Journal of Economic Behavior & Organization 112 (2015) 266–288

ELSEVIER

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

The evolution of social learning and its economic consequences☆

Benjamin Bossan[a,*], Ole Jann[b], Peter Hammerstein[a]

[a] Humboldt University Berlin, Institute for Theoretical Biology, Invalidenstrasse 43, 10115 Berlin, Germany
[b] University of Copenhagen, Department of Economics, Øster Farimagsgade 5, 1353 Copenhagen K, Denmark

ARTICLE INFO

Article history:
Received 26 March 2013
Received in revised form 26 January 2015
Accepted 28 January 2015
Available online 14 February 2015

Keywords:
Imitation
Learning strategies
Evolution
Natural selection
Market inefficiencies

ABSTRACT

We use an evolutionary model to simulate agents who choose between two options with stochastically varying payoffs. Two types of agents are considered: individual learners, who rely on trial-and-error methods, and social learners, who imitate the wealthiest sampled individual. Agents adapt to changing environments within one generation by using their respective learning strategy. The frequency of the agent types adapts between generations according to the agents' acquired wealth. During the course of evolution, social learning becomes dominant, resulting in three major effects: First, for better or worse, the decisions of social learners are more exaggerated than those of individual learners. Second, social learners react with a delay to changes in the environment. Third, the behavior of social learners becomes more and more detached from reality. We argue that our model gives insights into economic systems and markets.

1. Introduction

A large part of what separates us from most other animals is arguably the ability to learn from one another, and not just from nature. Precious little of our knowledge is strictly gained from personal experience, and much is learned from someone else (who, in turn, probably also had learned it from others) – either by direct instruction or by imitation. Focusing on the latter, we can see that there are fields in which our behavior is governed purely by imitation: Very few people invent their own articles of clothing and food without ever having seen anyone else wear or consume them before, and a whole range of 'appropriate' behaviors that govern our daily routine was picked up through observation and imitation.

Such imitative learning is also part of economic behavior. One in three institutional investors, who often receive considerable remuneration for their knowledge and independent judgment, say that their stockbroker was "influential" in their decision to buy a stock. Only one in four said that they had "conducted their own analysis" (Shiller and Pound, 1989).[1] Imitation influences the competitiveness in Cournot markets (Huck et al., 1999; Apesteguia et al., 2007).

When humans learn from others, they imitate the successful. Even three- and four-year olds preferentially learn from successful models (Birch et al., 2008). Investors place a disproportionate amount of money in funds whose returns in the past 12 months were exceptional – even though these funds tend to do much worse than average in the following period

☆ We are grateful to Robert Boyd, Peter Norman Sørensen, Nicolaas Vriend (the editor) and two anonymous referees for helpful comments. B.B. and P.H. thank the Deutsche Forschungsgemeinschaft (SFB 618) for funding.

* Corresponding author. Tel.: +49 30 20938649.
E-mail addresses: b.bossan@gmail.com (B. Bossan), ole.jann@econ.ku.dk (O. Jann), p.hammerstein@biologie.hu-berlin.de (P. Hammerstein).

[1] Since the survey was conducted by questionnaire, social learning is likely under- and individual research over-reported.

http://dx.doi.org/10.1016/j.jebo.2015.01.010
0167-2681/

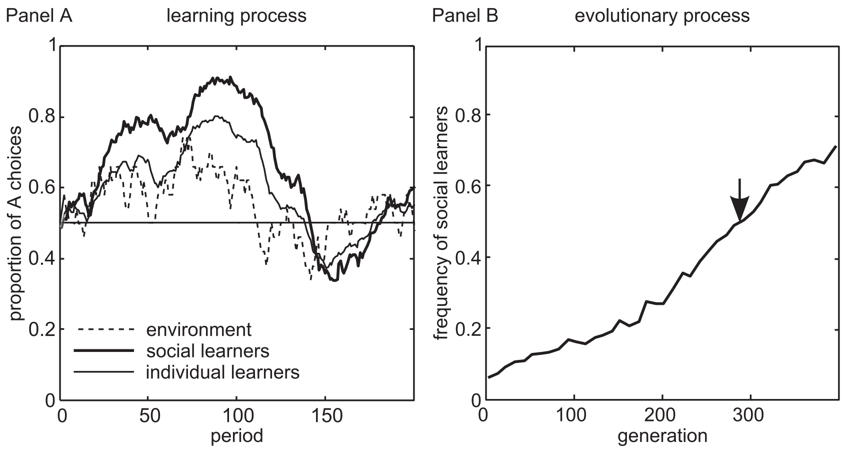

Fig. 1. (A) Exemplary simulation run with a population consisting of 50% social and 50% individual learners. (B) Exemplary evolutionary run showing the frequency of social learners.

(Economist, 2011). And the history of bubbles, from tulipomania to subprime, is the story of outsiders becoming market participants because they watch others make huge profits (Shiller, 2003; Tett, 2009).

But given the tendency of people to adopt successful behaviors or strategies, these cultural traits should become more frequent over time. New variations of learning strategies may enter the market through creative processes or mistakes, whereas scarcity of capital creates competition that will weed out unsuccessful learning strategies. These three ingredients – variation, competition, and preferential replication of successful strategies – are sufficient to give rise to an evolutionary process. E.g. the investment strategy invented by a successful portfolio manager survives and is copied by others, while unsuccessful strategies disappear – either because their agents now have less capital at their disposal, leave the market, or choose to imitate the successful agents. Although this process is not natural (Darwin, 1859) but cultural, it will still have similar features (cf. Weibull, 1995; Gintis, 2000; Hammerstein and Hagen, 2005; Mesoudi et al., 2006). This leads to the evolutionary theory of economic behavior which Alchian (1950) has described in his seminal article in 1950.[2]

The model we have in mind is that of a society in which thousands of individuals engage in different learning strategies, and the frequency of different strategies is itself dependent on how successful these strategies have been in the past. This model, with stochastic underlying values, independent choices of all individuals, and the interconnectedness through learning from others, is too complex to lend itself to analytical solution methods. That is why we engage in agent-based simulations. We thoroughly check these simulations for robustness to changes in the set-up and parameter values.

Our work is based on the assumption that imitation of success gives rise to the forces of evolution. Instead of introducing a presumably rational or irrational strategy and study the resulting phenomena, we observe which learning strategy prevails and how abundant it becomes in equilibrium. Then we analyze how the behavior of a population in equilibrium relates to the environment. Our model is discussed in Section 2, a detailed elaboration of our findings is given in Section 3, the connection of these findings with real world phenomena and comparisons with existing models is part of Section 4.

Using the evolutionary framework, we assume that agents with distinct learning strategies adopt different behaviors, interact with other agents in the population, and reproduce according to the fitness (or "wealth") they aggregate during their lifetime. A quick presentation of the model can be found in Fig. 1. The environment is characterized by two options, *A* and *B*, whose payoffs cannot be directly observed and fluctuate over time. Consequently, none of the options will always be better than the other. Within each generation, individuals adapt by repeatedly choosing one of the options according to a fixed learning strategy. Between generations, those individuals with the most success replicate at a higher rate, which leads to an adjustment of the frequency of the learning strategies. Therefore, our model has a nested structure: *What* an individual learns varies within each generation and is adapted to the environment; *how* the individuals learn varies between generations and is adapted to the composition of the population. The model thus differs from other models in that the behavior of the individuals flexibly adjusts according to the learning strategy (panel A), whereas the learning strategies adjust over a larger time scale (panel B). A detailed model description can be found in Section 2.

Imitative behavior has been studied previously in the context of informational cascades (Banerjee, 1992; Bikhchandani et al., 1992; Anderson and Holt, 1997; Ziegelmeyer et al., 2010). Informational cascades may arise if choices are made sequentially and previous choices by others can be observed. In most of the models, one individual after the other has to decide between two options. The decision can be based on a private signal that is noisy and on observation of the choice of previous individuals. After a certain amount of individuals have chosen one option over the other, the reliability of the information

[2] Alchian himself has described his approach as "reverting to a Marshallian type of analysis combined with the essentials of Darwinian evolutionary natural selection" (Alchian, 1950, fn 7 on p. 113).

derived from observing previous choices will exceed the reliability of the private signal. It is then individually rational for all subsequent individuals to disregard their personal information and follow the others, triggering an informational cascade.

The model we propose is not just another theory of herding or informational cascades. There are several choices to be made over time, and in-between observing the behavior of others and choosing yourself, the environment may have changed, rendering the previous observations obsolete. Thus, in our model, it is not trivially true that one should blindly follow the majority despite its choice conflicting with personal judgment. Whether the benefits acquired by imitation outweigh the costs depends on the composition of the population. We find that the more imitators there are, the less it pays to imitate. In this way, we mirror earlier findings (Rogers, 1988).

Informational cascades can lead to bad aggregate behavior if the first few individuals accidentally chose the worse option. Still it is more likely than not that a cascade ends up with the better option. Therefore, if the game is repeated several times, we should expect the better option to be chosen more frequently than expected by chance. In our model, it may easily happen that the majority of the population chooses the worse option for a long stretch of time.

Imitation is a common and powerful learning strategy for coping with changing environments, a finding that has repeatedly been made (see Boyd and Richerson, 1985; Richerson and Boyd, 2005, and references therein) and which is supported by our model. In an economic context, this has also been studied by Vriend (2002), who finds that simulated agents who learn socially by various strategies can also exhibit, as an overall population, behavior that is detached from reality.

The question we address here is how we should expect aggregate behavior to differ if imitation is a major source of information acquisition. This has, for instance, implications for the efficiency of markets.

2. Model description

2.1. Basic structure

A number of simulated agents faces a task that requires them to make decisions based on available information. More precisely, they have to choose repeatedly between two options with different payoff prospects,[3] *A* and *B*. These could be thought of as stocks and bonds with risky returns or as two consumption goods from different brands whose payoff is measured in utility derived from their use (say, PC vs Mac). Any two types of options whose performance is uncorrelated will do.

In a given period *t*, options from class *A* give a return of 1 with probability p_A and a return of 0 otherwise, while options from class *B* give a return of 1 with probability p_B. Expected payoffs are hence p_A and p_B. The experiences of the agents are independent, even if they have chosen the same option. Thus one agent could receive a return of 1 after choosing *A* in period *t*, while another agent receives 0 despite choosing the same option.

