Market selection and learning under model misspecification

Giulio Bottazzi[1], Daniele Giachini*[1], and Matteo Ottaviani[2,3]

[1]Institute of Economics & Department EMbeDS, Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, 56127 Pisa - Italy
[2]German Centre for Higher Education Research and Science Studies, Schützenstraße 6a, 10117 Berlin - Germany
[3]Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa - Italy

May 8, 2023

Abstract

This paper studies market selection in an Arrow-Debreu economy with complete markets where agents learn over misspecified models. Under model misspecification, standard Bayesian learning loses its formal justification and biased learning processes may provide a selection advantage. Given the natural connection between selection outcomes and long-run asset prices, understanding which biased learning processes are evolutionary fit is instrumental to build a parsimonious long-run asset valuation model robust to misspecification. Leveraging two cases of model misspecification and four learning processes, our analysis reveals a general difficulty in ranking learning behaviors with respect to their survival prospects. For example, the advantage of predictions averaging disappears when the true data generating process does not belong to the same family of models agents use to learn. Rules that generically guarantee survival, appear to require an unreasonable amount of knowledge about all the agents that compose the market ecology. The goal of a parsimonious long-run asset valuation model robust to model misspecification remains out of reach.

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

Keywords: Learning, Market Selection, Model Misspecification, Financial Markets

*Corresponding author, email: daniele.giachini@santannapisa.it

1 Introduction

The market selection hypothesis, applied to competitive environments where agents are able to *learn* (i.e., update their beliefs according to given rules), implies that only those who incorporate evidence into their probabilistic predictions according to Bayes rule are able to survive and, thus, influence assets' long-run evaluation (see e.g. Blume and Easley, 2006, 2009a,b). Such a statement relies upon the assumption that the learning problem is correctly specified (i.e., the true data generating process belongs to the set of models over which agents are learning) or, at least, a version of the complete class theorem holds. In those situations, traders who learn in a Bayesian way are, indeed, able to drive non-Bayesian traders out of the market (Sandroni, 2005). However, as reported by Gigerenzer and Gaissmaier (2011), Savage (1954) – the founder of Bayesian decision theory – used to distinguish between situations in which there is perfect information (*small worlds*) and situations in which relevant pieces of information are not available to decision makers (*large worlds*). Gigerenzer and Gaissmaier (2011), among many others, argue that real decision makers mostly face large-world situations and, thus, their learning problems are seldom correctly specified. Indeed, the models on which they have to rely are approximations or simplified versions of the real data generating process, thus, they face *model misspecification*. In such a case, Bayesian learning loses its formal justification and whether its selection advantage is conserved is not clear.

A recent study by Massari (2020) shows that a learning bias known as *under-reaction* – i.e., giving larger weight to the prior than what Bayesian learning prescribes (Epstein et al., 2010)– provides a selection advantage over Bayesian learning in Arrow-Debreu economies characterized by model misspecification. That is, while updating beliefs according to Bayes rule allows an agent to asymptotically be as accurate as the best model in its support (Berk, 1966), under-reaction produces either the same beliefs of a Bayesian agent or more accurate ones generated by a persistent (but not fixed) mixture of models. Since competitive (and complete) markets favor those who make accurate predictions (Sandroni, 2000), an under-reacting agent facing a Bayesian one is always able to maintain a positive consumption share and, in generic cases, it can even asymptotically consume all the aggregate endowment. Thus, under model misspecification, moving away from Bayesian learning can be beneficial and under-reaction emerges as a robust learning behavior in terms of survival. In a similar setting, Antico et al. (2023) investigate the evolutionary fitness of a trader behaving according to the *sentiment investor* learning model of Barberis et al. (1998) when competing against a Bayesian agent under model misspecification. The authors find that long-run selection outcomes are profoundly related to the characteristics of the agents and of the economy: depending on parameter settings, one can observe either that the

sentiment investor let the Bayesian trader vanish or vice-versa. Notwithstanding those results, one may argue that, since under model misspecification Bayesian updating is no-longer formally justified, selection outcomes derived from competition against a Bayesian trader may not be very informative about which learning behavior is actually observed in the long-run and, thus, has a persistent influence on asset evaluation. Indeed, one can imagine that the introduction of an ecology of different learning rules may actually generate non-trivial selection results. Once one acknowledges that, a related issue that immediately emerges is whether one can devise a ranking of learning processes in terms of their fitness to survive under model misspecification, since that would be instrumental to build a parsimonious (i.e., that considers only a limited number of learning behaviors) long-run asset valuation model that is fully robust to model misspecification.

In this paper we investigate those issues considering an Arrow-Debreu pure-exchange economy with complete markets, four different learning processes, and two cases of model misspecification. With respect to learning processes, we extend the framework of Massari (2020) adding to Bayesian learning and under-reaction two new processes: limited memory Bayesian learning and moving average. The first one consists in continuously resetting the Bayesian learning process. The second one consists in averaging the predictions of a reference learning process. With respect to the two cases of model of model misspecification, we consider parametric and structural misspecification. The first case is obtained assuming that the true probability measure belongs to the same class of probabilistic models the agents use to learn, but with different parameter values. For simplicity, we focus on i.i.d. true process and models. The second case consists in assuming that the true probability measure has a more complex probabilistic structure than the models agents use to learn. In our case, we consider a Markov true probability measure and i.i.d. models.

Our analysis draws from two approaches to the study of market selection. The first is characterized by general equilibrium, intertemporal utility maximization, and complete markets (see e.g. Sandroni, 2000; Blume and Easley, 2006, 2009a; Jouini and Napp, 2011; Kogan et al., 2006, 2017; Massari, 2017; Dindo and Massari, 2020; Beddock and Jouini, 2021; Bottazzi and Giachini, 2022). The second, instead, relies upon temporary equilibrium, bounded rationality, evolutionary dynamics among investment rules (see e.g. Hens and Schenk-Hoppé, 2005; Evstigneev et al., 2009, 2016; Holtfort, 2019; Bottazzi and Dindo, 2013, 2014; Bottazzi et al., 2018, 2019; Bottazzi and Giachini, 2017, 2019b,a; Elmiger, 2020).[1] Indeed, we combine the complete market Arrow-Debreu economy, characterizing most of the

[1]The two approaches are not separated, indeed there generically exist evolutionary models that produce the same wealth dynamics of general equilibrium ones. The link is built by means of *effective beliefs*, see Bottazzi et al. (2018), Dindo (2019), Giachini (2021).

contributions belonging to the first approach, with biased learning schemes, which are closer to the second approach. Our choice is motivated by avoiding compensation effects between non-optimality in investment rules and misspecification in beliefs, as described in Bottazzi et al. (2018) and Giachini (2021).

Our analysis shows that the ecology of learning behaviors operating in the market and the type of model misspecification characterizing the economy is of crucial importance for selection results. Thus, a general difficulty emerges in providing a ranking of learning processes in terms of their evolutionary robustness to model misspecification. For instance, under-reaction shows a *generic advantage* in terms of selection under parametric misspecification when the true probability measure is a convex combination of the i.i.d. models the agents use to learn: increasing the level of under-reaction lets the agent increase its accuracy. Indeed, a sufficiently high under-reaction degree makes any agent whose beliefs are bounded away from the truth – such as, for instance, the limited memory Bayesian learner – disappear. Such a particular selection advantage is shared by a moving average agent that leverages under-reaction. Indeed, our analysis suggest that, under parametric misspecification, averaging predictions appears as a key mechanism in generating a selection advantage. Such a mechanism, however, partially breaks down when structural misspecification occurs. More specifically, a trade-off between how well beliefs approximate the true Markov chain's invariant distribution (i.e. the best i.i.d. model) and how well fluctuations in conditional probabilities are captured, seems to appear. For instance, on the one hand, averaging past predictions allows an agent to better approximate the invariant distribution, but, on the other hand, decreases the relevance of fresh information and the speed of adaptation to changes. As a consequence, the averaging approaches can be generically outperformed by the limited memory Bayesian learning process. Finally, we discuss some examples of learning rules that can survive no matter the form of model misspecification. Our discussion points out that the key feature of those rules is that they exploit information about all the other market participants rather than efficiently combining information about past realizations of the true data generating process. This is, however, disruptive for the quest of a parsimonious long-run asset evaluation model robust to model misspecification: one cannot dispense from knowing important features of all the agents in the economy.

2 The Model

Consider an Arrow-Debreu economy with infinite horizon and discrete time (indexed by $t = 0, 1, \ldots$). There is a homogeneous consumption good and markets are complete. Call $s_t \in \{1, 2, \ldots, S\}$ the state realized at time $t > 0$. We indicate with $\sigma = (s_1, s_2, \ldots, s_t, \ldots)$ a path and with $\sigma_t = (s_1, s_2, \ldots, s_t)$ a partial history until

time t. The set of all the possible paths is Σ while Σ_t indicates the set of all partial histories until time t. Let $\mathcal{C}(\sigma_t) = \{\sigma \in \Sigma | \sigma = (\sigma_t, \ldots)\}$ be the cylinder with base σ_t, $\mathcal{F}_t$ is the σ-algebra generated by the cylinders $\mathcal{C}(\sigma_t)$. Then, by construction, $(\mathcal{F}_t)_{t=0}^{\infty}$ is a filtration and we indicate with $\mathcal{F}$ the σ-algebra generated by the union of filtrations. We indicate with p the true probability measure on $(\Sigma, \mathcal{F})$, such that $(\Sigma, \mathcal{F}, p)$ is a well-defined probability space. We assume that any partial history has a positive probability of being realized, $p(\sigma_t) > 0$, $\forall \sigma_t$. Expectation is denoted with E and, when there is no subscript or superscript, it is computed with respect to p.

The economy is populated by N agents indexed by $i = 1, 2, \ldots, N$. Every agent i is endowed with a stream of non-zero and uniformly bounded consumption good for any path σ, $(e_i(\sigma_t))_{t=0}^{\infty}$. Agent i has a subjective probability measure p_i on $(\Sigma, \mathcal{F})$. Denote with $p_i(s_t|\sigma_{t-1})$ the (subjective) conditional probability attached to the realization of s_t after a partial history σ_{t-1} and with $p_i(\sigma_t) = \prod_{\tau=1}^{t} p_i(s_\tau|\sigma_{\tau-1})$ the (subjective) likelihood of partial history σ_t. Agent i chooses its consumption plan $(c_i(\sigma_t))_{t=0}^{\infty}$ solving

$$\max_{\{c_i(\sigma_t),\, \forall t, \sigma\}} \mathrm{E}_{p_i}\left[\sum_{t=0}^{\infty} \beta_i^t u_i(c_i(\sigma_t))\right] \text{ s.t. } \sum_{t=0}^{\infty} \sum_{\sigma_t \in \Sigma_t} q(\sigma_t)\left(e_i(\sigma_t) - c_i(\sigma_t)\right) \geq 0,$$

where $\beta_i \in (0,1)$ is agent i's discount factor, u_i is the Bernoulli utility of consumption of agent i, and $q(\sigma_t)$ is the price of the Arrow-Debreu security paying one if partial history σ_t is realized and zero otherwise. We will further assume that individual probabilities p_i are absolute continuous with respect to p and that the Bernoulli utilities are continuously differentiable, increasing, strictly concave, and satisfies the Inada condition at zero. With these hypotheses, a competitive equilibrium exists unique and $\forall \sigma_t$, $q(\sigma_t) > 0$, $\sum_{i=1}^{N} c_i(\sigma_t) = \sum_{i=1}^{N} e_i(\sigma_t) = e(\sigma_t)$.

2.1 Consumption asymptotic behavior

Our main goal is to evaluate the selection dynamics taking place in competitive markets under different learning protocols. We need the following.

Definition 2.1. An agent i:

- *vanishes* if $\lim_{t\to\infty} c_i(\sigma_t) = 0$, p-almost surely;
- *survives* if $\lim_{t\to\infty} \sup_t c_i(\sigma_t) > 0$, p-almost surely;
- *dominates* if $\lim_{t\to\infty} c_i(\sigma_t)/e_t(\sigma) = 1$, p-almost surely.

