The equity premium puzzle: an application of an agent-based evolutionary model

Luca Gerotto[1], Paolo Pellizzari[2], and Marco Tolotti[3]

[1] Università Cattolica del Sacro Cuore, Department of Economics and Finance, Milan, Italy
luca.gerotto@unicatt.it
https://sites.google.com/view/gerottoluca/home-page
[2] Ca' Foscari University of Venice, Department of Economics, Venice, Italy
[3] Ca' Foscari University of Venice, Department of Management, Venice, Italy

Abstract. We describe an agent-based model of a financial market with a stock and a bond. Agents compete in repeated rounds, decide whether to acquire costly information and can pick one of 16 strategies to allocate their investments, under evolutionary pressure driven by the comparison of the realized short-term revenues from trading. We show that, while informed traders survive in some cases, the equilibrium shares are strongly biased in favor of strategies that make little use of information and systematically overestimate the riskiness of the stock. As a consequence, the majority of the population ends up in buying fewer stocks than would be otherwise expected or deemed rational.
This evolutionary dynamics offers a novel way to explain the equity premium puzzle first described by Mehra and Prescott (*The equity premium: A puzzle*. Journal of Monetary Economics 1985), according to which it's hard to find reasons for the widespread lack of investment in risky assets. Evolution based on a straightforward comparison of revenues is a simple and cognitively appealing avenue to reach a population of traders using (over-)cautious strategies to curb the risk of long-term "financial extinction". Simulations run in NetLogo also demonstrate that very little information may be used in noisy markets or when the cost of information is substantial.

Keywords: Agent-based models, Evolutionary game theory, Equity premium puzzle.

1 Introduction

As famously pointed out in Mehra and Prescott [6] it is difficult to reconcile standard financial economic models with the observation that investors purchase relatively small amounts of stocks, whose average returns are historically much bigger than the safe rate (obtained with highly-rated bonds or bills). This conundrum has been know as the "equity premium puzzle" (EPP), see the page of the Federal Reserve Economic Data website myf.red/g/6LsS for a visual representation of the premium in the last 25 years (difference between stock and BBB corporate yields).

Economists have tried to single out ways to explain why people invest far less in stocks that would be implied by their risk aversion, as measured in other (personal or social) situations. In Benartzi and Thaler [2] it is argued that “the combination of a high equity premium, a low risk-free rate, and smooth consumption is difficult to explain with plausible levels of investor risk aversion” and “myopic loss aversion” is introduced as a possible justification. Barberis et al. [1], in a somewhat similar vein, need (dis)utility from fluctuation of financial wealth. DeLong and Magin [4] survey other approaches, including the use of prospect theory, the value of liquidity or the role of taxation to account for the puzzle.

Agent-based modeling is a methodology to build computational models of real-world systems where autonomous agents (individuals, traders, households, firms, software agents, robots...) interact in various forms, learn, sense the environment and often use fast and frugal heuristics that do not need unrealistic degrees of rationality or processing capability. A good introduction is in Railsback and Grimm [8], that also includes a thorough treatment of the NetLogo programming platform that was used to develop the model presented in this paper. See Steinbacher et al. [11] for a recent review.

Among the features that can be used in agent-based models to analyze possible paths to generate a sizeable equity premium, we investigate the role of direct and indirect interaction among traders. Quite naturally in a financial setup, agents collectively contribute to form the market price of the stock and indirectly affect –and are affected by– the strategies used by the other traders. Moreover, agents are occasionally paired with random peers and contrast the profitability of their strategy, switching to the best-performing one in the quest for improvement. This direct learning scheme is based on pure imitation of successful examples with no need to gather, or elaborate data or (try to) compute sophisticated conditional equilibria.

The conceptual framework of this paper is inspired by evolutionary game theory. Originally introduced by biologists to analyze with formal tools long-term adaptation of biological populations, the idea that the reproductive fitness depends on the genotypes was later extended to economics, see Sandholm [10] and Newton [7]. Of course, in this setup, agents are not assumed to be genetically pre-programmed but are able to adjust their behaviors or strategies favoring larger payoffs (as opposed to Darwinian fitness). Recently, in Robson and Orr [9] it is argued by means of an evolutionary model that the EPP is due to agents’ greater aversion to aggregate risk, such as the one faced in financial markets, with respect to idiosyncratic risk (of more personal nature).