Choices have to be made repeatedly. After every period, the option is discarded and the agents have to pick again. But p_A and p_B also change after every period. They increase by an increment (0.02 in the standard parameter setting) with probability $1 - p_i$ and decrease by the same increment otherwise. The probabilities therefore have a tendency to decrease if larger and increase if smaller than 0.5 – they revert to the mean.

p_A and p_B are unknown to the agents at all time. However, after each period, they can observe whether they received a payment or not, and from that they can draw conclusions as to how large the success probability of the chosen type of option might be. The agents hence find themselves in a situation of uncertainty: They do not know the expected returns from either option, and they could only correctly determine them by sampling from each many times. During that time, however, the underlying probabilities will already have changed. Agents therefore never perfectly know the underlying probabilities that characterize the environment and have to find other ways to deal with this situation of Knightian uncertainty (Knight, 1921).

We model two types of agents, who follow different strategies:

Individual Learners base their decision only on their own experience and not on the behavior of others. In particular, they practice reinforcement learning (cf. Kaelbling et al., 1996), i.e. they pick their assets according to a trial-and-error-method. Since p_A and p_B are fluctuating, the choice between *A* and *B* has to be updated constantly. Individual learners, therefore, have a memory that allows them to draw on earlier experiences, discounting it exponentially to account for gradual changes in the environment. The discount rate was set to 0.9, which preliminary analysis showed to be optimal.

Social Learners do not rely on their own experience, but learn strictly from other agents. For this, they sample a number of agents in every period and pick the option that the wealthiest among those sampled chose – a strategy that we call "imitate the wealthiest" (ITW).

Another approach to model individual learners would be to use Bayesian learning instead. We could show, however, that even if Bayesian learners were given a huge informational advantage, they were at best very slightly better than individual

[3] This is similar to previous works (Rogers, 1988; Feldman et al., 1996; Henrich and Boyd, 1998; Wakano et al., 2004; McElreath et al., 2008; Kendal et al., 2009), which saw *A* and *B* as options which represented the natural environment.

learners. The reason for this is that the Bayesian learning process is confronted with too much noise to form the reasonably accurate beliefs compared to the speed at which the environment changes. As Bayesian learning was not found to be a huge improvement over reinforcement learning by exponential discounting, all the while being much more complex and unrealistic, we used reinforcement learning as the individual learning strategy. More details on reinforcement learning are given below in Section 2.2.1; more information on Bayesian individual learning is in section A of the Appendix.

We decided to have social learners imitate the wealthiest individuals and not those with the highest short-run payoff for two reasons. First, since there are only two payoffs in our model (0 and 1), simply imitating individuals who have been successful in the last period does not really provide much information, as they might have been lucky or unlucky. In any case, there would always be several lucky and unlucky individuals among those observed, so that social learning would become almost purely stochastic, instead of really allowing for the imitation of success. Second, we argue that wealth (as the sum of past successes) can actually be easier to observe in the real economy than short-term success. Who knows whether a stock trader had a good or bad day yesterday, or whether a large company will make money on its latest production decision? But the stock trader's penthouse and six-seven-figure bonus, and the companies' capital surplus and stock market value, are there for all to see.

We also studied other social learning strategies as well as strategies that mix individual and social learning that were proposed in the literature. Overall, we found that pure social learning was the most successful strategy. More on those forms of social learning can be found in section B of the Appendix.

Our model is evolutionary in the sense that the frequency of the different learning strategies is adjusted after each generation according to how well the strategies perform. A generation consists of 50 periods, during which the strategies learn and adjust their choice. Depending on their success, they accumulate wealth, which is translated to fitness. When a generation ends, each agent contributes a number of offspring to the offspring pool that is equal to the agent's fitness. Of this offspring pool, 10,000 offspring are drawn at random, so that the total population size remains constant at 10,000.[4] This algorithm is called Fisher-Wright process (Fisher, 1930; Wright, 1931) and is commonly used for evolutionary modeling. Offspring inherit from their parent only the learning strategy and their first period's choice, mutations do not occur.

There are thus two sources of stochasticity in our model. First, the individual's probabilities to succeed depend on p_A and p_B but are independent, so that two individuals with the same choice can still face different outcomes. When two individuals choose different options, the one who chooses the better option is more likely to succeed than the other but the opposite may also occur. Second, natural selection itself has a random component. An individual with a higher lifetime success than another individual has a higher number of expected offspring but may at the end contribute less offspring to the next generation. By choosing a high population size (10,000) and a low base fitness, we make sure, though, that performance on the choice task and not random drift (Ohta, 1992) is the main determinant in guiding selection.

All Simulations were run on Matlab 7.0 and 8.0 (The MathWorks, Inc.).

2.2. The agents' learning strategies

2.2.1. Individual learning – reinforcement

There are several ways to simulate a strategy of reinforcement learning. In its most simplistic form, it would amount to a win-stay lose-shift strategy (Robbins, 1952). As the name suggests, this strategy would consist of sticking with an asset type when it had a positive return in the previous period and switching otherwise. However, performance would be low: As we model our environment as stochastic, a failure with asset type A after a sequence of successes with it should not promptly lead to a switch. The fact that A was good in the past periods is an indicator that it is good in the present period; switching may be premature. To allow individual learners to take account of that, we give them a memory of past events. If A led to many successes in the previous periods, the propensity towards choosing A increases and a single miss does not suffice to induce a switch.

For the purpose of our model, we assume that individuals have a propensity, P_i, for each option i. This propensity is increased if the option is reinforced; in our case, that means it is reinforced by the amount R_i if the option yields success. Here R_i is equal to 1 if the option was chosen and yielded success and −1 otherwise. In addition, the propensity of each period depends on the propensity of the last period but discounted by a decay factor, q. That gives the propensity for option i in period t as

$$P_i(t) = q \cdot P_i(t-1) + R_i(t-1)$$

under the assumption that, as in our model, choices are made at the end of a period and determine success in the next period.

The probability $\Pr_A(t)$ of choosing A in period t then is given by the function

$$Pr_{\mathrm{A}}(t) = \frac{\exp(P_{\mathrm{A}}(t))}{\exp(P_{\mathrm{A}}(t)) + \exp(P_{\mathrm{B}}(t))} = \frac{1}{1 + \exp(-\Delta P(t))} \quad (1)$$

[4] This number is estimated to be the effective population size of humans during the last 1 million years (Takahata, 1993; Harding et al., 1997).

where

$$\begin{aligned}\Delta P(t) &\equiv P_A(t) - P_B(t) \\ &\equiv q \cdot \Delta P(t-1) + R_A(t-1) - R_B(t-1).\end{aligned} \tag{2}$$

When A yielded a success, $R_A(t-1) = 1$, while a failure with A results in $R_A(t-1) = -1$. The same is true for $R_B(t-1)$ when B is chosen. The propensity of the option that was not chosen remains the same. Therefore, $R_A(t-1) - R_B(t-1)$ equals 1 if A yielded success or B did not and it equals -1 if A did not yield success or B did. We can thus define:

$$R(t-1) \equiv \begin{cases} 1 & \text{A successful or B not} \\ -1 & \text{B successful or A not} \end{cases}$$

Inserting this in the definition of ΔP (Eq. (2)), we get:

$$\Delta P(t) = q \cdot \Delta P(t-1) + R(t-1) \tag{3}$$

Inserting Eq. (3) in Eq. (1), we get:

$$\Pr_A(t) = 1 - \Pr_B(t) = \frac{1}{1+\exp(-\Delta P(t))} = \frac{1}{1+\exp(-q \cdot \Delta P(t-1) - R(t-1))}$$

An additional modification of reinforcement learning can be implemented. The sensitivity factor, λ, alters the steepness of the probability of choosing an option as a function of the propensity of that option (see e.g. in Klucharev et al., 2009):

$$\Pr_A(t) = 1 - \Pr_B(t) = \frac{1}{1+\exp(-\lambda \cdot \Delta P(t))}$$

The higher λ, the steeper the adjustment of the probability.

When $\lambda = \infty$, this results in a simple threshold function that prescribes to choose A if propensity for A exceeds propensity for B and vice versa:

$$\Pr_A(t) = \begin{cases} 1 & \Delta P(t) > 0 \\ 0.5 & \Delta P(t) = 0 \\ 0 & \Delta P(t) < 0 \end{cases}$$

We found performance not to increase when using a $\lambda \neq \infty$ in contrast to just using the threshold function. Therefore, we just used the threshold function and ignored the additional parameter λ.

2.2.2. Social learning – imitate the wealthiest

Social learners make their decision after observing the (last-period) decisions of others. In our model, we not only allow social learners to observe the decisions of others (which would only allow some sort of conformist strategy), but also let them observe the number of successes of the individuals – i.e. their wealth. The social learners then imitate the last-period choice of the most successful individual they observe – a strategy we call "imitate the wealthiest" (ITW). Unless otherwise stated, the particular realization of this strategy that we choose for our work is to sample seven individuals and choose the same option as the wealthiest among them chose in the previous period. In reality, information about the aggregated payoff could be derived by observing another individual's lifetime reproductive success, her prestige in the population, or simply her total monetary wealth.

Observing the choices of others would also a conformist strategy to exist (Boyd and Richerson, 1985; Henrich and Boyd, 1998; Giraldeau et al., 2002). Conformists observe what options other individuals choose and adopt the most frequent option with a probability greater than what would be expected from random sampling. In robustness checks, we found conformism to perform much worse than ITW. Earlier works have studied strategies called "imitate the best" (Vega-Redondo, 1997; Schlag, 1999; Huck et al., 1999; Apesteguia et al., 2007) but such strategies correspond more closely to payoff-biased social learning, as only the last period's payoff and not the total payoff is used as information. ITW is instead more akin to prestige-biased social learning (Henrich and Gil-White, 2001), but note that our model does not include idiosyncratic differences except for learning strategies and personal histories of decisions and outcomes, so that prestige cannot, e.g. depend on the personal skill level.

Payoff-biased social learning may at first seem to be a simpler alternative to ITW. It requires the observation of the short-term success (income) of others instead of the long-term success (wealth). Information about the long-term success is, however, more readily available than information about the short-term success. Whether someone drives a big car, wears designer suits, or owns a big house, i.e. facts that can be easily observed, is more dependent on wealth than on income. The permanent income hypothesis (Friedman, 1957) further supports the view that short-term fluctuations in income should have little impact on consumption decisions; consumption can therefore not be used to infer short-term income changes. Firms like banks, on the other hand, are often only required to disclose their financial data on a quarterly basis or even less frequently; privately held companies often do not disclose them at all. Therefore, it is difficult for competitors to infer the short-term impact of business strategies.