The study of the asymptotic dynamics of the relative consumption of agents can be reduced to the analysis of their individual probability measures and discount factors by the following mathematical passage (Blume and Easley, 2006). From the F.O.C. of the optimal consumption problem, $\forall i, j \in 1, \ldots, n$,

$$\frac{u_i'(c_i(\sigma_t))}{u_j'(c_j(\sigma_t))} = \left(\frac{\beta_j}{\beta_i}\right)^t \frac{p_j(\sigma_t)}{p_i(\sigma_t)} \frac{u_i'(c_i(\sigma_0))}{u_j'(c_j(\sigma_0))},$$

that is

$$\frac{1}{t}\log\frac{u_i'(c_i(\sigma_t))}{u_j'(c_j(\sigma_t))} = \log\frac{\beta_j}{\beta_i} + \frac{1}{t}\log\frac{p(\sigma_t)}{p_i(\sigma_t)} - \frac{1}{t}\log\frac{p(\sigma_t)}{p_j(\sigma_t)} + \frac{1}{t}\log\frac{u_i'(c_i(\sigma_0))}{u_j'(c_j(\sigma_0))}. \quad (1)$$

To describe the agent's individual probabilities, consider K i.i.d. measure whose conditional probabilities are the vectors $\boldsymbol{\pi}_1, \ldots, \boldsymbol{\pi}_K$, belong to the topological interior of the $(S-1)$-simplex, $\boldsymbol{\pi}_k = (\pi_k(1), \pi_k(2), \ldots, \pi(S)) \in \Delta_+^{S-1}$. These vectors are uniformly bounded away from zero and diverse, that is $\exists \epsilon, d\pi > 0$ such that $\pi_k(s) > \epsilon$ and $\|\boldsymbol{\pi}_k - \boldsymbol{\pi}_h\| > d\pi$, $\forall s, k, h$. To simplify our investigation we assume the following.

Assumption 1. Agents' individual conditional probabilities belong to the convex hull H_K generated by the conditional probabilities of the K models,

$$p_i(s \mid \sigma_t) \in H_K = \left\{ \sum_{k=1}^K \eta_k \boldsymbol{\pi}_k \mid \sum_{k=1}^K \eta_k = 1, \eta_k \geq 0 \right\} \subseteq \Delta_+^{S-1}, \forall s, \sigma_t.$$

Note that the previous assumption guarantees that individual probabilities have conditionals uniformly bounded away from zero, $p_i(s \mid \sigma_t) > \epsilon, \forall \sigma_t, s$. Denote the (conditional) relative entropy of the individual probability measure p_i with respect to the truth p given partial history σ_t and its partial average as

$$D_{p|p_i}(\sigma_t) = \sum_{s=1}^S p(s \mid \sigma_t) \log\frac{p(s \mid \sigma_t)}{p_i(s \mid \sigma_t)} \text{ and } \overline{D}_{p|p_i}(\sigma_t) = \frac{1}{t+1}\sum_{\tau=0}^t D_{p|p_i}(\sigma_\tau).$$

By Assumption 1, these quantities are bounded, $D_{p|p_i}(\sigma_t), \overline{D}_{p|p_i}(\sigma_t) \in [0, -\log\epsilon]$. We shall use the partial average of the relative entropy as a measure of accuracy, for instance, we will say that an agent j is more accurate than an agent i at σ_t if $\overline{D}_{p|p_j}(\sigma_t) < \overline{D}_{p|p_i}(\sigma_t)$. Moreover, we have the following.

Theorem 2.1. *Under Assumption 1, $\forall i = 1, \ldots, N$ and $\forall \alpha < 1/2$, p-almost surely,*

$$\frac{1}{t}\log\frac{p(\sigma_t)}{p_i(\sigma_t)} = \overline{D}_{p|p_i}(\sigma_{t-1}) + o\left(\frac{1}{t^\alpha}\right).$$

Proof. Define $z_i(s \mid \sigma_{\tau-1}) = \log\left(p(s|\sigma_{\tau-1})/p_i(s|\sigma_{\tau-1})\right) - D_{p|p_i}(\sigma_{\tau-1})$, so that

$$\log \frac{p(\sigma_t)}{p_i(\sigma_t)} = \sum_{\tau=1}^{t} \log \frac{p(s_\tau \mid \sigma_{\tau-1})}{p_i(s_\tau \mid \sigma_{\tau-1})} = \sum_{\tau=1}^{t} z_i(s_\tau \mid \sigma_{\tau-1}) + \sum_{\tau=1}^{t} D_{p|p_i}(\sigma_{\tau-1}).$$

By Assumption 1, and by the fact that the maximum of the function $x \log^2 x$ for $x \in [0,1]$ is $4e^{-2}$,

$$\mathrm{E}[\boldsymbol{z}_i^2 \mid \sigma_t] = \sum_{s=1}^{S} p(s \mid \sigma_t) \, \log^2 \frac{p(s \mid \sigma_t)}{p_i(s \mid \sigma_t)} - D_{p|p_i}(\sigma_t)^2 \leq 4e^{-2} S - \log \epsilon.$$

Hence, if $\alpha < 1/2$, $\sum_{t=1}^{\infty} t^{2\alpha-2} E[\boldsymbol{z}_i^2 \mid \sigma_{t-1}] < +\infty$. Since $\mathrm{E}[\boldsymbol{z}_i \mid \sigma_{t-1}] = 0$, $\forall \sigma_t$, by Theorem 3, p. 243, in Feller (1971), p-almost surely, $\lim_{t\to\infty} t^{\alpha-1} \sum_{\tau=1}^{t} z_i(s_\tau \mid \sigma_{\tau-1}) = 0$. This implies that, p-almost surely, $\log p(\sigma_t)/p_i(\sigma_t) - t\overline{D}_{p|p_i}(\sigma_{t-1}) = o(t^{1-\alpha})$. Dividing by t proves the assertion. □

In particular, the previous Theorem applies to the K i.i.d. models of Assumption 1. Substituting the statement of Theorem 2.1 in (1),

$$\frac{1}{t} \log \frac{u_i'(c_i(\sigma_t))}{u_j'(c_j(\sigma_t))} = \left(\log \beta_j - \overline{D}_{p|p_j}(\sigma_{t-1})\right) - \left(\log \beta_i - \overline{D}_{p|p_i}(\sigma_{t-1})\right) + o\left(\frac{1}{t^\alpha}\right).$$

The asymptotic behavior of the quantities inside the parentheses in the right-hand side determine the asymptotic behavior of the relative marginal utilities of the two agents. Some results of the literature trivially follow (Sandroni, 2000; Blume and Easley, 2006; Dindo and Massari, 2020).

Corollary 2.1. *Assume there exist two agents i and j such that, p-almost surely, exist $\overline{D}_{p|p_i}(\sigma) = \lim_{t\to\infty} \overline{D}_{p|p_i}(\sigma_t)$ and $\overline{D}_{p|p_j}(\sigma) = \lim_{t\to\infty} \overline{D}_{p|p_j}(\sigma_t)$. Then, if Assumptions 1 applies and, p-almost surely,*

$$\log \beta_j - \overline{D}_{p|p_j}(\sigma) > \log \beta_i - \overline{D}_{p|p_i}(\sigma),$$

agent i vanishes.

Proof. By hypothesis, we can apply Theorem 2.1, so that

$$\lim_{t\to\infty} \frac{1}{t} \log \frac{u_i'(c_i(\sigma_t))}{u_j'(c_j(\sigma_t))} = \left(\log \beta_j - \overline{D}_{p|p_i}(\sigma)\right) - \left(\log \beta_i - \overline{D}_{p|p_i}(\sigma)\right) > 0,$$

which implies $\lim_{t\to\infty} \log u_i'(c_i(\sigma_t))/u_j'(c_j(\sigma_t)) = +\infty$ From the boundedness of the endowment, the quantities $u_i'(c_i(\sigma_t))$ are bounded from above. Thus, it must be $\lim_{t\to\infty} \log u_i'(c_i(\sigma_t)) = +\infty$. According to the Inada condition, this, in turn, implies that $\lim_{t\to\infty} c_i(\sigma_t) = 0$. □

Note that the hypothesis of Corollary 2.1 are not trivial, as Assumption 1 is not sufficient to guarantee the existence of the limit of $\overline{D}_{p|p_i}(\sigma_t)$ and $\overline{D}_{p|p_j}(\sigma_t)$. However, the existence of these limits is not necessary. One can, for example, realize that agent i vanishes if, p-a.s., $\beta_j - \overline{D}_{p|p_j}(\sigma_t) > \log \beta_i - \overline{D}_{p|p_i}(\sigma_t)$ for sufficiently large t.

3 Learning processes

According to Assumption 1, $\forall \sigma_t$ agents individual probabilities can be written as

$$p_i(s|\sigma_t) = \sum_{k=1}^{K} w_{i,k}(\sigma_t)\, \pi_k(s), \quad w_{i,k}(\sigma_t) \geq 0, \forall k, \quad \sum_{k=1}^{K} w_{i,k}(\sigma_t) = 1, \tag{2}$$

were $w_{i,k}(\sigma_t)$ denotes the weight agent i attaches to model k after having observed the partial history σ_t. Agents differ on how they compute their wights. We will consider four learning processes: Bayesian learning, learning with under-reaction, limited memory Bayesian learning, and moving average of an underlying model.

Bayesian learning The Bayesian learning process can be considered the cornerstone of online learning. Weights are updated according to Bayes rule,

$$w_{i,k}(\sigma_t) = \frac{\pi_k(s_t)\, w_{i,k}(\sigma_{t-1})}{p_i(s_t|\sigma_{t-1})} = \frac{\pi_k(\sigma_t)}{p_i(\sigma_t)}\, w_{i,k}(\sigma_0) \quad \forall k, t, \sigma\,. \tag{3}$$

The weight $w_{i,k}(\sigma_t)$ can be considered the probability agent i attaches to the event "model k is the true one" conditional upon the observation of partial history σ_t. The key property of Bayesian learning is that it makes an agent as accurate as the model with highest likelihood in its set at any t sufficiently large (see also Berk, 1966, for further detail).

Proposition 3.1. *Define $k_t^* = argmax_{k\in\{1,\dots,K\}}\{\pi_k(\sigma_t)\}$. For any Bayesian agent i and $\forall \alpha < 1/2$, p-almost surely*

$$\left|\overline{D}_{p|p_i}(\sigma_{t-1}) - \overline{D}_{p|\pi_{k_t^*}}(\sigma_{t-1})\right| \leq o\left(\frac{1}{t^\alpha}\right).$$

Proof. By iteratively substituting (3) in (2),

$$p_i(\sigma_t) = \sum_{k=1}^{K} \pi_k(\sigma_t) w_{i,k}(\sigma_0).$$

Thus, $\pi_{k_t^*}(\sigma_t) w_{i,k_t^*}(\sigma_0) \leq p_i(\sigma_t) \leq \pi_{k_t^*}(\sigma_t)$ and, as a consequence,

$$\frac{1}{t} \log \frac{p(\sigma_t)}{\pi_{k_t^*}(\sigma_t)} \leq \frac{1}{t} \log \frac{p(\sigma_t)}{p_i(\sigma_t)} \leq \frac{1}{t} \log \frac{p(\sigma_t)}{\pi_{k_t^*}(\sigma_t)} - \frac{1}{t} \log w_{i,k_t^*}(\sigma_0) \,.$$

The statement follows by applying Theorem 2.1 to the individual probability p_i and the measure $\pi_{k_t^*}$. □

Notice that, in general, the model with highest likelihood at a date t depends on the specific realization σ_t. Moreover, without further assumptions on the true probability p, we are not guaranteed that the Bayesian agent will asymptotically follow one specific model in the set.