Our agent-based evolutionary model converges to an equilibrium with a overwhelming presence of demand functions (or, if you wish, strategies) which systematically overestimate the variance of the risky asset. As a consequence, a large share of market population hold relatively small amount of stocks. The option to buy a costly information signal to predict return can reduce the effect. However, this holds only if information is accurate, cheap and used in non-volatile markets. Generally speaking, the overestimation of the variance in the long-run, and its related implications, are observed in many of the instances examined in a detailed robustness test. Moreover, in such instances, most of the (survived) equilibrium strategies appear to use little or no information. These results shed a novel light on the puzzle and point to a potential new channel to explain this long-debated anomaly.

Next Section presents the ABM model, describing the market, the agents and the learning protocol that is naturally used as an evolutionary device to favor trading strategies with higher payoffs. Sect. 3 presents simulations' results obtained in a benchmark case. Some key parameters are then varied in the following section that shows that results are remarkably robust. We finally conclude with some discussion.

2 The model

This section describes a simple market with a risky stock and a riskless asset. The setup is minimal to keep the focus on the co-evolution of a population of traders who compete for high profits and decide which information and risk factors to take into account in their decisions.

2.1 The market

We assume N heterogeneous agents are given an initial endowment w_0 at the beginning of every period, place orders and collect revenues that are immediately consumed at the end of the period. Some agents are then allowed, with some probability, to change their trading strategy using an imitation mechanism that favors the ones with larger revenues. A new population, with a different distribution of strategies, is formed and the game is repeated T times, $t = 1, ..., T$.

The riskless asset has unit cost and pays $R = 1 + r$ after one period. There is also a stock in zero net supply with random payoff $\tilde{D}_t = d + \tilde{\theta}_t + \tilde{\epsilon}_t$, where d is a known deterministic component of revenues, $\tilde{\theta}_t$ is an informative signal that can be acquired at a cost of c per period and $\tilde{\epsilon}_t$ is an unobservable noise term (unknown to everyone). We omit t and occurrences of tilde, unless needed, in what follows and notice that θ can be referred as *information*, as it truly affects the random revenue D. Some agents, however, may believe that an uninformative signal $\tilde{\gamma}_t \equiv \gamma$ also affects the outcome. γ can be obtained at no cost, if desired, and it can be considered as pure *misinformation* having nothing to do with D (even though agents regard it as helpful).

We assume that θ, ϵ, γ are normally and independently distributed:

$$\theta \sim N(0, v_\theta), \epsilon \sim N(0, v_\epsilon), \gamma \sim N(0, v_\gamma), \theta \perp \epsilon \perp \gamma,$$

where v_θ, v_ϵ and v_γ are the variances of θ, ϵ and γ.

The equilibrium price at any time t is determined by (net) demands of agents solving the equation

$$\sum_{i=1}^{N} x_i(p_t|\mathbf{b}_i, \epsilon, \gamma) = 0, \tag{1}$$

where $x_i(p_t|\mathbf{b}_i)$ is the demand of the i-th agent at price p_t and $\mathbf{b}_i$ is a vector of heterogenous parameters differentiating individual strategic behavior and will be described in detail in the next subsection.

2.2 The agents

The demand of the risky asset is consistent with the idea that agents, as a whole, are aware that D depends on some of (but not necessarily all) the variables mentioned above: for a price p the demand function of the i-th agent is

$$x_i(p, \mathbf{b}_i) = \frac{d + b_1\theta + b_2\gamma - pR}{a(b_3 v_\theta + b_4 v_\gamma + v_\epsilon)}, \tag{2}$$

where $\mathbf{b}_i = (b_1, b_2, b_3, b_4), i = 1, ..., N$ is a vector of individual bits (i.e., $\mathbf{b}_i \in \{0,1\}^4$) that can evolve in time due to imitation (and, hence, should be formally denoted as $\mathbf{b}_{it}$ when referring to the i-th agent at time t). Again, for the sake of exposition, we omit individual and temporal indexes. As $\mathbf{b}$ shapes and determines the trading behavior of the agents, we will refer to it using the term "strategy".

