INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

WORKING PAPER SERIES

Strategically biased learning in market interactions

Giulio Bottazzi [a]
Daniele Giachini [a]

[a] Institute of Economics and EMbeDS Department, Scuola Superiore Sant'Anna, Pisa, Italy.

2022/02 January 2022
ISSN(ONLINE) 2284-0400

Strategically biased learning in market interactions

Giulio Bottazzi[1] and Daniele Giachini[1]

[1]Institute of Economics & Department EMbeDS, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56127 Pisa - Italy

January 15, 2022

Abstract

We consider a market economy where two rational agents are able to learn the distribution of future events. In this context, we study whether moving away from the standard Bayesian belief updating, in the sense of under-reaction to some degree to new information, may be strategically convenient for traders. We show that, in equilibrium, strong under-reaction occurs, thus rational agents may strategically want to bias their learning process. Our analysis points out that the underlying mechanism driving ex-ante strategical decisions is diversity seeking. Finally, we show that, even if robust with respect to strategy selection, strong under-reaction can generate low realized welfare levels because of a long transient phase in which the agent makes poor predictions.

JEL Classification: C60, D53, D81, D83, G11, G12

Keywords: Learning, Strategic interaction, Behavioral Bias, Financial Markets

1 Introduction

In this paper we study whether non-bayesian belief updating can persist in financial markets when agents can strategically choose the degree of bias in their learning process. We focus on a particular type of behavioral bias that has been widely studied in the literature, both theoretical and empirical: *under-reaction* (see e.g. Barberis et al., 1998; Epstein et al., 2010; Massari, 2020; Jiang et al., 2021). Indeed, Bayesian learning – i.e. updating beliefs according to Bayes rule – is commonly identified as the "rational" way of incorporating new evidence into probabilistic predictions: under suitable conditions it makes beliefs eventually accurate (Epstein et al., 2010). According to this analysis, thus, deviations from the Bayesian benchmark, such as under-reaction, have to be considered biases. The conditions for the rationality of Bayesian updating include having the true data generating process in the prior set of models over which the agent is learning or, equivalently, assigning a non negative initial prior probability to the true model. If this is not the case, we face model misspecification and, in such a situation, Bayesian learning generically lets the agent predict according to the most accurate model in its set (Berk, 1966). Thus, even if not perfect, Bayesian learning still provides a reasonably good accuracy outcome and can be considered the benchmark also in misspecified settings.

In complex environments – *large worlds*, using the language of Savage (1954) as reported by Gigerenzer and Gaissmaier (2011) – being able to learn the correct data generating process appears unrealistic and model misspecification should be considered as a defining characteristic. Under model misspecification, a trader who under-reacts to information cannot be driven out of the market by a trader that learns in a Bayesian fashion (Massari, 2020). Thus, under-reaction provides an evolutionary advantage over Bayesian learning in complex environments characterized by model misspecification. A straightforward consequence is that a Bayesian trader who wanted to increase its prospects to survive the market selection struggle should bias its learning process toward under-reaction. Still, is deliberately biasing learning somehow *rational* for the trader? Jouini and Napp (2016) show that market selection outcome and individual welfare may be negatively related: vanishing agents can experience larger expected utility than dominating ones. Hence, it is not clear whether the advantage in selection under-reaction has over Bayesian learning in misspecified contexts translates in larger expected utility. Thus, supposing that agents could strategically choose how much under-reaction show in their belief updating, would they bias their learning process? In this paper we answer these questions considering a standard economy where agents trade intertemporally. Since the focus of this paper is on strategically biasing learning, we sterilize other possible factors that may affect the choices – and, thus, the welfare – of an agent assuming that, apart from learning protocols, agents are ex-ante identical. In our setting this translates in assuming an homogeneous discount factor, logarithmic instantaneous utility, and constant and homogeneous endowments. Moreover, we focus on the case of two agents, such that the strategic environment is streamlined. Strategic interaction is introduced letting traders choose (ex-ante) their degree of under-reaction. Hence, before engaging in trading, agents play a simultaneous game where they evaluate their prospects and select the degree of under-reaction that provides the larger welfare in terms of expected utility.

Our results show that in any pure-strategy Nash equilibrium there is a trader that shows a high degree of under-reaction and a trader that learns in a Bayesian manner. In the mixed-strategy Nash equilibrium the agents are Bayesian or strongly under-react with positive and almost equal probabilities. Thus, under-reaction emerges and persists even in misspecified contexts where agents can strategically choose how much biased their learning process has to be. At the same time, it pays to be "irrational" – i.e. non Bayesian – only when the opponent is "rational" – i.e. Bayesian – and vice-versa. That is, the strategical advantage under-reaction may have stems from the *gains of trade* that can only be obtained when agents' beliefs are sufficiently diverse. Finally, we investigate agents' realized utilities assuming a specific exogenous "true" process driving the state of nature. Our analysis shows that the far away the true process is from the initial prior of the agents, and the larger is their degree of under-reaction, the lower their realized utility. The reason of this finding is the presence of a long transient phase in which the predictions of a strong under-reaction agent are very inaccurate. Thus, even if under-reaction proves to be robust both in evolutionary and strategic sense when facing model misspecification, it may be quite detrimental in terms of level of welfare eventually realized by the trader.