Learning with under-reaction This learning process consists in a modification of (3) according to the notion of under-reaction in Epstein et al. (2010) and Massari (2020). This learning protocol can be considered a form of "moderate" Bayesian 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:

$$w_{i,k}(\sigma_t) = \lambda_i \, w_{i,k}(\sigma_{t-1}) + (1 - \lambda_i) \, \frac{\pi_k(s_t) \, w_{i,k}(\sigma_{t-1})}{p_i(s_t | \sigma_{t-1})} \quad \forall k, t, \sigma \,, \tag{4}$$

with $\lambda_i \in [0, 1)$. Setting $\lambda_i = 0$, Bayesian learning is recovered. Learning with under-reaction entails a form of averaging. The probabilistic prediction of an under-reacting agent i for state s_{t+1} after a partial history σ_t can be seen as the convex combination of the probabilistic prediction agent i would make after the partial history σ_{t-1} and the Bayesian prediction given a prior $\boldsymbol{w}_i(\sigma_{t-1}) = (w_{i,1}(\sigma_{t-1}), \ldots, w_{i,K}(\sigma_{t-1}))$ and the observation of state s_t (Epstein et al., 2010; Giachini, 2021). Under-reaction represents a robust learning strategy. In case of model misspecification it can outperform Bayesian learning (Massari, 2020). Moreover, this rule is equivalent to the *Soft-Bayes* algorithm of Orseau et al. (2017), match the dynamics of prices and wealth in the prediction market model of Bottazzi and Giachini (2017, 2019b), and describe the risk neutral probabilities and consumption shares in the pure exchange economy model analyzed by Dindo and Massari (2020). The following proposition adapts a result on under-reaction by Massari (2020) to our framework.[2] It shows that an under-reacting agent is at least as accurate as the most accurate model in its set for t sufficiently large.

Proposition 3.2. *For any under-reacting agent i and $\forall \alpha < 1/2$, it is p-almost surely*

[2]The result by Massari (2020) makes use of the notion of *empirical distribution of states* while we state it with respect to the true measure p. If p describes an i.i.d. process, then the two statements are equivalent.

i) $\overline{D}_{p|p_i}(\sigma_{t-1}) \leq \overline{D}_{p|\rho_{i,k}}(\sigma_{t-1}) + o\left(t^{-\alpha}\right)\ \forall k \in \{1, 2, \ldots, K\}$,

ii) $\overline{D}_{p|\rho_{i,k}}(\sigma_{t-1}) \leq \overline{D}_{p|\pi_k}(\sigma_{t-1}) + o\left(t^{-\alpha}\right)\ \forall k \in \{1, 2, \ldots, K\}$;

where $\rho_{i,k}(s_{t+1}|\sigma_t) = \lambda_i\, p_i(s_{t+1}|\sigma_t) + (1-\lambda_i)\,\pi_k(s_{t+1})$.

Proof. Note that $w_{i,k}(\sigma_t) = w_{i,k}(\sigma_{t-1})\rho_{i,k}(s_t|\sigma_{t-1})/p_i(s_t|\sigma_{t-1})$. Iterative substitution with the previous equation gives

$$p_i(\sigma_t) = p_i(\sigma_{t-1}) \sum_{k=1}^{K} \rho_{i,k}(s_t|\sigma_{t-1}) w_{i,k}(\sigma_{t-1}) = \cdots = \sum_{k=1}^{K} \rho_{i,k}(\sigma_t)\, w_{i,k}(\sigma_0),$$

where $\rho_{i,k}(\sigma_t) = \prod_{\tau=1}^{t-1} \rho_{i,k}(s_{\tau+1}|\sigma_\tau)$. Hence, $\forall k$, $p_i(\sigma_t) \geq \rho_{i,k}(\sigma_t)\, w_{i,k}(\sigma_0)$, and

$$\log \frac{p(\sigma_t)}{p_i(\sigma_t)} \leq \log \frac{p(\sigma_t)}{\rho_{i,k}(\sigma_t)} - \log w_{i,k}(\sigma_0) \leq \\ \lambda_i \log \frac{p(\sigma_t)}{p_i(\sigma_t)} + (1-\lambda_i) \log \frac{p(\sigma_t)}{\pi_k(\sigma_t)} - \log w_{i,k}(\sigma_0),$$

where we have used the inequality $\log \rho_{i,k}(\sigma_t) \geq \lambda_i \log p_i(\sigma_t) + (1-\lambda_i) \log \pi_k(\sigma_t)$. Note that, by definition, $\rho_{i,k}(s \mid \sigma_t) \in H_K$, thus applying Theorem 2.1 to p_i, $\rho_{i,k}$ and π_k, the statements are recovered. □

Again, the most accurate model at a date t depends on the specific realization σ_t and Proposition 3.2 does not imply the asymptotic convergence of the under-reacting agent to a single i.i.d. model.

Limited memory Bayesian learning The limited memory Bayesian learning is a version of the standard Bayesian learning process in which the agent deliberately forgets observations in the past. Here we consider the version with the shortest possible memory, that is a memory of one. In this case, the weight assigned to model k after a partial history σ_t reads

$$w_{i,k}(\sigma_t) = \frac{\pi_k(s_t) w_{i,k}(\sigma_0)}{\sum_{k'=1}^{K} \pi_{k'}(s_t) w_{i,k'}(\sigma_0)} \tag{5}$$

In any period t, agent i is forgetting all the sequence of states occurred until $t-2$ (included) and restarts its Bayesian learning procedure simply considering the previous state and the initial prior distribution of weights $\boldsymbol{w}_i(\sigma_0) = (w_{i,1}(\sigma_0), w_{i,2}(\sigma_0), \ldots, w_{i,K}(\sigma_0))$. The i.i.d. nature of the models on which the agent learns makes the limited memory Bayesian learner have a simple Markov structure. This process also displays a strong dependence on initial weights.

Moving average learning The moving average learning process consists in taking a reference learning process p^* and applying a moving average to the sequence of probabilistic predictions generated for every state. Assume agent i adopts a moving average learning with memory M_i, then

$$p_i(s \mid \sigma_t) = \begin{cases} p^*(s \mid \sigma_t) & \text{if } t < M_i - 1\,, \\ M_i^{-1} \sum_{m=1}^{M_i} p^*(s \mid \sigma_{t-m+1}) & \text{if } t \geq M_i - 1\,. \end{cases} \tag{6}$$

If the underlying learning process follows Assumption 1, the same thing can be restated in terms of weights with $w_{i,k}(\sigma_t) = M_i^{-1} \sum_{m=1}^{M_i} w_k^*(\sigma_{t-m+1})$ if $t \geq M_i - 1$. The moving average learning represents a further layer of "smoothing" over the predictions of the underlying process.

4 Misspecified models

In what follows, we study the performance of the learning processes described above in a competitive environment where the K i.i.d models they use are misspecified. We investigate two specific cases of mispecification. We start with an i.i.d. true measure that does not belong to the set of models the agents can learn. In this case the i.i.d. models the agent use belong to the same class of the true measure, but their parameters are, generically, not correct. We call this case *parametric misspecification*. In the second case, the proper *structural misspecification* case, we assume that the true measure is Markov. In this case, the models employed by the agents belong to a different, and less general, class than the truth.

4.1 Parametric misspecification

We assume that states of nature follow an i.i.d. process, such that the models agents use belong to the same family of the truth but have misspecified parameters. Formally,

Assumption 2. The true measure p is an i.i.d. process whose conditional distribution are described by the vector $\boldsymbol{\pi} = (\pi(1), \pi(2), \ldots, \pi(S)) \in \Delta_+^{S-1}$, such that $p(s_t \mid \sigma_{t-1}) = \pi(s_t)$. Agents' models are misspecified: $\|\boldsymbol{\pi}_k - \boldsymbol{\pi}\| > 0$ $\forall k \in \{1, 2, \ldots, K\}$.

We first present some analytic results about the behavior of the different learning models under this assumption. Then, we propose some numerical simulation to clarify their relative performances.

Bayesian learning. Under Assumption 2, p-almost surely

$$\lim_{t\to\infty} \overline{D}_{p|\pi_k}(\sigma_t) = D_{\pi|\pi_k} = \sum_{s=1}^{S} \pi(s) \log \frac{\pi(s)}{\pi_k(s)} > 0.$$

Assuming that model k, on which the Bayesian agent i learns, has the lowest relative entropy of all K models, then, according to Proposition 3.1, $\lim_{t\to\infty} \overline{D}_{p|p_i}(\sigma_t) = D_{\pi|\pi_k}$, which implies that $\lim_{t\to\infty} w_{i,k}(\sigma_t) = 1$. The convergence of the weights, however, is not otherwise guaranteed. In any case, the Bayesian learner is as accurate as the most accurate model available.

Learning with under-reaction. From Proposition 3.2 it follows that also the under-reacting agent is never less accurate than a Bayesian learner. Massari (2020) proves that when the parameter λ is large enough and the conditional probability of the true i.i.d. model belongs to H_K, an agent that learns with under-reaction has a selection advantage over a Bayesian learner. This represents a *specific* advantage of the under-reacting versus the Bayesian agent. Inspired by the results of Bottazzi and Giachini (2017) and Dindo and Massari (2020), the following Proposition extends the analysis proving the existence of a *generic* advantage of an under-reacting agent.

Proposition 4.1. *Suppose there exists a K-dimensional vector $\boldsymbol{\zeta}$ with $\zeta_k \geq 0$ and $\sum_{k=1}^{K} \zeta_k = 1$, such that $\boldsymbol{\pi} = \sum_{k=1}^{K} \zeta_k \boldsymbol{\pi}_k$. Then, for any under-reacting agent i and $\forall \alpha < 1/2$, it is p-almost surely*

$$\overline{D}_{p|p_i}(\sigma_t) \leq \frac{1-\lambda_i}{2(\lambda_i+\epsilon)^2} + \frac{o\left(t^{-\alpha}\right)}{1-\lambda_i}.$$

Proof. Setting $\theta = 1 - \lambda_i > 0$ and using the first order Taylor expansion with Lagrange remainder of the logarithmic function, for each realization σ_τ,

$$\begin{aligned} D_{p|p_i}(\sigma_\tau) - \sum_{k=1}^{K} \zeta_k D_{p|\rho_{i,k}}(\sigma_\tau) &= \theta \sum_{s=1}^{S} \pi(s) \left(\frac{\pi(s)}{p_i(s|\sigma_t)} - 1 \right) \\ &\quad - \frac{\theta^2}{2} \sum_{k=1}^{K} \zeta_k \sum_{s=1}^{S} \pi(s) \frac{(\pi_k(s) - p_i(s|\sigma_t))^2}{\left(\eta_{k,s}(\sigma_\tau)\pi_k(s) + (1-\eta_{k,s}(\sigma_\tau))\, p_i(s|\sigma_t)\right)^2}, \end{aligned}$$

for some $\eta_{k,s}(\sigma_\tau) \in [0, \theta]$. Thus, by Assumption 1,

$$\sum_{k=1}^{K} \zeta_k \sum_{s=1}^{S} \pi(s) \frac{(\pi_k(s) - p_i(s|\sigma_t))^2}{\left(\eta_{k,s}(\sigma_\tau)\, \pi_k(s) + (1-\eta_{k,s}(\sigma_\tau))\, p_i(s|\sigma_t)\right)^2} \leq \frac{1}{(1-\theta+\epsilon)^2},$$

and because $x - 1 \geq \log x$,

$$\sum_{s=1}^{S} \pi(s) \left(\frac{\pi(s)}{p_i(s|\sigma_\tau)} - 1 \right) \geq D_{p|p_i}(\sigma_\tau),$$

so that

$$D_{p|p_i}(\sigma_\tau) - \sum_{k=1}^{K} \zeta_k \, D_{p|\rho_{i,k}}(\sigma_\tau) \geq (1 - \lambda_i) D_{p|p_i}(\sigma_\tau) - \frac{(1-\lambda_i)^2}{2(\lambda_i + \epsilon)^2}.$$

Summing on τ from 0 to $t-1$ and dividing by t, according to point i) of Proposition 3.2, the left-hand side is lower then $o\left(t^{-\alpha}\right)$, and the statement follows. □

Hence, when the truth belongs to H_K, an under-reacting agent becomes extremely accurate as its level of under-reaction increases and, as a consequence, obtains a survival advantage with respect to other traders. To see it, consider a market in which there is discount factor homogeneity and agent 1 is under-reacting with parameter λ_1. Any trader i for which it is p-almost surely $\overline{D}_{p|p_i}(\sigma_t) > (1-\lambda_1)/(2(\lambda_1+\epsilon)^2)$ for t sufficiently large will vanish. If all traders apart the under-reacting one are bounded away from the truth, $\overline{D}_{p|p_i}(\sigma_t) > \delta > 0$, for $i > 1$, then the under-reacting agent dominates if its parameter is sufficiently large. For instance, a sufficient condition is

$$\lambda_1 \in \left(\frac{\sqrt{1+8\delta} - 1}{4\delta}, 1 \right).$$

Limited memory Bayesian learning. Under Assumption 2, the limited memory Bayesian agent's conditionals only depend on the last realized state of nature. Thus, one has $D_{p|p_i}(\sigma_t) = D_{p|p_i}(s_t)$ and a straightforward application of the Strong Law of Large Numbers delivers the following.