From a modeling perspective, payoff-biased social learning is also more complicated than ITW. As, after each period, each observed individual is either successful or not (payoff of 1 or 0), many individuals will be tied for the highest payoff. A simple decision mechanism as "imitate the most successful in this period" is thus impossible. Instead, some form of integration of the observations is necessary. For example, the average income of each option could be calculated (McElreath et al., 2008); or the social learner could sum up the total success of each option and then compare them. In any case, these mechanisms require more sophisticated cognitive processes and are therefore less likely to be applied. Nevertheless, we tested such payoff-biased social learning strategies as well. We discuss those in more detail in the Appendix.

We assume that information on the history of choices of other individuals is unavailable, as well as information on the strategy used by other individuals. That means that social learners cannot base their choice on the last three choices of a sampled individual, just as it is usually possible to observe someone's current behavior, but not necessarily past behaviors in real world situations (in a changing environment such as ours, it would usually also not be advisable to look back for too many periods). Social learners are also not able to exclusively sample individual learners, since we assume that while the current choices of others are observable, their underlying strategy is not. As the number of social learners in the population increases, it therefore becomes more likely that social learners imitate others who, themselves, have not learned by experience but have only been imitating others.

Most economic models of learning have assumed that individuals pay a certain cost to learn which of the options is better (cf. Grossman and Stiglitz, 1976, 1980). The same is true for most gene-culture coevolution models (Rogers, 1988; Feldman et al., 1996; Kameda and Nakanishi, 2002; Wakano et al., 2004; Enquist et al., 2007; Kendal et al., 2009), which are similar in nature to our model. In such models, the cost of social learning is assumed to be lower than the cost of learning individually. The main advantage of social learning is then the lower cost of acquiring the information itself. We do not implement exogenous costs of learning. This is not because we think they do not exist, but rather because we want to focus on the *informational* advantage of learning socially. For instance, individual learning might result in a certain probability to choose the better of the two options. If social learning can improve this probability, the social learner would have an advantage because she faces reduced uncertainty and not because of lower learning costs. Our analysis thus shifts the emphasis from saving the costs of learning individually to the benefits of learning socially.

2.3. Description of a simulation run

This section describes the chronology of a simulation. Pseudocode for the simulations can be found in section E of the Appendix. The parameters used in the description are the standard parameters used to generate most of the results in this paper. We have conducted extensive checks to make sure that our results are robust to changes in the parameters. Robustness checks were done with Latin hypercube sampling (see section B of the Appendix for robustness checks).

2.3.1. Initiation

10,000 agents are generated, 9000 of them individual learners and 1000 social learners. They all receive an endowment ("base fitness") of 10. p_A and p_B, the success probabilities of the two options, are start at 0.5.

2.3.2. What happens in period t?

At the beginning of every period, the agents observe whether their choice was successful in the last period, i.e. whether they received a payoff of 1 or 0 in period $t-1$. This payoff is then added to their wealth. After each agent has observed her own payoff, she makes another choice between *A* and *B*.

The individual learners make this choice based on their previous experience (see explanations on reinforcement learning above). Each social learner randomly samples seven other agents (if not stated otherwise), and adopts the last-period choice of the wealthiest among these agents as her own choice. In the first period of each new generation, the newborn agents, not yet having made any observation, make the same choice as their "parent" in the last period. In the very first period of the simulation, all agents make a random choice between the two options, since they have no information upon which they could act.

Parallelly (and unobservably for the agents), p_A and p_B change according to the mean-reverting random-walk process described above. This means that the success probabilities of the two agents have changed slightly from when the agents made their choices.

Now, after each agent has chosen either *A* or *B* and the new probabilities p_A and p_B are set, the payoff of each agent is realized. If an agent has chosen *A*, for example, she will receive a payoff of 1 with probability p_A and a payoff of 0 with probability $1-p_A$. Then the next period, $t+1$, begins. During a generation, agents never change their strategy (individual vs social learning), only their choices (*A* or *B*) change.

2.3.3. What happens between generations?

Every 50 periods (if not stated otherwise), all agents "die" and get replaced by 10,000 new agents. The information that the agents have collected so far is eliminated with them, and the new agents do not possess any information – just like the new-born agents in the first period. Wealth is also not inherited, all agents start with the same base endowment.

The new-born agents are also either individual or social learners, and the frequency of each type is determined by an evolutionary process. Specifically, every agent in the new generation is the "offspring" of an agent from the old generation,

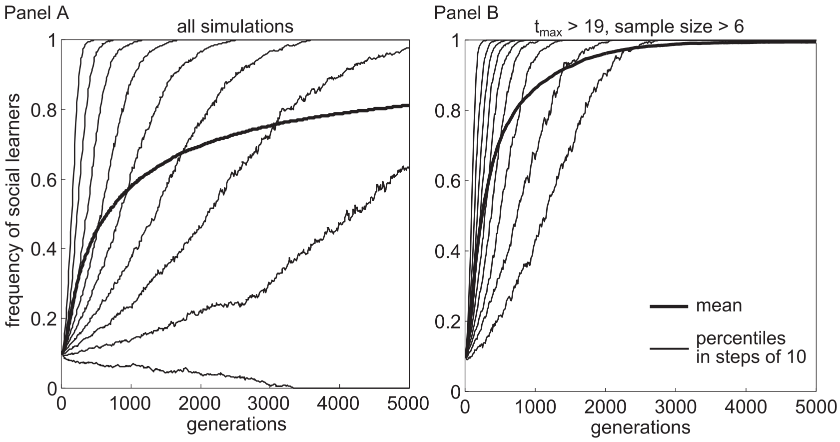

Fig. 2. Mean (thick line) and percentiles in steps of 10 (thin lines) of the frequency of social learners over time. Panel A: all 1000 simulations. Panel B: only for the parameter space with generation length $t_{max} > 19$ and sample size >6.

where the parent's type (i.e. strategy) and first-period choice but not her information or wealth is inherited by her offspring. The more wealth an agent has assembled at the end of the generation, the more offspring she will produce. Evolution occurs according to a Fisher–Wright process (Fisher, 1930; Wright, 1931): Every agent has one offspring for each unit of wealth she possesses at the end of her lifetime. From this combined pool of offspring (usually consisting of several hundred thousand agents), 10,000 agents are randomly chosen to comprise the next generation.

As a simple example, assume that there are 5000 individual and social learners each, but that after 50 periods the social learners have on average 50% more wealth than the individual learners. Then the expected distribution in the next generation is 4000 individual and 6000 social learners; the real distribution can differ because the offspring are drawn randomly from the offspring pool.

So while the type of every agent (and therefore the frequency of agent-types in the population) stays constant throughout a period, the next generation will usually have a different composition of individual and social learners. Especially if there are very few agents of one type left, and their wealth is not hugely superior to that of the agents of the other type, these agents can be wiped out entirely by the stochastic evolutionary process.

2.3.4. End of the simulation

The simulation ends after 5000 generations (i.e. 250,000 periods), or when one type of agent dies out and all 10,000 agents have the same type.

3. Results

3.1. Evolution of social learners

Before analyzing the consequences of abundant social learning, we have to establish whether social learning would indeed be favored by an evolutionary process. To do this, we ran 1000 simulations with parameter values randomly drawn through Latin hypercube sampling. The parameters we changed were the mean value of p_A and p_B, the reversion factor of the environment, the size of steps at which the environment changes, the number of periods per generation, the sample size of social learners, and the influence of skill (see section B of the Appendix for more details). The main results are shown in Fig. 2.

Panel A shows the mean frequency of social learners, as well as percentiles in steps of 10, from all 1000 simulations. The mean frequency reaches 81.2% after 5000 generations, with a slight upwards trend even at the end of the 5000 generations. The frequency of social learners at the end of the simulation rarely takes intermediate values. Instead, their frequency after 5000 generations was less than 1% in 118 of the 1000 simulations and more than 99% for 689 simulations. This means that either social learners or individual learners dominate, with few cases of intermediate equilibria.

The apparent scarcity of stable equilibria might seem puzzling at first. Theoretically, individual and social learners should coexist if their performance is roughly identical. The performance of individual learners is frequency-independent, since they will be right or wrong the same number of times regardless of how many other individual or social learners there are. The performance of social learners, on the other hand, is frequency-dependent, since a higher frequency of social learners in the population makes it more likely that a social learner samples other social learners. At a high frequency, social learners, mostly imitating other social learners, should start to lose touch with reality and perform worse than individual learners. Therefore, stable equilibria should be expected.

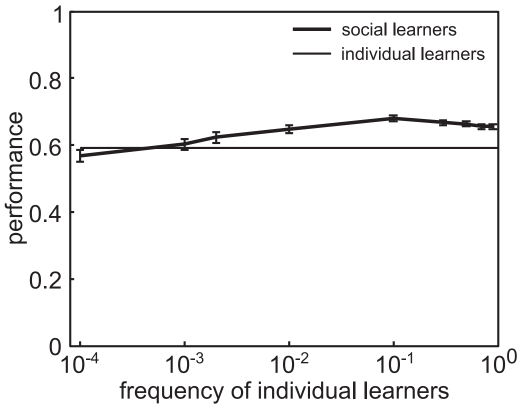

Fig. 3. Performance of individual and social learners as a function of their frequencies (each frequency was simulated 1000 times).

To explain why we find few cases of coexistence, have a look at Fig. 3. It shows the mean performance of individual and social learners from 1000 simulations (standard errors of the mean are shown for social learners) using the default parameter settings. An equilibrium is to be expected at the intersection of the lines representing the performance of social and of individual learners. It lies between 0.2% and 0.01% individual learners. This equilibrium is technically stable, because social learners perform worse when more frequent and better when less frequent. However, at such a low frequency, individual learners are frail to become extinct due to stochastic effects. The apparent scarcity of stable equilibria is thus caused by the equilibrium frequency of social learners being so close to 1 that stochastic effects lead to extinction of individual learners.[5]

Among the 7 parameters that we changed to check our results and obtain the results presented in Fig. 2, the two most important were the number of periods per generation, t_{max}, and the sample size (see Appendix, section B). As we show in the Appendix, a small t_{max} (below 20) leads to lower frequencies of social learners because there is simply less time (measured in periods) for social learners to become more frequent, and random processes instead of correct choices determine success in the evolutionary process. Relative fitness differences, which determine the speed of evolution, are also necessarily smaller when there is a base fitness and the number of periods per generation is small.

The sample size of social learners, on the other hand, is the only tested parameter that is rather a property of the agent and not of the environment. Although we included it in the robustness analysis, arguably evolutionary forces would fix it at the optimal level, given the constraints.[6] To have a more clear picture of how sample size and t_{max} affect the frequency of social learners, we again show mean and percentiles, but this time for the parameter space of $t_{max} \geq 20$ and sample size ≥ 7 (Fig. 2, Panel B). We find the mean frequency of social learners to be 99.5%, and in 97,8% of all such simulations individual learners had died out completely by the end of the simulation. In this parameter space, in which generations are long enough to allow for differences between the strategies to make a real difference and in which social learners observe the choices of enough individual learners, social learners almost always completely replace individual learners regardless of how we set the other tested parameters of the model.

