\tau)\Theta^{Q'},$$

where $\tau \ll 1$ controls the update rate.

Our implementation follows standard practices in deep reinforcement learning. Both the actor and critic networks consists of two fully connected hidden layers with 400 and 300 neurons, respectively. The final layer of the actor network employs a *softmax* activation function, ensuring that chosen actions form a valid probability distribution. To enhance exploration, we introduce Dirichlet noise to the policy, balancing stochasticity and directed search for optimal strategies. The network weights are initialized using the Glorot normal distribution. The agent learns from mini-batches of size 500 and to ensure sufficient variability in the data we start the learning process when the replay buffer includes at least 10,000 transitions.

The choice of DDPG aligns with our goal of modeling sophisticated AI trading strategies in continuous, high-dimensional environments. Unlike tabular Q-learning methods used in prior studies (e.g., Calvano et al. 2020), which require discretized action spaces and exhibit slow convergence, DDPG efficiently scales to realistic trading scenarios, allowing AI traders to discover optimal policies dynamically. Table 3 summarizes the key hyperparameters used in our implementation.

Table 3: Model parameters used in the implementation, following Lillicrap et al. (2015) and Yang et al. (2020).

| | | |
|---|---|---|
| $\tau$ | 0.001 | Target update rate |
| $\alpha$ | 0.0001 | Actor learning rate |
| $\beta$ | 0.001 | Critic learning rate |
| $\mathcal{B}$ | 500 | Batch size |
| $\lvert B\rvert$ | $10^5$ | Replay buffer size |
| $\gamma$ | 0.99 | Discount rate |

*Notes:* This table reports the hyperparameters of the reinforcement learning model.

### 3.6 Investigation strategy

In our simulations (experiments) we focus on the 10 large-cap U.S. stocks in Table 2; this choice keeps the dimensionality tractable yet reflect a diverse range of fundamentals and latent-demand dynamics. Let $\mathcal{N} = \{\text{IBM}, \ldots, \text{XRX}\}$ denote the set of stocks considered in our simulations. The set $\mathcal{J} = \{1, 2, 5\}$ represents the set of AI traders, and $\Omega = \{1\%, 5\%, 10\%\}$ specifies the set AI trader's initial wealth relative to the beginning-of-period market capitalization of the asset. For each triple $(n, J, \omega) \in \mathcal{N} \times \mathcal{J} \times \Omega$ we run a total of $S$ independent simulations.[8] In each simulation, the AI traders face $H$ independent episodes of $T$ periods each; the time series of stock characteristics and latent demand is simulated as explained in Section 3.3. AI traders live each episode $\tau$ times. We set $S = 50, H = 100, T = 95, \tau = 5$.

In each period, a trader's information set in asset $n$ is $\mathcal{I}_{n,t} = \{me^*_{n,t}, \{\boldsymbol{x}_{n,t}\}\}$, where $\boldsymbol{x}_{n,t}$ is the vector of stock characteristics for stock $n$ at time $t$. Given $\mathcal{I}_{n,t}$, the AI trader decides the fraction $\theta^j_{n,t} \in [0, 1]$ of its wealth to allocate to the stock. Given all AIs' portfolio decisions, the market clears and $P_{n,t}$ is determined. Then, the AI observes its reward given by the portfolio return $R^j_{p,t+1}$ and the next-period state $\mathcal{I}_{n,t+1}$. The tuple $(\mathcal{I}_{n,t}, \theta^j_{n,t}, R^j_{p,t+1}, \mathcal{I}_{n,t+1})$ constitutes a transition, which is the basis for algorithmic learning as explained in Section 3.5.

For each simulation $s$ and triple $(n, J, \omega)$ we obtain a set of $J$ portfolio policy functions, one per AI trader, such that $\theta^j_{n_s}(\cdot; \omega, J) : \mathcal{I} \to [0, 1]$. We define the average policy as $\theta^{AI}_{n_s}(\cdot; \omega, J) = \frac{\sum_{j=1}^{J} \theta^j_{n_s}(\cdot; \omega, J)}{J}$.

We employ a variety of statistics to assess the AI trader's ability to learn how to trade and the market outcome of AI trading. First, for each $(n, J, \omega)$ we compute the distance between the AIs' average portfolio policy and the equilibrium portfolio policy in the rational benchmark, $\mathcal{D}^n_{\{AI,b\}}(\omega, J)$, defined as

$$\mathcal{D}^n_{\{AI,b\}}(J, \omega) := \frac{\sum_{s=1}^{S}}{S} \frac{\sum_{h=1}^{H}}{H} \frac{\sum_{t_h=1}^{T}}{T} |\theta^{AI}_{n_s}(\mathcal{I}_{t_h}; \omega, J) - \theta^b_n(\mathcal{I}_{t_h}; \omega, J)|. \tag{29}$$

This measure represents the average distance between the investment in the stock as a fraction of AI traders' total wealth and the corresponding value in the rational benchmark, where the average

[8] Running multiple independent simulations mitigates the effect of random neural network weight initialization. Additionally, randomness in exploration and mini-batch selection remains independent across simulations.

is computed over all time periods and simulations.

In a similar way, for each $(n, J, \omega)$ whit $J > 1$, we compute the average distance between among AIs' portfolios, $\mathcal{D}_{\{AI\}}$, defined as

$$\mathcal{D}_{\{AI\}}(\omega, J) := \frac{\sum_{s=1}^{S}}{S} \frac{\sum_{j=1}^{J} \sum_{i=1, i\neq j}^{J}}{J(J-1)} \frac{\sum_{h=1}^{H}}{H} \frac{\sum_{t_h=1}^{T}}{T} |\theta_{n_s}^{j}(\mathcal{I}_{t_h}; \omega, J) - \theta_{n_s}^{i}(\mathcal{I}_{t_h}; \omega, J)| \tag{30}$$

This measure represents the average pairwise distance between AIs' portfolios, where the average is computed over all time periods and simulations. We remark that equilibrium portfolio policies in the rational benchmark are symmetric, so this measure is zero in the benchmark.

Our measures of market efficiency $\mathcal{ME}$ and liquidity $\mathcal{L}$ correspond to our theoretical definitions in Section 2.3. Market efficiency measures the fraction of return variation that is not predictable by public information. Omitting the $(J, \omega)$ notation for simplicity, $\mathcal{ME}$ is computed as the average across simulations of the ratio

$$\frac{Var(e_{n,t+1})}{Var\left[g_{n_s}(\mathcal{I}_{n,t})\right] + Var(e_{n,t+1})},$$

where $g_{n_s}(\mathcal{I}_{n,t}) = E_t(R^e_{n,t+1})$ is the conditional expectation of future excess returns that accounts for the endogenous treading response by AI traders in simulation $s$ (see Eq. (A19)) and $e_{n,t+1}$ is the unpredictable component of returns (see Eq. (A20)).

Similarly, liquidity $\mathcal{L}$ is computed as the average value of the price impact of the supply shock $|\frac{\partial p_{n,t}}{\partial \sigma}|_{\sigma=0}$ computed numerically across all time periods and simulations (see Eqs. (A27)-(A30) for details on the numerical approximation).

# 4 Experimental results

This section presents the outcomes of our numerical simulations, highlighting how AI traders' portfolio choices align or diverge from the rational benchmark, and how their actions influence market efficiency and liquidity. First, we analyze AI traders' portfolio policies and returns, examining how these are impacted by trader size and competition intensity. Second, we investigate how these trading behaviors translate into overall market efficiency and liquidity.

## 4.1 Portfolio choice and returns

### 4.1.1 Portfolio policies

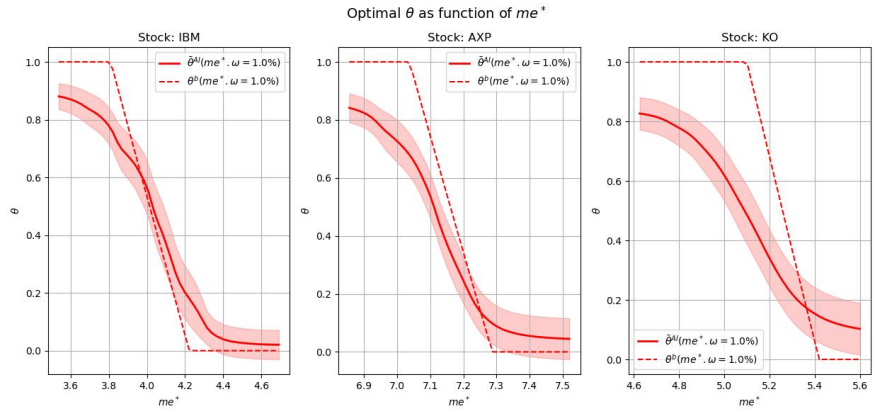
**Comparative statics in $me$.** Figures 2–4 show that the AI traders' policy functions are strictly decreasing in $me^*$. This aligns with Proposition 1 from the theoretical benchmark, which states that the optimal portfolio weight should decline with adjusted market equity. Intuitively, a higher $me^*$ (for given stock characteristics) reflects a larger inferred latent demand, which, due to mean reversion, lowers expected capital gains. Both the AI solution and the benchmark solution exhibit this downward slope, suggesting that our DDPG-based agents internalize the fundamental relationship between $me$, latent demand, and future returns in a manner consistent with theory.

**The effect of trader size (larger $\omega$) for fixed $J$.** As we increase $J\omega$ while keeping $J$ fixed—effectively raising the overall fraction of market capitalization managed by AI—the portfolio policy becomes less aggressive, reflecting a greater internalization of price impact. This result is qualitatively consistent with Proposition 1. However, compared to the rational benchmark, the AI does not reduce its holdings as sharply, suggesting only partial, rather than complete, internalization of its own price impact.

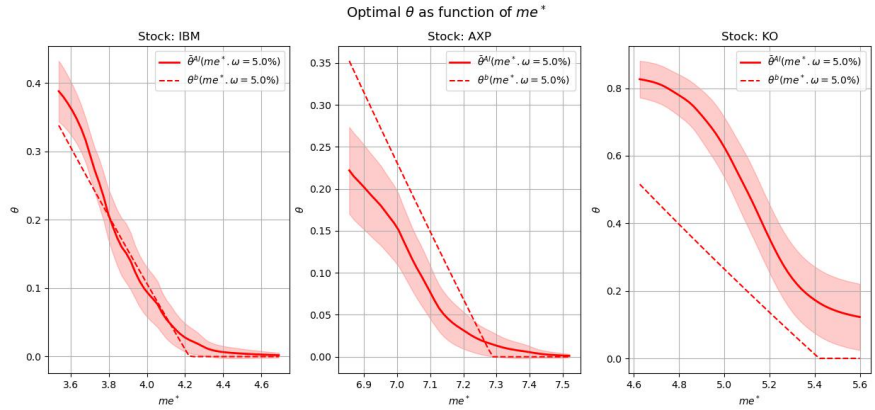
**The effect of competition (larger $J$ for fixed $J\omega$).** In the rational benchmark, an increase in $J$ for a fixed $J\omega$ leads individual traders to adopt more aggressive strategies. Figures 2–4 confirm that the AI solutions exhibit a similar qualitative pattern. However, with higher $J$, the red solid line typically stands above the benchmark more than in the $J = 1$ case, particularly at moderate $me^*$. This suggests that AI traders only partially internalize the collective price impact of their competitors. Moreover, the AI's deviation from the rational benchmark becomes more pronounced as $J$ increases, implying that when competition intensifies, the ability of AI traders to learn deteriorates.

Figure 2: Portfolio policy for $J = 1$: comparative statics in $me^*$

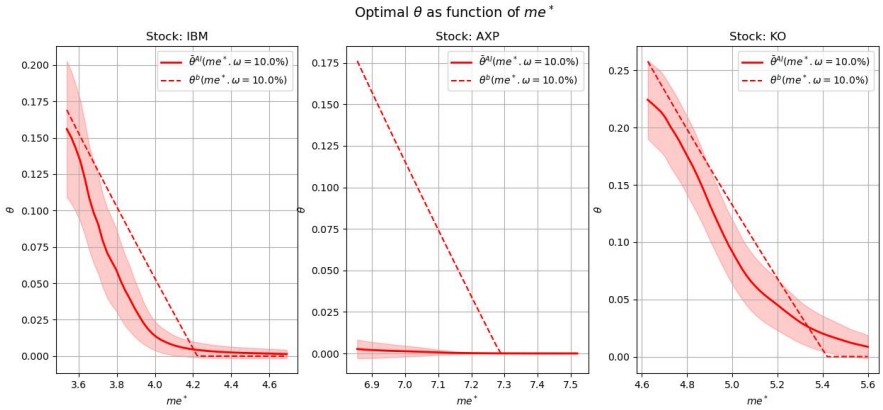
(a) $J = 1,\ \omega = 1\%$




(b) $J = 1,\ \omega = 5\%$



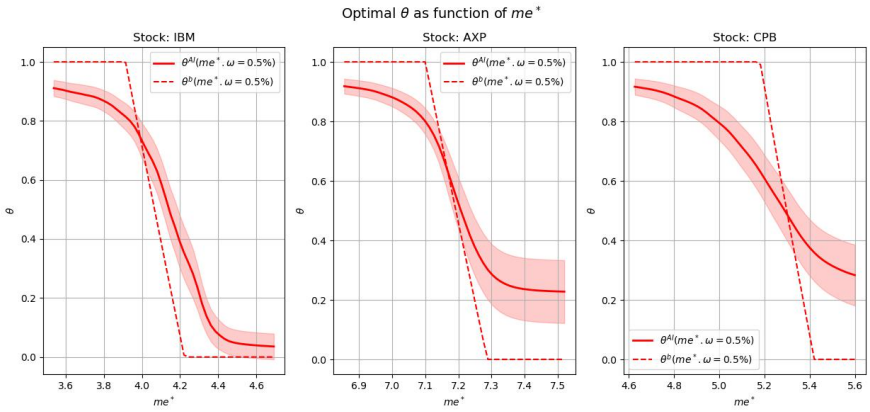

(c) $J = 1,\ \omega = 10\%$



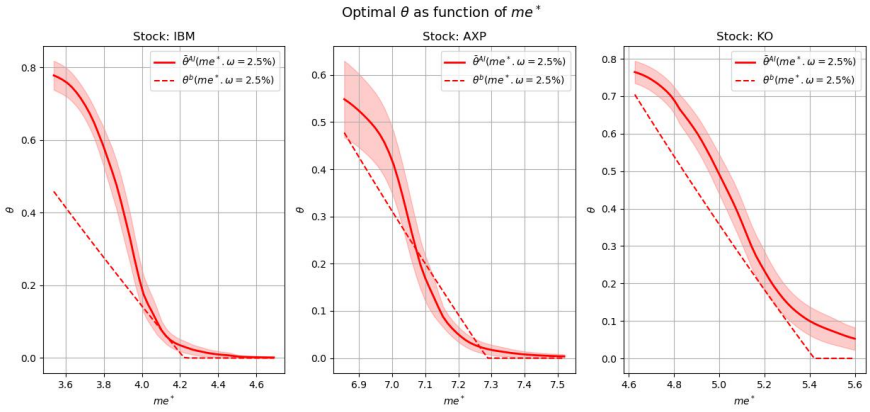

*Notes:* Each panel shows the average portfolio weight on the risky asset of the AI trader across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the adjusted market equity ($me$*). The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the trader's assets-under-management share $\omega$ of the market size (1%, 5%, and 10%).