Corollary 4.1. *For any agent i that uses the limited memory Bayesian learning process, it is $\overline{D}_{p|p_i}(\sigma) = \lim_{t\to\infty} \overline{D}_{p|p_i}(\sigma_t) = \sum_{s=1}^{S} \pi(s) D_{p|p_i}(s)$ p-almost surely.*

Hence, alternating among different convex combinations of models depending on the last realized state, the accuracy of the limited memory Bayesian agent depends upon how accurate those convex combinations are on average. A straightforward implication of Corollary 4.1 is that $D_{p|p_i}(s_m) \leq \overline{D}_{p|p_i}(\sigma) \leq D_{p|p_i}(s_M)$ where $s_m = \text{argmin}_s\{D_{p|p_i}(s)\}$ and $s_M = \text{argmax}_s\{D_{p|p_i}(s)\}$. Thus, a limited memory Bayesian agent can be maximally accurate only when all of its convex combinations match the true probabilities. This is never possible, for instance, if the K models on which the agent learns are linearly independent.

Moving average learning. Under Assumption 2, the smoothing in conditionals applied by an agent following the moving average learning process reflects into its average relative entropy. Define the moving window average $\mu(s,\sigma_t) = \sum_{m=0}^{M-1} p^*(s \mid \sigma_{t-m})/M$ and variance $\sigma^2(s,\sigma_t) = \sum_{m=0}^{M-1} \left(p^*(s \mid \sigma_{t-m}) - \mu(s,\sigma_t)\right)^2/M$ of the conditional probability of the underlying model.

Proposition 4.2. *Consider the conditional expected variance of the prediction* $\sigma^2(\sigma_t) = \sum_{s=1}^{S} \pi(s)\sigma^2(s,\sigma_t)$*, then,* $\forall \sigma_t$ *and for any agent* i *using the moving average learning process, it is*

$$\frac{1}{M}\sum_{m=0}^{M-1} D_{p|p^*}(\sigma_{t-m}) - D_{p|p_i}(\sigma_t) \geq \frac{\sigma^2(\sigma_t)}{2\,(1-\epsilon)}\,.$$

If $\overline{D}_{p|p^*}(\sigma) = \lim_{t\to\infty} \overline{D}_{p|p^*}(\sigma_t)$ *exists, then* $\limsup_{t\to\infty} \overline{D}_{p|p_i}(\sigma_t) \leq \overline{D}_{p|p^*}(\sigma)$*, with strict inequality if* $\exists \varepsilon > 0$ *such that* $\sigma^2(\sigma_t) > \varepsilon$*.*

Proof. Using the result on the bounds of the arithmetic and geometric means inequality in Perisastry and Murty (1982), and the bounds on the probability models in Assumption 1, $\forall s, \sigma_\tau$,

$$\frac{\sigma^2(s,\sigma_\tau)}{2\,(1-\epsilon)} \leq \log\left(\frac{1}{M}\sum_{m=0}^{M-1} p^*(s \mid \sigma_{\tau-m})\right) - \frac{1}{M}\sum_{m=0}^{M-1} \log p^*(s \mid \sigma_{\tau-m}) \leq \frac{\sigma^2(s,\sigma_\tau)}{2\,\epsilon}\,.$$

Focusing on the inequality on the left, adding and subtracting $\log \pi(s)$ to the central member, multiplying by $\pi(s)$, and summing over s, one obtains the first assertion.

Averaging from $\tau = M-1$ to $\tau = t-1$ and adjusting in order to obtain the definitions of average relative entropy, one has

$$\frac{1}{M}\sum_{m=0}^{M-1} \frac{t-m}{t}\,\overline{D}_{p|p^*}(\sigma_{t-m-1}) - \overline{D}_{p|p_i}(\sigma_{t-1}) \geq \frac{1}{t}\sum_{\tau=M}^{t} \frac{\sigma^2(\sigma_{\tau-1})}{2(1-\epsilon)} - o\left(\frac{1}{t}\right),$$

that proves the second assertion. □

Proposition 4.2 suggests that an evolutionary advantage can be extracted from averaging predictions. In this case, the moving average learning process is never less accurate than the learning process it leverages if $\overline{D}_{p|p^*}(\sigma)$ exists. In fact, if the underlying model p^* converges p-almost surely to a constant conditional probability $p^*(s|\sigma_t)$, like Bayesian learning in the presence of a best model, then from Proposition 4.2 one has p-almost surely that $\lim_{t\to\infty} \overline{D}_{p|p_i}(\sigma_t) = \overline{D}_{p|p^*}(\sigma)$. Thus, assuming that the underlying reference model is adopted by other traders in the market, averaging does not lead to any long-run advantage over them. If, instead,

the underlying model entails some sort of persistent fluctuation in conditionals, moving average brings a definite advantage over other traders adopting the reference model. In any case, the moving average learning model never under-performs when compared to the underlying model on which its predictions are built.

Despite providing useful information about the inner workings of the different learning models, the previous results do not allow to devise any general ranking among them. The generic advantage that under-reaction learning enjoys depends upon the degree of under-reaction and the truth belonging to the convex hull of models. However, Proposition 4.1 does not imply, for instance, that increasing λ one increases accuracy. The moving average approach has an advantage when the reference learning process does not converge. Thus, it may be effective when it leverages (and competes against) an under-reacting agent characterized by persistent fluctuations in beliefs (Massari, 2020). However that advantage generically disappears when it competes against a Bayesian learner. For generic initial weights, the limited memory Bayesian learner can never be maximally accurate. This does not mean that it is the worst learning approach. Indeed, by resetting the learning process every time, the agent constantly mixes the misspecified i.i.d. models and never converges to a single one. This might represent an advantage. The numerical exercises proposed in the next section exemplify the difficulties in ranking and provide some new insights about the relative performances of the models matter of study.

4.1.1 Numerical exploration

We consider an economy with two possible states of the world, $S = 2$, driven by an i.i.d. true process. Hence, slightly abusing notation, we set $p(1|\sigma_t) = p \ \forall t, \sigma$ with $p \in (0, 1)$, that is $\boldsymbol{\pi} = (p, 1 - p)$. Agents learn on two models (i.e. $K = 2$) with respective probabilities $\boldsymbol{\pi}_1 = (\pi_1, 1-\pi_1)$ and $\boldsymbol{\pi}_2 = (\pi_2, 1-\pi_2)$; $\pi_1, \pi_2 \in (0, 1)$. The performances of the different learning models are expressed in terms of their average relative entropy $\overline{D}_{p|p_i}(\sigma_t)$ and are reported in Figure 1. The Bayesian model $\lambda_i = 0$ (darker and thicker solid line) always converge to the best model (c.f. Proposition 3.1) and the average relative entropy can be computed analytically. Its value is zero when the true probability matches one of the two underlying models, i.e. $p = 0.3$ and $p = 0.8$. In the case of an under-reacting agent an analytical expression is not available. Thus, the average value of $\overline{D}_{p|p_i}(\sigma_t)$ is computed for different values of λ_i over 10^2 independent random partial histories of length $t = 2\times 10^4$ (thinner and lighter lines).[3] It is worth to remark that, following Theorem 2.1, the number we report for each given combination of parameters can be understood

[3]The first 10^4 steps of each independent replication have been discarded to mitigate the initial condition bias.

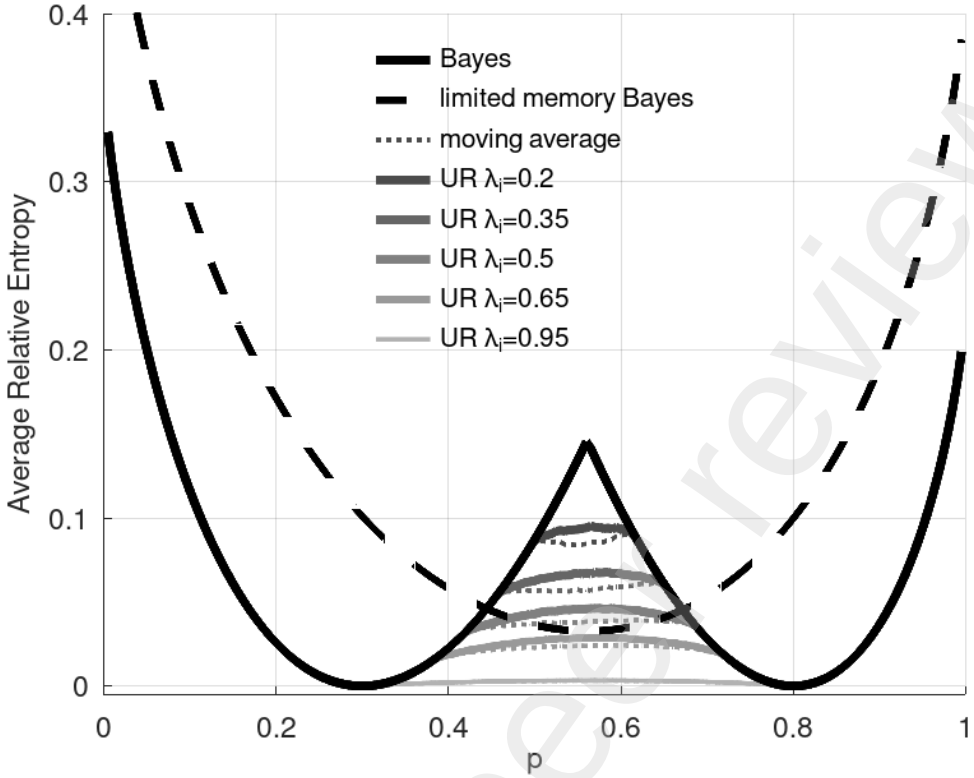

Figure 1: Average relative entropy of the different learning models as a function of p, i.e. the true probability of the realization of state 1. Parameter settings are $\pi_1 = 0.3$, $\pi_2 = 0.8$, and, for the moving average model, $M_j = 10$. For estimated values, standard errors are in the order of 10^{-4} or smaller.

as an estimate of the value to which $t^{-1}\log(p(\sigma_t)/p_i(\sigma_t))$ is close for t sufficiently large. Moreover, $\overline{D}_{p|p_i}(\sigma_t)$ appears extremely stable across the independent replicas of the cases we consider, suggesting that the selection argument of Corollary 2.1 can be applied. When $p \leq \pi_1$ or $p \geq \pi_2$, under-reaction learning is equivalent to Bayesian learning, irrespective of the value of λ_i. In the parameter domain where Proposition 4.1 applies, $p \in (\pi_1, \pi_2)$, the average relative entropy decreases as the degree of under-reaction increases. This monotonic decreasing relationship is a novelty; it is a fresh new feature, not prescribed by the results of the previous Section. The performances of the moving average learning process, built on top of the considered under-reaction processes, are reported as dotted lines in Figure 1. When the under-reacting reference persistently mixes the two models, the usage of a moving average learning process is fruitful. On the contrary, when the under-reacting reference settles on the best i.i.d. model, the moving average process does