Highlighting strategic and welfare issues that may occur in economies populated by logarithmic traders,[1] this paper complements the analyses of evolutionary dynamics performed in Bottazzi and Dindo (2013, 2014), Bottazzi and Giachini (2017, 2019a,b) Bottazzi et al. (2018, 2019), Dindo and Massari (2020), Giachini (2021). Indeed, exploiting the selection results in those contributions, our paper paves the way to further explorations on the link between heuristic behavioral rules, evolutionary dynamics, and strategic choices.

2 The Model

Consider a discrete-time pure exchange economy with an infinite horizon. At each time $t \geq 0$ the economy can be in one of S finite states. We denote with 0 the

[1] Notice that, in many of the aforementioned contributions, agents use behavioral rules. Nonetheless, those rules can be rationalized using the arguments reported in Bottazzi et al. (2018), Dindo (2019), and Giachini (2021)

certain state of the economy at date 0. Let $s_t \in [S] = \{1, 2, \ldots, S\}$ be the state at time $t > 0$, $\sigma_t = (s_1, s_2, \ldots, s_t) \in \Sigma_t = [S]^t$ a partial history of the economy until time t (a sequence of t elements of the set $[S]$) and with $\sigma = (s_1, s_2, \ldots, s_t, \ldots) \in \Sigma$ a generic infinite history. The economy is populated by N agents that want to satisfy their consumption need. At each date t a single good is available and each agent $i \in [N]$ receives an amount $e_i(\sigma_t) > 0$ of that good. Let $c_i(\sigma_t)$ be the consumption of agent i on the partial history σ_t. Agent i optimal consumption is found solving

$$\begin{aligned} \max_{\{c_i(\sigma_t),\, \forall t, \sigma\}} U_i =& (1-\beta_i) \sum_{t=0}^{\infty} \sum_{\sigma_t \in \sigma_t} \beta_i^t p_i(\sigma_t) \log(c_i(\sigma_t)/e_i(\sigma_t)) \\ \text{subject to} \quad & \sum_{t=0}^{\infty} \sum_{\sigma_t \in \sigma_t} q(\sigma_t)\,(e_i(\sigma_t) - c_i(\sigma_t)) \geq 0, \end{aligned} \tag{1}$$

where $\beta_i \in (0, 1)$ is agent i's discount factor and $q(\sigma_t)$ is the price of the consumption good at date t if σ_t is realized. All agents adopt the same, logarithmic, Bernoulli utility to judge the attained level of consumption at eact time step. In turn, the price of the consumption good contingent on the realization of a given partial history σ_t is set by the market clearing condition

$$\sum_{i=1}^{N} c_i(\sigma_t) = \sum_{i=1}^{N} e_i(\sigma_t) = e(\sigma_t)\ . \tag{2}$$

We assume that all agents receive the same constant endowment: $e_i(\sigma_t) = e > 0$ for each $i \in [N]$ and for all σ_t. Under these assumptions we can solve (details are in Appendix 1) for agents equilibrium consumption

$$c_i(\sigma_t) = \frac{(1-\beta_i)\,\beta_i^t\, p_i(\sigma_t)}{\sum_{j=1}^{N}(1-\beta_j)\,\beta_j^t\, p_j(\sigma_t)}\, N\, e, \tag{3}$$

and consumption good prices

$$q(\sigma_t) = \frac{\sum_{i=1}^{N}(1-\beta_i)\,\beta_i^t\, p_i(\sigma_t)}{\sum_{i=1}^{N}(1-\beta_i)}. \tag{4}$$

The assumption of identical Bernoulli utilities and endowments make possible the comparison of the utility levels attained by agents at equilibrium. Indeed for each agent i, at equilibrium, it is $0 \leq U_i \leq \log(N)$. The first inequality comes from the observation that consuming the endowment is feasible for each agent, and it entails zero utility. Since the individual consumption choice derives from an optimization, at equilibrium it must be $U_i \geq 0$. At the same time, the best possible outcome for each agent is consuming the whole aggregate endowments, thus it must be $U_i \leq \log N$. Substituting the consumption (3) in the agent's utility one obtains the expected utility at equilibrium

$$U_i = (1-\beta_i) \sum_{t=0}^{\infty} \sum_{\sigma_t \in \sigma_t} \beta_i^t \, p_i(\sigma_t) \, \log \left(\frac{(1-\beta_i) \, \beta_i^t \, p_i(\sigma_t)}{\sum_{j=1}^N (1-\beta_j) \, \beta_j^t \, p_j(\sigma_t)/N} \right). \tag{5}$$

If agents share the same probabilistic model p and have the same discount factor β, we are in a no-trade situation: all agents consume their endowment and $U_i = 0$ for any $i \in [N]$. This is the worst case scenario for all agents. More generally, notice that the quantities $P_i(\sigma_t) = (1-\beta_i) \, \beta_i^t \, p_i(\sigma_t)$ define a probability measure on the space of all histories, indeed $P_i(\sigma_t) > 0$ for any σ_t and $\lim_{T \to +\infty} \sum_{t=0}^T \sum_{\sigma_t \in \Sigma_t} P_i(\sigma_t) = 1$. Defining the average population probability measure starting from agents' individual measures $\bar{P} = \sum_{i=1}^N P_i/N$, the expected utility can be rewritten in terms of the Kullback-Leibler divergence of the agent's individual measure with respect to the population average measure

$$U_i = D_{KL}\left(P_i \mid \bar{P}\right) = \lim_{T \to +\infty} \sum_{t=0}^{T} \sum_{\sigma_t \in \Sigma_t} P_i(\sigma_t) \log \left(\frac{P_i(\sigma_t)}{\bar{P}(\sigma_t)} \right) .$$