The demand in (2) increases with the perceived average revenue in excess of what would be gained with a riskless investment (see the numerator) and is corrected for the perceived variance, up to the relative risk aversion coefficient a, held constant across the population of traders. Each bit $b_j, j = 1, ..., 4$ can be thought as a way to switch on and off some random variable in Equation (2). Take, for instance, an agent with strategy $\mathbf{b} = (1, 0, 0, 0)$: she acquires and uses information θ in the numerator and perceives a residual variance (in the denominator) depending on ϵ alone. Such an agent can be regarded as *informed*, as she employs θ and discard γ, as well as *rational*, as she correctly realizes that the variance of D is not affected by misinformation γ or by θ, that is known, but only depends on the noisy and unobservable component ϵ. Indeed, in this paper, rationality refers to the correct understanding of the data-generating process of D, and would be achieved in the model when b_1 is either 0 or 1, $b_3 = 1 - b_1$ and $b_2 = b_4 = 0$. In other words, a rational agent would ignore bits related to γ and would either buy the information or, if not, include it in the denominator of Eq. (2).

By contrast, someone using the strategy $\mathbf{b} = (0, 0, 1, 1)$ may be considered quite *prudent*: indeed, none of the signals θ or γ is used and the perceived variance is large as it includes both the summands v_θ, v_γ, as well as the ubiquitous v_ϵ. As a consequence, such an agent would trade a much smaller x, for any given p, than an informed trader. In this specific case, clearly, the strategy is not fully rational as γ is erroneously affecting the demand.

Table 1 lists some relevant strategies that can, to some extent, be interpreted in terms of their ability to correctly identify the conditional expected revenue and variance.

While it's not always possible to provide a behavioral description of every strategy encoded in $\mathbf{b}$, Table 1 features a few meaningful examples. For instance, uninformed agents discard useful information (avoiding the cost), but are rational in that they correctly understand the way returns are generated and take

Table 1. Description of several strategies encoded in the vector $\mathbf{b} = (b_1, b_2, b_3, b_4)$. The first two bits are related to the use of the information and misinformation in the prediction of the mean returns; the third and fourth bits are used to compute the perceived risk.

Nickname	b_1	b_2	b_3	b_4	Description
Informed	1	0	0	0	*informed, rational*
Prudent	0	0	1	1	*uninformed, irrational, small demand*
Uninformed	0	0	1	0	*uninformed, rational*
Fearless	0	0	0	0	*uninformed, irrational, relatively large and stable demand*
Confused	1	1	1	1	*informed and misinformed, irrational*

into account the uncertainty arising from the unknown θ; traders with a null $\mathbf{b}$, in the last row of the table, are dubbed *fearless* to stress the lack of any risk adjustment in the denominator of (2), an action leading often to relatively large orders.

2.3 Learning

At the end of any period t, the equilibrium price p_t is computed using Eq. (1). Clearly, p_t is a function of the distribution of the strategies in the population and of realized random variables θ_t, γ_t, that are known to the agents whose first and second bits are set to 1. After the unobservable shock ϵ_t is drawn and uncertainty is resolved, the realized profit for an agent is

$$w_{it} = x_{it} D_t + (w_0 - x_{it} p_t) R - c \cdot b_1,$$

where the first component is the revenues arising from x units of the risky stock, the second part comes from investing in the riskless asset all the cash that was not used to get the stocks, and cb_1 is the cost of getting the information.

Learning is based on agents' pairwise comparisons of the profits. In detail, we form $h < N/2$ random couples of traders and, letting individuals i and j be one such couple, the vectors $\mathbf{b}$ are updated using:

$$\begin{aligned}&\text{If } w_{it} > w_{jt},\ \mathbf{b}_{j,t+1} = \mathbf{b}_{i,t};\\ &\text{If } w_{it} < w_{jt},\ \mathbf{b}_{i,t+1} = \mathbf{b}_{j,t};\\ &\text{If } w_{it} = w_{jt}, \text{ no change.}\end{aligned}$$

The interpretation of this learning scheme is immediate in terms of evolutionary game theory: agents using possibly different strategies obtain different payoffs; they occasionally meet another peer and revise their strategy switching to the one with bigger profits (pure imitation of better strategies); as a consequence, strategies with better payoffs tend to increase their relative frequency (which, in turn, may alter their future success).