The emphasis of our work is the consequence for aggregate behavior when imitation is the dominant form of learning. Three major consequences are shown in the following sections: exaggeration, delayed response to changes, and detachment of behavior from reality.

3.2. Exaggeration

The behavior of individual learners is independent of the population that surrounds them. The behavior of social learners, however, depends on that of other individuals – their decisions (as well as their performance) are *frequency-dependent*. To study the behavior of social learners, it is thus necessary to indicate their frequency in the population. Fig. 4 illustrates the behavior of individual learners and social learners at different frequencies in a randomly generated environment. Panel A shows the behavior of individual learners (solid line, left hand scale) in a population consisting only of individual learners. We see that their choices quickly catch up with changes in $p_A - p_B$, the difference in the success probabilities of options *A* and *B* (dashed line, right hand scale). When this difference is close to zero, approximately half of the individual learners choose *A* and the other half *B*. If p_A exceeds p_B by 10%, approximately 60% of individual learners choose *A*. Linear regressions show that

[5] Of course, the simulations that led to the extinction of social learners are most easily explained: Individual learners just outperformed them on the whole range of frequencies.

[6] We put an upper limit of 11 on the sample size.

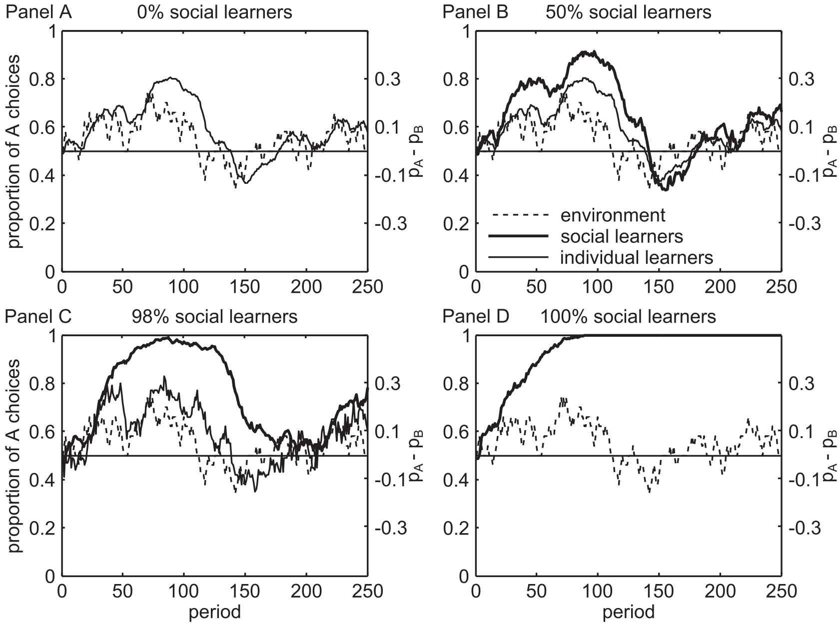

Fig. 4. Illustration of the behavior of individual and social learners.

the proportion of A choices x_A is best approximated by $x_A = 0.5 + 1.036 \cdot (p_A - p_B)$, an almost one to one correspondence. We thus say that the individual learners "match" the probabilities (Baum, 1979).[7]

Selection will act to increase the frequency of social learners to very high levels. The resulting behavior is illustrated in the other panels of Fig. 4. Panel B shows the behavior of a population consisting of 50% individual learners and 50% social learners (thick line, left hand scale). The proportion of A choices made by social learners stays close to the proportion of A choices made by individual learners. However, instead of matching the probabilities, the social learners "exaggerate".[8] Linear regression shows that a 1% point increase in $p_A - p_B$ is matched by a 1.795% point increase in A choices. Panel C displays a population consisting of 2% individual learners and of 98% social learners. Here, social learners strongly exaggerate. In the most extreme case, the population consists solely of social learners, as shown in Panel D. Lacking input from individual learners, eventually, all social learners will choose the same option and never again switch away from it. This shows that social learners require individual learners to be able to adapt to the environment.

Theoretically, the best strategy is to always choose A if $p_A > p_B$ and B otherwise. The best strategy should thus completely exaggerate even small differences in probabilities. Exaggeration is therefore efficient, as long as the direction is correct. Individual learners match probabilities quite closely, meaning that their performance will at best be moderate. In contrast, social learners tend to exaggerate, especially when they are very frequent. Social learners are thus closer to optimal behavior than individual learners, which explains why they have, at least initially, a higher fitness and become more frequent in the population.

3.3. Delay

It is already clear from Fig. 4 that individual learners react to changes in the environment with a small delay (or lag) and that this delay is larger for social learners. This section elaborates the influence of imitation on delays.

The delay in the adaption of social learners to changes in the environment is presumably frequency-dependent – the more social learners there are, the more they will learn from one another and hence the greater the delay with which they process 'actual' information provided by individual learners. To examine this, we determined the cross-correlation between behavior, measured as the proportion of A choices, and environment, measured as $p_A - p_B$. The peak of the cross-correlation indicates the average lag between behavior and environment; the later the peak, the slower the strategy is in adopting. The magnitude of the correlation indicates how strict the strategy adheres to the environment. The lower the magnitude, the less strict is the correspondence between behavior and environment.

Simulation results confirmed our expectations, as shown in Fig. 5. Individual learners start with a high cross-correlation for small delays that decreases rather fast for higher delays (confidence intervals are of the magnitude ±0.009.). This closely

[7] The reader should be aware that this is not the exact definition of matching as in the psychological literature (Herrnstein, 1961; Baum, 1979) and that it is contested how robust matching is (Vulkan, 2000).

[8] Baum (1979) refers to a similar effect as "overmatching".

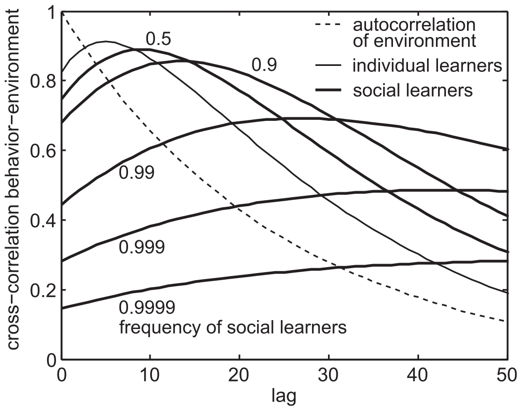

Fig. 5. Cross-correlation between behavior and environment, based on 50,000 periods; frequency of social learners as indicated.

corresponds to the autocorrelation of the environment, meaning that on average, individual learners track environmental changes very closely. For a frequency of 50%, social learners have a cross-correlation that is close to that of individual learners but already lags behind by several more periods and tracks the environment less closely. While selection works to increase the frequency of social learners, the cross-correlation of their behavior with the environment declines steadily and the peak is shifted to higher delays. Thus the behavior of social learners becomes more and more detached from environmental changes until it is hardly connected with the environment at all.

This development also entails that social learners' performance becomes worse as their overall frequency grows – not because they stop exaggerating but because their behavior starts to seriously lag behind the environmental fluctuations. Exaggeration is only efficient if the choice is exaggerated in the correct direction, which becomes more and more unlikely with increasing lag.

3.4. Relation between behavior and reality

The fundamental information that is crucial in our model is whether *A* or *B* is the better option to choose, which depends on whether p_A or p_B is currently greater. If a clever observer were to draw conclusions from the observed individuals' behavior on the underlying fundamentals, she might use the heuristic "choose *A* if more than *x* percent of market participants choose *A* and choose *B* otherwise". If *x* is chosen wisely and behavior always reflects the environmental state, this strategy should lead to correct inferences at all time. In contrast, if behavior does not reflect the environmental state, we should expect a low accuracy.

Using the described heuristic and a threshold *x* that ex post maximizes the number of correct inferences, we analyzed how often these inferences were correct as a function of the studied strategy (for more details on this analysis, see section D of the Appendix). For individual learners, we found that behavior indeed allows to infer the environmental state with high accuracy (88.19% correct). For social learners at 50% frequency, inferences were still quite accurate (82.87% correct). However, when social learners become more abundant, inference becomes less accurate (68.69% for a frequency of 99%, 61.56% for a frequency of 99.9%), reaching a level only hardly above that of guessing. In the short run, behavior of social learners at high frequency does thus not reflect the environment very well.

We were interested in whether these inferences become more accurate when short-term trends in the environment are neglected. To test this, we used a moving average filter that smoothes out short run trends in the environment. The greater the filter size, the longer the remaining trends are. Re-analysis of the inference accuracy showed that indeed the behavior of social learners reflects changes in the environment when long-term trends are considered (Fig. 6). For social learners at 99% frequency, accuracy can reach up to 90.70% for trends of 88 periods length and for social learners at 99.9% frequency, accuracy can reach up to 85.62% for trends of 224 periods length. In summary, the behavior of social learners does indeed reflect changes in the environment quite well, but only if short-term trends are ignored.

4. Discussion

4.1. Methods and assumptions

The plausibility of the model we have developed here and the conclusions we will draw is based on a chain of three main arguments. First, economic agents imitate one another, and especially imitate successful (i.e. wealthy) individuals. Second, this is the case because imitation is in itself successful and will hence become more common over time through the evolutionary forces of the market. Third, the effects of this cannot be understood in the aggregate, but rather each agent must be modeled on her own to show the potentially unexpected macro-effects of micro-behavior.

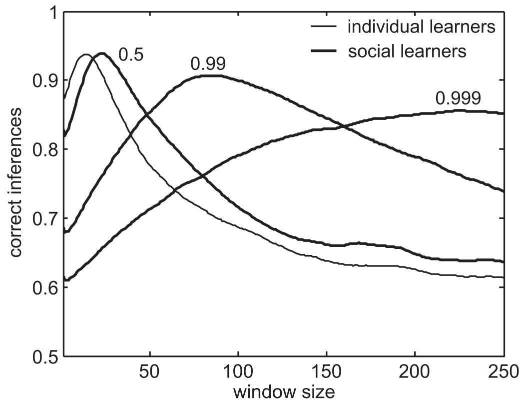

Fig. 6. Percentage of correct inferences from behavior to environment as function of moving average filter length applied to the environment, based on 50,000 periods; frequency of social learners as indicated.