Because, in general, $D_{KL}\left(P_i \mid \bar{P}\right) \neq D_{KL}\left(P_j \mid \bar{P}\right)$, the gain from trade is different for different agents, despite the fact that they have the same endowment and the same Bernoulli utility. Agents with an individual measure diverging more from the average are those with a higher expected utility. This consideration introduce a strategic dimension into the usual GE framework: given the ecology of agents present in the market, there are individual measures (and intertemporal discount factors) that assure a higher gain from trade. In the next sections we will investigate the relative performance, in terms of utility, of agents adopting different probabilistic models and we will explore which models are characterized by the best performance in an heterogeneous framework. For definiteness, we will focus on the role of individual probabilities and we will assume that agents share the same intertemporal discount factor $\beta_i = \beta$ for all $i \in [N]$. In this way, agent's preferences are identical and they only differ with respect to the probabilities they assign to the different possible futures. There is however a practical question to address before.

2.1 Estimation strategy

Even for relatively simple models, the analytical computation of the expected utility in (5) can result impossible. Thus, we shall rely upon numerical estimates. The sum defining U_i is infinite but in the case of homogeneous discount factor we can derive an upper bound to the error incurred in truncating it. Let $U_i(T)$ be the partial sum of the first T terms in (5), the truncation error reads

$$\varepsilon_{i,T} = |U_i - U_{i,T}| \leq (1-\beta) \sum_{t=T+1}^{\infty} \sum_{\sigma_t \in \sigma_t} \beta^t p_i(\sigma_t) \left| \log \left(\frac{p_i(\sigma_t)}{\sum_{j=1}^{N} p_j(\sigma_t)/N} \right) \right| .$$

Assume there exists a $\underline{\pi}$ and an $\overline{\pi}$ such that $\underline{\pi}^t \leq p_i(\sigma_t) \leq \overline{\pi}^t$ for any i and any σ_t. This assumption is compatible with a large class of probabilistic models, as it is sufficient to assume that the conditional probabilities are uniformly bounded away from zero and one. Then $|\log(p_i(\sigma_t)/(\sum_{j=1}^{N} p_j(\sigma_t)/N))| \leq t \log \overline{\pi}/\underline{\pi}$. Substituting in the previous equation and summing the geometric series one gets

$$\varepsilon_{i,T} \leq \beta^{T+1} \frac{T(1-\beta)+1}{1-\beta} \log \frac{\overline{\pi}}{\underline{\pi}} \sim T \frac{\beta^{T+1}}{1-\beta} \log \frac{\overline{\pi}}{\underline{\pi}} . \qquad (6)$$

If the discount factor is sufficiently small, one can obtain a good approximation also with relatively small T. However, the number of realizations that must be computed grows as S^T, making the task daunting even for a moderate value of T. We exploit the fact that $U_i(T)$ involve an expectation, hence we can easily provide estimates of those expectations by means of Monte Carlo approach. We generate a sample of M independent sequences of realizations $\sigma_{T,r} = (s_1, s_2, \ldots, s_T)$ according to the probability measure p_i, we compute the quantity

$$u_{i,r}(T) = (1-\beta) \sum_{t=0}^{T} \beta^t \log \left(\frac{p_i(\sigma_{t,r})}{\sum_{j=1}^{N} p_j(\sigma_{t,r})/N} \right) \qquad (7)$$

on each replica r, obtaining a weighted sample of M independent replicas of agent i (truncated) utility. Then we estimate agent i expected utility as $\hat{U}^i(T) = \sum_{r=1}^{M} u_{i,r}(T)/M$ and the associated standard error dividing the standard deviation of the sample of (7) by $\sqrt{M}$.

The same procedure can be applied to measure the realized utility, with respect to the "true" model, simply replacing the measure p_i used in obtaining the sequences of realizations $(s_1, s_2, \ldots, s_T)$ with the measure implied by the actual process driving the states of the economy. To avoid confusion, we shall call $\hat{R}^i(T)$ the estimated realized utility of agent i.

3 Numerical analysis

We investigate the strategic dimension of model selection in GE utility maximization by performing theree exercises. In the next Section, we coompare the performance of agents with fixed models. This provides a benchmark for the following