Figure 3: Portfolio policy for $J = 2$: comparative statics in $me^*$

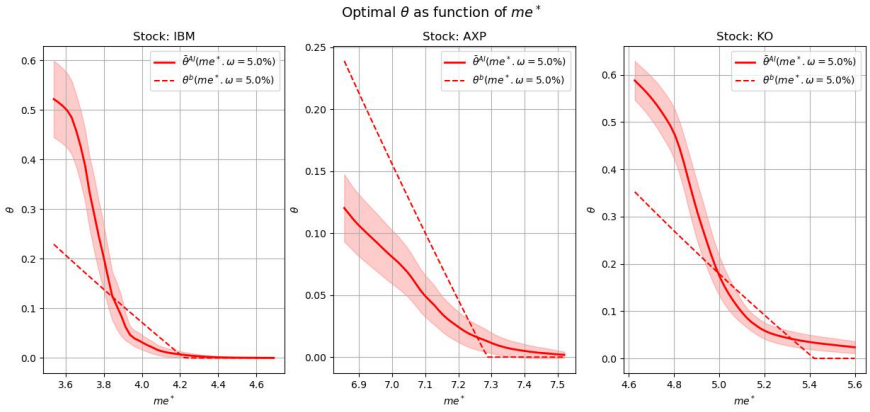
(a) $J = 2,\ J\omega = 1\%$




(b) $J = 2,\ J\omega = 5\%$



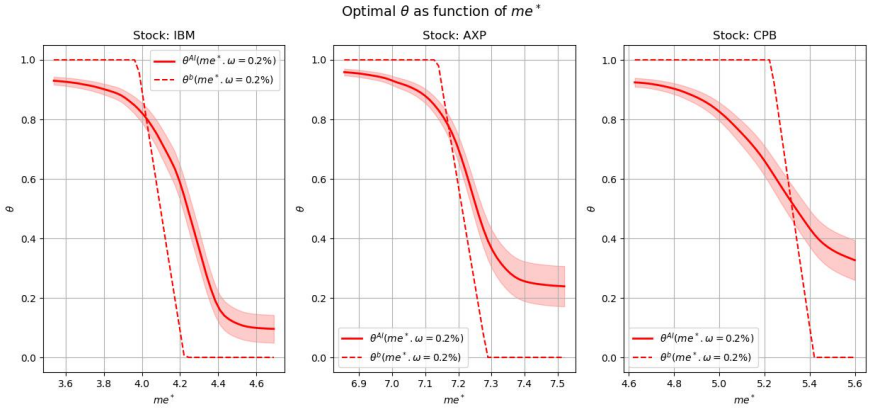

(c) $J = 2,\ J\omega = 10\%$



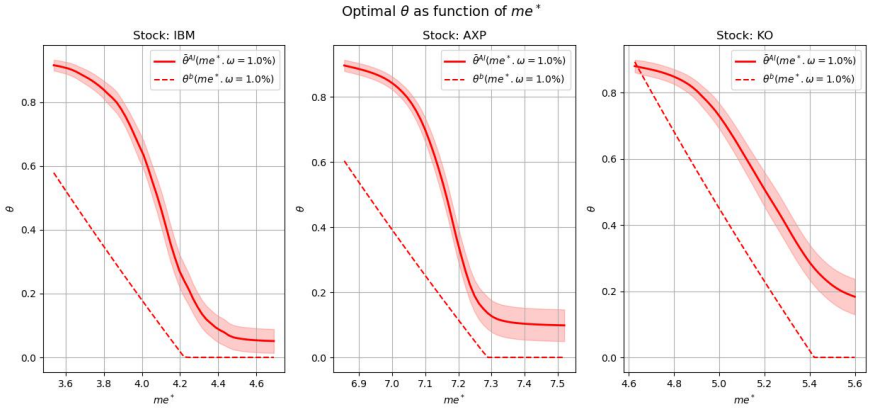

*Notes:* Each panel shows the average portfolio weight on the risky asset of the two AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the adjusted market equity ($me$*). The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure 4: Portfolio policy for $J = 5$: comparative statics in $me^*$

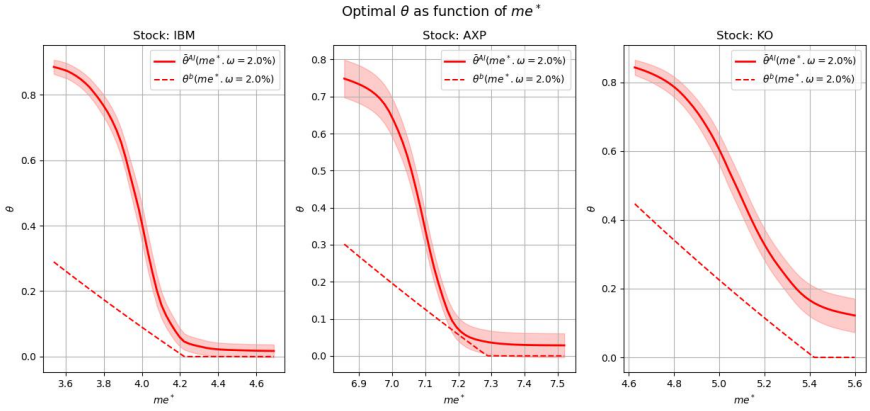
(a) $J = 5,\ J\omega = 1\%$




(b) $J = 5,\ J\omega = 5\%$




(c) $J = 5,\ J\omega = 10\%$




*Notes:* Each panel shows the average portfolio weight on the risky asset of the five AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the adjusted market equity ($me$*). The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

**Comparative statics in stock characteristics.** Figures A2-A16 In Appendix B illustrate the AI portfolio comparative statics in the publicly observable stock characteristics (log book equity, profitability, investment growth, dividend-to-book and market beta).

AI traders exhibit a high degree of qualitative alignment with the rational benchmark when the reacting to dividend-to-book, book-to-equity, and investment growth. In these cases, AI traders correctly recognize the directional effects predicted by theory: they allocate more to stocks with lower dividend-to-book ratios and investment growth, reflecting their role in predicting returns. The AI traders show more systematic deviations in their response to market beta and profitability. A possible reason for this is that the optimal portfolio choice is not very sensitive to these variables, reflecting a weaker predictive signal in the underlying data generating process. Finally, the spread with the rational benchmark tends to widen for larger $J$, consistent with the results on $me^*$.

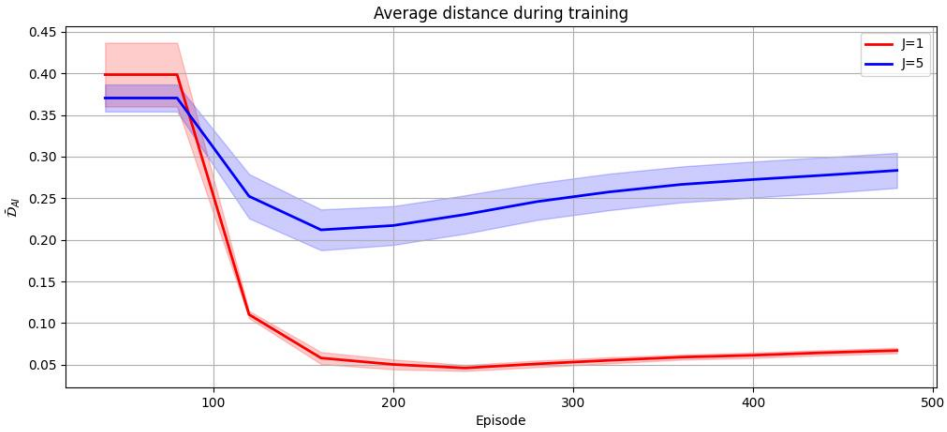
**Explaining the quantitative gap: a learning externality.** While the AI's qualitative behavior aligns well, qualitatively, with the theoretical predictions in Proposition 1, it deviates quantitatively: the AI tends to trade too aggressively when more traders are present and does not sufficiently scale down its holdings as size increases. This discrepancy suggests a limited internalization of price impact, an effect that becomes more pronounced as more AI agents simultaneously learn. When multiple agents are training in parallel, each experiences a noisier mapping from actions to rewards, making it harder to internalize their true market impact.

A key driver of these deviations is the model-free nature of DDPG. The algorithm lacks an explicit model linking large trades to subsequent price changes. Instead, it learns purely from reward signals, which can be distorted by the random exploration of other AI traders. When $J$ is large, these distortions intensify, as each AI struggles to disentangle its own impact from that of others. As a result, AI-driven strategies deviate further from the rational benchmark as $J$ increases.

These findings suggest a negative learning externality among AI traders. This externality not only weakens their ability to internalize price impact but also hampers their capacity to extract predictive information from signals. The root of this issue is that the AI does not maintain separate models for return predictability and price impact. Instead, it learns the entire mapping from states and actions to rewards in a single process, making it more susceptible to interference from competing agents.

To illustrate this learning externality, Figure 5 compares the average distance between the AI's portfolio policy and the benchmark during the learning process for $J = 1$ and $J = 5$. The figure shows slower convergence and a larger average deviation from the benchmark in the $J = 5$ case compared to the $J = 1$ case, supporting the interpretation that interactions among multiple AIs introduce noise into each agent's learning process.

Figure 5: Negative learning externality




*Notes:* This figure plots the average distance between AI portfolios and benchmark, $\mathcal{D}_{\{AI,b\}}$, computed every 40 episodes during the training phase. The stock is IBM and $J\omega = 5\%$

We now examine how these policy choices translate into the portfolio distance metrics, returns, market efficiency, and liquidity.

### 4.1.2 Portfolio Distances and Returns

**Distance from the Benchmark.** Table 4 shows that when $J\omega$ is fixed, the distance $\mathcal{D}_{AI,b}$ tends to increase with $J$. This pattern reflects our earlier observation that as $J$ rises, the AI trades more aggressively than the benchmark—particularly around moderate values of $me^*$. Consequently, the absolute deviation from the rational policy becomes larger.

In contrast, when there is only a single AI trader ($J = 1$), $\mathcal{D}_{AI,b}$ is often much smaller. This indicates that the AI can more closely align with the benchmark in the absence of additional external noise introduced by multiple learning agents.

Table 4: Average distance between AIs and benchmark, $\mathcal{D}_{\{AI,b\}}$

| | $J\omega = 1\%$ | | | $J\omega = 5\%$ | | | $J\omega = 10\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | IBM | AMX | KO | IBM | AMX | KO | IBM | AMX | KO |
| J=1 | 0.149 | 0.224 | 0.199 | 0.068 | 0.118 | 0.239 | 0.046 | 0.088 | 0.07 |
| | (0.07) | (0.064) | (0.057) | (0.009) | (0.033) | (0.124) | (0.011) | (0.008) | (0.016) |
| J=2 | 0.155 | 0.256 | 0.219 | 0.136 | 0.14 | 0.12 | 0.087 | 0.086 | 0.103 |
| | (0.062) | (0.086) | (0.065) | (0.025) | (0.022) | (0.039) | (0.031) | (0.018) | (0.017) |
| J=5 | 0.186 | 0.248 | 0.219 | 0.273 | 0.231 | 0.196 | 0.264 | 0.186 | 0.235 |
| | (0.060) | (0.030) | (0.038) | (0.058) | (0.064) | (0.062) | (0.042) | (0.067) | (0.07) |

*Notes:* Std. Dev across simulations in parenthesis

Table 5: Average distance among AIs, $\mathcal{D}_{\{AI\}}$ :

| | $J\omega = 1\%$ | | | $J\omega = 5\%$ | | | $J\omega = 10\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | IBM | AMX | KO | IBM | AMX | KO | IBM | AMX | KO |
| J=2 | 0.205 | 0.257 | 0.312 | 0.192 | 0.174 | 0.204 | 0.152 | 0.107 | 0.196 |
| | (0.075) | (0.193) | (0.159) | (0.040) | (0.060) | (0.076) | (0.035) | (0.058) | (0.04) |
| J=5 | 0.218 | 0.268 | 0.304 | 0.212 | 0.236 | 0.236 | 0.209 | 0.188 | 0.226 |
| | (0.102) | (0.166) | (0.158) | (0.098) | (0.183) | (0.132) | (0.094) | (0.101) | (0.111) |

*Notes:* Std. Dev across simulations in parenthesis

**Average Returns.** Table 6 shows that in the $J = 1$ case, the AI's average returns closely match the benchmark's. This suggests that when a single agent learns with a moderate $\omega$, it achieves near-rational performance.

However, as $J$ increases, average returns decline more sharply than in the benchmark. In some cases, particularly for $J = 5$ and large $\omega$, the AI's mean returns turn negative. This indicates that overly aggressive strategies, combined with mutual interference in the learning process, can significantly degrade performance.

Table 6: Average returns for AIs vs. benchmark

| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
|---|---|---|---|---|---|---|
| Panel A: IBM | AI | RB | AI | RB | AI | RB |
| J=1 | 0.02128 (0.00609) | 0.0238 | 0.01300 (0.00051) | 0.0146 | 0.0102 (0.00101) | 0.01201 |
| J=2 | 0.02036 (0.00434) | 0.02285 | 0.00464 (0.00268) | 0.01362 | 0.0035 (0.00494) | 0.01177 |
| J=5 | 0.01816 (0.00428) | 0.02218 | -0.01277 (0.00666) | 0.01158 | -0.03019 (0.00702) | 0.01085 |
| Panel B: AXP | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.02762 (0.0088) | 0.03304 | 0.01443 (0.0031) | 0.01857 | 0.00886 (0.00081) | 0.0143 |
| J=2 | 0.02144 (0.01207) | 0.03193 | 0.01268 (0.00269) | 0.01652 | 0.01183 (0.00176) | 0.01342 |
| J=5 | 0.02185 (0.00423) | 0.03116 | -0.00697 (0.00878) | 0.01345 | -0.01426 (0.01348) | 0.01123 |
| Panel C: KO | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.04195 (0.01168) | 0.04623 | 0.00076 (0.0291) | 0.02782 | 0.01874 (0.00163) | 0.02138 |
| J=2 | 0.03546 (0.01203) | 0.04506 | 0.01848 (0.00585) | 0.02254 | 0.00712 (0.00336) | 0.01782 |
| J=5 | 0.03634 (0.0075) | 0.04420 | 0.00128 (0.01067) | 0.01687 | -0.02627 (0.01505) | 0.01172 |

*Notes:* Std. Dev across simulations in parenthesis

## 4.2 Market outcomes

**Market Efficiency.** Table 7 shows that the presence of AI traders consistently increases market efficiency compared to the scenario with only representative investors and no AIs. Moreover, for a fixed $J\omega$, efficiency generally rises with $J$ (with one exception), aligning with the rational benchmark and the prediction of Proposition 2: as $J$ increases, more aggressive trading by speculators reduces return predictability in equilibrium.

Interestingly, in one case ($J = 2$), market efficiency with AIs slightly exceeds the rational benchmark. This can occur if the AI trades more aggressively on signals than the rational model predicts. In equilibrium, that aggressive behavior leads to an over-elimination of predictable

components, albeit at the cost of lower individual returns.

Table 7: Market efficiency $\mathcal{M}^{AI}(\omega, J)$

| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
|---|---|---|---|---|---|---|
| Panel A: IBM, $\mathcal{M}(\omega,0)=0.942$ | | | | | | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.955 (0.002) | 0.958 | 0.965 (0.003) | 0.969 | 0.958 (0.01) | 0.969 |
| J=2 | 0.956 (0.002) | 0.968 | 0.979 (0.002) | 0.976 | 0.972 (0.005) | 0.976 |
| J=5 | 0.955 (0.002) | 0.980 | 0.984 (0.003) | 0.980 | 0.973 (0.005) | 0.981 |
| Panel B: AXP, $\mathcal{M}(\omega,0)=0.889$ | | | | | | |
| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.905 (0.003) | 0.912 | 0.912 (0.013) | 0.940 | 0.890 (0.006) | 0.941 |
| J=2 | 0.903 (0.005) | 0.929 | 0.931 (0.013) | 0.951 | 0.914 (0.013) | 0.954 |
| J=5 | 0.903 (0.002) | 0.957 | 0.943 (0.006) | 0.957 | 0.946 (0.004) | 0.963 |
| Panel C: KO, $\mathcal{M}(\omega,0)=0.906$ | | | | | | |
| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.917 (0.002) | 0.920 | 0.946 (0.010) | 0.952 | 0.940 (0.006) | 0.955 |
| J=2 | 0.915 (0.003) | 0.933 | 0.950 (0.004) | 0.957 | 0.965 (0.002) | 0.966 |
| J=5 | 0.915 (0.002) | 0.951 | 0.947 (0.007) | 0.959 | 0.970 (0.007) | 0.973 |

*Notes:* Std. Dev across simulations in parenthesis

**Market Liquidity.** Table 8 shows that liquidity is an order of magnitude lower for the AI relative to the benchmark. The explanation for this result is that in the rational benchmark, traders understand the temporary nature of the shock and anticipate the price reversal in the following period, so they buy aggressively to profit from the anticipated reversal. By contrast, the AI interprets the initial price drop as a sign of reduced latent demand, which is only partly reversed by mean reversion. Thus, the AI perceives the price drop as a weaker signal of price reversal and

it does not fully exploit such dips, resulting in lower liquidity provision than the rational scenario.

In other words, the AI has only partial knowledge of the structural supply-shock induced price reversion. It sees "price down" → "latent demand down," which is not a guaranteed short-run arbitrage opportunity. Consequently, it sits out more than it should, leading to lower liquidity provision.