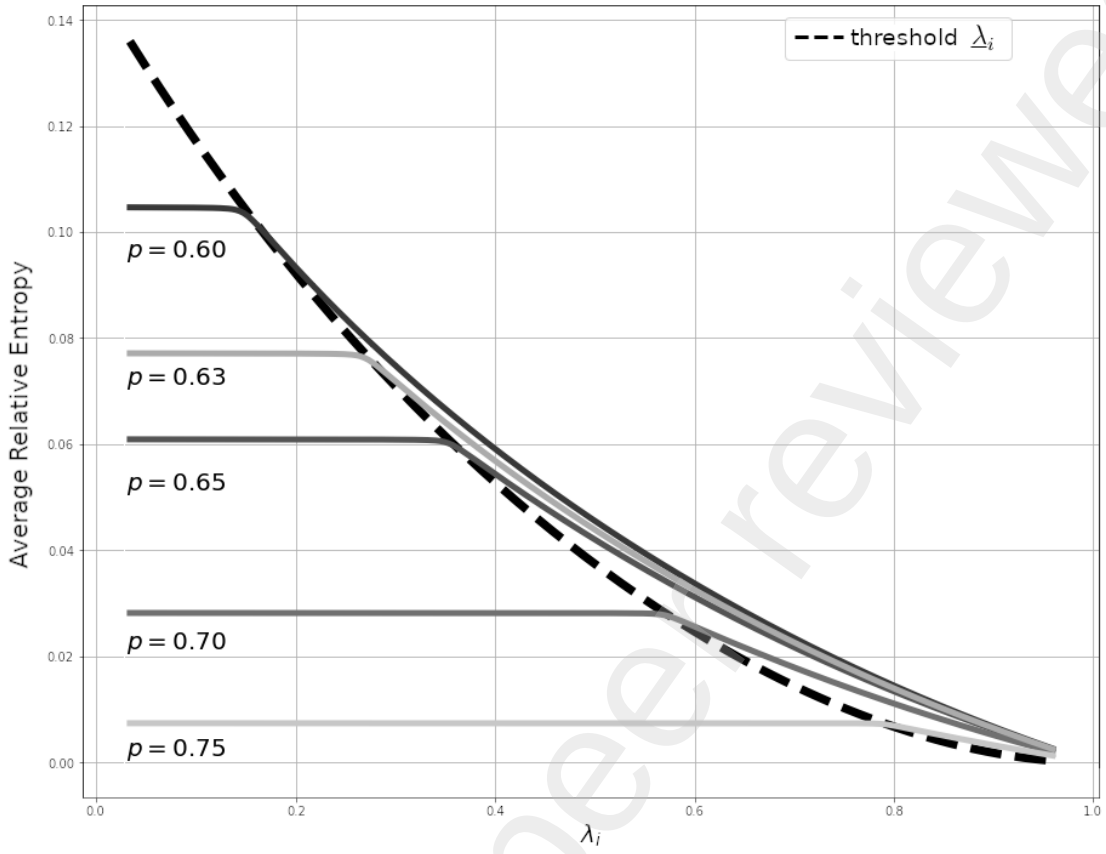

Figure 2: Average relative entropy of an under-reacting agent as a function of λ_i and for different values of p. Parameter settings: $\pi_1 = 0.3$, $\pi_2 = 0.8$. The values of p has been chosen such that $\overline{D}_{p|\pi_2}(\sigma) < \overline{D}_{p|\pi_1}(\sigma)$ holds. Standard errors are in the order of 10^{-4} or smaller.

not deliver any selection advantage. Finally, because the two considered models are linearly independent, the average relative entropy of the limited memory Bayesian learner is bounded away from zero (black dashed line in Figure 1). It is worth to point out that, in case p belongs to a specific sub-interval of (π_1, π_2) and discount factor homogeneity holds, this simple model is able to make a Bayesian agent vanish. Anyhow, it succumbs to an agent showing a sufficiently high level of under-reaction.

If the value of the parameter p characterizing the true measure is close to the one of the two models, π_1 or π_2, the reduction of the relative entropy due to under-reaction does not seem to click-in immediately when reducing λ. To further investigate this point, in Figure 2 we report the average relative entropy of an under-reacting agent i as a function of λ_i, for different values of p. We consider

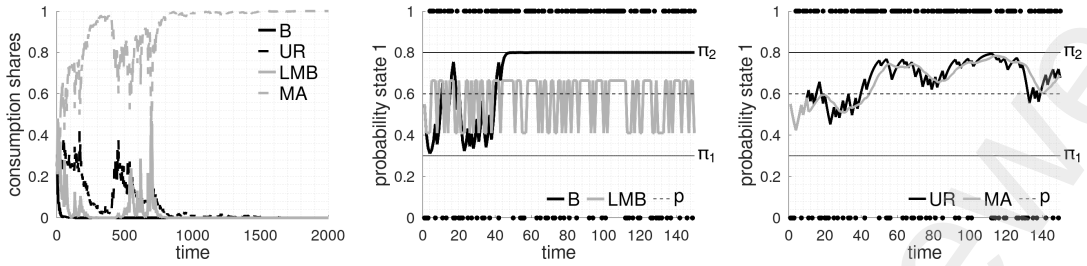

Figure 3: **Left**: consumption share dynamics of agent 1 (B), agent 2 (UR), agent 3 (LMB), agent 4 (MA). **Center**: $p_1(1|\sigma_t)$ (B) and $p_3(1|\sigma_t)$ (LMB) for the first 150 time steps. **Right**: $p_2(1|\sigma_t)$ (UR) and $p_4(1|\sigma_t)$ (MA) for the first 150 time steps. Black dots represent on 1 represent the occurrence of $s_t = 1$, while black dots on 0 represent $s_t = 2$.

the same number and length of partial histories used for Figure 1. For any value of p, it seems that there exists a threshold value $\underline{\lambda}_i$ such that, as λ_i increases beyond it, the monotonically decreasing behavior appears. For $\lambda_i \leq \underline{\lambda}_i$, the under-reacting model behaves exactly as the Bayesian one (i.e. $\lambda_i = 0$). Indeed, when $\lambda_i \to 0$, the under-reacting agent behaves as a Bayesian one and converges to a single model. Thus, the threshold $\underline{\lambda}_i$ represents the point in which under-reaction starts to play a role and the agent starts persistently mixing both models. Intuitively (see also the discussion in Massari, 2020), this should happen when the mixing coefficient λ_i is large enough for the mixture of the two models to start having a lower average entropy then the best model. Thus, considering the case in which $D_{p|\pi_2} < D_{p|\pi_1}$, the threshold value $\underline{\lambda}_i$ should solve the equation $D_{p|\underline{\lambda}_i \pi_2 + (1-\underline{\lambda}_i)\pi_1} = D_{p|\pi_2}$. Since $\pi_1 < p$, $D_{p|\pi_2}$ is a decreasing function of π_1, thus there exists a number $\tilde{\pi}_1 \in (\pi_1, p)$ such that $D_{p|\tilde{\pi}_1} = D_{p|\pi_2}$. By direct substitution one can verify that the value $\underline{\lambda}_i = (\tilde{\pi}_1 - \pi_1)/(\pi_2 - \pi_1)$, which is reported as a dashed line in Figure 2, fulfills the requirement.

4.1.2 Dynamics of consumption shares and subjective probabilities

To study how the characteristics of the learning process we have seen above shape the dynamics of consumption shares we conduct a market selection exercise. As in the previous Subsection, we set $K = S = 2$, $\pi_1 = 0.3$, $\pi_2 = 0.8$. The market is populated by 4 agents: agent 1 is Bayesian; agent 2 under-reacts with $\lambda_2 = 0.65$, agent 3 is a limited memory Bayesian learner, agent 4 uses the moving average learning process with $M_4 = 10$ exploiting the predictions of the under-reaction learning process of agent 2. Following Bottazzi and Giachini (2022), we assume $e_i(\sigma_t) = e > 0 \ \forall i, t, \sigma$ and $u_i(c) = (1-\beta)\log(c/e) \ \forall i$, with β the homogeneous

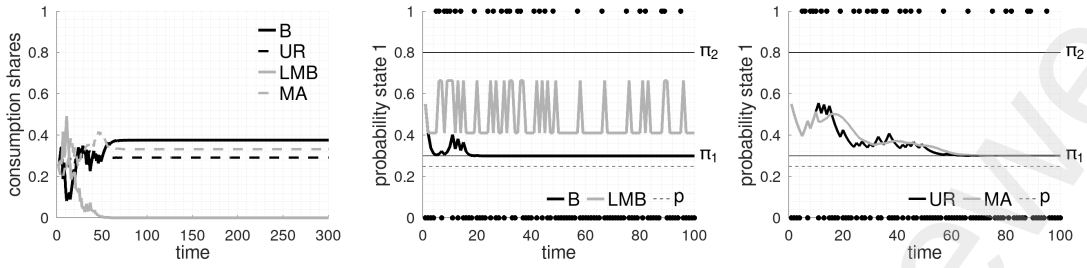

Figure 4: **Left**: consumption share dynamics of agent 1 (B), agent 2 (UR), agent 3 (LMB), agent 4 (MA). **Center**:$p_1(1|\sigma_t)$ (B) and $p_3(1|\sigma_t)$ (LMB) for the first 100 time steps. **Right**: $p_2(1|\sigma_t)$ (UR) and $p_4(1|\sigma_t)$ (MA) for the first 100 time steps. Black dots represent on 1 represent the occurrence of $s_t = 1$, while black dots on 0 represent $s_t = 2$.

discount factor. Hence, the consumption share of agent i at σ_{t+1} is

$$\tilde{c}_i(\sigma_{t+1}) = \frac{c_i(\sigma_{t+1})}{4e} = \frac{p_i(s_{t+1}|\sigma_t)c_i(\sigma_t)}{\sum_{j=1}^{4} p_j(s_{t+1}|\sigma_t)c_j(\sigma_t)} \quad \forall i, t, \sigma\,, \tag{7}$$

with $\tilde{c}_i(\sigma_0) = 0.25$. A random sequence of states is drawn assuming that the true probability process is i.i.d. with $p(1|\sigma_t) = p \in (0, 1)\ \forall t, \sigma$. For our first exercise, we set $p = 0.6$, such that, from Figure 1, one has that agent 4 (moving average) is the most accurate trader in the market. Accordingly, Figure 3 left panel shows the convergence towards 1 of agent 4's consumption share. On the other hand, agent 1 is the first to approach a null consumption share, while the last to vanish appears to be agent 2. Looking at subjective probabilities attached to state 1 (Figure 3 center and right panels), one notices that agent 1 converges to model 2 quite quickly. Agents 2, 3, and 4, instead, persistently fluctuate. However, while agents 2 and 4 tend to stay between the truth and the best model displaying a rather smooth path, agent 3 strongly jumps between its two levels.

For the second exercise, we draw the random sequence of states setting $p = 0.25$. In this case, Figure 1 shows that agents 1, 2, and 4 achieve the same level of average relative entropy, while agent 3 is less accurate than them. As one can observe in Figure 4, consumption shares stabilize quite quickly on their long-run level and, while agents 1, 2, and 4 show a strictly positive share, agent 3 vanishes. Looking at subjective probabilities, one notices that agents 1, 2, and 4 converge to model 1, while agent 3 fluctuate between its two levels in the (π_1, π_2) interval. Since $p < \pi_1$, selecting model 1 is the best possible choice and only those agents able to do that survive.

4.2 Structural misspecification

A more general process for the true probability measure is now considered. Indeed, here we assume that the states of nature follow a Markov process with transition matrix composed by strictly positive entries.

Assumption 3. The true measure p follows a discrete-time Markov chain with transition matrix P: $p(s_{t+1}|\sigma_t) = P_{s_t,s_{t+1}}$ $\forall t, \sigma$ and $p(s|\sigma_0) = p_{s,0}$ with $p_{s,0} > 0$ $\forall s \in \{1, 2, \ldots, S\}$. For any $(s, s') \in \{1, 2, \ldots, S\} \times \{1, 2, \ldots, S\}$, it is $P_{s,s'} > 0$.

The strict positiveness of the transition matrix's entries implies that the Markov chain defining the true probability measure p is irreducible and, as a consequence, the invariant probability distribution $\boldsymbol{\pi} = (\pi(1), \pi(2), \ldots, \pi(S))$, with $\pi(s) > 0$ $\forall s$, exists unique (see Feller, 1968, page 393). The invariant distribution emerges in the computation of models' average relative entropy and, in turn, this is useful to understand the accuracy of Bayesian learning.

Bayesian Learning. When the truth follows a Markov chain as in Assumption 3, one can explicitly compute the limiting value of the average relative entropy of any i.i.d. model. Indeed, the following holds.