All of these arguments are debatable. There might not be that much imitation in our societies, or it might not be focused that much on the especially wealthy. But additionally to the empirical examples cited in the introduction, our story is also plausible if we consider the incentives of economic agents: Whether placing the same bet as some of the most successful players is optimal or not, it will certainly serve to protect oneself against too harsh consequences if one should be wrong – after all, how can one be blamed for erring along with the crowd or even some of the top names?

Indeed, if one looks closer at speculative market episodes in history, it almost seems as if imitation had been key to every one of them. Carswell (1960, p. 161) describes the case of a banker at the time of the south sea bubble who got into the market saying, “when the rest of the world are mad, we must imitate them in some measure.” And Stewart (1991, p. 97) concludes that “what really fueled the takeover boom [during the 1980s] was the sight of other people making money, big money, by buying and selling companies.”

Secondly, how do we know that the “evolutionary” way of thinking is applicable to the field of economics at all? In nature, variation arises randomly and selection is natural (Darwin, 1859), while in economics, both might be the result of conscious (and social) processes. But we think that the ingredients are sufficient to allow a process analogous to natural selection to take hold. As Alchian (1950, p. 220) put it:

> *the economic counterparts of genetic heredity, mutations, and natural selection are imitation, innovation, and positive profits.*

The instruments of evolutionary theory can be applied with caution. As a matter of fact, this way of reasoning even appears in what has come to be seen as a cornerstone of modern economic methods: In the chapter “The Case for Flexible Exchange Rates” in his “Essays in Positive Economics”, Friedman (1953) argues that since a certain kind of speculator would continually lose money, she would be forced to either leave or change her strategy. Or, in other words: since such speculators could not survive, they could not exist.

Further critique could be raised with regard to our model choice. We used agent-based modeling, meaning that each individual is simulated separately. Therefore, each individual's personal history – what she chose, whether she was successful, whom she sampled, whom she imitated – combined with her learning rule, determines her behavior. The alternative approach would be to aggregate individuals with the same learning strategy. This would require us to calculate probability distributions over all possible states, which is impossible in such a complex system. As important variables, for example, the propensities to choose certain options have no simple or fixed distributions, especially when an individual's history matters. As small differences can lead to divergent behavior, approximating these variables would not lead to meaningful results.

According to Bonabeau (2002), agent-based modeling is appropriate when (1) there are non-linearities in behavioral rules, (2) behavior is path-dependent, (3) there is spatial structure, or (4) basins of attraction are small with regard to fluctuations. Except for (3), all points apply or may sometimes apply to our system, supporting the use of agent-based modeling.

Finally, for placing the arguments of this paper in the history of economic thought, consider also the view of F.A. Hayek, who wrote that “all enduring structures . . . up to the brain and society are the results of, and can be explained only in terms of, processes of selective evolution” Hayek (1979, p. 158). Not only that, but Hayek was, as Vriend (2002) argues, a proponent of agent-based modeling well before it was possible, and might have engaged in tinkering with Matlab codes were he alive today.[9]

[9] For another well-known example of agent-based modeling before sufficient computing power became available see the famous segregation experiment in Schelling (1978).

4.2. Connection to real world phenomena

We proposed a rather simple evolutionary model that on the one hand contains individual learners, who learn about two possible options by relying on own experience, and social learners, who imitate the choice of the most wealthy individual they observe. Surprisingly, the behavior that ensues reflects market anomalies that are used to justify the claim that markets are not efficient. Therefore, social learning could be a simple explanation for market inefficiencies. The details of the observed anomalies are discussed below.

4.2.1. Exaggeration

We found that social learning leads to exaggeration. There are many instances of such exaggeration in economic markets. Shiller (1981) finds that stock price volatility is “five to thirteen times too high to be attributed to new information.” In the short run, Shiller et al. (1984) also find evidence for “excess volatility” in the overreaction of stock prices to dividends. Perhaps the most visible, however, is the medium-to-long-run ‘overshooting’ of markets. The typical stock-market index in the developed world more than doubles from the low to the high point of a business cycle – while few economists would argue that the value of the economy’s largest companies undergoes a similar development within just a few years. But we need not even look as far as to financial markets to find examples of exaggeration through imitation. The fact that some tourist destinations that used to be sanctuaries of the wealthy only years ago are completely overrun now is regularly bemoaned, and it certainly exaggerates rather than matches real (or perceived) quality differences. And well-known investors who have previously acquired wealth through good investments have at times profited from the “Buffet effect”, in which prices of a security they buy are later driven up by imitators.

4.2.2. Delay

A higher frequency of social learners leads to a larger delay in incorporating new information into the market. This seems to contradict the widely held belief, spawned by the efficient-market hypothesis, that markets react to new information not only correctly, but also immediately. While that is certainly the case for obvious and unambiguous information – if a refinery explodes, the oil price goes up – it is far from being a universal truth. The reader might only look at the recent financial crisis, where problems in the subprime-market took more than a year to reach the stock prices of financial institutions that were holding the assets in question. Obviously, very few deviations from the norm of immediate information diffusion are enough to make a great difference.

In contrast to the oversimplified model of markets where equilibrium price reacts to exogenous events with no time delay, we can offer another interpretation. In our model, information about changes in the environment slowly flows through the economy, reaching more and more market participants each period, who subsequently influence more market participants. The higher the degree of imitation, the longer it will take until a critical mass reacts. In times of high social influence, this means that deviations from the hypothetical fundamental value can be upheld for a very long time. To paraphrase a remark that is commonly attributed to Keynes: Markets can take much longer to react to the real environment than you and I can remain solvent.

4.2.3. Relation between behavior and fundamental values

For the efficiency of markets, it is crucial that the choices of market participants reflect fundamentals in some way – that “security prices fully reflect all available information” (Fama, 1991). We do not model prices here, but if we take the behavior of market participants (which is based on their beliefs about the underlying success probabilities of several options, which is also what would drive prices), we can ask: How much does the behavior of the agents in our simulations reflect the real environment? How reliable can the state of nature be inferred from observing the choices made by the agents? We found that such inferences were possible with high accuracy when they are based on the behavior of individual learners or on the behavior of social learners if those were infrequent. When social learning became abundant, however, aggregate behavior hardly reflected the environmental state. This corresponds to the finding that dividends and asset prices are not highly correlated in the short run (Shiller, 1990).

In addition, we found that when social learning was very or extremely frequent – as it was bound to become through evolutionary processes – inferences were inaccurate for short-term trends and best for exceedingly long trends (between 88 and 224 periods). This means that the behavior of social learners does reflect environmental changes, but only for long-term trends. This corresponds to the finding of high correlation between asset prices and dividends if long but not short time periods are considered (Shiller, 1990) and could explain an under-reaction to short- and over-reaction to long-term trends in stock prices (Barberis et al., 1998).

In our simulation, the accuracy of trying to infer fundamentals from agents’ behavior falls if the frequency of social learners rises. The recent financial crisis of 2008 – if one is still allowed to propose another view on its emergence – showed a similar mechanism. To calculate risk correlations necessary to correctly price collateralized debt obligations (CDOs), finance professionals used the spreads of credit default swaps (CDS) as a basis (Li, 2000; Salmon, 2009). Investment bankers charged with evaluating CDOs hence relied as “imitative learners” upon others who priced and traded CDSs, assuming that they were collecting actual information (or acting as individual learners). But it is not clear whether CDSs where priced correctly, as historical data on countrywide default correlation of mortgages, especially of subprime mortgages, were very scarce (Tett, 2009). The market thus started to imitate imitations of decisions that were based on unreliable data in the first place, and clearly exaggerated the safety (or success probabilities) of these assets.

Then markets began their delayed realization of the underlying values of CDOs. In 2007, Citigroup and Merrill Lynch had to write down a combined \$48.5 billion, mostly because of CDOs on their accounts (The Financial Crisis Inquiry Report). Until February 2009, total write-downs due to CDOs accumulated to \$218.2 billion (Benmelech and Dlugosz, 2009).

Grossman and Stiglitz (1976, 1980) pioneered the work on how social learning might crowd out individual learning from financial markets. Informed and uninformed traders have to decide whether to invest in a risky or a safe asset. Informed traders can buy a signal (that may be noisy) which conveys information with regard to the risky asset, whereas uninformed traders base their decision on the price of the assets. The prices in turn are influenced partly by the behavior of the informed traders and thus reflect the fundamentals to some degree. Most importantly, the authors assume that the decision to become informed or stay uninformed is endogenous, meaning that traders switch strategies based on which one leads to the highest utility. If the noise of the signal goes to zero, the equilibrium frequency of informed traders goes to zero. But without informed traders, decisions become detached from reality and prices become completely uninformative.

Some of our results are similar to those of Grossman and Stiglitz. In their model, under rare circumstances, a state can be achieved in which hardly any or no individual at all learns about the environment. In our model, this also happens but occurs under very general circumstances. We neither have to assume that individual learning is more costly than social learning, nor that individual learners can completely eliminate uncertainty.

5. Conclusion

We have shown, by means of evolutionary simulations, that imitating others in decisions whose success depends on the environment can be a good strategy, and can even return more favorable results than trying to learn about the environment by oneself. This is because imitators will often decidedly choose the better option, while individual learners choose the worse option with a moderately high probability. However, when the environment changes, social learners are also quite likely to decidedly choose the wrong option for some time, while individual learners will adjust their choice more quickly.

Keynes (1936, p.158) seemed to have a model not entirely dissimilar to ours in mind when he distinguished between "speculators", who try to predict "the psychology of the markets", and "enterprise", meaning those who actually try their hand at "forecasting the prospective yields of assets." The analogy to social and individual learners is not too far-fetched. Keynes famously goes on to write that,

> *Speculators may do no harm as bubbles on a steady stream of enterprise. But the position is serious when enterprise becomes the bubble on a whirlpool of speculation.*

In the simulation runs of our model, we have observed that the imitators, even if they start out as bubbles on a stream of individual learners, will soon become the torrent which drowns out what little information the individual learners gather. Therefore, even if our method is unusual for an economic study, the argument is not. Markets are powerful information aggregators which can help in predicting anything, from sporting events (Spann and Skiera, 2009) to presidential elections (Wolfers and Zitzewitz, 2004). But they can only fulfill their task if enough information "flows" into them. We have seen that imitating the successful can enhance an individual's performance but may at the same time worsen group performance. If, as our analysis suggests, information inflow is very low in equilibrium, aggregate behavior tends to become self-referential. Thus, if market participants rely on others to procure information and assume that all their knowledge is accurately reflected in their behavior, market reactions will be exaggerated and delayed, and correspondence between behavior and reality will be weak.