Table 8: Market liquidity $\mathcal{L}^{AI}(\omega, J)$

| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
|---|---|---|---|---|---|---|
| Panel A: IBM | | | | | | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.006 (0.001) | 0.052 | 0.013 (0.002) | 0.096 | 0.009 (0.006) | 0.096 |
| J=2 | 0.006 (0.001) | 0.054 | 0.02 (0.004) | 0.128 | 0.006 (0.005) | 0.129 |
| J=5 | 0.006 (0.001) | 0.054 | 0.011 (0.008) | 0.154 | -0.002 (0.005) | 0.16 |
| Panel B: AXP | | | | | | |
| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.010 (0.002) | 0.037 | 0.015 (0.008) | 0.091 | 0.001 (0.004) | 0.092 |
| J=2 | 0.009 (0.004) | 0.038 | 0.028 (0.01) | 0.116 | 0.014 (0.007) | 0.123 |
| J=5 | 0.009 (0.001) | 0.039 | 0.015 (0.015 ) | 0.133 | 0.003 (0.015) | 0.153 |
| Panel C: KO | | | | | | |
| | $J\omega$=1% | | $J\omega$=5% | | $J\omega$=10% | |
| | AI | RB | AI | RB | AI | RB |
| J=1 | 0.004 (0.001) | 0.026 | 0.019 (0.005) | 0.095 | 0.015 (0.004) | 0.105 |
| J=2 | 0.003 (0.001) | 0.026 | 0.013 (0.003) | 0.110 | 0.001 (0.003) | 0.139 |
| J=5 | 0.006 (0.001) | 0.026 | 0.003 (0.008) | 0.116 | -0.006 (0.005) | 0.165 |

*Notes:* Std. Dev across simulations in parenthesis

# 5 Conclusion

This paper studies how AI-driven investors—modeled with Deep Reinforcement Learning (DRL)—trade in an empirically-calibrated demand-based asset pricing model with price impact and return predictability. Our work helps fill two gaps in the literature. First, we are the first to study the behavior of AI-based traders in an equilibrium environment with empirically plausible return predictability and price impact. Second, while recent studies often employ tabular Q-learning schemes with discrete action sets, we adopt a continuous-control DRL method. This allows the trading agent to learn flexibly and avoid coarse discretization of portfolio weights.

Our experimental results deliver several key insights. In line with a rational expectations benchmark, the AI agents learn to exploit return predictability arising from mean-reverting fundamentals and latent demand. In particular, the AI policies qualitatively reproduce the comparative statics derived in a rational-benchmark model, indicating that that they learn to decode latent demand shocks from prices and internalize the price impact of their trades.

Yet quantitative differences emerge. With multiple AI traders, their actions introduce additional noise that distorts each agent's learning process. Each AI finds it more difficult to isolate its own price impact from that of its competitors, which leads to only partial internalization of price impact. We view this as a "negative learning externality" that degrades performance relative to the rational benchmark—particularly when trader size and competition intensify. Relatedly, while AI-driven trading consistently raises market efficiency by reducing predictable return components, albeit to a lower extent compared to the rational benchmark, the AI agents do not provide the same liquidity as rational speculators would. Their model-free algorithms do not fully anticipate short-lived arbitrage opportunities from transient supply shocks, leading to more pronounced price movements.

From a practical standpoint, these results underscore how DRL-based traders can significantly transform market outcomes. They do reduce return predictability, thus enhancing a form of semi-strong efficiency. At the same time, they provide less liquidity than predicted by rational models, reflecting incomplete internalization of mean reversion that is not directly observable. This tension highlights the importance of modeling the interplay between traders' learning processes and market dynamics, rather than treating AI strategies as atomistic or frictionless.

# References


Abada, I. and X. Lambin (2023): "Artificial Intelligence: Can Seemingly Collusive Outcomes Be Avoided?" *Management Science*, 69, 5042–5065.

Bai, J., T. Philippon, and A. Savov (2016): "Have Financial Markets Become More Informative?" *Journal of Financial Economics*, 122, 625–654.

Barberis, N. and L. Jin (2023): "Model-free and Model-based Learning as Joint Drivers of Investor Behavior," Available at https://www.law.nyu.edu/sites/default/files/DGJ_AI_Trading_Collusion_Market%20Manipulation_Penn_NYU_Law.pdf.

Biais, B., T. Foucault, and S. Moinas (2015): "Equilibrium Fast Trading," *Journal of Financial Economics*, 116, 292–313.

Brogaard, J., T. Hendershott, and R. Riordan (2015): "Quantifying the High-Frequency Trading "Arms Race"," *Quarterly Journal of Economics*, 130, 1547–1621.

Calvano, E., G. Calzolari, V. Denicolò, and S. Pastorello (2020): "Artificial Intelligence, Algorithmic Pricing, and Collusion," *American Economic Review*, 110, 3267–3297.

Colliard, J.-E., T. Focault, and S. Lovo (2023): "Algorithmic Pricing and Liquidity in Securities Markets," Working Paper.

Dou, W. W., I. Goldstein, and Y. Ji (2023): "AI-Powered Trading, Algorithmic Collusion, and Price Efficiency," Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4452704.

Dugast, J. and T. Foucault (2018): "Data abundance and asset price informativeness," *Journal of Financial Economics*, 130, 367–391.

Farbood, M., A. Matray, L. Veldkamp, and V. Venkateswaran (forthcoming): "Where Has All the Data Gone?" *Review of Financial Studies*.

Foucault, T., J. Hombert, and I. Roşu (2016): "News Trading and Speed," *Journal of Finance*, 71, 335–382.

Fuchs, W., S. Fukuda, and D. Neuhann (2024): "Demand-System Asset Pricing: Theoretical Foundations," Working paper.

Johnson, J. P., A. Rhodes, and M. Wildenbeest (2023): "Platform Design When Sellers Use Pricing Algorithms," *Econometrica*, 91, 1841–1879.

Koijen, R. S. J. and M. Yogo (2019): "A Demand System Approach to Asset Pricing," Forthcoming at *Journal of Political Economy*.

Lillicrap, T. P., J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra (2015): "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*.

van der Beck, P. (2022): "Short- versus Long-run Demand Elasticities in Asset Pricing," Swiss Finance Institute Research Paper 22-67, Swiss Finance Institute.

Vyetrenko, S. and S. Xu (2019): "Risk-Sensitive Compact Decision Trees for Autonomous Execution in Presence of Simulated Market Response," *ArXiv*, abs/1906.02312.

Watkins, C. J. C. H. (1989): "Learning from Delayed Rewards," Ph.D. thesis, King's College, Cambridge.

Yang, H. (2024): "AI Coordination and Self-Fulfilling Financial Crises," Available at https://drive.google.com/file/d/1SCb68JxyXN3Zq35dXrU3ZHs8Ruoxrce-/view.

Yang, H., X.-Y. Liu, S. Zhong, and A. Walid (2020): "Deep Reinforcement Learning for Automated Stock Trading: An Ensemble Strategy," Tech. rep.

# A Appendix

## A.1 Model solution

**Derivation of Eq. (7)**

Using the portfolio formula for the representative investor in Eq. (1) together with the representative investor's budget constraint $\sum_{m=1}^{N} w_{m,t} + w_{0,t} + \gamma_t = 1$, we obtain

$$w_{n,t} = \frac{\delta_{n,t}}{1 + \sum_{m=1}^{N} \delta_{m,t}}.$$

The market clearing condition Eq. (6) for each stock can be equivalently written as

$$A_t w_{n,t} = P_{n,t} \tilde{S}_{n,t}. \tag{A1}$$

Using Eq. (1) into the market clearing condition Eq. (A1) we obtain

$$\delta_{n,t} A_t = \left(1 + \sum_{m=1}^{N} \delta_{m,t}\right) P_{n,t} \tilde{S}_{n,t}.$$

Summing the last equation over $n$ and rearranging, we obtain

$$\delta_{n,t} = \frac{P_{n,t} \tilde{S}_{n,t}}{A_t - \sum_{m=1}^{N} P_{m,t} \tilde{S}_{m,t}}.$$

Using Eq. (1) to substitute for $\delta_{n,t}$ in the last equation and taking logs, we obtain

$$\beta_{0,t}(p_{n,t} + s_n) + \sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t} = p_{n,t} + \tilde{s}_{n,t} - \log\left(A_t - \sum_{m=1}^{N} P_{m,t} \tilde{S}_{m,t}\right). \tag{A2}$$

AI investors' the budget constraint is

$$A_t^j = \sum_{m=1}^{N} S_{m,t}^j P_{m,t} + S_{0,t}^j. \tag{A3}$$

Combining Eq. (A3) with the AI investors' asset dynamics in Eq. (5) we obtain

$$\sum_{m=1}^{N} S_{m,t}^j P_{m,t} + S_{0,t}^j = \sum_{m=1}^{N} S_{m,t-1}^j \left(P_{m,t} + D_{m,t}\right) + S_{0,t-1}^j R_f. \tag{A4}$$

Using Eqs. (4) and (A4) into Eq. (A2) and rearranging, we obtain Eq. (7) in the text. □

**Derivation of Eq. (8)**

Under Assumption 1, the argument of the logarithm in the r.h.s. of Eq. (7) reduces to

$D_{M,t} + S_{0,t-1}$. With Assumption 2, the representative investor's budget constraint reads

$$A_t = \sum_{m=1}^{N} \tilde{S}_{m,t} P_{m,t} + S_{0,t} + \lambda(D_{M,t} + S_{0,t-1} R_f),$$

which, together with Eq. (4) and Assumption 1, imply

$$S_{0,t} = (1-\lambda)(D_{M,t} + S_{0,t-1} R_f).$$

Using this expression recursively, we obtain

$$D_{M,t} + S_{0,t-1} R_f = \sum_{\tau=0}^{t} D_{M,t-\tau} (1-\lambda)^\tau R_f^\tau. \tag{A5}$$

Assumptions 3 and 4 imply

$$\sum_{\tau=0}^{t} D_{M,t-\tau} (1-\lambda)^\tau R_f^\tau = D_{M,t} \frac{(1+g)\left[1 - \left(\frac{R_f(1-\lambda)}{1+g}\right)^{t-1}\right]}{1+g-R_f(1-\lambda)} \approx D_{M,t} \frac{(1+g)}{1+g-R_f(1-\lambda)},$$

where the approximation is exact for that as $t \to \infty$. Together, Assumptions 1-4 imply that the price equation Eq. (7) simplifies to

$$p_{n,t} = \frac{\beta_0 s_n - \tilde{s}_{n,t} + \sum_{k=1}^{K-1} \beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t} + \log(D_{M,t}) + \phi}{1-\beta_0}, \tag{A6}$$

where

$$\phi = \log\left(\frac{(1+g)}{1+g-R_f(1-\lambda)}\right).$$

Using the definition $\alpha_{n,t}^{AI} = \frac{S_{n,t}^{AI}}{S_n}$ into Eq. (A6) and simplifying we obtain Eq. (8) in the text. □

## A.2 Proofs of propositions

### Proof of Proposition 1

**Part-(i): equilibrium existence and uniqueness** We begin by showing that speculators' objective functions are concave in portfolio shares. Using Eqs. (16)-(17) we obtain

$$\begin{aligned}\frac{\partial^2 E_t(R_{p,t}^j)}{(\partial \theta_{n,t}^j)^2} = E_t\Bigg\{ &- \frac{P_{n,t+1}}{P_{n,t}} \Bigg[ \left(\frac{1}{P_{n,t}} \frac{\partial P_{n,t}}{\partial \theta_{n,t}^j}\right)\left(1 - \frac{\theta_{n,t}^j}{P_{n,t}} \frac{\partial P_{n,t}}{\partial \theta_{n,t}^j}\right) \\ &+ \frac{A_t^s \left[(1-\beta_0) S_n P_{n,t} \left(1 - \frac{\theta_{n,t}^j}{P_{n,t}} \frac{\partial P_{n,t}}{\partial \theta_{n,t}^j}\right) + \beta_0 \sum_{i \neq j}^{J} \theta_{n,t}^i A_t^s\right]}{\left[(1-\beta_0) S_n P_{n,t} + \beta_0 \sum_{i=1}^{J} \theta_{n,t}^i A_t^s\right]^2} \Bigg]\Bigg\},\end{aligned} \tag{A7}$$

which is indeed negative.

Next, combine Eqs. (16)-(17) as follows:

$$\frac{\partial E_t(R^j_{p,t})}{\partial \theta^j_{n,t}} = E_t\left[\frac{P_{n,t+1}}{P_{n,t}}\left(1 - \frac{\theta^j_{n,t}A^s_t}{(1-\beta_0)S_nP_{n,t} + \beta_0\sum_{i=1}^J \theta^i_{n,t}A^s_t}\right)\right] + E_t(DY_{n,t+1}) - R_f := \delta^j \quad \text{(A8)}$$

Eq. (A8) implies that if an equilibrium exists, it must be unique. To see this, assume $\theta^j_{n,t} > \theta^i_{n,t}$. Eq. (A8) implies $\delta^j < \delta^i$. In case both $\theta^j_{n,t}, \theta^i_{n,t} \in (0,1)$, optimality requires $\delta^j = \delta^i = 0$, a contradiction. In case $\theta^j_{n,t} = 1, \theta^i_{n,t} \in [0,1)$, optimality requires $\delta^j \geqslant 0, \delta^i \leqslant 0$, again a contradiction.

Using the definitions $\theta^j_{n,t} = \frac{S^j_{n,t}P_{n,t}}{A^s_t}$, $\alpha^j_{n,t} = \frac{S^j_{n,t}}{S_n}$ and $\alpha^S_{n,t} = \sum_{i=1}^J \alpha^i_{n,t}$, we can rewrite

$$\frac{\theta^j_{n,t}}{P_{n,t}}\frac{\partial P_{n,t}}{\partial \theta^j_{n,t}} = \frac{\theta^j_{n,t}A^j_t}{(1-\beta_0)S_nP_{n,t} + \beta_0\sum_{i=1}^J \theta^i_{n,t}A^i_t} = \frac{\alpha^j_{n,t}}{(1-\beta_0) + \beta_0\alpha^S_{n,t}}. \quad \text{(A9)}$$

We proceed to show that a symmetric equilibrium exists and is unique. We will first solve for an equilibrium in terms of $\alpha^S_{n,t}$. Given $\alpha^S_{n,t}$, and using the definitions $\theta_{n,t} = \frac{P_{n,t}S^j_{n,t}}{A^s}$ and $A^s_t = \omega P^*_{n,t}S_n$, and the fact that $P^*_{n,t}/P_{n,t} = \left(1-\alpha^S_{n,t}\right)^{\frac{1}{1-\beta_0}}$, we can uniquely determine the equilibrium portfolio share of each speculator as

$$\theta^j_{n,t} = \frac{\alpha^S_{n,t}}{J\omega\left(1-\alpha^S_{n,t}\right)^{\frac{1}{1-\beta_0}}}. \quad \text{(A10)}$$

Since each $\theta_{n,t} \in [0,1]$, the equilibrium must satisfy $\alpha^S_{n,t} \in [0,\bar{\alpha}]$ where $\bar{\alpha}$ solves

$$1 = \frac{\bar{\alpha}}{J\omega\left(1-\bar{\alpha}\right)^{\frac{1}{1-\beta_0}}}. \quad \text{(A11)}$$

Using Eqs. (2)-(3) and the price functions in Eq. (14)-(15) we can compute the expected capital gain as

$$E_t\left(\frac{P_{n,t+1}}{P_{n,t}}\right) = \Pi\left(\mathcal{I}_t\right)\left(1-\alpha^S_{n,t}\right)^{\frac{1}{(1-\beta_0)}}\Phi, \quad \text{(A12)}$$

where

$$\Pi\left(\mathcal{I}_t\right) = \exp\left(-\frac{\sum_{k=1}^{K-1}\beta_k(1-\rho_{k,n})x_{k,n,t} + (1-\rho_{\epsilon_n})E_t\left(\epsilon_{n,t}\right)}{1-\beta_0}\right), \quad \text{(A13)}$$

and

$$E_t\left(\epsilon_{n,t}\right) = (1-\beta_0)me^*_{n,t} - \left(\sum_{k=1}^{K-1}\beta_k x_{k,n,t} + \beta_K + \phi\right), \quad \text{(A14)}$$

and

$$\Phi = \exp\left(\frac{\sum_{k=1}^{K-1}\beta_k c_{n,k} + c_{n,\epsilon} + \log(1+g)}{1-\beta_0} + \frac{\sum_{k=1}^{K-1}\beta_k^2\sigma^2_{\eta_{k,n}} + \sigma^2_{\xi n}}{2(1-\beta_0)^2)}\right). \quad \text{(A15)}$$

Define the first derivative of a speculator's objective function when all $J$ speculators choose the

same strategy resutling in $\alpha^S_{n,t} = x$ as

$$G\left(x;\mathcal{I}_t\right) = \Pi\left(\mathcal{I}_t\right)\left(1-x\right)^{\frac{1}{(1-\beta_0)}}\Phi\left(1-\frac{x/J}{(1-\beta_0)+\beta_0 x}\right) + E(DY_{n,t+1}) - R_f. \tag{A16}$$

With this formulation, a symmetric equilibrium of the trading game is a fraction $\alpha^S_{n,t} \in [0,\bar{\alpha}]$ such that either (i) $\alpha^S_{n,t} = 0$ and $G\left(0;\mathcal{I}_t\right) \leqslant 0$, or (ii) $\alpha^S_{n,t} \in (0,1)$ and $G\left(\alpha^S_{n,t};\mathcal{I}_t\right) = 0$, or (iii) $\alpha^S_{n,t} = \bar{\alpha}$ and $G\left(\bar{\alpha};\mathcal{I}_t\right) \geqslant 0$. Since $G\left(x;\mathcal{I}_t\right)$ is strictly decreasing in $x$, an equilibrium exists and is unique for all $\mathcal{I}_t$. □

**Part-(ii): comparative statics** In an interior equilibrium $\alpha^S_{n,t}$ solves

$$G\left(\alpha^S_{n,t};\mathcal{I}_t\right) = 0, \tag{A17}$$

where $G$ is as in Eq. (A16). The corresponding equilibrium portfolio share is determined by Eq. (A9), which is strictly increasing in $\alpha^S_{n,t}$.