Slightly more formally, a population $\mathcal{B}_t = \{\mathbf{b}_{it} : i = 1, ..., N\}$ of agents (or strategies) at time t evolves using the above revision protocol to obtain $\mathcal{B}_{t+1}$, which has at most h differences with respect to $\mathcal{B}_t$. The relative frequencies of each of the 16 strategies are then investigated letting t reach T, for large T.

3 Results

This section discusses the results obtained simulating in NetLogo [12] the agent-based model described in the previous section.[4] We first illustrate the outcomes in a benchmark case and then show how results change varying systematically some of the parameters of the model.

3.1 The benchmark case

Table 2 lists the values taken by the parameters of the model in a benchmark configuration.

Table 2. Values of the parameters of the model used in the benchmark case, with a brief description.

Param.	Value	Description	Param.	Value	Description
N	1000	Number of agents	R	1.01	Gross return of riskless asset
T	10000	Trading periods	a	2	Risk aversion coefficient
c	0.03	Cost the informative θ	v_θ	0.01	Variance of information θ
h	15	Learning couples	v_ϵ	0.04	Variance of noise ϵ
d	1.1	Part of stock revenue	v_γ	0.01	Var. of misinformation γ

The values are roughly representative of a market where, for instance, one period is one year, the riskless rate is 1%, the standard deviation of the revenues of the risky asset is 20%=$\sqrt{0.04}$, the standard deviation of information (and misinformation) is 10% and the cost of acquiring the informative signal is 3% (that may be a reasonable approximation of the fees of a financial professional providing valuable advice).

Figure 1 shows how the fractions of informed (1000), prudent (0011) and uninformed (0010) agents evolve in 10000 periods in one standard simulation run.[5] It can be seen that about 2000 periods suffice to reach a homeostatic equilibrium where the share of prudent traders hovers around 90%, informed agents are 10% and we observe the extinction of the uninformed (as well as any **b** other then with 1000 and 0011.

The result that Strategies 1000 and 0011 are the only survivors in the long run is a first and important regularity of the model for this parameters' choice. Figure 2 is based on 100 simulations (of 10000 periods) and depicts the mean share of all the strategies (the box) equipped with standard deviations (equal to the length of the vertical line extending over the bars). On average, equilibrium is reached when about nine tenths of agents are prudent and the remaining ones are informed.

Quite remarkably, Figure 2 shows that when equilibrium is reached evolutionary pressure has obliterated all other strategic variations. Notice that, at

[4] The code is available on the website of the authors.

[5] All simulations are initialized setting the bits in **b** randomly in $\{0, 1\}$.

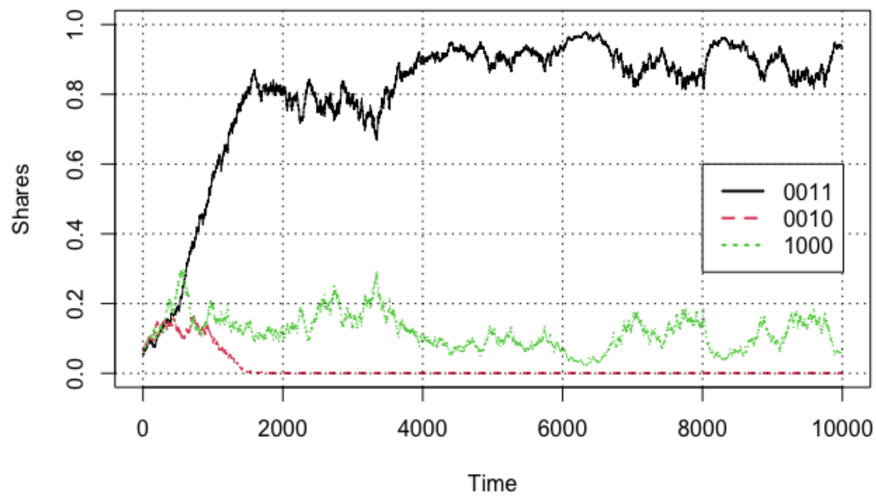

Fig. 1. Time series of the shares of strategies 0011 (prudent), 1000 (informed) and 0010 (uninformed) in a standard simulation run of the benchmark case lasting 10000 periods. In particular, $c = 0.03, v_\epsilon = 0.04$.

equilibrium, the probability that the profits of the informed are bigger than the ones of the prudent is 50% (this holds because, essentially, learning forces the surviving strategies to have the same median profits and, if this were not the case, the shares would have drifted away from that equilibrium in the presence of a tendency to prefer one of the two strategies).