Appendix A. Individual Bayesian learning

In an attempt to make the individual learners even stronger and behave like the classical rational economic agent, we also simulated the evolutionary outcome when social learners compete with rational, Bayesian individual learners. For this, we have to make an assumption about the priors of a Bayesian learner. A generous assumption is to allow the Bayesian learner to know the total prior distribution of p_A and p_B – i.e. to assume that the Bayesian learner understands the process and can rationally infer from her observation, but does not know in which "generation" she is. To derive this distribution, we simulated p_A and p_B over the time of 10^7 periods and used the resulting distribution (panel A of Fig. 7) as the prior. This is a very generous approach, since other strategies do not have access to this vast amount of information; nor, in fact, do they even know that there is a fixed probability distribution at all. This should put Bayesian learners in a very favorable position.

Having the priors, a Bayesian learner chooses the option with the higher expected value (or chooses randomly if she is indifferent). After choosing, the outcome is realized, meaning that the chosen option either leads to success or failure. In any case, the Bayesian learner updates her belief about the chosen option according to this new bit of information and the cycle continues. After each generation, parents die and offspring are born. These offspring inherit their first period's choice but not the priors from their parents; this is analogous to reinforcement learners (the individual learners in our model) who also inherit their parent's last choice but not their parent's propensities.

Even though the probability distribution of p_A and p_B in the first generation or so is usually narrower than the Bayesian learner's priors suggest, in later generations their prior distribution quite accurately reflects the real distribution. In other

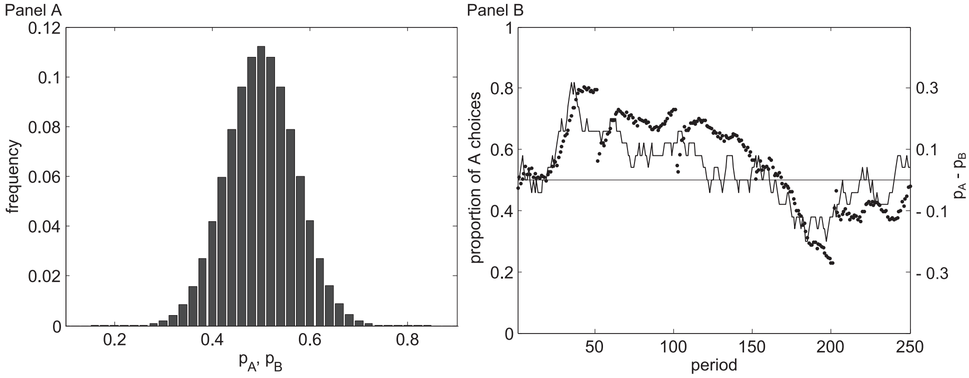

Fig. 7. Panel A: Probability distribution of p_A and p_B. Panel B: Proportion of A choices made by Bayesian learners (left y-axis) as a function of $p_A - p_B$ (right y-axis) over time.

words, every new generation of Bayesian learners starts off with a quite accurate picture of how p_A and p_B are distributed, and then proceeds to learn from trial and error what their current realizations are. The latter knowledge is not passed from generation to generation. Bayesian learners who use the prior distribution as the basis for learning, then, behave rationally, as would be expected if it is unknown in which generation one currently is.

Furthermore, we endowed Bayesian learners with the knowledge about the process that generates the environment. This way, a Bayesian learner cannot only optimally compute the probability distributions of p_A and p_B but also predict how these distributions will change from one period to the other. No other strategy that we tested was given such deep knowledge about the environment it is faced with.

The resulting behavior of Bayesian learners can be seen in panel B of Fig. 7. Bayesian learners engage in probability matching. Discontinuities in behavior after each generation, i.e. after 50 periods, can be observed, reflecting the reset of the priors. With the default parameter values, the performance of Bayesian learners reached 60.15% (±0.67%, S.E.M.), only slightly above the performance of reinforcement learners.

We are interested in how well Bayesian learners perform. For this, we simulated their performance over 1000 generations using the default parameter values. Again, we started with 9000 individual learners and 1000 social learners. The picture that emerges is the same as in our main setting with reinforcement learning: Social learners quickly rise in frequency and eventually displace individual learners from the population (Fig. 8). Therefore, Bayesian individual learning does not prevent the dominance of social learning. To check for robustness, we varied the number of periods per generation from 25 to 250. This had little influence on performance.

The reason why Bayesian learners cannot reach a higher performance is that the feedback from the environment is too noisy to form correct beliefs. Bayesian learning is not fast enough to keep up with the constantly changing environment. In conclusion, Bayesian learners, even if given a large informational advantage, can just barely beat reinforcement learning by exponential discounting. On the other hand, Bayesian learning is much more demanding and less parsimonious than reinforcement learning. A Bayesian learning strategy is thus inappropriate for our task.

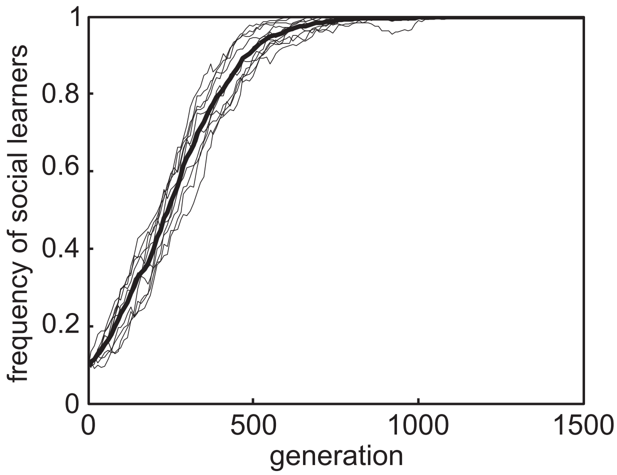

Fig. 8. Simulated frequency of social learners competing with Bayesian individual learners from ten evolutionary simulations (thin lines), mean (thick line).

Appendix B. Robustness to parameter changes

To check the robustness of our finding that social learning by imitating the wealthiest is dominant, we repeated the evolutionary simulations while varying the parameter values. We adopted a Latin hypercube design to sample the parameters. Overall, we made 1000 simulation runs.

B.1. Tested parameters

Most parameters we tested influence the environment. It consists of two options, A and B, whose success probabilities are p_A and p_B, respectively. p_A and p_B change after each period by a certain increment according to a process akin to a random walk, except that they are bound between 0 and 1 and that they tend to revert to the mean. Several parameters affect how the environment changes:

- The mean success rate, p_{mean}: The arithmetic mean of p_i, $i = \{A, B\}$, is 0.5 in the default condition. There is, however, no particular reason why that should be so. Higher p_{mean} imply that successes are more common and lower p_{mean} that failures are more common. The wealth of a strategy is the most distinct feature if successes are neither too common nor too rare. Therefore, we changed the mean success rate and checked whether ITW would still dominate.
- Reversion factor r: p_i tends to revert to the mean. For the default condition, that means that if, say, p_i is 10/20/30 percentage points greater than the mean, it sinks in the next period with probability 60/70/80%, and if it is 10/20/30 percentage points less than the mean, it increases in the next period with probability 60/70/80%. In general, the probability that p_i increases is equal to $0.5 - r \times (p_i - p_{mean})$, with r equaling 1 in the default condition. Higher r imply that there are more switches between A and B being the better option, and thus that strategies have to adjust faster; ITW should fare worse with higher r, since it is slower to adapt than individual learners.
- Probability of environmental change, p_{incr}: In the default condition, the environment changes after each period. It could, however, be possible that in some periods, it remains constant. p_{incr} gives the probability that p_A or p_B changes after each period (they move independently). Lower p_{incr} should favor ITW, since it is the slower strategy and gets more time to react to environmental changes.
- Step size k_{incr}: When p_i changes, it increases or decreases by an amount k_{incr}. The default value is $k_{incr} = 0.02$. Higher k_{incr} imply that the environment changes faster, and since it takes fewer steps to move away from the mean, higher k_{incr} also imply the environment reverts more aggressively to the mean. In sum, both effects should act to decrease the performance of ITW.

In sum, this is the process that describes how the environment changes from one period to the other:

$$p_i(t+1) = \begin{cases} p_i(t); & 1 - p_{incr} \\ p_i(t) + k_{incr}; & p_{incr}(0.5 - r \times (p_i(t) - p_{mean})) \\ p_i(t) - k_{incr}; & p_{incr}(0.5 + r \times (p_i(t) - p_{mean})) \end{cases}$$

Apart from the factors that affect the environment, there are other factors that might have an influence on the performance of ITW.

- Sample size ss: In the default condition, it was determined that the right sample size for ITW is the "magical number seven" (Miller, 1956), but this number is more or less arbitrary. It may be possible that ITW cannot beat individual learners if they could only sample less individuals, which we decided to test.
- Length of a generation t_{max}: In the default condition, one generation lasts 50 periods. During this time, strategies learn and accumulate wealth. After each generation, strategies reproduce and start fresh. When the generation length is very short, the spread in wealth between the actors could be too low to be informative for ITW. If generations last too long, presently high wealth could reflect successful behavior in the past but not necessarily successful behavior in the present. Therefore, both too low and too high t_{max} could be detrimental for the performance of ITW.
- The skill parameter p_{skill}: In our simulations, each individual is the same a priori; nobody has an advantage in performing the required task. It could be possible, however, that in the real world, some individuals are inherently more skilled than others and that consequently wealth is not only determined by making good choices but also by being skilled. We thus introduced skill. If $p_{skill} = 0.1$, for example, it means that for each individual, at the beginning of a generation, a skill is drawn from the uniform distribution $U(-0.1, 0.1)$. This skill is added to the probability to succeed, so if $p_i = 0.5$ and a skill level of 0.05 is drawn, the actual probability to succeed is 0.55. Skill is determined independently from strategy, cannot be observed, and cannot be inherited. Increasing p_{skill} implies that wealth depends to a higher degree on skill, which ITW can neither copy nor account for, and to a lesser degree on a history of making good choices. This should lead to ITW having a lower performance.

Table 1
Parameters that were tested for the robustness analysis. The default value was used outside of the robustness analysis.

Parameter	Explanation	Default	Distribution	Range
p_{mean}	Mean value of p_A, p_B	0.5	$0.5+(x-0.55)/2$	0.2754–0.7247
r	Reversion factor	1	$\exp(2x)-1$	0.221–6.389
p_{incr}	Probability of environmental change	1	x	0.1–1
k_{incr}	Step size	0.02	$\exp(x/20)-1$	0.005–0.051
ss	Sample size	7	round$(10x+1)$	2–11
t_{max}	Length of a generation	50	round$(\exp(5.5x))$	2–245
p_{skill}	Skill factor	0	$(x-0.1)/3$	0–0.3[a]

[a] A values of p_{skill} of 0.3 corresponds to an R^2 of the correlation between wealth and skill of to 0.99.