Eq. (A16) implies that $G\left(\alpha^S_{n,t};\mathcal{I}_t\right)$ is strictly decreasing in $\alpha^S_{n,t}$ and strictly increasing in $\Pi(\mathcal{I}_t)$. Eqs. (A13)-(A26) imply that $\Pi(\mathcal{I}_t)$ in increasing in $me^*_{n,t}$. Therefore, implicit differentiation of Eq. (A17) implies that $\alpha^S_{n,t}$ is decreasing in $me^*_{n,t}$, and so is the equilibrium portfolio share by Eq. (A9).

Eqs. (A13)-(A26) imply that $\Pi(\mathcal{I}_t)$ in increasing in $x_{k,n,t}$ if and only if $\beta_k(\rho_{k,n} - \rho_\epsilon) > 0$. Therefore, implicit differentiation of Eq. (A17) implies that $\alpha^S_{n,t}$ is increasing in $x_{k,n,t}$ if and only if $\beta_k(\rho_{k,n} - \rho_\epsilon) > 0$, and so is the equilibrium portfolio share by Eq. (A9).

Since $G\left(\alpha^S_{n,t};\mathcal{I}_t\right)$ is independent of $\omega$, then so is $\alpha^S_{n,t}$. Therefore, the equilibrium portfolio share is decreasing in $\omega$ by Eq. (A9). Finally, since $G\left(\alpha^S_{n,t};\mathcal{I}_t\right)$ is increasing in $J$, implicit differentiation of Eq. (A17) implies that $\alpha^S_{n,t}$ is also increasing in $J$. For fixed $\omega J$, Eq. (A9) implies that the equilibrium portfolio share is increasing in $J$. □

**Proof of Proposition 2**

Using Eqs. (2)-(3) and the price functions in Eq. (14)-(15), we can write the data generating process for the (log) capital gain as

$$p_{n,t+1} - p_{n,t} = g(\mathcal{I}_t) + e_{t+1}, \tag{A18}$$

where

$$g(\mathcal{I}_t) = \frac{\log\left(1-\alpha^S_n(\mathcal{I}_t)\right) + \sum_{k=1}^{K-1}\beta_k(c_{k,n} - (1-\rho_{k,n})x_{k,n,t}) + c_{\epsilon,n} - (1-\rho_{\epsilon,n})\epsilon_{n,t} + \log\left(1+g\right)}{1-\beta_0}, \tag{A19}$$

and

$$e_{t+1} = \frac{\sum_{k=1}^{K-1}\beta_k\eta_{k,n,t+1} + \xi_{n,t+1}}{1-\beta_0}. \tag{A20}$$

In case equilibrium, $\alpha^S(\mathcal{I}_t)$ is interior, it solves Eq. (A17), which using Eq. (A16), we can

use to rewrite Eq. (A19) as

$$g(\mathcal{I}_t) = \log\left(R_f - E(DY_{n,t+1})\right) - \frac{\sum_{k=1}^{K-1}\beta_k^2\sigma_{\eta_{k,n}}^2 + \sigma_{\xi n}^2}{2(1-\beta_0)^2} - \log\left(1 - \frac{\alpha_n^S(\mathcal{I}_t)/J}{(1-\beta_0)+\beta_0\alpha_n^S(\mathcal{I}_t)}\right). \quad \text{(A21)}$$

Plugging Eq. (A21) into Eq. (A18) and simplifying, we can write the expected excess return in the security as

$$E_t\left(\exp\left(p_{n,t+1} - p_{n,t}\right) + DY_{n,t+1} - R_f\right) = \left(1 - \frac{\alpha_n^S(\mathcal{I}_t)/J}{(1-\beta_0)+\beta_0\alpha_n^S(\mathcal{I}_t)}\right)^{-1} \quad \text{(A22)}$$

By Eq. (A17) and Eq. (A16), the ratio $\frac{\alpha_n^S(\mathcal{I}_t)/J}{(1-\beta_0)+\beta_0\alpha_n^S(\mathcal{I}_t)}$ is decreasing in $J$, and so is the expected excess return by Eq. (A22). On the other hand, $\omega$ has no impact on excess returns in this case.

In case equilibrium, $\alpha^S(\mathcal{I}_t)$ is not interior, we have two cases. If $\alpha^S(\mathcal{I}_t) = 0$, then expected excess returns are negative and the short-selling constraint binds, so changes in $\omega, J$ have no impact. If, instead, $\alpha^S(\mathcal{I}_t) = \bar{\alpha}$, then all speculators are fully invested in the security, expected excess returns are strictly positive, and the borrowing constraint binds. In this case an increase in $\omega$ increases $\bar{\alpha}$ by Eq. (A11), therey reducing expected returns by Eqs. (A18)-(A19). On the other hand, keeping $\omega J$ fixed, a marginal increase in $J$ has no impact.

Since $|E_t(R_{n,t+1}^e)|$ decreases in $\omega$ and in $J$ keeping $\omega J$ fixed, so does that $\mathcal{P} = E\left(|E_t(R_{n,t+1}^e)|\right)$. □

**Proof of Proposition 3** The supply shock implies that $p_{n,t}$ equals

$$p_{n,t} = -(s_n + \sigma) + \frac{-\log\left(1-\alpha_{n,t}^S\right) + \sum_{k=1}^{K-1}\beta_k x_{k,n,t} + \beta_K + \epsilon_{n,t} + \log\left(D_{M,t}\right) + \phi}{1-\beta_0}, \quad \text{(A23)}$$

And therefore the expected capital gain changes to

$$E_t\left(\frac{P_{n,t+1}}{P_{n,t}}\right) = \Pi\left(\mathcal{I}_t;\sigma\right)\left(1-\alpha_{n,t}^S\right)^{\frac{1}{(1-\beta_0)}}\Phi, \quad \text{(A24)}$$

where

$$\Pi\left(\mathcal{I}_t;\sigma\right) = \exp\left(\sigma - \frac{\sum_{k=1}^{K-1}\beta_k(1-\rho_{k,n})x_{k,n,t} + (1-\rho_{\epsilon_n})E_t\left(\epsilon_{n,t}\right)}{1-\beta_0}\right), \quad \text{(A25)}$$

and

$$E_t\left(\epsilon_{n,t}\right) = (1-\beta_0)(me_{n,t}^* + \sigma) - \left(\sum_{k=1}^{K-1}\beta_k x_{k,n,t} + \beta_K + \phi\right), \quad \text{(A26)}$$

and $\Phi$ is unchanged. The equilibrium condition is as in the proof of Proposition 1 with $\Pi\left(\mathcal{I}_t;\sigma\right)$ replacing $\Pi\left(\mathcal{I}_t\right)$. Denote the resulting equilibrium value $\alpha^S(\mathcal{I}_t;\sigma)$. Therefore, the definition of

liquidity implies

$$\mathcal{L} = 1 + E\left(\left.\frac{\partial p_{n,t}}{\partial \sigma}\right|_{\sigma=0}\right) = E\left(\left.\frac{\partial}{\partial \sigma}\big[-\log\big(1-\alpha^S(\mathcal{I}_t;\sigma)\big)\big]\right|_{\sigma=0}\right). \tag{A27}$$

For the case with AI trading, the supply shock changes the state variable $me^*_{n,t}$ to $me^*_{n,t} - \sigma$. Thus, defining the aggregate share of supply held by the AI traders as $\alpha^{AI}\big(me^*_{n,t}, \{\boldsymbol{x}_{n,t}\}\big) = \sum_{j=1}^{J}\alpha^j\big(\mathcal{I}_t\big)$, the definition of liquidity implies

$$\mathcal{L} = E\left(\left.\frac{\partial}{\partial \sigma}\left[-\log\left(1-\alpha^{AI}\big(me^*_{n,t}-\sigma,\{\boldsymbol{x}_{n,t}\}\big)\right)\right]\right|_{\sigma=0}\right). \tag{A28}$$

Numerically, we approximate the derivatives

$$\left.\frac{\partial}{\partial \sigma}\big[-\log\big(1-\alpha^S(\mathcal{I}_t;\sigma)\big)\big]\right|_{\sigma=0} \approx -\frac{1}{\sigma}\log\left(\frac{1-\alpha^S(\mathcal{I}_t;\sigma)}{1-\alpha^S(\mathcal{I}_t;0)}\right), \tag{A29}$$

and

$$\left.\frac{\partial}{\partial \sigma}\left[-\log\left(1-\alpha^{AI}\big(me^*_{n,t}-\sigma,\{\boldsymbol{x}_{n,t}\}\big)\right)\right]\right|_{\sigma=0} \approx -\frac{1}{\sigma}\log\left(\frac{1-\alpha^{AI}\big(me^*_{n,t}-\sigma,\{\boldsymbol{x}_{n,t}\}\big)}{1-\alpha^{AI}\big(me^*_{n,t},\{\boldsymbol{x}_{n,t}\}\big)}\right) \tag{A30}$$

for a numerical value of $\sigma = \log(1.01)$, corresponding to a 1% increase in the supply of the asset.

### A.3 Simulating stock characteristics and latent demand

The stock characteristics log book equity, profitability, investment, dividend-to-book equity, market beta and latent demand are simulated from an AR(1) process such that:

$$y_t = c + \rho y_{t-1} + \varepsilon \tag{A31}$$

where $y_t$ is alternatively one of the stock's characteristics, $c$ is the intercept, $\rho$ is the autoregressive coefficient and $\varepsilon$ is the error term. The AR(1) process is estimated by simulated method of moments, by imposing the following moment conditions on the data:

$$\begin{aligned}
\mathbb{E}_t[y_t] &= \frac{c}{1-\rho} \\
Var_t[y_t] &= \frac{\sigma_\varepsilon^2}{1-\rho^2} \\
\frac{\mathbb{E}_t[y_t y_{t-1}]}{Var_t[y_t]} &= \rho
\end{aligned}$$

By exploiting empirical moments, we recover the parameters of the AR(1) process. Stock characteristics are then simulated accordingly using the empirical mean as initial condition.

We construct the dividend-to-$P_{t-1}$ by sampling with replacement from the data. Finally,

we simulate returns by solving the system of Eq. **??**, given the simulated characteristics, dividends and the consumption rate of the representative investor.

Table A1: Autoregressive coefficients

| **Characteristic / Asset** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
|---|---|---|---|---|---|---|---|---|---|---|
| **Log book equity** | 0.97 | 0.952 | 0.821 | 0.844 | 0.955 | 0.982 | 0.851 | 0.918 | 0.914 | 0.975 |
| **Profitability** | 0.923 | 0.661 | 0.641 | 0.505 | 0.821 | 0.854 | 0.184 | 0.809 | 0.575 | 0.963 |
| **Investment** | 0.817 | 0.763 | 0.852 | 0.718 | 0.842 | 0.848 | 0.769 | 0.783 | 0.735 | 0.777 |
| **Dividend-to-book equity** | 0.965 | 0.336 | 0.62 | 0.774 | 0.969 | 0.763 | 0.71 | 0.948 | 0.421 | 0.88 |
| **Market beta** | 0.886 | 0.919 | 0.855 | 0.892 | 0.9 | 0.843 | 0.898 | 0.873 | 0.931 | 0.923 |

*Notes:* This table reports the estimate coefficient of stock characteristics for each stock. The autoregressive coefficient is estimated using simulated method of moments described in Section A.3

Table A2 reports the summary statistics, mean and standard deviation of the stock characteristics, dividend, latent demand and returns of the data and the simulation.

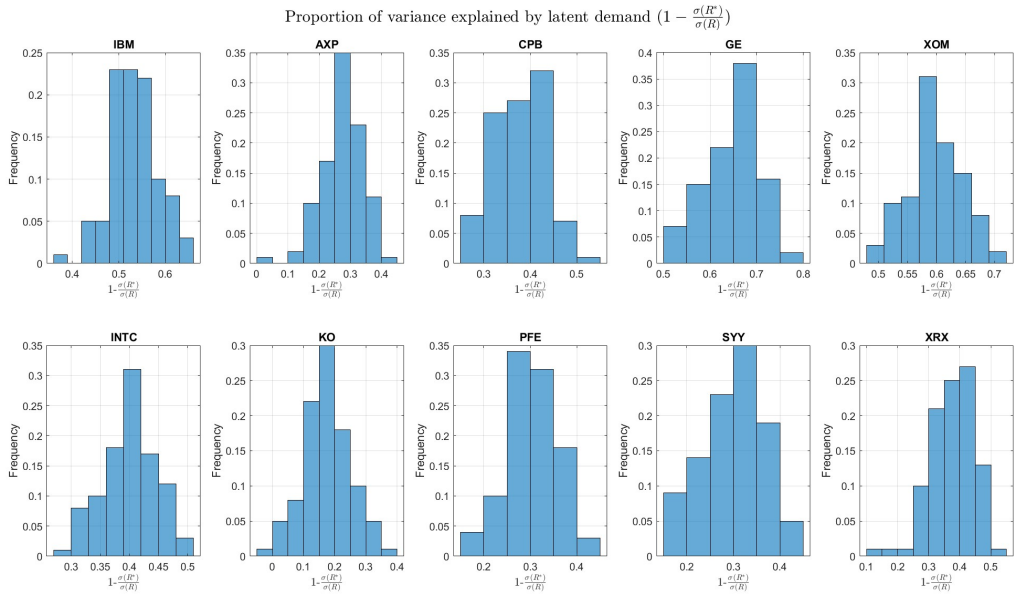
Table A2: Simulated characteristics and returns