This outcome suggests a novel explanation of the EPP from the bottom up. The large majority of prudent traders underinvest in the risky asset, being their demand particularly low as observed in the previous section. Indeed, this is due to a systematic overestimation of the variance of the gains of the stock that, in turn, reduces the demand of the risky asset and favors more conservative savings in the safe bond. Even if the prudent strategy reduces the average profit of investment, nonetheless the median revenues of the prudent are the same as the ones of the informed agents (at the end of each period, when consumption takes place).

Put differently, traders demanding small amounts of the risky asset become very popular in a market where they "compete" according to the (sharp) rule described in Subsec. 3.2 and occasionally compare their profits, achieving a reduction of the risk of being pushed out of the market in the long run. Such a majority of prudent traders fits very well the puzzling observation that fewer agents than expected invest in equities. Assuming an unrealistic level of sophistication, it could perhaps be argued that more sophisticated agents would maximize utility or realize that larger mean profits can be traded for the smaller

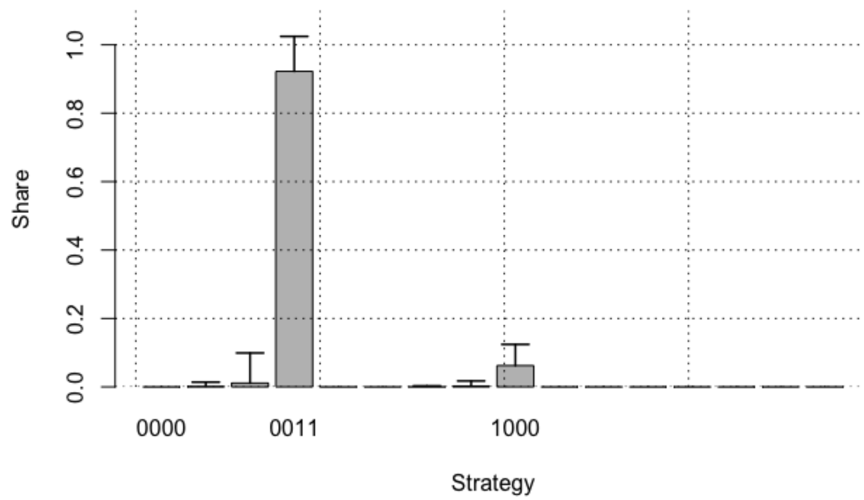

Fig. 2. Average shares and related standard deviation of different strategies in the benchmark case (with $c = 0.03, v_\epsilon = 0.04$). Mean values are shown based on 100 simulations of 10000 periods. Essentially, only strategies 0011 and 1000 survive.

gains obtained in about 50% of time. However, simple and straightforward comparisons based on the question "does $\mathbf{b}_i$ produce larger gains than $\mathbf{b}_j$?" are strong calls to immediate action and more convincing behavioral drivers.

3.2 Robustness tests

As expected, the outcomes described previously are sensitive to the parameters of the market. In this subsection, we explore the robustness of the results with respect to changes in the cost c of information and in the size v_ϵ of the unobservable and idiosyncratic shock. We use `BehaviorSpace`, a NetLogo's tool that allows to run experiments and gather data systematically "sweeping" (portions of) the parameters' space. Table 3 shows, for $c \in \{0.01, ..., 0.05\}$ and $v_\epsilon \in \{0.03, ..., 0.06\}$, the largest average share at equilibrium, based on 100 simulations of 10000 periods for each of the 20 couples (c, v_ϵ). For instance, corresponding to the benchmark parameters (boldfaced in Table 3), we see that the largest share (92%) is made of prudent investors (with $\mathbf{b}$ =0011).

Table 3 demonstrates that prudent investors dominate the scene in many cases, with shares varying from 37 to 93%. Therefore, to a great extent, quite some underinvestment in stocks is natural. Informed agents are prevalent at equilibrium only when the cost of information and the variance of the noise are low, in the top-left corner of the Table where $c = 0.01$ and $v_\epsilon = 0.03$.