Proposition 4.3. *Ror any i.i.d. model k, it is p-almost surely*

$$\lim_{t\to\infty} \overline{D}_{p|\pi_k}(\sigma_t) = \overline{D}_{p|\pi_k}(\sigma) = \sum_{s=1}^{S} \pi(s) \log \frac{\pi(s)}{\pi_k(s)} + \sum_{s'=1}^{S} \pi(s') \sum_{s=1}^{S} P_{s',s} \log \frac{P_{s',s}}{\pi(s)} \,. \quad (8)$$

Proof. From the definition of $\overline{D}_{p|\pi_k}(\sigma_t)$ one has p-almost surely

$$\overline{D}_{p|\pi_k}(\sigma) = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} \sum_{s=1}^{S} p(s|\sigma_\tau) \log \frac{p(s|\sigma_\tau)}{\pi_k(s)} = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} \sum_{s=1}^{S} P_{s_\tau,s} \log \frac{P_{s_\tau,s}}{\pi_k(s)} =$$

$$= \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=1}^{t} \sum_{s'=1}^{S} \mathbb{1}_{s',s_\tau} \sum_{s=1}^{S} P_{s',s} \log \frac{P_{s',s}}{\pi_k(s)} = \sum_{s'=1}^{S} \pi(s') \sum_{s=1}^{S} P_{s',s} \log \frac{P_{s',s}}{\pi_k(s)} \,,$$

where $\mathbb{1}_{s',s}$ represents the indicator function ($\mathbb{1}_{s',s} = 1$ if and only if $s' = s$ and 0 otherwise) and the last equality is an application of the Strong Law of Large Numbers. The equation in the statement directly follows adding and subtracting $\sum_{s=1}^{S} \pi(s) \log \pi(s)$ and exploiting the properties of the invariant distribution, i.e. $\pi(s) = \sum_{s'=1}^{S} P_{s',s} \pi(s')$ $\forall s$. □

Proposition 4.3 shows that the average relative entropy of an i.i.d. model in a Markov world results from the sum of two components: the relative entropy of the model with respect to the "best" i.i.d. distribution – the invariant distribution of the chain – and the average relative entropy of the "best" i.i.d. model with respect to the transition probabilities. Thus, from Proposition 3.1, one has that a Bayesian agent is asymptotically as accurate as the i.i.d. model with the lowest relative entropy with respect to the invariant distribution. Moreover, a Bayesian becomes increasingly inaccurate as the true Markov model is increasingly divergent from the invariant distribution. This can eb understood as the loss of accuracy a Bayesian agent suffers because of structural misspecification.

Learning with under-reaction. Concerning under-reaction, an application of Proposition 3.2 combined with the results of Massari (2020) delivers that the under-reacting agent maintains a *specific* advantage over the Bayesian one. However, we cannot provide an extension to the Markovian case of its *generic* advantage delivered by Proposition 4.1 in the parametric misspecification case. The intuition here is that unconditionally averaging predictions becomes less fruitful when the truth is Markov. Indeed, in so doing one is combining predictions without discriminating the fact that probabilities change depending on the realized state. Thus, a trade-off emerges: on the one hand, some form of averaging may allow the learner to get closer to the best i.i.d model (i.e. the invariant distribution), but, on the other hand, dampening fluctuations may be counterproductive when the true probabilities naturally fluctuate.

Limited memory Bayesian learning. Concerning the limited memory Bayesian learning process, the situation is rather different. Indeed, the limitation in the number of observations the agent adopts makes its predictions show the Markovian property. Thus, in those cases in which the i.i.d. models and the initial weights are such that the resulting probabilistic prediction are close to true transition probabilities, it can show an high level of accuracy. More specifically, under Assumption 3 and for an agent i that uses the limited memory Bayes protocol in eq. (5), it is p-almost surely

$$\lim_{t\to\infty} \overline{D}_{p|p_i}(\sigma_t) = \overline{D}_{p|p_i}(\sigma) = \sum_{s'=1}^{S} \pi(s') \sum_{s=1}^{S} P_{s',s} \log \frac{P_{s',s}}{\sum_{k=1}^{K} \pi_k(s) \frac{\pi_k(s') w_{i,k}(\sigma_0)}{\sum_{k'=1}^{K} \pi_{k'}(s') w_{i,k'}(\sigma_0)}} . \tag{9}$$

Hence, if the transition probabilities can be written as a particular convex combination of the i.i.d. models, then the average relative entropy of the limited memory

Bayesian agent is zero. That is, $\overline{D}_{p|p_i}(\sigma) = 0$ if $P_{s',s} = \sum_{k=1}^{K} z_{k,s'} \pi_k(s) \; \forall s', s$, with

$$z_{k,s'} = \frac{\pi_k(s') w_{i,k}(\sigma_0)}{\sum_{k'=1}^{K} \pi_{k'}(s') w_{i,k'}(\sigma_0)} \quad \forall s', k \,.$$

As a consequence, learning processes that do not provide relevant selection advantages in the parameter misspecification case, may become effective as structural misspecification occurs. In the case of the limited memory Bayesian, the key is its continuous resetting of the learning process. Such a peculiar behavior conveys a structure to predictions that cannot be recovered otherwise.

Moving average learning. The argument used in Proposition 4.2 to show the advantage of the moving average approach under parametric misspecification cannot be generically extended when Assumption 3 holds. Since the basic mechanism underlying moving average learning is, again, smoothing, the intuition follows the same logic of the one provided for learning with under-reaction: a trade-off between getting closer to the the invariant distribution and matching fluctuations emerges. Hence, it could be generically possible that smoothing the conditional of the underlying model, a moving average learning agent under-performs with respect to its reference.

Summarizing, structural misspecification makes the task of ranking learning processes in terms of survival prospects even harder. Indeed, the aforementioned trade-offs let the selection picture become much blurrier than the case of parametric misspecification. To support, validate, and better understand the intuitions provided here, in the following section we perform a numerical exercise. It clearly shows that selection outcomes can generically and profoundly change depending on how one sets the parameters of the true Markov chain.

4.2.1 Numerical exploration

For our numerical exercise, we consider the same settings used in subsection 4.1.1 with the exception of the true probability. That is, we set $K = S = 2$, $\boldsymbol{\pi}_1 = (\pi_1, 1 - \pi_1)$, $\boldsymbol{\pi}_2 = (\pi_2, 1 - \pi_2)$, and states of nature appear according to a Markov chain with transition matrix

$$P = \begin{bmatrix} P_{1,1} & 1 - P_{1,1} \\ P_{2,1} & 1 - P_{2,1} \end{bmatrix} .$$

Thus, $p(1|\sigma_t) = p(1|s_t) = P_{s_t,1} \; \forall t, \sigma$. The invariant distribution reads

$$\boldsymbol{\pi} = \left(\frac{P_{2,1}}{1 - P_{1,1} + P_{2,1}} , \frac{1 - P_{1,1}}{1 - P_{1,1} + P_{2,1}} \right)$$

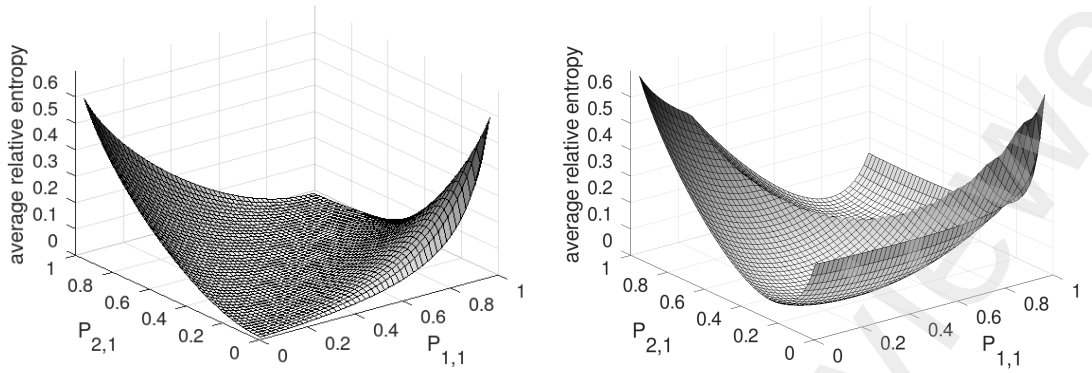

Figure 5: Average relative entropy of the invariant distribution (left) and of a Bayesian agent (right) for different combinations of $P_{1,1}$ and $P_{2,1}$. Parameter settings: $\pi_1 = 0.3$, $\pi_2 = 0.8$.

and we shall use its average relative entropy as a reference point throughout the analysis. Such a quantity can be analytically computed and in left panel of Figure 5 one can read its values for different combinations of $(P_{1,1}, P_{2,1})$. In the right panel, instead, the average relative entropy of a Bayesian agent is showed. One immediately notices that the loss in accuracy one suffers by using the best i.i.d. model when the truth is Markov progressively grows as we move away from the $P_{2,1} = P_{1,1}$ line. In the case of a Bayesian agent, one recovers the shape of the solid black line in Figure 1 along the $P_{2,1} = P_{1,1}$ line. Moving towards the corners $(P_{1,1}, P_{2,1}) = (0, 1)$ and $(P_{1,1}, P_{2,1}) = (1, 0)$, the average relative entropy progressively grows as a consequence of structural misspecification.

Next, we show in Figure 6 the difference between the average relative entropy of each learning process and the average relative entropy of the invariant distribution. To compute the average relative entropy of the under-reacting agent and of the moving average agent, we rely upon a numerical estimation, details are provided in the caption of each Figure. Even in this case, the reported numbers can be understood as estimates of the values to which $t^{-1} \log(p(\sigma_t)/p_i(\sigma_t))$ is close for t sufficiently large and, given the extremely stability across replicas, the selection argument of Corollary 2.1 can be applied. As expected, the Bayesian agent cannot be more accurate than the invariant distribution and, thus, the surface shown in the top-left panel is completely in the positive part of the graph. Comparing under-reaction (top-right panel) with moving average built on it (bottom-left panel), one appreciates the trade-off between dampening fluctuations in predictions to get closer to the best i.i.d. model and keeping changing predictions in order to match transition probabilities. Focusing on the regions of the $(P_{1,1}, P_{2,1})$ space where the Markov chain favors switching, $P_{1,1} \simeq 0$ and $P_{2,1} \simeq 1$, one notices that

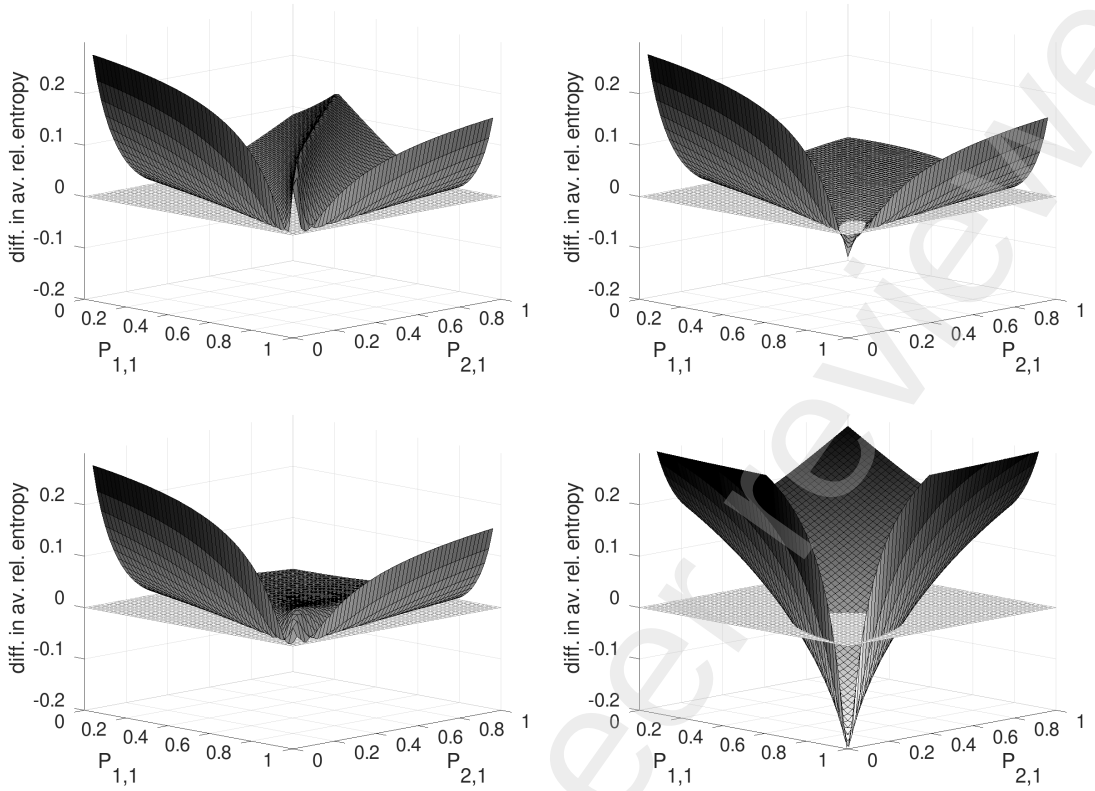

Figure 6: Differences between average relative entropy of the learning process and the average relative entropy of the invariant distribution. **Top-left**: Bayesian learning. **Top-right**: under-reaction with $\lambda = 0.65$. **Bottom-left**: moving average agent with $M = 20$ exploiting the under-reaction with $\lambda = 0.65$. **Bottom-right**: limited memory Bayesian learning. Parameter settings: $\pi_1 = 0.3$, $\pi_2 = 0.8$, $w_{i,1}(\sigma_0) = w_{i,2}(\sigma_0) = 0.5$. The average relative entropy of under-reaction and moving average have been estimated over 200 independent realizations of 2500 steps each. For estimated values, standard errors are in the order of 10^{-4} or smaller. The plots are rotated of 90° clockwise with respect to Figure 5 in order to improve the visualization of results.