| **log book equity** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
|---|---|---|---|---|---|---|---|---|---|---|
| Data (mean) | 6.42073 | 10.9827 | 6.81952 | 7.87984 | 5.75179 | 5.10696 | 8.97407 | 8.93513 | 7.99588 | 4.44588 |
| Simulation (mean) | 6.46414 | 10.9789 | 6.8238 | 7.84665 | 5.73352 | 5.00938 | 9.02358 | 8.96768 | 7.96014 | 4.34766 |
| Data (std. dev.) | 0.92798 | 0.93983 | 0.56049 | 0.60944 | 0.84505 | 1.68108 | 0.93541 | 0.57757 | 1.11229 | 0.89225 |
| Simulation (std. dev.) | 0.85289 | 0.87291 | 0.54566 | 0.61171 | 0.81112 | 1.46504 | 0.93727 | 0.5636 | 1.12084 | 0.80957 |
| t-stat | -0.64577 | 0.05572 | -0.09978 | 0.68964 | 0.28612 | 0.84433 | -0.67119 | -0.73375 | 0.40523 | 1.53923 |
| **Profitability** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | 0.21889 | 0.20844 | 0.14551 | 0.50808 | 0.17318 | 0.20784 | 0.117 | 0.4146 | 0.13856 | 0.47824 |
| Simulation (mean) | 0.2221 | 0.20919 | 0.14387 | 0.5058 | 0.17715 | 0.20316 | 0.11704 | 0.41487 | 0.14037 | 0.48269 |
| Data (std. dev.) | 0.09555 | 0.08512 | 0.07434 | 0.16604 | 0.13235 | 0.097 | 0.08928 | 0.16781 | 0.06285 | 0.17513 |
| Simulation (std. dev.) | 0.09238 | 0.08621 | 0.07323 | 0.16224 | 0.13061 | 0.0955 | 0.08979 | 0.15963 | 0.06186 | 0.16465 |
| t-stat | -0.44199 | -0.10976 | 0.28474 | 0.1782 | -0.38663 | 0.62346 | -0.00661 | -0.02192 | -0.37191 | -0.34292 |
| **Investment** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | 0.07682 | 0.05851 | 0.05192 | 0.06657 | 0.05114 | 0.12337 | 0.06906 | 0.07023 | 0.0775 | 0.08431 |
| Simulation (mean) | 0.06908 | 0.05874 | 0.04449 | 0.06727 | 0.05047 | 0.1284 | 0.06614 | 0.06723 | 0.08183 | 0.08766 |
| Data (std. dev.) | 0.11684 | 0.12934 | 0.11384 | 0.1294 | 0.19332 | 0.16178 | 0.21201 | 0.0956 | 0.15134 | 0.13762 |
| Simulation (std. dev.) | 0.1167 | 0.1295 | 0.11141 | 0.12641 | 0.18515 | 0.15656 | 0.2072 | 0.0929 | 0.15112 | 0.13835 |
| t-stat | 0.84242 | -0.02257 | 0.84672 | -0.06976 | 0.04613 | -0.40766 | 0.17938 | 0.41066 | -0.36418 | -0.30788 |
| **Dividend-to-book equity** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | 0.02268 | 0.04472 | 0.02944 | 0.10286 | 0.00783 | 0.00756 | 0.03598 | 0.10865 | 0.03853 | 0.09137 |
| Simulation (mean) | 0.02363 | 0.04456 | 0.02981 | 0.10335 | 0.01347 | 0.00966 | 0.04018 | 0.10598 | 0.03867 | 0.09151 |
| Data (std. dev.) | 0.00984 | 0.01128 | 0.01342 | 0.03632 | 0.01661 | 0.00899 | 0.03402 | 0.04004 | 0.0162 | 0.03017 |
| Simulation (std. dev.) | 0.00865 | 0.01121 | 0.01334 | 0.03516 | 0.01069 | 0.00736 | 0.02897 | 0.0383 | 0.01588 | 0.02951 |
| t-stat | -1.3954 | 0.18111 | -0.35434 | -0.17849 | -6.6358 | -3.608 | -1.8377 | 0.88526 | -0.1158 | -0.0613 |
| **Market beta** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | -0.00115 | 0.00373 | -0.00093 | 0.00217 | 0.00823 | -0.00203 | -0.00242 | 0.00078 | 0.00478 | -0.00235 |
| Simulation (mean) | -0.00143 | 0.00473 | -0.00173 | 0.00249 | 0.00833 | -0.00225 | -0.00327 | 0.00089 | 0.00438 | -0.00272 |
| Data (std. dev.) | 0.01495 | 0.01105 | 0.01522 | 0.0085 | 0.01989 | 0.018 | 0.01631 | 0.00943 | 0.00912 | 0.01507 |
| Simulation (std. dev.) | 0.01491 | 0.0107 | 0.01492 | 0.00829 | 0.02005 | 0.01812 | 0.0156 | 0.00905 | 0.009 | 0.01479 |
| t-stat | 0.23857 | -1.182 | 0.68041 | -0.4921 | -0.06178 | 0.14962 | 0.69637 | -0.1462 | 0.55339 | 0.31628 |
| **Dividend-to-P_{t-1}** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | 0.0009 | 0 | 3e-05 | 0.00039 | 0 | 0.00083 | 0.00568 | 0 | 0 | 3e-05 |
| Simulation (mean) | 0.00089 | 0 | 4e-05 | 0.00037 | 0 | 0.00091 | 0.00561 | 0 | 0 | 4e-05 |
| Data (std. dev.) | 0.00222 | 0 | 0.00036 | 0.00172 | 0 | 0.00464 | 0.00683 | 0 | 0 | 0.00044 |
| Simulation (std. dev.) | 0.00221 | 0 | 0.00045 | 0.00169 | 0 | 0.0051 | 0.00666 | 0 | 0 | 0.00046 |
| t-stat | 0.06318 | 0 | -0.5459 | 0.09748 | 0 | -0.22745 | 0.13035 | 0 | 0 | -0.08595 |
| **Latent demand** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | -0.20346 | 0.04419 | -0.38442 | -0.43084 | -0.32923 | -0.10603 | -0.33884 | -0.03691 | -0.19298 | -0.123 |
| Simulation (mean) | -0.1842 | 0.04298 | -0.36584 | -0.44043 | -0.31928 | -0.11385 | -0.32919 | -0.03669 | -0.17637 | -0.11621 |
| Data (std. dev.) | 0.23989 | 0.13805 | 0.20277 | 0.25763 | 0.20547 | 0.22737 | 0.19511 | 0.13898 | 0.22252 | 0.24316 |
| Simulation (std. dev.) | 0.16619 | 0.09547 | 0.14723 | 0.18202 | 0.15092 | 0.16246 | 0.13751 | 0.09718 | 0.16043 | 0.17208 |
| t-stat | -1.4611 | 0.1602 | -1.5932 | 0.66408 | -0.83275 | 0.60718 | -0.88485 | -0.02902 | -1.306 | -0.49788 |
| **Returns (%)** | **IBM** | **AXP** | **CPB** | **GE** | **XOM** | **INTC** | **KO** | **PFE** | **SYY** | **XRX** |
| Data (mean) | 0.91451 | 1.08485 | 0.09177 | 0.75742 | -0.21003 | 2.04382 | 0.88848 | 1.14193 | 2.00448 | 0.99955 |
| Simulation (mean) | 0.48439 | 1.00017 | 0.47566 | 1.34435 | 0.39247 | 0.76183 | 1.06302 | 1.18892 | 0.39994 | 0.5064 |
| Data (std. dev.) | 20.9169 | 12.7746 | 20.9741 | 12.5096 | 30.0809 | 21.2705 | 24.0883 | 14.1679 | 13.4409 | 14.4851 |
| Simulation (std. dev.) | 26.5417 | 25.2129 | 30.3453 | 59.572 | 35.9562 | 28.8055 | 39.2624 | 22.485 | 37.6887 | 25.9773 |
| t-stat | 0.20622 | 0.04287 | -0.16119 | -0.12596 | -0.21318 | 0.56698 | -0.05667 | -0.02664 | 0.54416 | 0.24232 |

*Notes:* This table reports the summary statistics (mean and standard deviation) of stock characteristics, dividend, latent demand and returns of the data and simulation. Stock characteristics and latent demand are modelled as stock-specific AR(1) processes fitted on data. Dividends are simulated from a normal distribution truncated below zero and the frequency of dividend payments is calibrated from the data. Returns are from prices computed according to Eq. 7

## A.4 Price variance decomposition

Figure A1: Latent demand contribution to price volatility




*Notes:* This figure reports the distribution of the proportion of returns variance explained by the latent demand across 100 independent simulated times series.

Table A3: Average contribution of latent demand

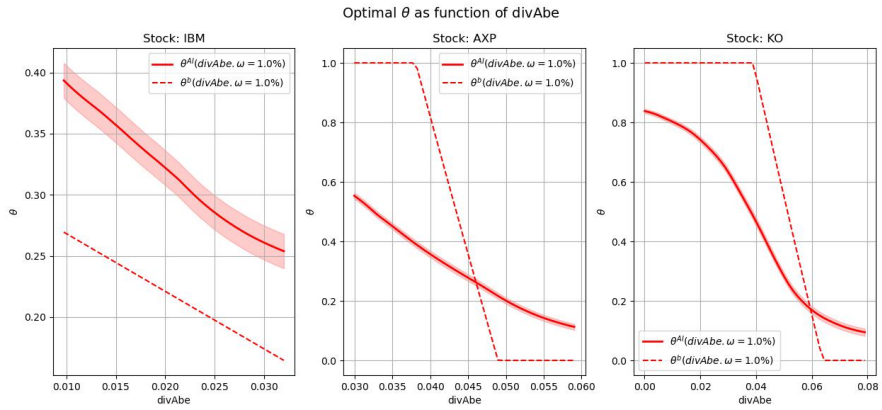
| | IBM | AXP | CPB | GE | XOM | INTC | KO | PFE | SYY | XRX |
|---|---|---|---|---|---|---|---|---|---|---|
| Variance explained by latent demand (%) | 0.53 | 0.28 | 0.4 | 0.7 | 0.6 | 0.4 | 0.18 | 0.31 | 0.3 | 0.37 |
| Autoregressive coefficient of latent demand | 0.87 | 0.79 | 0.8 | 0.4 | 0.66 | 0.87 | 0.83 | 0.82 | 0.82 | 0.91 |
| Variance of innovations of latent demand | 0.01 | 0.01 | 0 | 0.1 | 0.02 | 0.01 | 0.01 | 0.01 | 0.02 | 0.01 |

*Notes:* This table reports the proportion of returns variance explained by latent demand $\frac{\sigma(R^*)}{\sigma(R)}$ and the autoregressive coefficient and variance of innovation of the AR(1) process fitted on the data of latent demand. The autoregressive coefficient is estimated using simulated method of moments described in Section A.3.

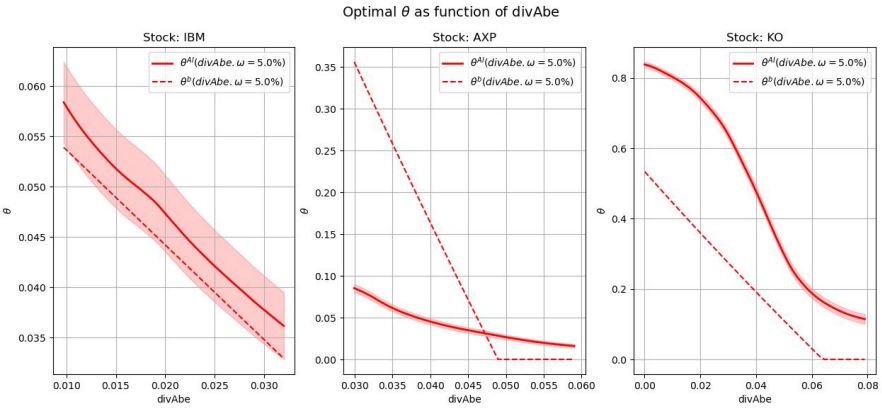
# B Portfolio comparative statics in stock characteristics

Figure A2: Portfolio policy for $J = 1$: comparative statics in $x_k$ =dividend-to-book

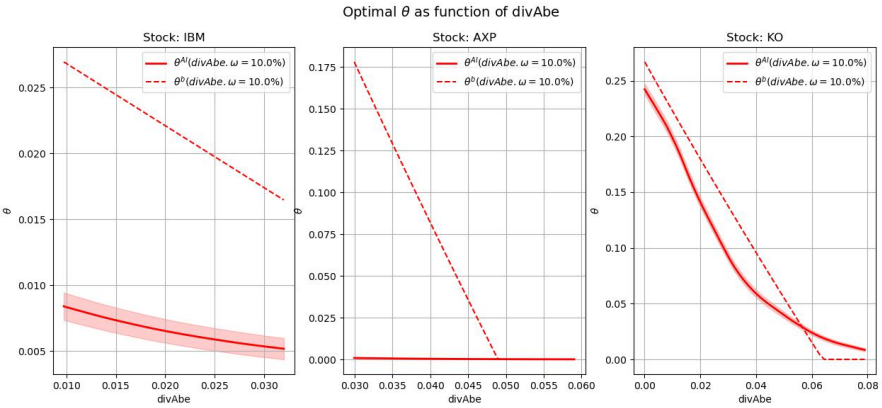
(a) $J = 1,\ J\omega = 1\%$




(b) $J = 1,\ J\omega = 5\%$



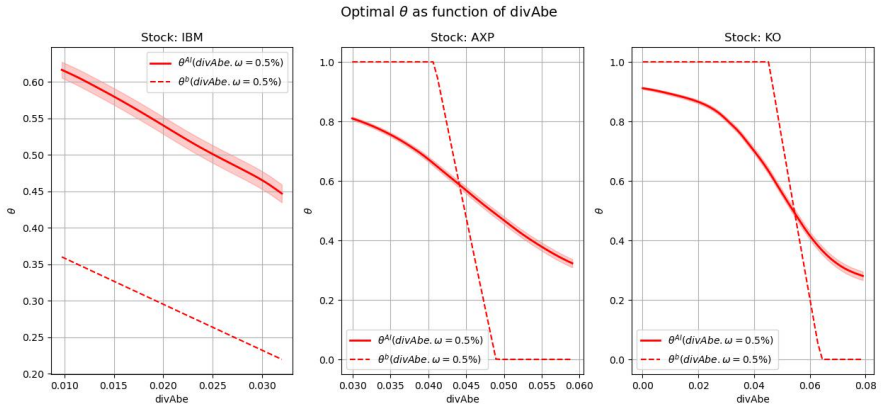

(c) $J = 1,\ J\omega = 10\%$



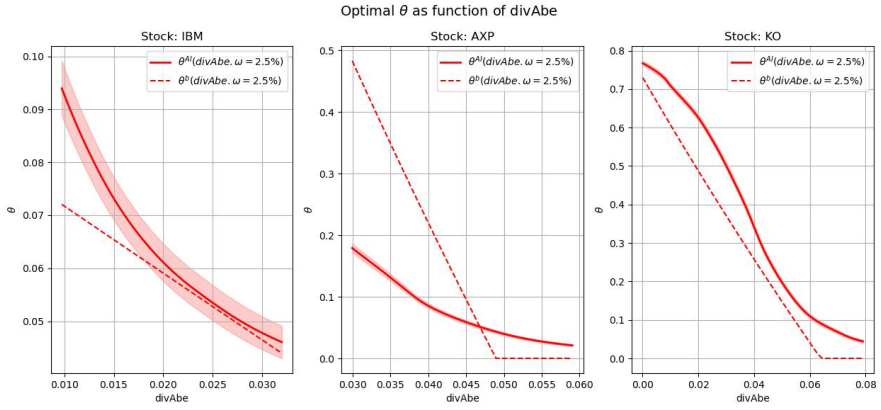

*Notes:* Each panel shows the average portfolio weight on the risky asset of the AI trader across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *dividend-to-book*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A3: Portfolio policy for $J = 2$: comparative statics in $x_k$ =dividend-to-book

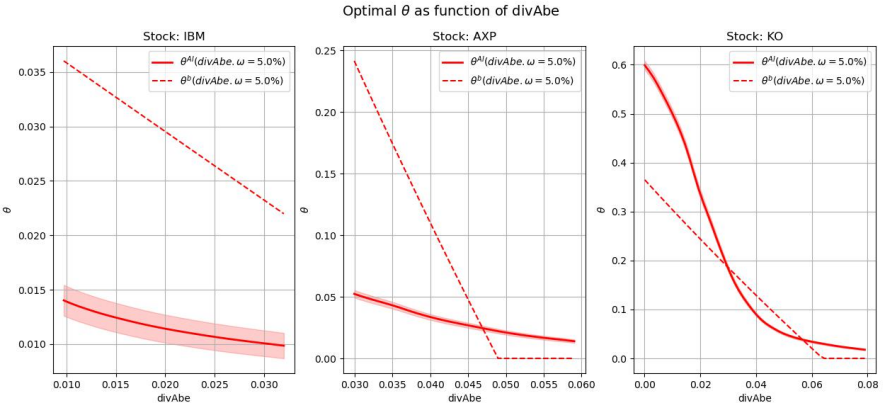
(a) $J = 2,\ J\omega = 1\%$




(b) $J = 2,\ J\omega = 5\%$



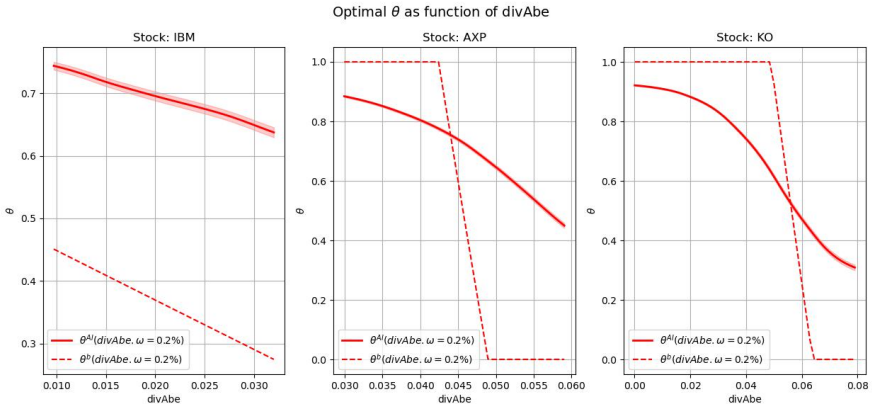

(c) $J = 5,\ J\omega = 10\%$



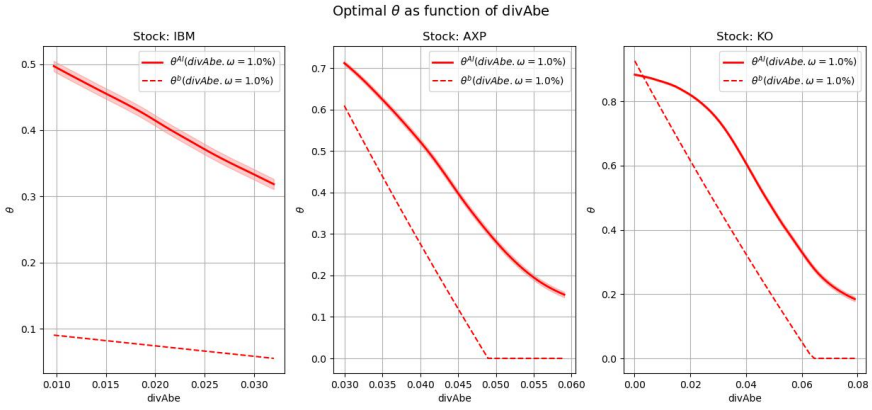

*Notes:* Each panel shows the average portfolio weight on the risky asset of the two AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *dividend-to-book*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A4: Portfolio policy for $J = 5$: comparative statics in $x_k$ =dividend-to-book