Table 3. Modal strategy at equilibrium, for several values of parameters c and v_ϵ. The entry $s/\mathbf{b}$ means that the modal share s was reached by agents using strategy $\mathbf{b}$. The boldfaced entry is relative to the benchmark configuration and the italicized one is discussed in the text,

	Cost c				
v_ϵ	0.01	0.02	**0.03**	*0.04*	0.05
0.03	0.54/1000	0.66/0011	0.80/0011	0.92/0011	0.90/0011
0.04	0.50/0011	0.73/0011	**0.92/0011**	0.75/0011	0.34/0000
0.05	0.59/0011	0.84/0011	0.78/0011	0.37/0011	0.51/0000
0.06	0.63/0011	0.93/0011	0.50/0011	*0.51/0000*	0.61/0000

Interestingly, the bottom-right corner of Table 3 shows that the majority share at equilibrium is made of *fearless* agents, as they are nicknamed in Table 1. When both the cost of information and the variance of ϵ are large, a fraction of agents thrive with no use of signals θ or γ and avoiding any adjustment for the variance of the risky profits. Figure 3 depicts the equilibrium shares in one of such market instances, when $c = 0.04, v_\epsilon = 0.06$.

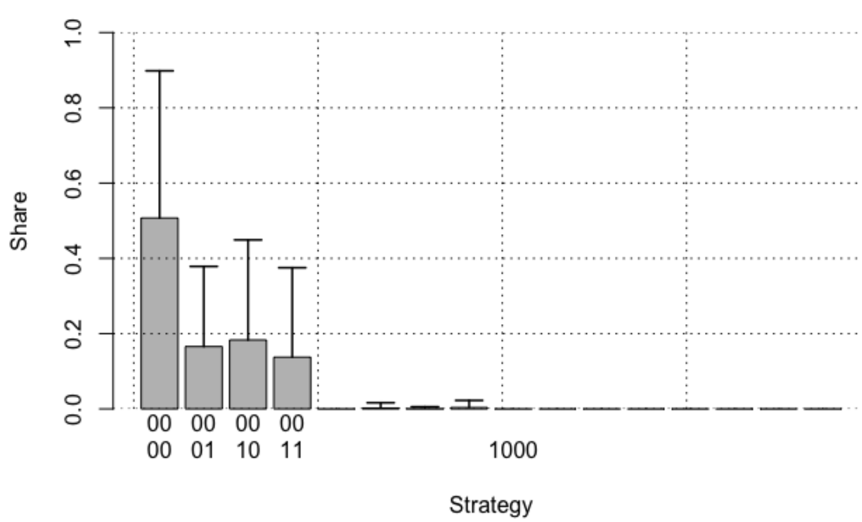

Fig. 3. Average shares and related standard deviation of different strategies when $c = 0.04, v_\epsilon = 0.06$). Mean values are based on 100 simulations of 10000 periods. Only Strategies 0000, 0001, 0010 and 0011 survive (with tiny exceptions), fearless traders are over 50% and very few agents use information, as discussed in the text.

Only strategies 0000, 0001, 0010, 0011 survive in the long run in this market where information is more costly and market returns are volatile. This is a plausible explanation of why Strategy 1000 died out but, truly, a reason for the observation that no survivor switch on bit b_1, that would imply information is bought and used, or b_2, that would imply misinformation is used (indeed, in this case we have that $b_1 = b_2 = 0$ for all agents, excluding a handful of outlying traders using 0110 or 0111, see the picture). Such a market is populated by many individuals who do not use any information (or misinformation, for what it matters), and who keep at 1 at least one of the bits b_3, b_4 located in the denominator of Eq. (2), reducing on average the quantity of stock kept in their equilibrium portfolio.

The main conclusion of our robustness test is that agents invest far less in the stock market than would be implied by a full-fledged (and probably unrealistic) model of rational allocation. This holds even under variations of several key parameters of the model (that, in some cases, lead altogether to the disappearance of information from the strategies).