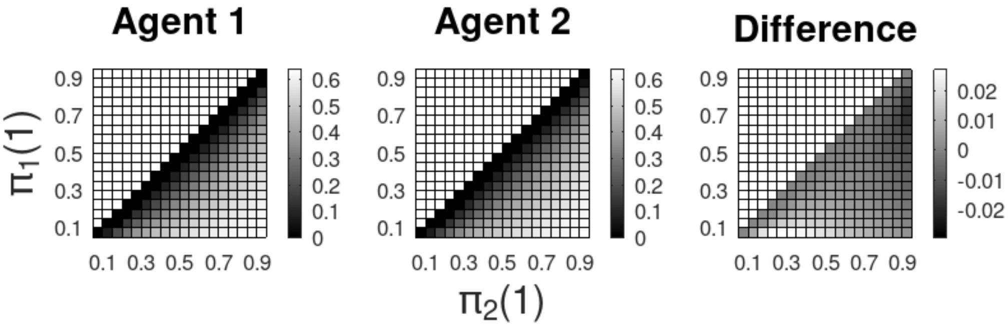

Figure 1: **Left**: $\hat{U}^1(T)$ for different combinations of $\pi_1(1)$ and $\pi_2(1)$. **Center**: $\hat{U}^2(T)$ for different combinations of $\pi_1(1)$ and $\pi_2(1)$. **Right**: $\hat{U}^1(T) - \hat{U}^2(T)$ for different combinations of $\pi_1(1)$ and $\pi_2(1)$. For each estimate we set $\beta = 0.95$, $T = 500$, and $M = 10^5$. Standard errors are $\propto 10^{-3}$ or smaller. The truncation error is $\varepsilon_{i,T} \leq 2.0348 \times 10^{-7}\ \forall i$.

analysis and will allow to identify the aspects of the probabilistic model that most contribute to the individual performances of the different agents. Section 3.2 contains our main exercise. We set up a simultaneous game in which players have the option to select their learning behavior and analyze the strategic implication of the different choices, identifying the game equilibria. Finally, in Section 3.3, we compare the expected utility levels attainded by various strategic choices with the realized ones.

3.1 Fixed models

For simplicity, consider the $S = 2$ and $N = 2$ case. Assumes agent have i.i.d. models: agent $i = 1, 2$ assign a constant probability $\pi_i(s)$, for $s = 1, 2$, to the realization of state s at each date, $p_i(s_t|\sigma_{\tau-1}) = \pi_i(s_t)$. Thus if t_1 is the number of time state 1 has been realized on the history σ_t, it is

$$p_i(\sigma_t) = \prod_{\tau=1}^{t} p_i(s_\tau|\sigma_{\tau-1}) = \pi_i(1)^n \pi_i(2)^{t-n} = \pi_i(1)^n (1 - \pi_i(1))^{t-n} .$$

We apply the numerical procedure described in Section 2.1 to estimate the agents' expected utilities at equilibrium for different values of $\pi_1(1)$ and $\pi_2(1)$. Results are reported in Figure 1, that shows the estimated utilities of the agents and the estimated difference between the two utilities. Notice that by moving towards the bottom right corner, agents' excess utility increases. Thus, agents experience the largest (expected) welfare when they have very different beliefs. This derives from

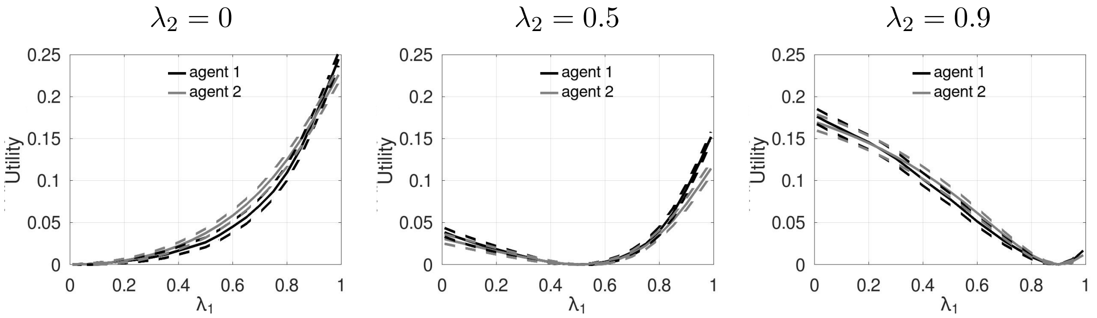

Figure 2: $\hat{U}^1(T)$ and $\hat{U}^2(T)$ as functions of λ_1 for $\lambda_2 \in \{0, 0.5, 0.9\}$. For each estimate we set $\pi_1(1) = 0.3$, $\pi_2(1) = 0.8$, $\beta = 0.95$, $T = 500$, and $M = 10000$. Confidence bands are computed adding and subtracting three times the standard errors. The truncation error is $\varepsilon_{i,T} \leq 9.5804 \times 10^{-8}\ \forall i$.

the fact that having nearly opposite evaluations about future states maximizes the benefit from trading. Indeed, agents are eager to move consumption from those states they believe less likely to those they believe more likely and the opposite evaluations allow them to trade at supposedly inexpensive prices. As expected, however, the gains are, in general, not evenly distributed. This is made apparent the right panel of Figure 1, where we notice significant differences in expected utility when one agent has rather extreme beliefs while the other shows milder evaluations.

The analysis of the case of fixed models confirm the importance of beliefs for (*ex-ante*) welfare evaluation. Beliefs do not have intrinsic merits, but they perform differently depending on the entire "ecology" of beliefs, that is on the models of the other agents in the model. The simplest possible model was analyzed here, however increasing the number of agents N or of states of nature S does not change the general picture: learning, that is the systematic way in which information is incorporated in beliefs, can be crucial in shaping the ex-ante welfare considerations of rational traders.