moving average is more accurate than the under-reacting process it is exploiting. Thus, averaging seems to provide a gain in accuracy when the underlying process jumps frequently between states. In such a situation, the trade-off is solved in favor of averaging. At the same time, the accuracy of the invariant distribution appears as an upper bound: the two learning processes cannot improve upon the best i.i.d. model. In the opposite case of a persistent Markov chain averaging appears less advantageous and letting predictions fluctuate can provide superior

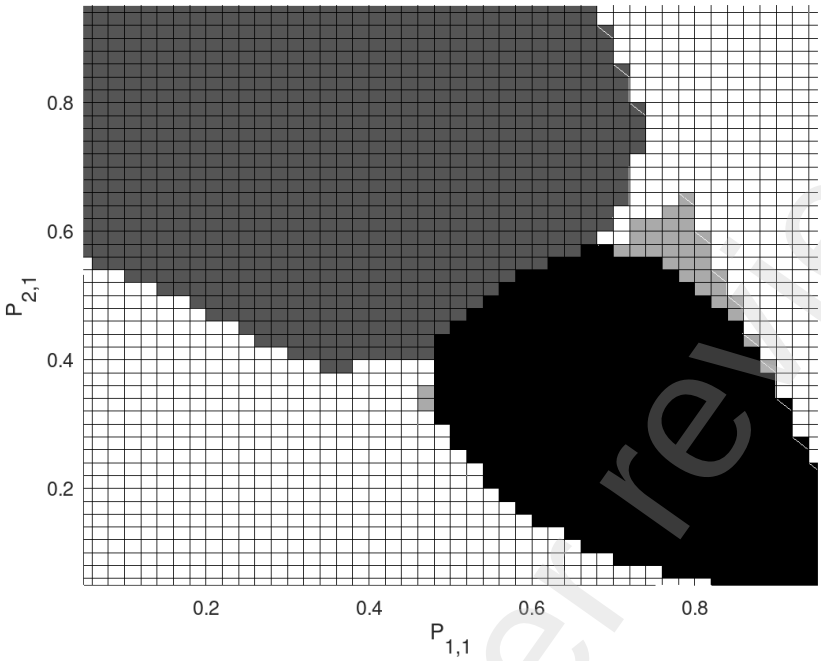

Figure 7: Most accurate processes over the $(P_{1,1}, P_{2,1})$ space. **White**: Multiple maximally accurate processes or the difference between the two lowest average relative entropy processes is not significant at $\sim 99\%$ confidence level. **Light gray**: under-reaction with $\lambda = 0.65$ is the most accurate. **Dark gray**: moving average with $M = 20$ exploiting under-reaction is the most accurate. **Black**: limited memory Bayesian learning is the most accurate.

outcomes. Indeed, when $P_{1,1} \simeq 1$ and $P_{2,1} \simeq 0$, under-reaction outperforms moving average and can even be more accurate than the invariant distribution. Hence, the trade-off is solved towards matching the fluctuations of true probabilities. This is even clearer looking at the performance of the limited memory Bayesian process (bottom-right panel). Such a learning rule provides the best outcome in terms of accuracy – outperforming each other learning process and the invariant distribution – when the underlying Markov chain is sufficiently (and symmetrically) persistent. At the same time, the limited memory Bayesian learning process presents the most extreme levels of average relative entropy: when the true Markov chain is not sufficiently persistent it performs worse than the other rules.

To complement and further support the previous analysis, Figure 7 shows the identity of the most accurate agent over the $(P_{1,1}, P_{2,1})$ space. As argued in advance, the moving average process exploiting under-reaction is the most accurate when the true Markov chain frequently switches between states, while the limited

memory Bayesian process is the most accurate when the Markov chain is persistent and close to be bi-stochastic. Interestingly, under-reaction prevails over a region in which $P_{1,1} \simeq 0.8$ and $P_{2,1} \simeq 0.6$. Following our intuition about the trade-off, in that region under-reaction is able to achieve the best combination between averaging to get close to the invariant and fluctuating to follow transition probabilities. In the regions around $P_{1,1} = P_{2,1} = 0$ and $P_{1,1} = P_{2,1} = 1$ and spanning for most of the white areas in the plot, from a direct comparison with Figure 6, one can conclude that Bayes, under-reaction, and moving average achieve the same (maximal) accuracy level.

4.2.2 Consumption shares dynamics

Finally, we propose some market selection exercises along the same lines of Subsection 4.1.2 but assuming that the true data generating process is a Markov chain as in Subsection 4.2.1. That is, we set $K = S = 2$, $\pi_1 = 0.3$, $\pi_2 = 0.8$ and we populate the market with the same 4 agents of Subsection 4.1.2. Thus, the consumption share of agent i at σ_{t+1} remains as in equation (7).

For our first exercise we consider a random sequence of states generated setting $P_{1,1} = 0.15$ and $P_{2,1} = 0.75$. From Figure 7, one immediately notice that such a point belongs to the region where the moving average agent (agent 4) has the lowest average relative entropy. As a consequence of Corollary 2.1 and given discount factor homogeneity, we shall observe that agent 4 dominates. The left panel of Figure 8 shows the evolution of agents' consumption shares in the first 300 steps. Consistently with our results, one observes that the consumption share of agent 4 approaches one around time-step 200 and stabilizes on such a level afterwards. One also notice that the order in which agents approach zero consumption is: first agent 3, second agent 1, and third agent 2. Looking at how the predictions for state 1 of the agents evolve over the first 100 time steps (central and right panel of Figure 8), one notices that agent 1 approaches model 1 quite quickly, agent 2 shows small fluctuations around a sort of long-run trend, agent 3 strongly fluctuates between its two predictions, agent 4 captures the long-run trend of agent 2.

Next, we draw our random sequence of states setting $P_{1,1} = 0.75$ and $P_{2,1} = 0.15$, such that the underlying Markov chain is highly persistent and close to be bistochastic. Figure 7 indicates that the limited memory Bayesian learning process generates the lowest average relative entropy. Indeed, as expected, in Figure 9 left panel, agent 3 attains a (almost) unitary consumption share in less than 100 steps. Looking at how beliefs evolve – Figure 9 center and right panels –, one notices that agent 1 settles on model 1 after few fluctuations; agent 2 and 4 persistently fluctuate but the persistence of states do not allow them to rapidly adapt when a switch occurs; agent 3, instead, quickly moves and remains close to the true probabilities as sequences of equal states alternate.

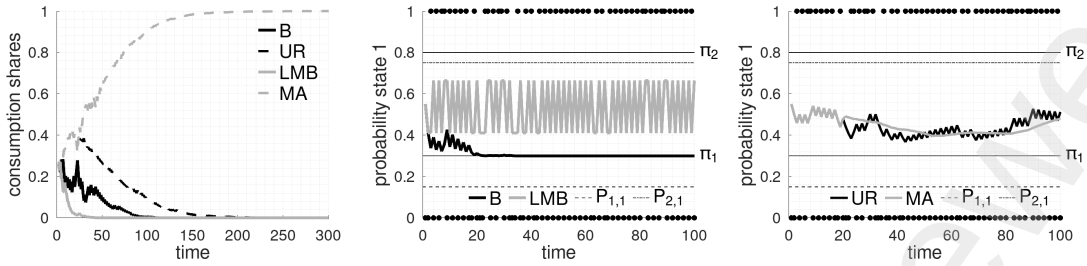

Figure 8: **Left**: consumption share dynamics of agent 1 (B), agent 2 (UR), agent 3 (LMB), agent 4 (MA). **Center**:$p_1(1|\sigma_t)$ (B) and $p_3(1|\sigma_t)$ (LMB) for the first 100 time steps. **Right**: $p_2(1|\sigma_t)$ (UR) and $p_4(1|\sigma_t)$ (MA) for the first 100 time steps. Black dots represent on 1 represent the occurrence of $s_t = 1$, while black dots on 0 represent $s_t = 2$.

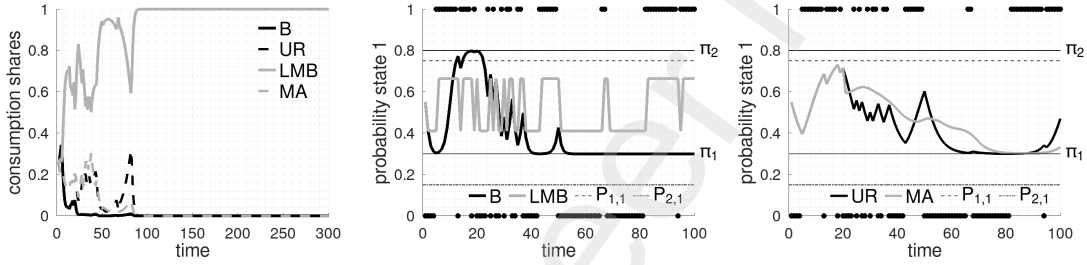

Figure 9: **Left**: consumption share dynamics of agent 1 (B), agent 2 (UR), agent 3 (LMB), agent 4 (MA). **Center**: $p_1(1|\sigma_t)$ (B) and $p_3(1|\sigma_t)$ (LMB) for the first 100 time steps. **Right**: $p_2(1|\sigma_t)$ (UR) and $p_4(1|\sigma_t)$ (MA) for the first 100 time steps. Black dots represent on 1 represent the occurrence of $s_t = 1$, while black dots on 0 represent $s_t = 2$.

Finally, we set $P_{1,1} = 0.35$ and $P_{2,1} = 0.2$ such that we are in a case of multiple survivors according to Figure 7. Figure 10 left panel confirms that: in less than 100 steps agent 1, 2, and 4 stabilize on positive and heterogeneous consumption shares. Interestingly, the under-reacting agent is the one achieving the highest share. Agent 3's consumption share, instead, goes to zero and the reason for that is evident from the center and right panels of Figure 10. Indeed, the probability that agent 3 assigns to state 1 continues to fluctuate between two levels that are outside the range defined by the transition probabilities. The other agents, instead, converge to model 1, whose predictions, lying in-between transition probabilities, turn out as the most accurate. This also explains why their consumption shares stabilize: they become identical.

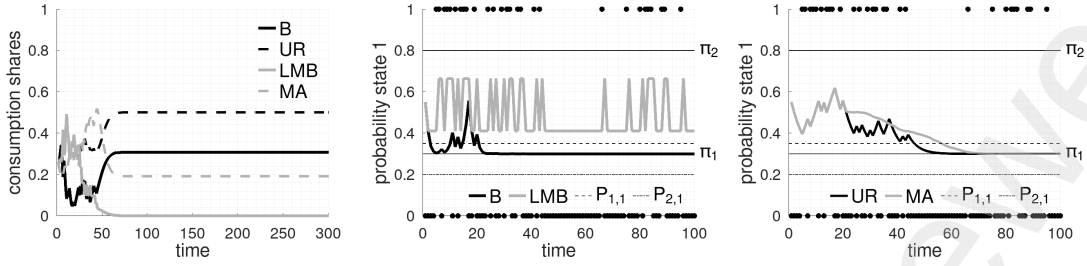

Figure 10: **Left**: consumption share dynamics of agent 1 (B), agent 2 (UR), agent 3 (LMB), agent 4 (MA). **Center**: $p_1(1|\sigma_t)$ (B) and $p_3(1|\sigma_t)$ (LMB) for the first 100 time steps. **Right**: $p_2(1|\sigma_t)$ (UR) and $p_4(1|\sigma_t)$ (MA) for the first 100 time steps. Black dots represent on 1 represent the occurrence of $s_t = 1$, while black dots on 0 represent $s_t = 2$.