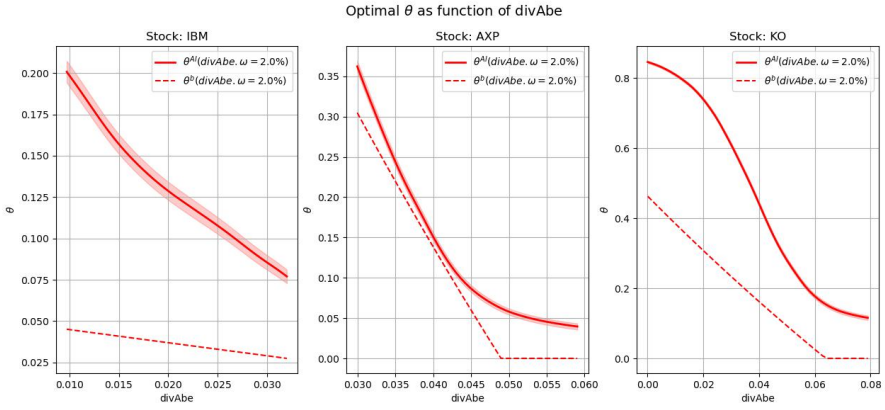
(a) $J = 5,\ J\omega = 1\%$




(b) $J = 5,\ J\omega = 5\%$



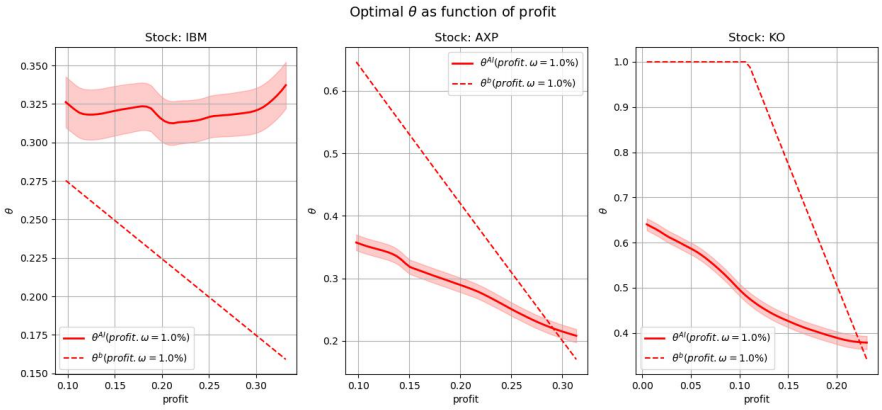

(c) $J = 5,\ J\omega = 10\%$



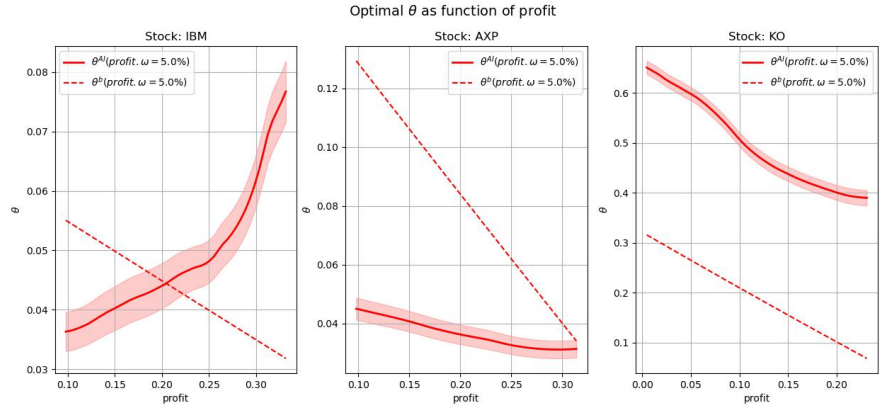

*Notes:* Each panel shows the average portfolio weight on the risky asset of the five AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *dividend-to-book*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A5: Portfolio policy for $J = 1$: comparative statics in $x_k$ =profitability

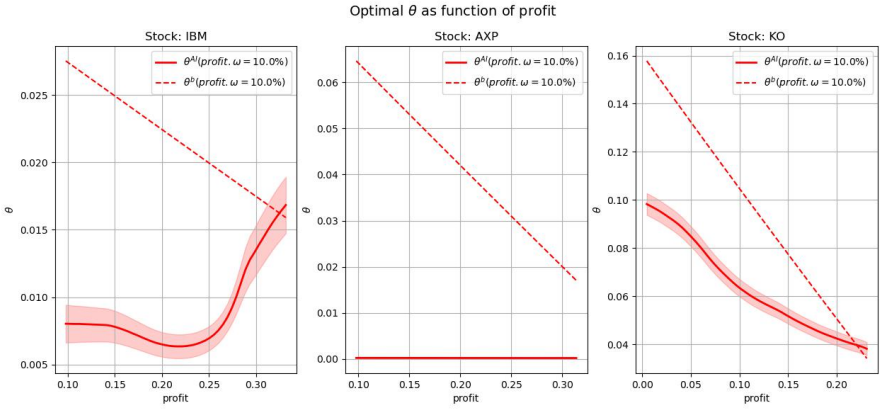
(a) $J = 1,\ J\omega = 1\%$




(b) $J = 1,\ J\omega = 5\%$



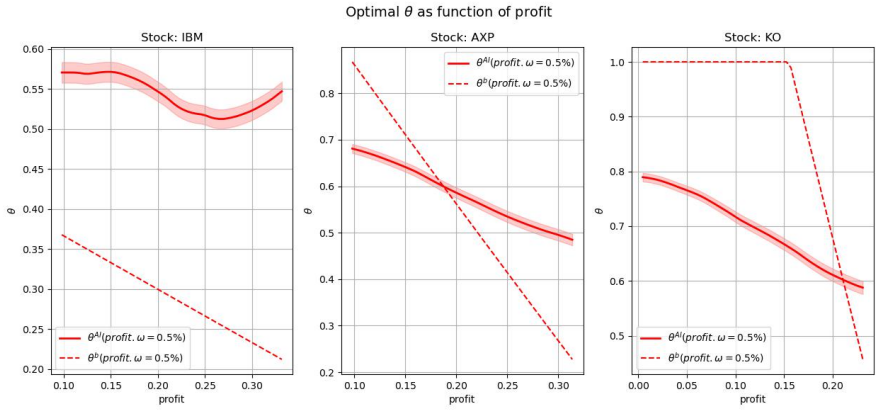

(c) $J = 1,\ J\omega = 10\%$



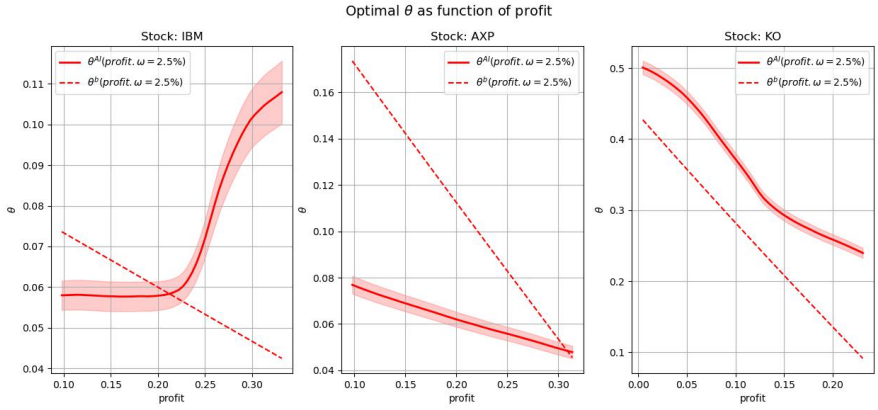

*Notes:* Each panel shows the average portfolio weight on the risky asset of the AI trader across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *profitability*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A6: Portfolio policy for $J = 2$: comparative statics in $x_k$ =profitability

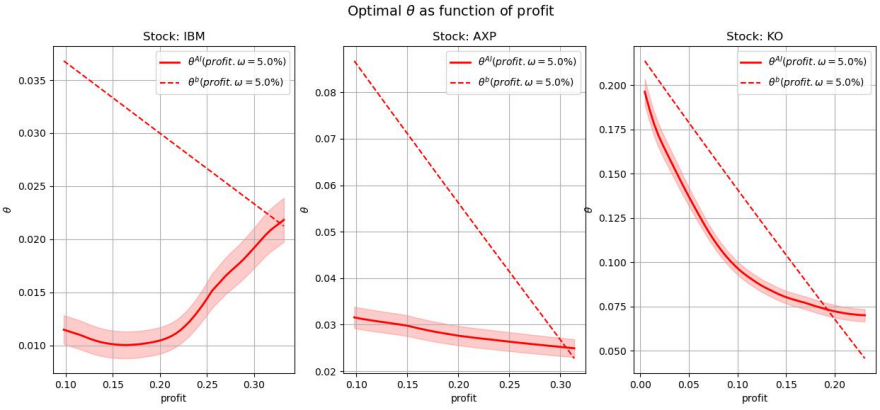
(a) $J = 2,\ J\omega = 1\%$




(b) $J = 2,\ J\omega = 5\%$



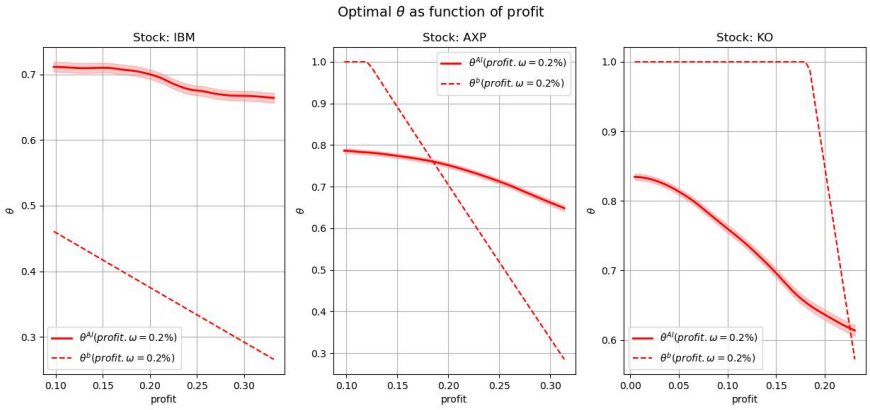

(c) $J = 2,\ J\omega = 10\%$



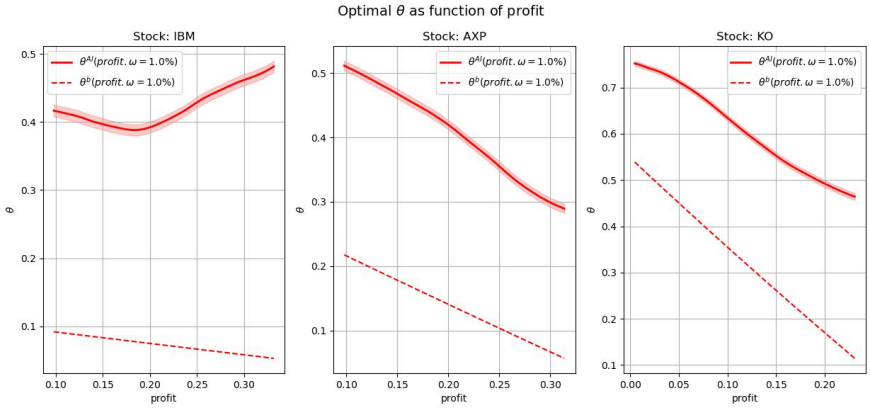

*Notes:* Each panel shows the average portfolio weight on the risky asset of the two AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *profitability*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A7: Portfolio policy for $J = 5$: comparative statics in $x_k$ =profitability

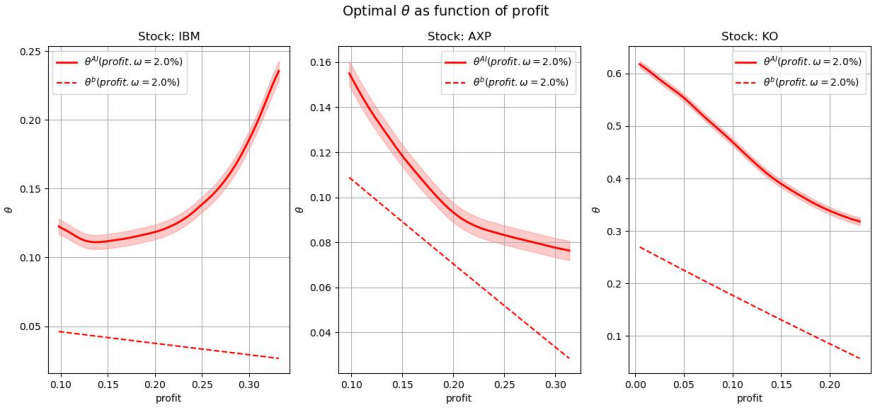
(a) $J = 5,\ J\omega = 1\%$




(b) $J = 5,\ J\omega = 5\%$



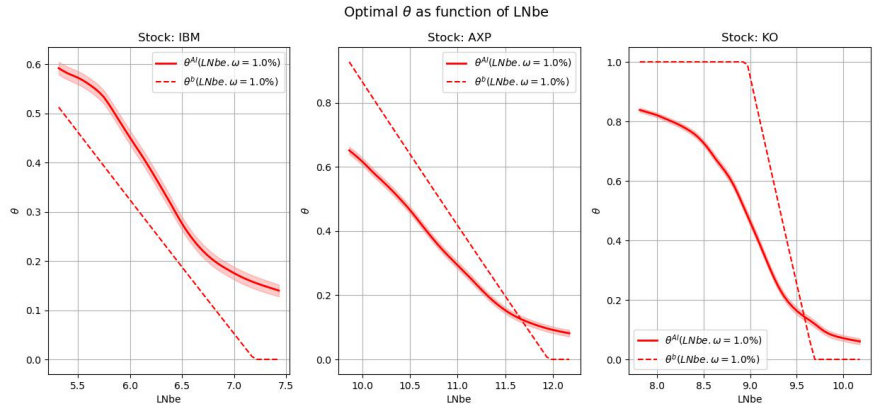

(c) $J = 5,\ J\omega = 10\%$



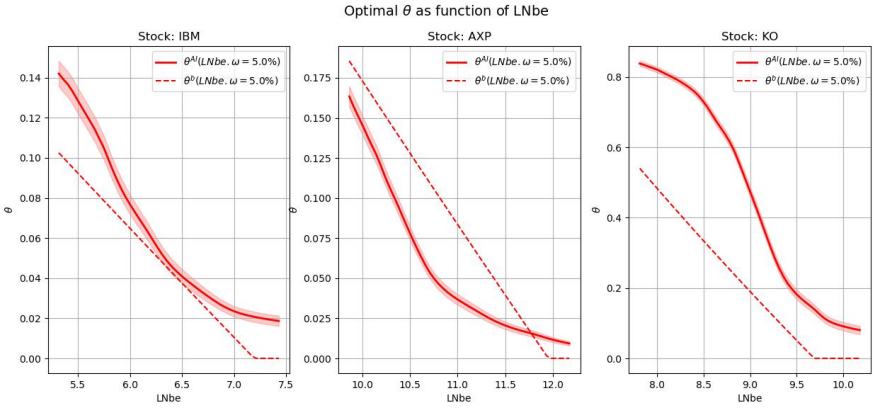

*Notes:* Each panel shows the average portfolio weight on the risky asset of the five AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *profitability*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A8: Portfolio policy for $J = 1$: comparative statics in $x_k$ =book-to-equity