3.2 Learning Processes

In this section we introduce a strategic dimension by defining a family of learning models and allow agents to select among them. We assume that agents have access to a set of K i.i.d. models $\pi_1, \ldots, \pi_K$. These models are defined analogously to the models of the previous section. Now assume that instead of keeping fixed their probabilistic model, agents try to learn the true one using all the K existing models. That is, agents' individual beliefs derive from a learning process. Agents' subjective conditional probabilities are build as a convex combination of models' predictions, in formal terms

$$p_i(s_t|\sigma_{t-1}) = \sum_{k=1}^{K} \pi_k(s_t) w_{i,k}(\sigma_{t-1}) \quad \forall t, \sigma\,, \tag{8}$$

with $w_{i,k}(\sigma_{t-1})$ denoting the weight agent i attaches to model k after having observed the partial history σ_{t-1}. Without loss of generality, we set $w_{i,k}(\sigma_0) = 1/K$ $\forall i, k$. Those weights evolve according to a generalization of Bayes rules that captures the notion of under-reaction as proposed in Epstein et al. (2010) and Massari (2020). Hence, one has

$$w_{i,k}(\sigma_{t+1}) = \lambda_i w_{i,k}(\sigma_t) + (1 - \lambda_i) \frac{\pi_k(s_{t+1}|\sigma_t) w_{i,k}(\sigma_t)}{p_i(s_{t+1}|\sigma_t)} \quad \forall k, t, \sigma\,, \tag{9}$$

with $\lambda_i \in [0, 1)$. Notice that setting $\lambda_i = 0$ one recovers Bayesian learning. Thus, such a learning protocol can be considered a form of "moderate" Baysian learning where the probability attached to the event "model k is the true one" in obtained taking a convex combination of Bayes rule with the prior probability. As in the previous section, we shall focus on the case $K = S = N = 2$.

Following the procedure in Section 2.1, we estimate the agents' utilities for different values of λ_1 and λ_2. In Figure 2 we report those estimates as functions of λ_1 and for $\lambda_2 \in \{0, 0.5, 0.9\}$. The case $\lambda_2 = 0$ identifies the setting of Massari (2020) in which an under-reacting agent is competing in the market with a Bayesian learner. In this case, increasing the value of λ_1 appears beneficial for both agents. The reason is again beliefs diversity. Indeed, a Bayesian converges rather quickly to the model with lower relative entropy with respect to true probability measure, while a under-reacting agent persistently mixes the two models (Massari, 2020). The higher the value of λ_1, the more the models are mixed. This creates opportunities for trade, supposedly convenient prices, and, as a consequence, higher expected utility for both agents. In this case, if agent 1 had the possibility of choosing its degree of under-reaction in an optimal way, it would set λ_1 to a value the closest possible to 1. The positive effects, both at the individual level and for the whole economy, of larger degrees of under-reaction by agent 1 emerge also in the case $\lambda_2 = 0.5$. When, instead, $\lambda_2 = 0.9$, one notices that lower levels of under-reaction by agent 1 appear preferable for both agents. This is the situation opposite to the one seen before: when agent 2 strongly under-reacts, the largest diversity – and, as a consequence, the largest expected welfare – occurs if we let agent 1 become Bayesian. Thus, in this case, if agent 1 had the possibility to choose its degree of under-reaction, it would set λ_1 to zero.

A straightforward *strategic* argument emerges from the previous considerations: suppose agents can strategically and optimally choose their degree of under-reaction, would such learning bias persist in equilibrium? That is, is biasing belief

		Agent 2		
		$\lambda_2 = 0$	$\lambda_2 = 0.5$	$\lambda_2 = 0.9$
	$\lambda_1 = 0$	$(0\,,0)$	$(\mathcal{U}_{0,0.5}\,,\mathcal{U}_{0.5,0})$	$(\mathcal{U}_{0,0.9}\,,\mathcal{U}_{0.9,0})$
Agent 1	$\lambda_1 = 0.5$	$(\mathcal{U}_{0.5,0}\,,\mathcal{U}_{0,0.5})$	$(0\,,0)$	$(\mathcal{U}_{0.5,0.9}\,,\mathcal{U}_{0.9,0.5})$
	$\lambda_1 = 0.9$	$(\mathcal{U}_{0.9,0}\,,\mathcal{U}_{0,0.9})$	$(\mathcal{U}_{0.9,0.5}\,,\mathcal{U}_{0.5,0.9})$	$(0\,,0)$

Table 1: Simultaneous game representation of the strategic choice of the degree of under-reaction among three given levels. For each entry, the first number in parenthesis represents the payoff of agent 1, while the second is the payoff of agent 2.

updating (for instance, moving away from Bayesian learning) rational in a strategic setting? To answer those questions, we exemplify our argument as follows. Suppose each agent $i \in \{1,2\}$ chooses $\lambda_i \in \{0, 0.5, 0.9\}$ such that to maximize its payoff in a simultaneous strategic interaction setting. Define $\mathcal{U}_i(\lambda_1, \lambda_2)$ as the payoff agent i gets as a function of the degrees of under-reaction of the two agents. We assume that agents' payoffs match agents' expected utilities as in (5) and we can safely and conveniently use the estimated values reported in Figure 2 to solve the game. Hence, $\mathcal{U}_1(x,y) = \mathcal{U}_2(y,x)$ and we can simplify the notation setting $\mathcal{U}_{x,y} = \mathcal{U}_1(x,y)$ and $\mathcal{U}_{x,x} = 0\ \forall x \in \{0, 0.5, 0.9\}$. The simultaneous game that emerges can be represented as in Table 1. The following inequalities hold:

$$0 < \mathcal{U}_{0.5,0} < \mathcal{U}_{0.9,0}\,;\ 0 < \mathcal{U}_{0,0.5} < \mathcal{U}_{0.9,0.5}\,;\ 0 < \mathcal{U}_{0.5,0.9} < \mathcal{U}_{0,0.9}\,.$$