5 General survival behaviors under model misspecification

The previous analyses show that the kind of model misspecification the agents face crucially affects the selection outcomes one shall observe in the long-run. Indeed, learning rules with low survival prospects under parametric misspecification, can dominate when structural misspecification is considered. However, this does not imply that examples of belief formation rules that persist in the market no matter the data generating process are nonexistent.[4] For instance, consider the averaging approach: in the previous sections, we investigated the performance of an agent who averages the predictions of one specific reference learning model, discussing the advantages and disadvantages it entails. However, if instead of averaging over time the conditional predictions of a given reference model, an agent averages the likelihoods of all the other learning processes in the market, such an agent survives on any path.

Proposition 5.1. *Assume that $\beta_i = \beta \ \forall i$ and that agent N assigns to partial history σ_t the weighted arithmetic mean of the probability of the other agents,*

$$p_N(\sigma_t) = \sum_{i=1}^{N-1} v_i p_i(\sigma_t), \forall \sigma_t, \tag{10}$$

with $v_i > 0$, $\forall i = 1, \ldots, N-1$ and $\sum_{i=1}^{N-1} v_i = 1$. Then, agent N survives on all σ.

[4]Intuitively, survival is guaranteed by any way of assigning probabilities to future events that makes an agent always consume its endowment.

Proof. Note that if $\lim_{t\to\infty} c_N(\sigma_t) = 0$, that is if $\lim_{t\to\infty} u'_N(c_N(\sigma_t))^{-1} = 0$, then $\lim_{t\to\infty} \sum_{i=1}^{N-1} c_i(\sigma_t) = e(\sigma_t) > 0$, that is $\liminf_{t\to\infty} \sum_{i=1}^{N-1} v_i u'_i(c_i(\sigma_t))^{-1} > 0$. Thus, if agent N vanishes on σ,

$$\lim_{t\to\infty} \frac{u'_N(c_N(\sigma_t))^{-1}}{\sum_{i=1}^{N-1} v_i u'_i(c_i(\sigma_t))^{-1}} = \lim_{t\to\infty} \frac{p_N(\sigma_t)}{\sum_{j=1}^{N-1} v_i p_i(\sigma_t)} = 0.$$

However, this is impossible, as $p_N(\sigma_t)/\sum_{i=1}^{N-1} w_i p_i(\sigma_t) = 1\ \forall t$. □

Recursively substituting (3) in (2), it is immediate to realize that (10) is just Bayesian learning with prior $(v_1, \ldots, v_{N-1})$ over the learning models of the other market participants. Moreover, if Assumption 1 holds for any agents $i = 1, \ldots, N-1$, then for agent N, behaving as in (10), it is

$$p_N(s_{t+1} \mid \sigma_t) = \sum_{k=1}^{K} \pi_k(s_{t+1}) w_{N,k}(\sigma_t) \text{ with } w_{N,k}(\sigma_t) = \sum_{i=i}^{N-1} w_{i,k}(\sigma_t) \frac{v_i p_i(\sigma_t)}{p_N(\sigma_t)}, \forall k, t.$$

Hence, Assumption 1 is also valid for agent N and the weights it used to construct its conditional probabilities are the likelihood-weighted averages of the weights used by the other agents in the market. Thus, behaving as in (10), an agent is able to let the accuracy of its beliefs asymptotically match the highest level achieved by market participants and this provides a survival advantage. Basically, it can adapt its predictions to the best performing ones on every given sequence of events. When the most accurate learning process is unique, this implies that such an agent learns the best learning process in the market. In the examples of the previous Section, for instance, an agent whose individual probabilities are as in (10) would asymptotically behave as the moving average agent in the market of Figure 8 and as the limited memory Bayesian in the market of Figure 9. In the market of Figure 10, where the 3 most accurate agents converge to predict according to model 1, an agent behaving as in (10) would asymptotically use model 1.

The selection result obtained here perfectly mirrors the necessary and sufficient condition for a trader to vanish that Massari (2017) recovers under homogeneous discount factors: an agent vanishes if and only if its beliefs are less accurate than those obtained via Bayesian learning over agents' models. Thus, if an agent behaves as in (10), it cannot vanish. The learning behavior generated by (10) shares some similarities with the *Follow the Market Strategy* (FMS) proposed by the same author. It prescribes to make condition predictions according to next-period risk neutral probabilities. Hence, defining $q(s_{t+1} \mid \sigma_t) = q(\sigma_{t+1})/q(\sigma_t)$, for an agent i forming beliefs according to the FMS, it is

$$p_i(s_t \mid \sigma_{t-1}) = \frac{q(s_t \mid \sigma_{t-1})}{\sum_{s=1}^{S} q(s \mid \sigma_{t-1})}. \tag{11}$$

Simple computations show that the behavior prescribed by (11) is equivalent to the one implied by (10) if all the agents in the economy have logarithmic Bernoulli utility. Otherwise, they are generically different. Massari (2017) shows that, in economies populated by CRRA traders with the same intertemporal discount factor and risk aversion parameter, if the trader with highest $p_i(\sigma_t)$ changes infinitely often and agents have distinct beliefs, the FMS trader: *i*) dominates if market participants are more risk averse than log, *ii*) survives without dominating if market participants have logarithmic utility; *iii*) vanishes if market participants are less risk averse than log. If there is a unique trader (or model) that assigns the highest likelihood to a sequence (as in the examples of the previous sections), then the FMS ensures survival, since it ends up imitating the most accurate behavior. In such a case, Massari (2017) argues that survival is guaranteed even by the *Follow the Leader Strategy* (FLS), that prescribes to imitate the most accurate trader in the market and for an agent i it is defined as

$$p_i(s_t \mid \sigma_{t-1}) = \begin{cases} p_j(s_t \mid \sigma_{t-1}) & j : p_j(\sigma_{t-1}) = \operatorname{argmax}_n \{p_n(\sigma_{t-1})\}, \\ \dfrac{\sum\limits_{j \in \mathcal{K}_{t-1}} p_j(s_t \mid \sigma_{t-1})}{|\mathcal{K}_{t-1}|} & \text{if ties occur}, \end{cases} \tag{12}$$

where $\mathcal{K}_{t-1}$ is the set of agents whose beliefs have the highest likelihood. If, instead, the trader with highest likelihood change infinitely often, then an agent i behaving as in (12) vanishes (Massari, 2017).

The selection results derived by Massari (2017) hold no matter the type of model misspecification affecting the market. Hence, comparing the learning behavior prescribed by (10), (11), and (12) one notices a common feature: to be implemented they require a trader to know (at least) the learning processes of all the other agents in the market.[5] This is in stark contrast with the learning processes investigated in the previous sections, where no information about the whole market ecology is needed to implement them. Hence, the examples and the discussion presented here indicate that high survival prospects under model misspecification could be more related to the amount of information one trader possesses about the other agents populating the market rather than to how sophisticated its learning mechanism is.

6 Conclusions

In this paper we study market selection in a complete-market Arrow-Debreu economy considering four learning processes and two cases of model misspecification:

[5] Actually, the FMS is more demanding than the other two: it also requires information about risk preferences, intertemporal discount factors, and endowments.

parametric and structural misspecification.

Our analysis shows that providing a general ranking of learning processes with respect to their survival prospects is difficult and the characteristics of the model misspecification problem strongly influence selection outcomes. Under parametric misspecification, learning processes build upon an averaging approach have a selection advantage over generic regions of the parameter space. Such an advantage partially disappears when structural misspecification occurs. Indeed, learning approaches that have lower fitness than averaging ones under parametric misspecification can become very accurate under structural misspecification. In the cases we analyze, the intuition is that a trade-off between approximating the best i.i.d. model and capturing the persistent fluctuations in true conditional probabilities emerges.

Even if one can generate examples of learning rules that allow an agent to survive no matter the kind of model misspecification characterizing the economy, those approaches appear to require the knowledge of fundamental details concerning all the other market participants. Given the natural connection between selection outcomes and long-run asset prices, a parsimonious long-run asset evaluation model robust to model misspecification appears out of reach, since one cannot dispense from knowing important features about all the agents in the economy.

Disclosure statement

Declarations of interest: none.

References

Antico, A., G. Bottazzi, and D. Giachini (2023). On the evolutionary stability of the sentiment investor. In D. Bourghelle, P. Grandin, F. Jawadi, and P. Rozin (Eds.), *Behavioral Finance and Asset Prices: The Influence of Investor's Emotions*, pp. 155–173. Cham: Springer International Publishing.

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

Beddock, A. and E. Jouini (2021). Live fast, die young: equilibrium and survival in large economies. *Economic Theory 71*(3), 961–996.

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

Blume, L. and D. Easley (2006). If you are so smart why aren't you rich? Belief selection in complete and incomplete markets. *Econometrica 74*, 929–966.

Blume, L. and D. Easley (2009a). The market organism: Long-run survival in markets with heterogenous traders. *Journal of Economic Dynamics and Control 33*, 1023–1035.

Blume, L. and D. Easley (2009b). Market selection and asset pricing. In T. Hens and K. Schenk-Hoppé (Eds.), *Handbook of Financial Markets: Dynamics and Evolution*. North-Holland (Handbooks in Economics Series).

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.

Bottazzi, G. and D. Giachini (2022). Strategically biased learning in market interactions. *Advances in Complex Systems 25*(02n03), 2250004.

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.

Elmiger, S. (2020). Market selection and the absence of arbitrage. *Available at SSRN 3746988*.

Epstein, L. G., J. Noor, and A. Sandroni (2010). Non-bayesian learning. *The BE Journal of Theoretical Economics 10*(1), 1–20.

Evstigneev, I., T. Hens, and K. Schenk-Hoppé (2009). Evolutionary finance. In T. Hens and K. Schenk-Hoppé (Eds.), *Handbook of Financial Markets: Dynamics and Evolution*. North-Holland (Handbooks in Economics Series).

Evstigneev, I., T. Hens, and K. R. Schenk-Hoppé (2016). Evolutionary behavioral finance. In *The handbook of post crisis financial modeling*, pp. 214–234. Springer.

Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*, Volume I. John Wiley and Sons: New York.

Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Volume II. John Wiley and Sons: New York.

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*(1), 451–482.

Hens, T. and K. Schenk-Hoppé (2005). Evolutionary finance: introduction to the special issue. *Journal of Mathematical Economics 41*, 1–5.

Holtfort, T. (2019). From standard to evolutionary finance: a literature survey. *Management Review Quarterly 69*(2), 207–232.

Jouini, E. and C. Napp (2011). Unbiased disagreement in financial markets, waves of pessimism and the risk-return trade-off. *Review of Finance 15*(3), 575–601.

Kogan, L., S. A. Ross, J. Wang, and M. M. Westerfield (2006). The price impact and survival of irrational traders. *The Journal of Finance 61*(1), 195–229.

Kogan, L., S. A. Ross, J. Wang, and M. M. Westerfield (2017). Market selection. *Journal of Economic Theory 168*, 209–236.

Massari, F. (2017). Markets with heterogeneous beliefs: A necessary and sufficient condition for a trader to vanish. *Journal of Economic Dynamics and Control 78*, 190–205.

Massari, F. (2020). Under-reaction: Irrational behavior or robust response to model misspecification? *Available at SSRN 3636136*.

Orseau, L., T. Lattimore, and S. Legg (2017). Soft-bayes: Prod for mixtures of experts with log-loss. In *International Conference on Algorithmic Learning Theory*, pp. 372–399. PMLR.

Perisastry, M. and V. M. Murty (1982). Bounds for the ratio of the arithmetic mean to the geometric mean. *The Two-Year College Mathematics Journal 13*(2), 160–161.

Sandroni, A. (2000). Do markets favor agents able to make accurate predictions. *Econometrica 68*(6), 1303–1341.

Sandroni, A. (2005). Efficient markets and bayes' rule. *Economic Theory 26*(4), 741–764.

Savage, L. J. (1954). *The foundations of statistics*. John Wiley and Sons.