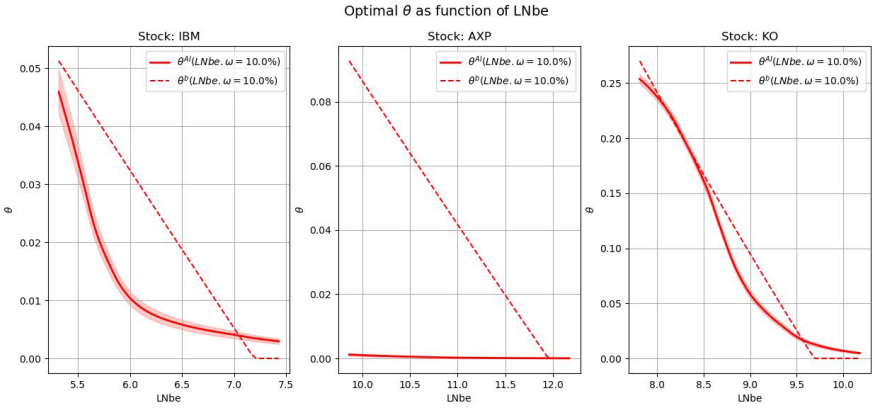
(a) $J = 1,\ J\omega = 1\%$




(b) $J = 1,\ J\omega = 5\%$



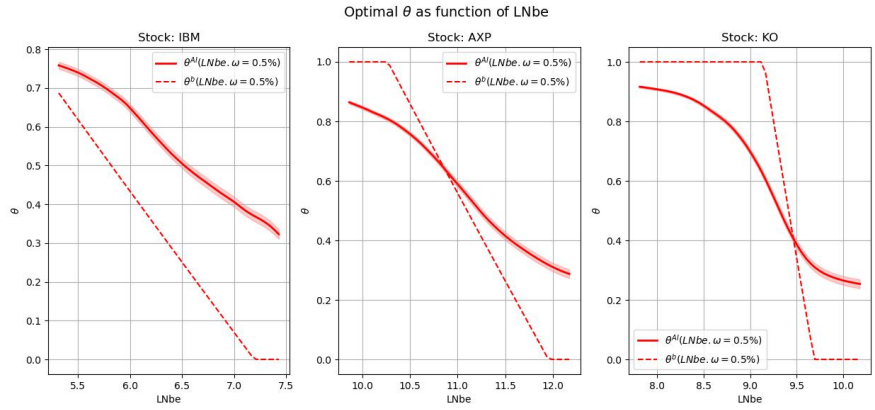

(c) $J = 1,\ J\omega = 10\%$



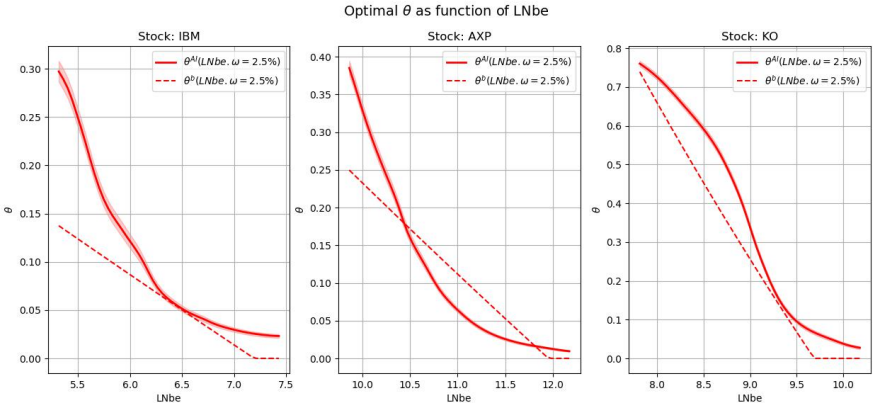

*Notes:* Each panel shows the average portfolio weight on the risky asset of the AI trader across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *book-to-equity*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A9: Portfolio policy for $J = 2$: comparative statics in $x_k$ =book-to-equity

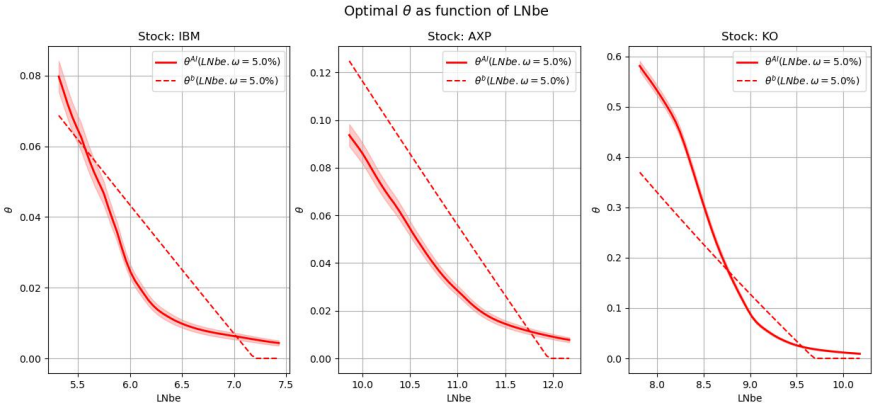
(a) $J = 2,\ J\omega = 1\%$




(b) $J = 2,\ J\omega = 5\%$



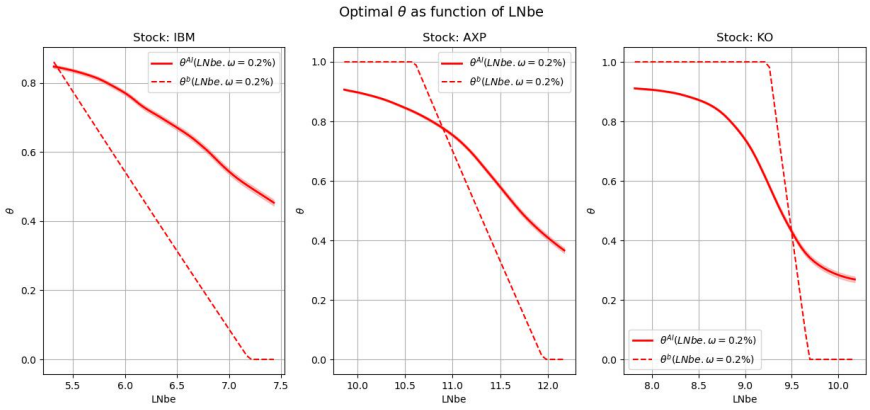

(c) $J = 2,\ J\omega = 10\%$



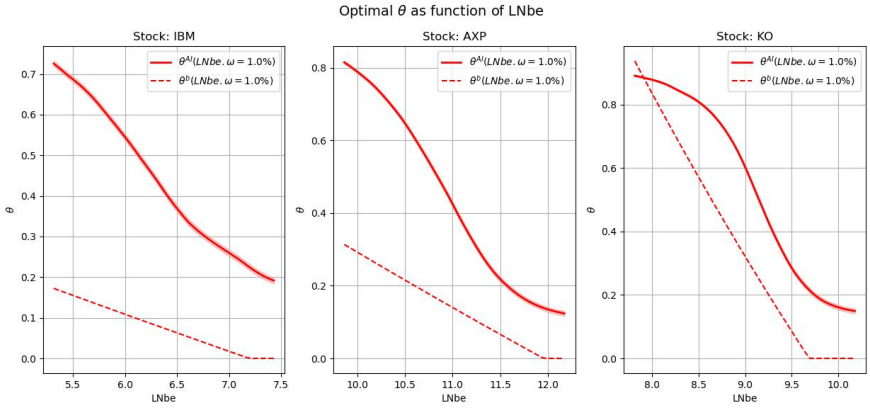

*Notes:* Each panel shows the average portfolio weight on the risky asset of the two AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *book-to-equity*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A10: Portfolio policy for $J = 5$: comparative statics in $x_k$ =book-to-equity

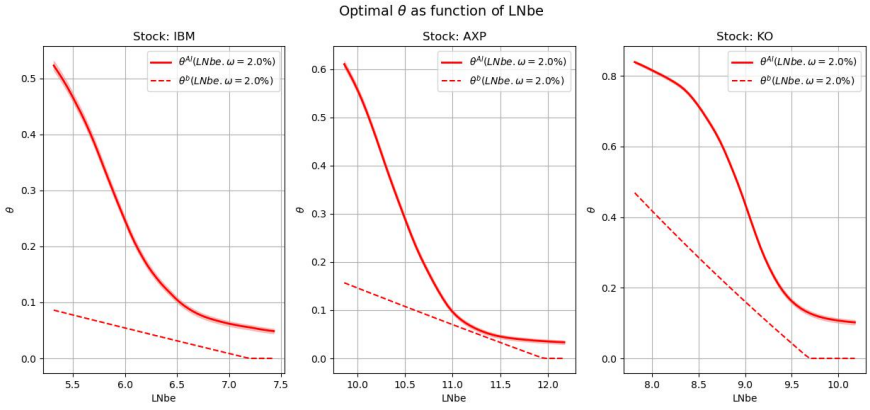
(a) $J = 5,\ J\omega = 1\%$




(b) $J = 5,\ J\omega = 5\%$



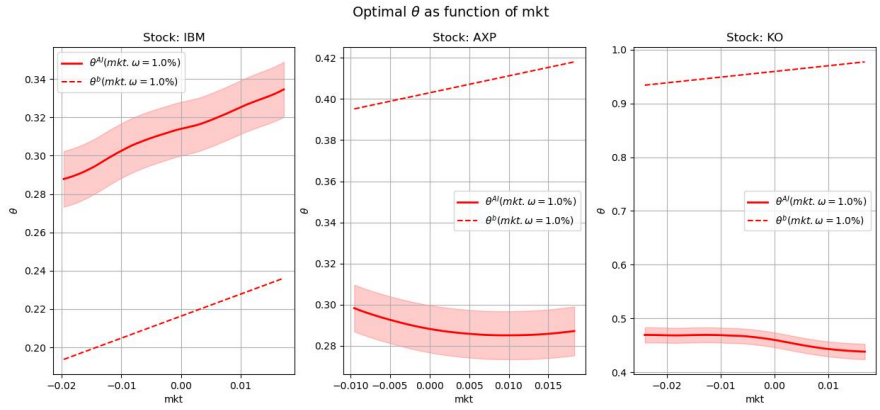

(c) $J = 5,\ J\omega = 10\%$



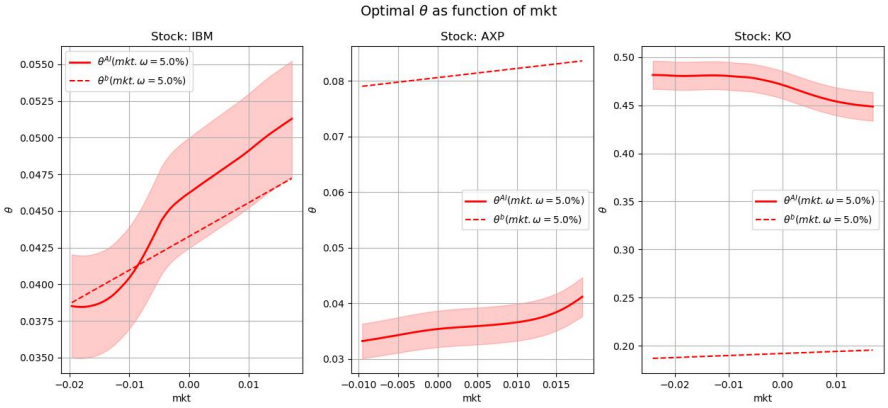

*Notes:* Each panel shows the average portfolio weight on the risky asset of the five AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *book-to-equity*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A11: Portfolio policy for $J = 1$: comparative statics in $x_k$ =market beta

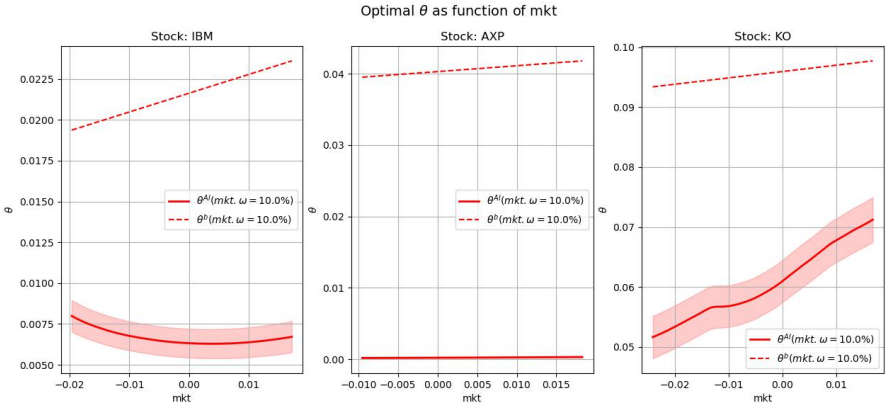
(a) $J = 1,\ J\omega = 1\%$




(b) $J = 1,\ J\omega = 5\%$



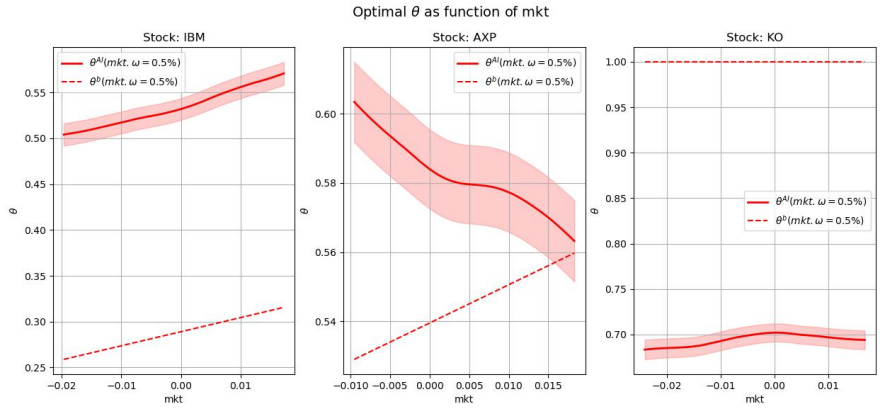

(c) $J = 1,\ J\omega = 10\%$



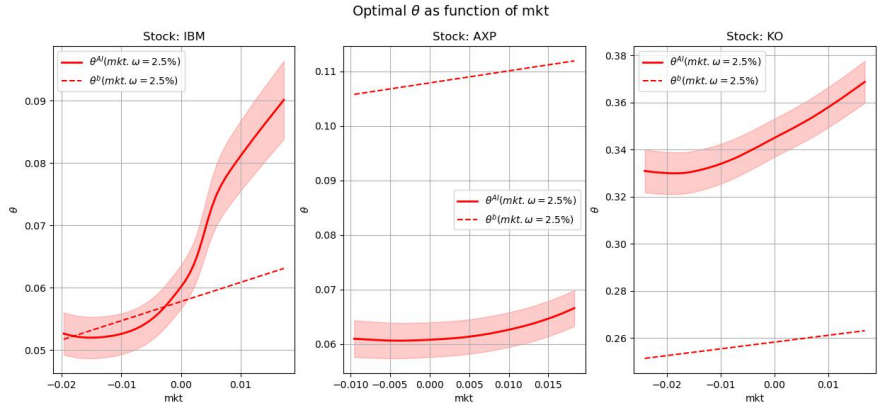

*Notes:* Each panel shows the average portfolio weight on the risky asset of the AI trader across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *market beta*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A12: Portfolio policy for $J = 2$: comparative statics in $x_k$ =market beta