By direct inspection of Table 1, one realizes that there exist two pure strategy asymmetric Nash equilibria, both characterized by one agent strongly under-reacting while the other learns in a Bayesian way. To investigate the existence of Nash equilibria in mixed strategies, we repeat our estimation exercise for $\lambda_1, \lambda_2 \in \{0, 0.5, 0.9\}$ considering $M = 10^6$ independent replications, such that we obtain standard errors in the order of 10^{-4} and, thus, more accurate estimates. We report our results in Table 2.

Focus, without loss of generality because of symmetry, on agent 1 and notice that a mixed strategy that prescribes to play $\lambda_1 = 0$ with probability 0.5 and $\lambda_1 = 0.9$ with probability 0.5 provides a strictly larger expected payoff than the strategy $\lambda_1 = 0.5$ for any possible choice of agent 2. Thus, $\lambda_1 = 0.5$ is dominated in mixed strategies and will not be played with positive probability in equilibrium. Hence, defining θ_i as the probability agent i attaches to play the strategy $\lambda_i = 0$, the unique mixed strategy equilibrium can be found solving the indifference condition $(1-\theta_i)\mathcal{U}_{0,0.9} = \theta_i \mathcal{U}_{0.9,0}$, that is

$$\theta_i = \frac{\mathcal{U}_{0,0.9}}{\mathcal{U}_{0,0.9} + \mathcal{U}_{0.9,0}}\,.$$

Agent 2

Agent 1		$\lambda_2 = 0$	$\lambda_2 = 0.5$	$\lambda_2 = 0.9$
	$\lambda_1 = 0$	$(0\,,0)$	$(0.0386\,,0.0289)$	$(0.1766\,,0.1705)$
	$\lambda_1 = 0.5$	$(0.0289\,,0.0386)$	$(0\,,0)$	$(0.0761\,,0.0868)$
	$\lambda_1 = 0.9$	$(0.1705\,,0.1766)$	$(0.0868\,,0.0761)$	$(0\,,0)$

Table 2: Simultaneous game representation of the strategic choice of the degree of under-reaction among three given levels with estimated payoffs. For each estimate we set $\pi_1(1) = 0.3$, $\pi_2(1) = 0.8$, $\beta = 0.95$, $T = 500$, and $M = 10^6$. Estimates are rounded to the 4th decimal digit and standard errors $\propto 10^{-4}$ are not reported.

Directly substituting with the estimated values, one gets $\theta_i = 0.5088$. We also provide an interval estimate by means of a Monte Carlo exercise. It is obtained repeating for 500 times the estimation of expected utilities (setting $M = 10^4$) and computing θ_i for each one of the 500 independent replicas. Our computations show that $\theta_i = 0.5092 \pm 0.0007$ with 99% statistical confidence.

The indications one can get from our game theoretic exercise are that in any pure-strategy equilibrium, there exists a single agent that strongly under-reacts, while in the mixed-strategy equilibrium the agents play Bayes or strong under-reaction with positive probabilities. Interestingly, the agents tend to play Bayes with slightly higher probability than strong under-reaction.

Overall, again, the largest benefits are generated by heavily differentiating from the opponent and this, together with the symmetric setting, shapes the strategic environment. Strategically biasing beliefs may be optimal because it allows one to differentiate from the opponent. Coupling our findings with those of Massari (2020), we can conclude that under-reaction is not only robust in an evolutionary perspective, but it is robust even in a strategical sense. Rational agents may optimally choose to show such a bias in equilibrium and market dynamics can fail to select against it. Thus, under-reaction emerges, persists, and can influences economic dynamics in the long-run.

Notice that our results and argument hold for any possible choice of the true data generating process. Indeed, so far, no assumption has been made on the true probability measure p governing the prcess of states of nature. This is because ex-ante expected utilities are based on subjective probabilistic evaluations. Thus, in principle, one does not suffer any ex-ante penalty in generating extremely biased beliefs.