A short recapitulation of all parameters, their default values, and the changes we made for test robustness, are found in Table 1. The variable x is drawn from a uniform distribution $U(0.1, 0.9)$ according to the Latin hypercube sampling mechanism implemented in Matlab.

B.2. ITW versus individual learners

We ran 1000 simulations with social learners using imitate the wealthiest competing against individual learners. The number of individuals was 10,000 in all simulations; social learners started with 1000 individuals, individual learners started with 9000 individuals. Every simulation was stopped after 5000 generations. Table 2 shows the estimated regression parameters (and confidence intervals) from a multivariate model of the type

$$\text{Frequency of social learners} = \beta_1 \text{intercept} + \beta_2 p_{mean} + \ldots,$$

where the frequency of social learners (in percentage points) at the end of the simulation is explained by the model parameters.

Only three model parameters do at all have a significant influence on the frequency of social learners. For those parameters, the sign of the β confirms the hypothesized trend. A higher r (and also a higher p_{incr}) implies that the environment changes more quickly, which should favor individual learners. Sample size, in contrast, correlates positively with the frequency of social learners, which is expected, at least until a certain degree.

The role of t_{max} is especially important to understand. This parameter correlates strongly and positively with the frequency of social learners. There are two explanations for that. First, if t_{max} is too small, i.e. when there are only few periods per generation, success is very noisy and thus wealth is a poor predictor for good choice. More periods allow more precise inference from wealth on good choices. However, one would also expect that if t_{max} is too large, wealth depends less and less on good choices in the last few periods and more and more on good choices made during periods long past. Since the environment probably changed in-between, wealth should become less good of a predictor of recent good decision making when t_{max} is too large. We will see that this holds true.

There is another reason why the frequency of social learners correlates so strongly with t_{max}. When t_{max} is short, there are only very few opportunities to accumulate fitness differences. Since there is always a fixed base fitness, the lower t_{max}, the lower the relative fitness differences between strategies. Therefore, even if social learners were consistently better, it would take them longer to increase in frequency when t_{max} is small.

Therefore, it is plausible that if we had simulated more generations (which was difficult due to the high computational run time of the simulations), social learners would have had more time to increase in frequency. This can be confirmed by looking at the increase of the frequency of social learners at the end of the simulations as a function of t_{max}. If social learners simply did not have enough time to reach equilibrium frequency when t_{max} is too small, we would expect the slope of their frequency to be higher for smaller t_{max}. This is indeed the case, as can be seen in Fig. 9. We determined the slope of the frequency of social learners during the last 1000 generations and found it to be greater the smaller t_{max} (confidence intervals are too small to be shown, $\pm 6 \cdot 10^{-4}$). This implies that social learners would have reached an even higher frequency, especially for small t_{max}, had the simulations run for more generations.

Table 2
β's and their 95% confidence intervals for the multiple linear regression of the frequency of social learners at the end of the simulations onto the parameters.

parameter	β	β, 95% CI
Intercept	0.572	0.074; 1.069
p_{mean}	−0.043	−0.208; 0.123
r	−0.217	−0.299; −0.134
p_{incr}	−0.120	−0.287; 0.047
k_{incr}	0.121	−0.045; 0.287
ss	0.235	0.071; 0.399
t_{max}	0.4473	0.282; 0.612
p_{skill}	−0.018	−0.182; 0.147

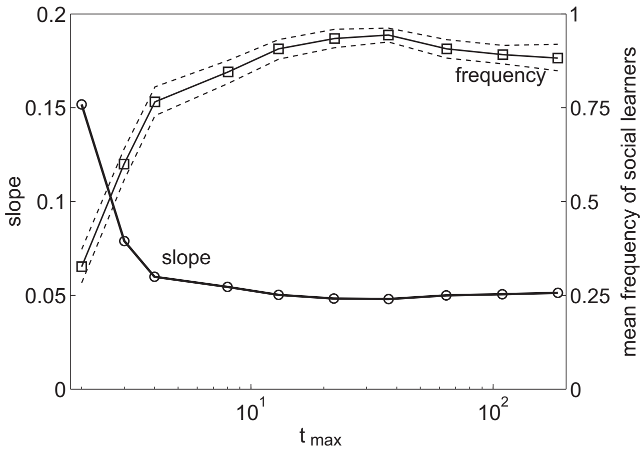

Fig. 9. Slope of the regression of frequency of social learners during the last 1000 generations (circles, left y-axis) as a function of t_{max} (x-axis, log scale). Mean frequencies of social learners (squares, right y-axis) and their standard errors of the mean (dashed line) are shown for comparison.

As a comparison, Fig. 9 also shows the mean frequency of social learners. It shows that there is an almost inverse relationship between mean and slope, which means that the lower the mean frequency of social learners, the more it would increase had the simulations run for longer. It also confirms that when t_{max} becomes too high, the mean frequency of social learners drops slightly; they do best for intermediate values of t_{max}.

B.3. Additional social learning strategies

Until now, we assumed that social learning is achieved by imitating the wealthiest. However, there are other possible forms of social learning. Those other social learning strategies could, hypothetically, prevent the total domination of social learning we saw until now and thus circumvent the more extreme behavioral responses we discussed in the main text. This can be achieved in two fashions. First, it is possible that the presence of other social learning strategies stabilizes individual learning and thus prevents it being marginalized. Second, if the social learning strategies rely, as part of their strategy, on sometimes learning individually, this could also prevent the extreme effects.

We added social learning strategies that were already discussed in the literature. These social learning strategies are:

- Conformism: Conformism has been discussed extensively in the past (Boyd and Richerson, 1982, 1985; Henrich and Boyd, 1998) In our simulation, this strategy relies on sampling a number of individuals and adapting the option that was most frequently observed in the sample. Initial tests established that conformists were the most efficient when sampling three individuals. We thus set the sample size of conformists to three.
- Opportunistic individual learning: This strategy is inspired by Enquist et al. (2007) and consists of learning socially by default but switching to individual learning if the last period's choice was not followed by a success (hence the name). In the original paper, social learning consisted simply of copying a random individual, but since conformism is superior to random copying, we instead used conformism as playing the part of social learning.
- Opportunistic conformists: This strategy is inspired by Rendell et al. (2010) and consists of using individual learning by default but switching to social learning if the last period's choice was not followed by a success. It is thus the converse of the opportunistic individual learning. Again, instead of random copying, we used conformism as the social learning mechanism.
- In doubt, conform: This strategy was first proposed by Boyd and Richerson (1995). It consists of relying on individual learning if the actor is sufficiently certain that she knows what option is better, and relying on social learning if not. Again, we used conformism for the social learning part. We defined an actor as certain about her choice if the difference in propensities between the option is sufficiently large. The required difference was determined in initial tests so as to maximize performance.
- Payoff-biased social learning is a social learning strategy that relies only on observations directly linked to the previous period. One type of payoff-biased social learning uses a scoring system reminiscent of the Borda count from social choice theory (Borda, 1781) but allowing for differential weighing of observed successes and failures. Scoring has the advantage that only this type obeys certain axioms of rational behavior (Young, 1975). Among all possible combinations of weights, it was found that weights of 1 on observed successes and 0 on observed failures, as well as 4 on observed successes and −1 on observed failures, prove particularly successful (Bossan, 2013), which is why we included these strategies here. Interestingly, conformism can be interpreted as a special case of payoff-biased social learning with equal positive weight on observed gains and failures.

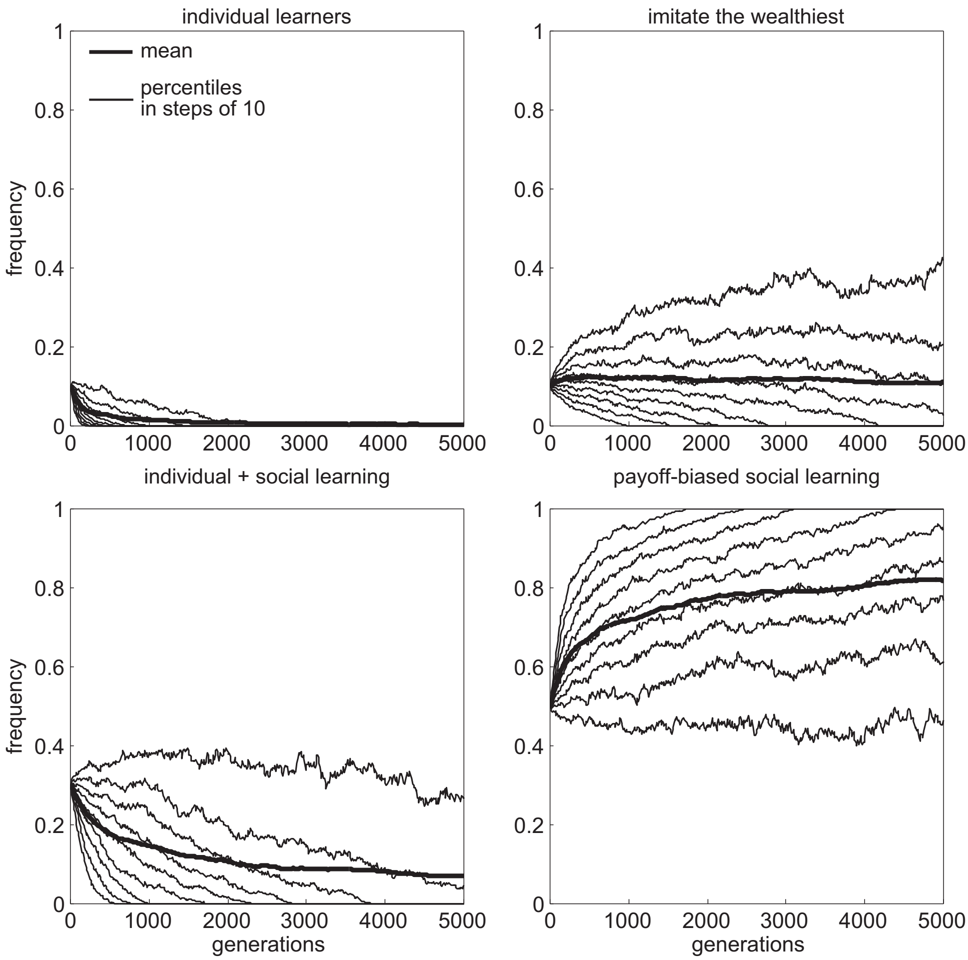

Fig. 10. Mean (thick line) and percentiles in steps of 10 (thin lines) of the frequencies of individual learners (top left), imitate the wealthiest (top right), the three strategies using a mix of individual and social learning (bottom left), and the five payoff-biased social learning strategies including conformism (bottom right).