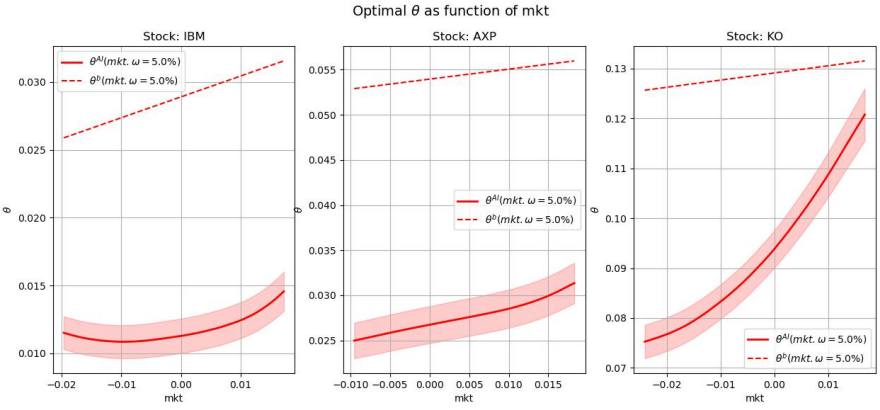
(a) $J = 2,\ J\omega = 1\%$




(b) $J = 2,\ J\omega = 5\%$



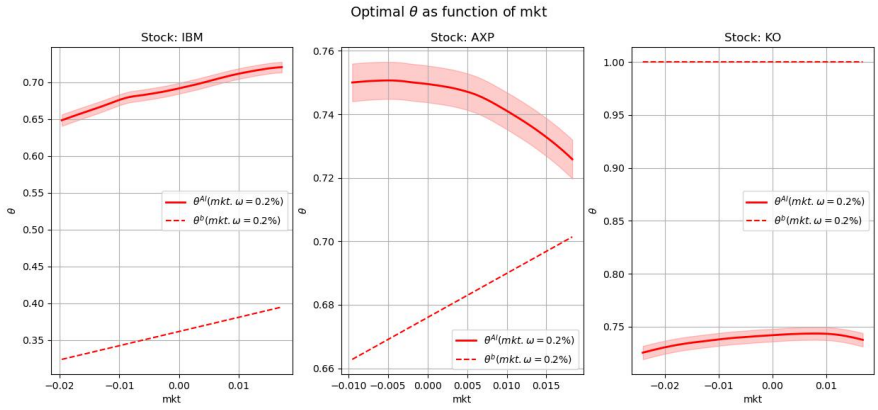

(c) $J = 2,\ J\omega = 10\%$



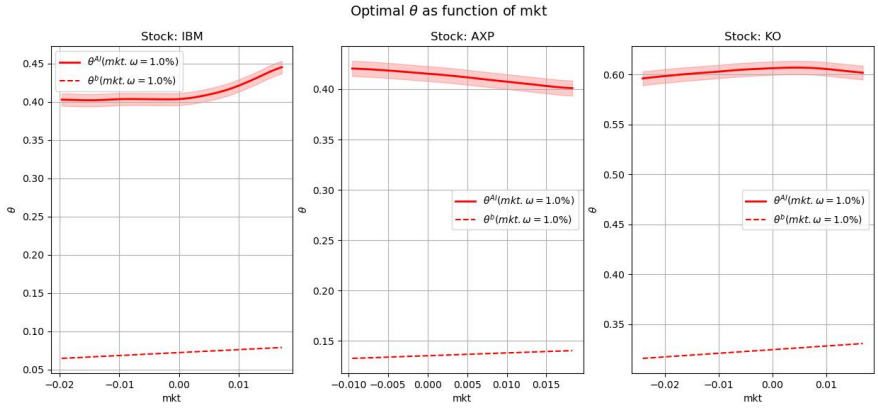

*Notes:* Each panel shows the average portfolio weight on the risky asset of the two AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *market beta*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A13: Portfolio policy for $J = 5$: comparative statics in $x_k$ =market beta

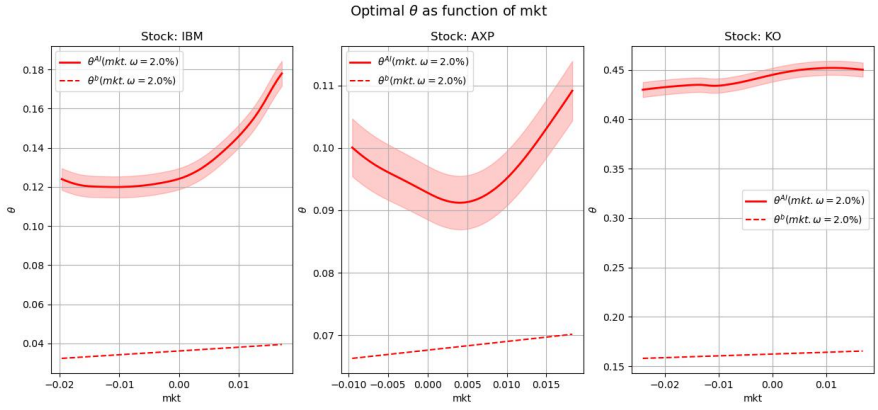
(a) $J = 5,\ J\omega = 1\%$




(b) $J = 5,\ J\omega = 5\%$



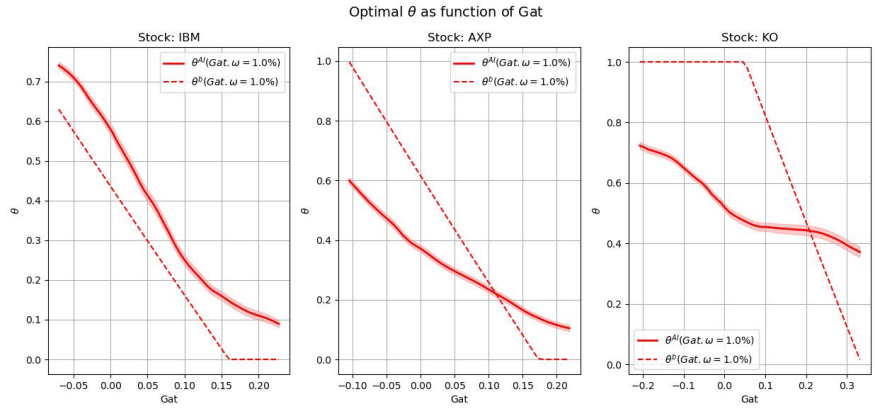

(c) $J = 5,\ J\omega = 10\%$



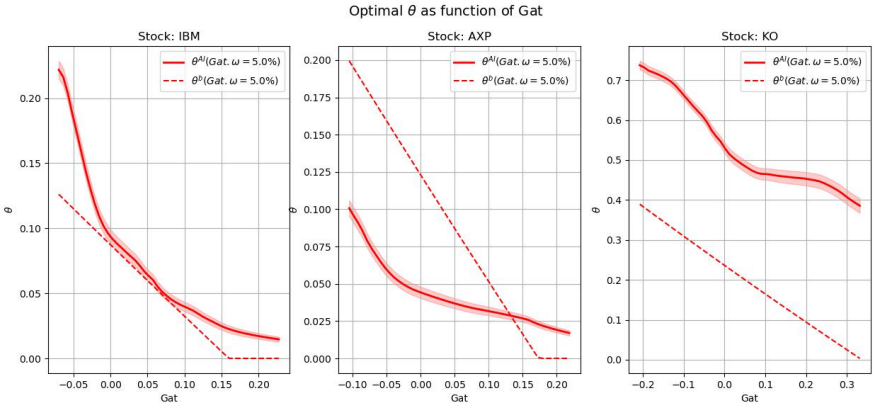

*Notes:* Each panel shows the average portfolio weight on the risky asset of the five AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *market beta*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A14: Portfolio policy for $J = 1$: comparative statics in $x_k$ =investment growth

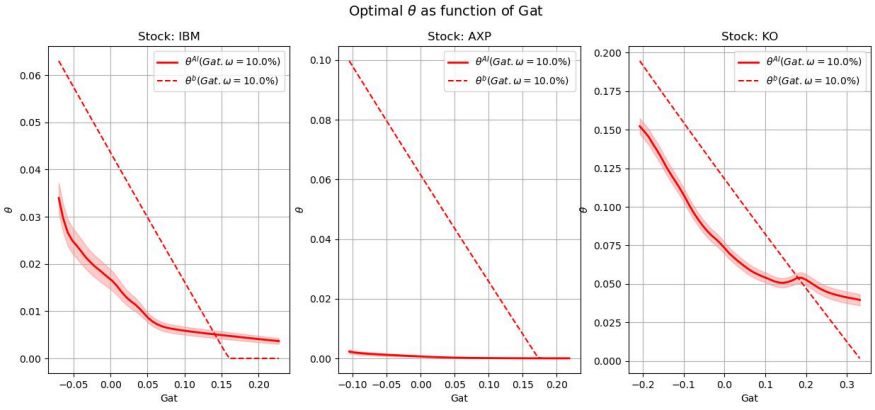
(a) $J = 1,\ J\omega = 1\%$




(b) $J = 1,\ J\omega = 5\%$



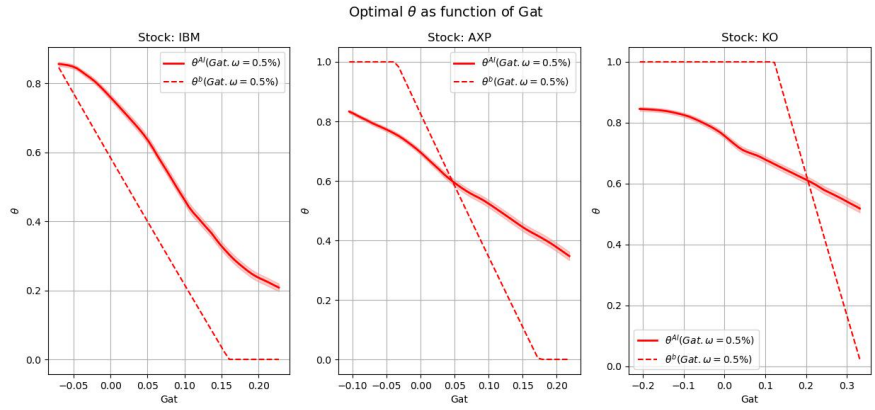

(c) $J = 1,\ J\omega = 10\%$



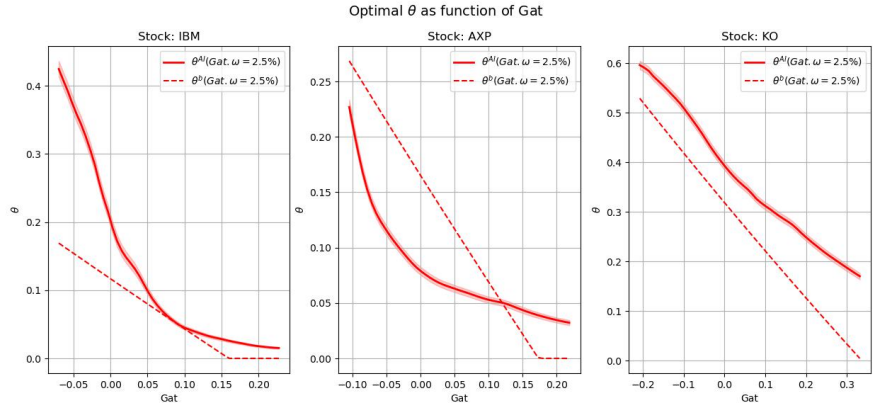

*Notes:* Each panel shows the average portfolio weight on the risky asset of the AI trader across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *investment growth*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A15: Portfolio policy for $J = 2$: comparative statics in $x_k$ =investment growth

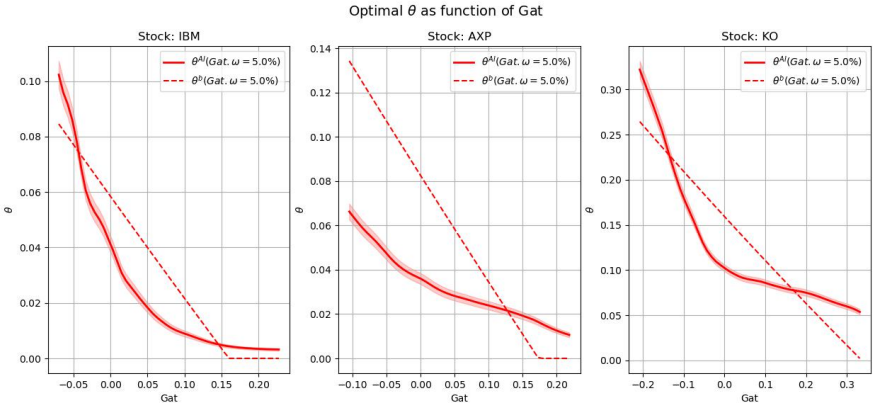
(a) $J = 2,\ J\omega = 1\%$




(b) $J = 2,\ J\omega = 5\%$



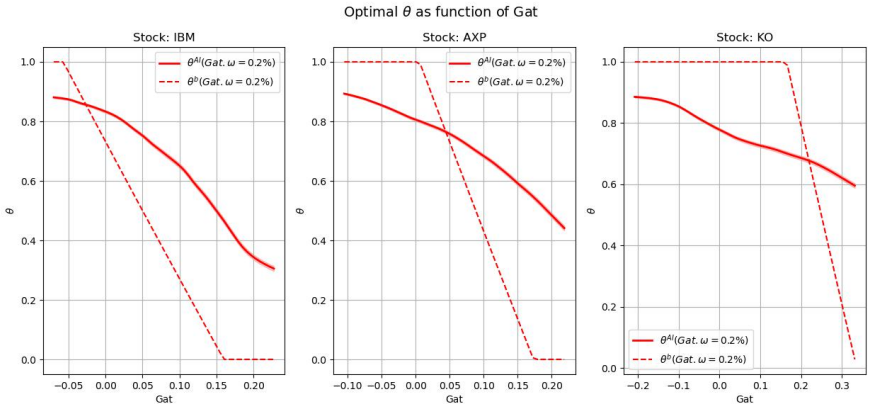

(c) $J = 2,\ J\omega = 10\%$



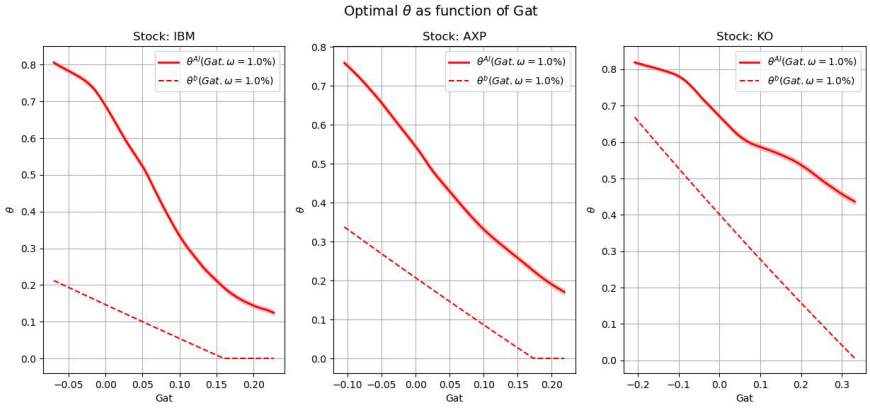

*Notes:* Each panel shows the average portfolio weight on the risky asset of the two AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *investment growth*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

Figure A16: Portfolio policy for $J = 5$: comparative statics in $x_k$ =investment growth

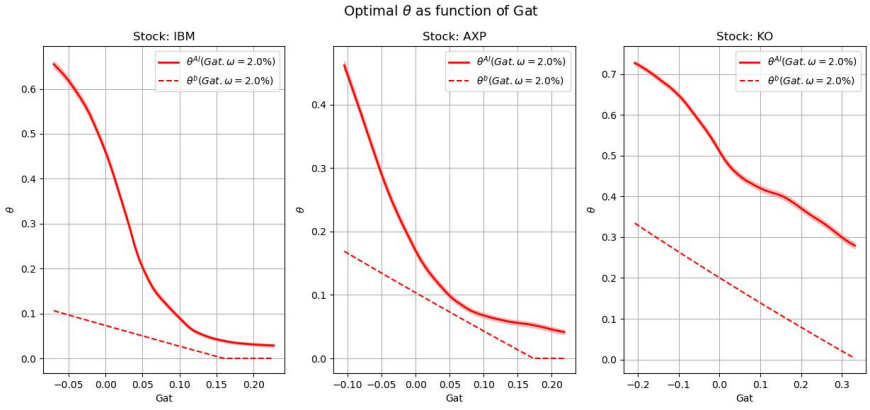
(a) $J = 5,\ J\omega = 1\%$




(b) $J = 5,\ J\omega = 5\%$




(c) $J = 5,\ J\omega = 10\%$




*Notes:* Each panel shows the average portfolio weight on the risky asset of the five AI traders across 50 simulations (red solid line), alongside the portfolio weight computed by solving the rational expectations model (red dashed line), plotted against the stock's *investment growth*. The red shaded areas indicate the 99% confidence interval. Other state variables are set to their sample averages. The three panels differ in the traders' total assets-under-management share $J\omega$ of the market size (1%, 5%, and 10%).

# C Additional tables

Table A4: Average correlation across AI traders $\rho^{AI}(\omega, J)$

| | $J\omega = 1\%$ | | | | | | $J\omega = 10\%$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | IBM | AMX | KO | IBM | AMX | KO | IBM | AMX | KO |
| J=2 | 0.881 | 0.885 | 0.862 | 0.817 | 0.64 | 0.889 | 0.57 | 0.213 | 0.821 |
| | (0.043) | (0.094) | (0.066) | (0.061) | (0.401) | (0.034) | (0.300) | (0.402) | (0.053) |
| J=5 | 0.877 | 0.874 | 0.872 | 0.854 | 0.865 | 0.873 | 0.819 | 0.717 | 0.852 |
| | (0.032) | (0.028) | (0.026) | (0.041) | (0.087) | (0.036) | (0.055) | (0.199) | (0.039) |

*Notes:* Std. Dev across simulations in parenthesis