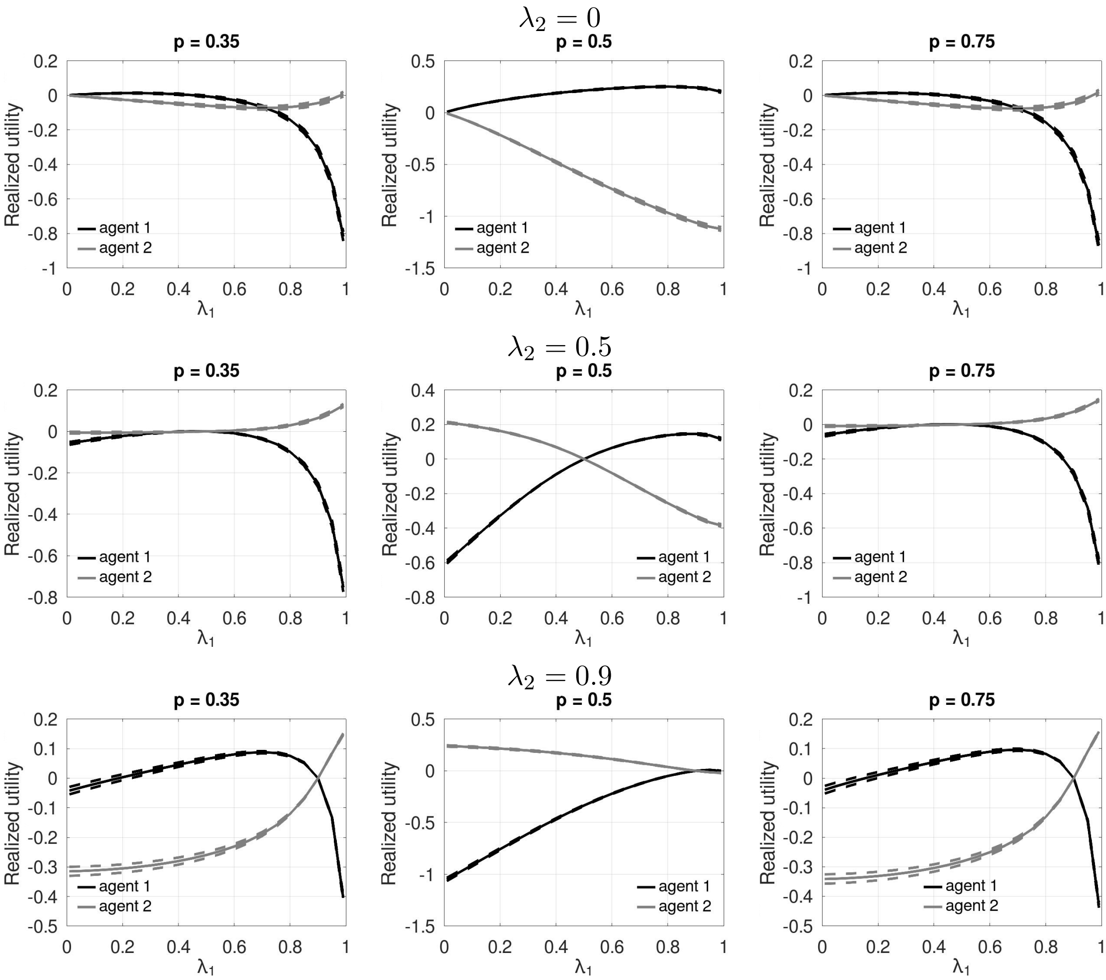

Figure 3: $\hat{R}^1(T)$ and $\hat{R}^2(T)$ as functions of λ_1 for $\lambda_2 \in \{0, 0.5, 0.9\}$ and $p \in \{0.35, 0.5, 0.75\}$. For each estimate we set $\pi_1(1) = 0.3$, $\pi_2(1) = 0.8$, $\beta = 0.95$, $T = 500$, and $M = 10^4$. Confidence bands are computed adding and subtracting three times the standard errors. The truncation error is $\varepsilon_{i,T} \leq 9.5804 \times 10^{-8}$ $\forall i$.

3.3 Realized utility

Now we try to assess whether biasing beliefs, even if strategically convenient, may become penalizing when facing true sequences of events. To do that we study the average *realized* utility, i.e. the utility level that, on average, agents receive on a sequence of events generated by the true process. We follow the numerical procedure outlined in advance, but, in contrast with the previous exercises, here we have to make assumptions on the true data generating process. In particular, we assume that the truth p is i.i.d. and, with a little abuse of notation, p shall be used to refer to the probability of observing state 1. For our numerical exercise, we choose $p \in \{0.35, 0.5, 0.75\}$.

In Figure 3 we report the results of our exercise and, as one can notice, when $p = 0.35$ or $p = 0.75$ large values of λ_1 appear detrimental for agent 1. This is a consequence of the long *transient* that high degrees of under-reaction entail: a λ_1 close to one means that asymptotically agent 1's beliefs are almost correct, but the speed of convergence is extremely low.[2] Thus, in the first periods agent 2 (with $\lambda_2 < \lambda_1$) gets close to the truth faster and takes advantage of that consuming more and enjoying larger utility, on average. Notice that the uninformative prior we assumed is crucial in driving such outcome. Indeed, when $p = 0.5$, initial beliefs are much closer to the truth than in the previous cases, thus the transient advantage of the agent with low degree of under-reaction quickly disappears. It follows that $\lambda_1 > \lambda_2$ becomes beneficial for agent 1 and detrimental for agent 2.

This exercise shows that ex-ante and ex-post (realized) levels of welfare may be very different. Thus, under-reaction, even if evolutionary and strategically robust, does not ensure that the agent showing it effectively experiences high levels of welfare. Obviously, this strongly depends on the true data generating process and different instances of model misspecification may severely affect the resulting welfare outcomes.