4 Discussion and conclusion

One of the most interesting features of ABMs is their ability to accommodate for heterogenous features of the agents. We have considered a bunch of strategies that differ in the content of information that is used to assess expected profits (in the numerator of Eq. 2), as well as in the risk factors that are considered (in the denominator). Even though, in principle, agents could use any combination of active bits, evolutionary pressure wipes out most of the strategies. As pointed out by a reviewer, whether only one strategy survives in the very long run is still an open question (and 10000 periods may not be enough to reveal the steady state). Further research should also consider the effects of the introduction of mutation or the replacement of some agents with new ones. In any case, many strategies may be assumed to be of little importance asymptotically. In particular, in the benchmark case only informed and prudent traders survive; with the latter keeping a prominent position in many other situations. Prudent and fearless strategies do not use information θ and this results in a small number of information users, as shown in the solid line of Figure 4 depicting the share of traders whose $b_1 = 1$ at equilibrium for $v_\epsilon = 0.04$. A similar tendency is reported in Gerotto el al. [5] in a setup with only two strategies.

The frequency of $b_4 = 1$ at equilibrium is even more relevant to explain the EPP and can be associated to a majority of traders who buy limited amounts of stocks because their strategy inflates the perception of risk and reduce traded quantities. The dashed line in Figure 4 shows the share of users with $b_4 = 1$: with the exception of a few markets where fearless traders prevail, most agents are extremely cautious for all levels of cost when $v_\epsilon = 0.04$ (and this fraction is substantial and rarely falls below 50% in the many parametric combinations we have investigated in Table 3).

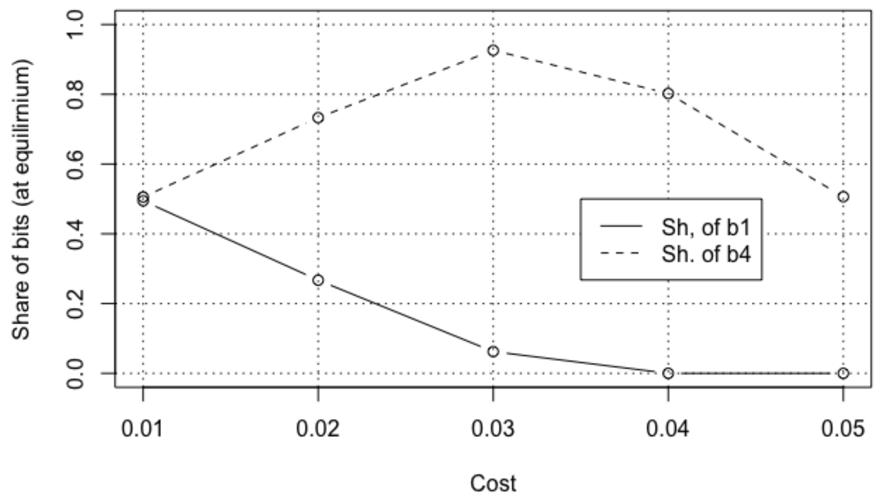

Fig. 4. Shares of agents with $b_1 = 1$ (solid line) and $b_4 = 1$ (dashed line), as a function of the cost c when $v_\epsilon = 0.04$. The former make some use of the information and the latter take misinformation γ as a risk factor.

Overall, our model suggest that the EPP stems, to some extent, from the evolutionary updating of strategies used by myopic traders, where the myopia mainly lies in the assumption that learning is performed with an eye on one-period performance only. While the introduction of long-term orientation may reduce the effect, the salience of recent rewards is well-documented, see Cosemans and Frehen [3] for a recent treatment, and can, in combination with strategy-switching, help in clarifying some of the issues raised by the EPP.

Acknowledgements We thank Giulia Iori for useful discussion. Two anonymous reviewers provided very useful suggestions but we are responsible for any error. Paolo Pellizzari is partially supported by ITN project EPOC "Economic Policies in Complex Environments" under the Marie Skłodowska-Curie grant agreement No 956107. The VERA Center of Ca' Foscari University of Venice funded Francesco Franz and Carlo Bellato who provided technical support. This work benefited from the inspiring FFSS facilities located in Trevignano-Signoressa (VE).