4 Conclusion

We investigate whether a seemingly irrational behavior – such as moving away from Bayesian updating via under-reaction – persists when agents competing in a market economy can strategically choose to bias their learning process. We find that in any pure-strategy Nash equilibrium a trader strongly under-reacts while the other follows Bayes rule. In the mixed-strategy equilibrium the agents randomize between Bayes and strong under-reaction giving slightly higher probability to Bayes. Thus, under-reaction does not disappear when agents can strategically choose their learning process. The underlying mechanism that drives our result is that agents seek to make their probabilistic evaluations diverse in order to generate gains from trade. Since strategic biasing involve expected utility evaluations, diversity seeking may let agents generate inaccurate beliefs. We investigate that estimating agents' average realized utility and our results show that strong under-reaction may cause low (realized) welfare because of a long transient phase in which the agent makes poor probabilistic evaluations.

[2] The statement is a straightforward consequence of the results in Bottazzi and Giachini (2017, 2019b) and Dindo and Massari (2020) coupled with the interpretation of market economies as learning algorithms provided by Massari (2021, 2020) and Bottazzi and Giachini (2019b).

References

Barberis, N., A. Shleifer, and R. Vishny (1998). A model of investor sentiment. *Journal of financial economics 49*(3), 307–343.

Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *The Annals of Mathematical Statistics 37*(1), 51–58.

Bottazzi, G. and P. Dindo (2013). Selection in asset markets: the good, the bad, and the unknown. *Journal of Evolutionary Economics 23*, 641–661.

Bottazzi, G. and P. Dindo (2014). Evolution and market behavior with endogenous investment rules. *Journal of Economic Dynamics and Control 48*, 121–146.

Bottazzi, G., P. Dindo, and D. Giachini (2018). Long-run heterogeneity in an exchange economy with fixed-mix traders. *Economic Theory 66*(2), 407–447.

Bottazzi, G., P. Dindo, and D. Giachini (2019). Momentum and reversal in financial markets with persistent heterogeneity. *Annals of Finance 15*(4), 455–487.

Bottazzi, G. and D. Giachini (2017). Wealth and price distribution by diffusive pproximation in a repeated prediction market. *Physica A: Statistical Mechanics and its Applications 471*, 473–479.

Bottazzi, G. and D. Giachini (2019a). Betting, selection, and luck: A long-run analysis of repeated betting markets. *Entropy 21*(6), 585.

Bottazzi, G. and D. Giachini (2019b). Far from the madding crowd: Collective wisdom in prediction markets. *Quantitative Finance 19*(9), 1461–1471.

Dindo, P. (2019). Survival in speculative markets. *Journal of Economic Theory 181*, 1–43.

Dindo, P. and F. Massari (2020). The wisdom of the crowd in dynamic economies. *Theoretical Economics 15*(4), 1627–1668.

Epstein, L. G., J. Noor, and A. Sandroni (2010). Non-bayesian learning. *The BE Journal of Theoretical Economics 10*(1).

Giachini, D. (2021). Rationality and asset prices under belief heterogeneity. *Journal of Evolutionary Economics 31*(1), 207–233.

Gigerenzer, G. and W. Gaissmaier (2011). Heuristic decision making. *Annual review of psychology 62*, 451–482.

Jiang, H., S. Z. Li, and H. Wang (2021). Pervasive underreaction: Evidence from high-frequency data. *Journal of Financial Economics 141*(2), 573–599.

Jouini, E. and C. Napp (2016). Live fast, die young. *Economic Theory 62*(1), 265–278.

Massari, F. (2020). Under-reaction: Irrational behavior or robust response to model misspecification?

Massari, F. (2021). Price probabilities: A class of bayesian and non-bayesian prediction rules. *Economic Theory 72*, 133–166.

Savage, L. (1954). *The foundations of statistics.*

A Appendix 1

Given the properties of the instantaneous utility, the individual problem of agent i reduces to maximizing with respect to consumption the Lagrangian

$$\mathcal{L}_i = U_i + \mu_i \sum_{t=0}^{\infty} \sum_{\sigma_t \in \sigma_t} q(\sigma_t) \left(e_i(\sigma_t) - c_i(\sigma_t) \right) .$$

Thus, the system of first order conditions is

$$\begin{cases} c_i(\sigma_0) = \dfrac{1-\beta_i}{\mu_i} \, , \\ c_i(\sigma_t) = \dfrac{(1-\beta_i)\beta_i^t \, p_i(\sigma_t)}{\mu_i \, q(\sigma_t)} \qquad \forall t, \sigma \, , \\ \sum_{t=0}^{\infty} \sum_{\sigma_t \in \sigma_t} q(\sigma_t) \left(e_i(\sigma_t) - c_i(\sigma_t) \right) = 0 \, , \end{cases} \tag{10}$$

while the equilibrium conditions read $1 = \sum_{i=1}^{N} c_i(\sigma_t) \;\; \forall t, \sigma$. Substituting in the budget constraint, one gets

$$\sum_{t=0}^{\infty} \sum_{\sigma_t \in \sigma_t} q(\sigma_t) e_i(\sigma_t) = c_i(\sigma_0) \sum_{t=0}^{\infty} \beta_i^t \sum_{\sigma_t \in \sigma_t} p_i(\sigma_t) = \frac{c_i(\sigma_0)}{1-\beta_i} \, , \tag{11}$$

while, from the equilibrium conditions, one obtains

$$q(\sigma_t) = \sum_{i=1}^{N} \beta_i^t p_i(\sigma_t) c_i(\sigma_0) \, . \tag{12}$$