References

1. Barberis, N., Huang, M., Santos, T.: Prospect theory and asset prices. The Quarterly Journal of Economics 116(1), 1–53 (2001), http://www.jstor.org/stable/2696442

2. Benartzi, S., Thaler, R.: Myopic Loss Aversion and the Equity Premium Puzzle. The Quarterly Journal of Economics 110(1), 73–92 (1995)
3. Cosemans, M., Frehen, R.: Salience theory and stock prices: Empirical evidence. Journal of Financial Economics 140(2), 460–483 (2021), https://www.sciencedirect.com/science/article/pii/S0304405X20303524
4. DeLong, J.B., Magin, K.: The u.s. equity return premium: Past, present, and future. The Journal of Economic Perspectives 23(1), 193–208 (2009), http://www.jstor.org/stable/27648300
5. Gerotto, L., Pellizzari, P., Tolotti, M.: Asymmetric information and learning by imitation in agent-based financial markets. In: De La Prieta, F., González-Briones, A., Pawleski, P., Calvaresi, D., Del Val, E., Lopes, F., Julian, V., Osaba, E., Sánchez-Iborra, R. (eds.) Highlights of Practical Applications of Survivable Agents and Multi-Agent Systems. The PAAMS Collection. pp. 164–175. Springer International Publishing, Cham (2019)
6. Mehra, R., Prescott, E.C.: The equity premium: A puzzle. Journal of Monetary Economics 15(2), 145–161 (1985), https://www.sciencedirect.com/science/article/pii/0304393285900613
7. Newton, J.: Evolutionary game theory: A renaissance. Games 9(2) (2018), https://www.mdpi.com/2073-4336/9/2/31
8. Railsback, S., Grimm, V.: Agent-Based and Individual-Based Modeling: A Practical Introduction, Second Edition. Princeton University Press (2019), https://books.google.it/books?id=X3SYDwAAQBAJ
9. Robson, A.J., Orr, H.A.: Evolved attitudes to risk and the demand for equity. Proceedings of the National Academy of Sciences 118(26), e2015569118 (2021), https://www.pnas.org/doi/abs/10.1073/pnas.2015569118
10. Sandholm, W.H.: Population Games And Evolutionary Dynamics. Economic learning and social evolution, MIT Press (2010)
11. Steinbacher, M., Raddant, M., Karimi, F., Camacho Cuena, E., Alfarano, S., Iori, G., Lux, T.: Advances in the agent-based modeling of economic and social behavior. SN Business & Economics 1(7), 99 (2021)
12. Wilensky, U.: NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999)

Reply to reviewers

We thank the reviewers for many valuable comments and suggestions. All of their remarks have been considered and incorporated in the revised text, keeping into account that space is so severely constrained that it was impossible, in some cases, to expand the treatment as requested.

In the following we list the amendments to the paper.

We have improved Figure 1 to enhance visibility if printed in black/white or for the color-blind reader.

First reviewer

1. We fixed the typos;
2. We now better describe how "rationality" should be interpreted in the context of the paper: "Rationality, in this paper refers to the correct understanding of the data-generating process of D, and would be achieved in the model when...".

   Moreover, we explain why 0010 is rational: "Even if uninformed agents discard useful information (avoiding the cost), they correctly understand the way returns are generated and take into account the uncertainty arising from the unknown θ.";
3. We have removed references to the Dec notation all over the paper (text, figures and tables);
4. We specify in a footnote that "All simulations are initialized setting the bits in $\mathbf{b}$ randomly in $\{0, 1\}$"
5. Absorbing state: this is an open issue at the moment and we admit it in the text where we added the sentence

   > As pointed out by a reviewer, whether only one strategy survives in the very long run is still an open question (and 10000 periods may not be enough to reveal the steady state). Further research should also consider the effects of the introduction of mutation or the replacement of some agents with new ones. In any case, many strategies may be assumed to be of little importance asymptotically.

Second reviewer

1. We fixed the typos;
2. We more carefully describe the puzzle,"Economists have tried to single out ways to explain why people invest far less in stocks that would be implied by their risk aversion, as measured in other (personal or social) situations.", and added an explanatory quote taken from Benartzi and Thaler (1995);

3. Simulations: we increase the number of the simulations to 100 for each configuration. All in all, we now run 2000 simulations of 10000 trading periods and Table 3 contains the exploration of 20 different parametric constellations. Results appear to be significant and quite robust to large variations of the level of noise in the market and of the cost of information.
   In the interest of space, we do not include standard deviations in the results. However, Figure 1 allows for visual inspection of the typical variations in shares; more importantly, both Figure 2 and 3 display the size of the standard deviations as vertical lines at the top of the bars depicting the average values (this feature was already present in the first version).
4. As requested, we mentioned in the text that the code can be downloaded from the websites of the authors.