- Payoff-biased social learning based on comparing averages has been shown to be quite successful (McElreath et al., 2008). We include this strategy, as well as a modification of it that trades off payoff and conformism, which was shown to be even more successful (see (Bossan, 2013), for more explanations). What is interesting about this latter strategy is that it produces effects that are very similar to those produced by ITW (Bossan, 2013).

In addition to those eight strategies, we added pure individual learning and social learning by imitating the wealthiest, the social learning strategy we focus on in this work. These ten strategies were again tested, using the different parameter settings described in the previous section and 200 simulations with Latin hypercube sampling. The main results are shown in Fig. 10.

On the top left, we see the frequency of individual learners. They quickly diminish in frequency, averaging at 0.27% after 5000 generations. We also tested three strategies that use a mix of individual and social learning, namely opportunistic individual learning, opportunistic conformism, and in doubt, conform. These strategies were also not very successful, averaging at 6.85%. Imitate the wealthiest did a little better, reaching 10.99% after 5000 generations. The most successful were the five payoff-biased social learning strategies, reaching 81.89%.

These results show that ITW does not necessarily perform so well against other, quite sophisticated social learning strategies, though it still performs a lot better than individual learning or other strategies that rely partly on individual learning. Instead, another form of pure social learners, payoff-biased social learners, dominate. Interestingly, among those, the strategy relying on averaging and payoff-conformism trade-off is the most frequent (46.01%). Though this suggests that ITW might not be the altogether best strategy to deal with a situation as described in our model, the point of this paper is not that ITW beats all other strategies, but that it beats individual learning strategies very clearly and thus leads to a dominance of social learner with the effects described above. As a matter of fact, simulations by Bossan (2013) have shown that the strategy relying on averaging and payoff-conformism trade-off produces very similar effects to the ones described

Table 3
Linear regression analysis of behavior on environment based on 50,000 data points; slopes including 95% confidence intervals, R^2, and sum of square errors are indicated, p-values too small to report.

Strategy	Frequency (%)	Slope (95% C.I.)	R^2	S.S.E.
Individual learners	–	1.047 (1.041–1.054)	0.6784	243.6
Social learners	10	1.743 (1.730–1.756)	0.5667	1088.3
Social learners	30	1.753 (1.740–1.767)	0.5648	1137.6
Social learners	50	1.795 (1.781–1.809)	0.5599	1233.2
Social learners	70	1.834 (1.819–1.849)	0.5370	1404.2
Social learners	90	1.934 (1.915–1.952)	0.4627	2170.4
Social learners	95	1.887 (1.866–1.907)	0.3980	2729.0
Social learners	99	1.612 (1.584–1.641)	0.1974	4950.7
Social learners	99.8	1.392 (1.359–1.425)	0.1191	7143.8
Social learners	99.9	1.184 (1.149–1.219)	0.0799	7850.3
Social learners	99.99	0.667 (0.628–0.707)	0.0214	9946.9

Table 4
Linear regression analysis of behavior on environment based almost 50,000 data points; slopes including 95% confidence intervals, R^2, and sum of square errors are indicated, p-values too small to report.

Strategy	Frequency	Lag	Slope (95% C.I.)	R^2	S.S.E.
Individual learners	–	5	1.159 (1.155–1.164)	0.8312	127.8
Social learners	10	8	2.063 (2.054–2.072)	0.7940	517.5
Social learners	30	9	2.079 (2.070–2.089)	0.7944	537.4
Social learners	50	9	2.135 (2.126–2.145)	0.7922	582.3
Social learners	70	10	2.225 (2.215–2.235)	0.7905	635.29
Social learners	90	13	2.434 (2.421–2.447)	0.7334	1076.8
Social learners	95	16	2.487 (2.472–2.501)	0.6904	1402.4
Social learners	99	27	2.511 (2.488–2.534)	0.4789	3214.1
Social learners	99.8	39	2.226 (2.197–2.256)	0.3047	5636.4
Social learners	99.9	43	2.037 (2.005–2.069)	0.2364	6513.2
Social learners	99.99	64	1.309 (1.270–1.347)	0.0820	9326.4

here. Among the two, however, ITW is vastly more plausible in many respects (as we also argue in the main text), which is why we focus on this strategy.

Appendix C. Linear regression analysis

Initial examination of our findings suggested that individual learners reflect environmental changes in a 1:1 fashion, whereas social learners exaggerate environmental changes. To analyze how strongly different agents react to environmental changes, we performed a linear regression analysis of the proportion of A choices made by the agents as a function of environmental changes in the form of $p_A - p_B$ (see Fig. 11, left column). The data for individual learners and for social learners of varying frequencies are shown in Table 3. Individual learners match the environment; they respond to a one percentage point increase in the difference $p_A - p_B$ with a 1.047% point increase in A choices. Social learners overmatch, with overmatching intensity peaking at around 90% frequency of social learner. At this frequency, they respond to a 1% point increase in the difference $p_A - p_B$ with a 1.934% point increase in A choices. In general, linear fits become weaker and weaker the more frequent social learners become in the population, meaning that their behavior becomes more difficult to predict using a linear regression.

In general, individual learners react with only a short delay to environmental changes, whereas social learners, especially when very frequent in the population, react with a strong delay. This could explain why the linear model generates bad fits to the behavior of social learners, making it seem unpredictable. To correct for this, we determined the cross-correlation of behavior and environment for delays of different lengths (see also Section 3.3 of the main text). Then we took the delay that maximized cross-correlation and fitted behavior to this lagged environment to achieve the best possible correlation between behavior and environment (see Fig. 11, right column). The results are shown in Table 4. In contrast to the previous findings, we now report overall higher slopes for the correlation between behavior and environment. Still individual learners match environmental changes (1.159:1), whereas social learners tend to overmatch (2.135:1 for 50% frequency of social learners, 2.511:1 for 99% frequency of social learners). Obviously, R^2 become better after correcting for delay but still remain quite small for higher frequencies of social learners. Even after accounting for delay, the behavior of social learners thus remains unpredictable.

Appendix D. How well behavior reflects reality

We were interested in whether the behavior of the agents in our model allows to infer the state of the environment. More precisely, we assume that there is a threshold value x that is chosen to infer that $p_A > p_B$ and thus that A should be chosen

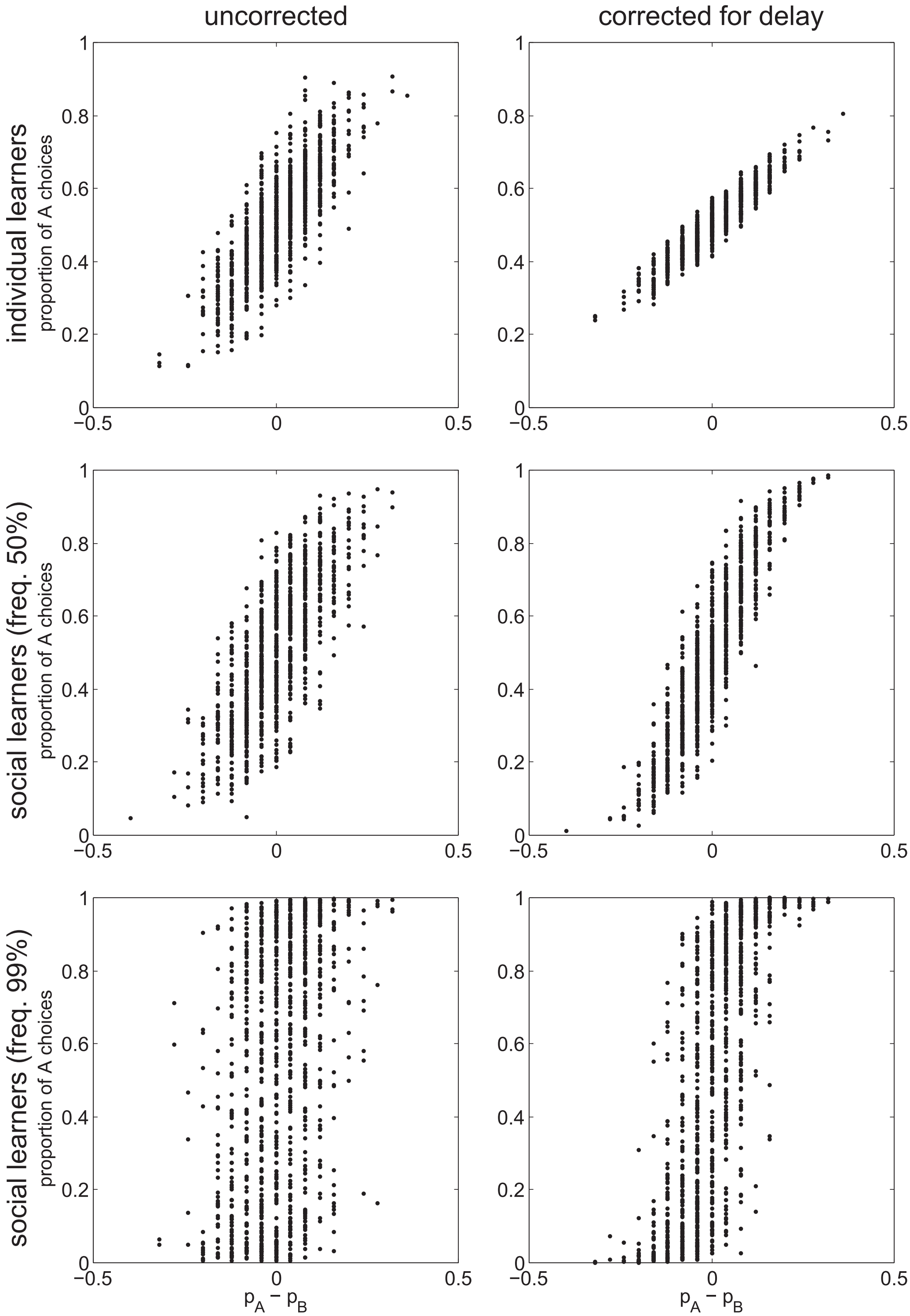

Fig. 11. Scatter plots of behavior (proportion of A choices) of different strategies as a function of the environmental state ($p_A - p_B$); only every 50th data point is shown to reduce file size.

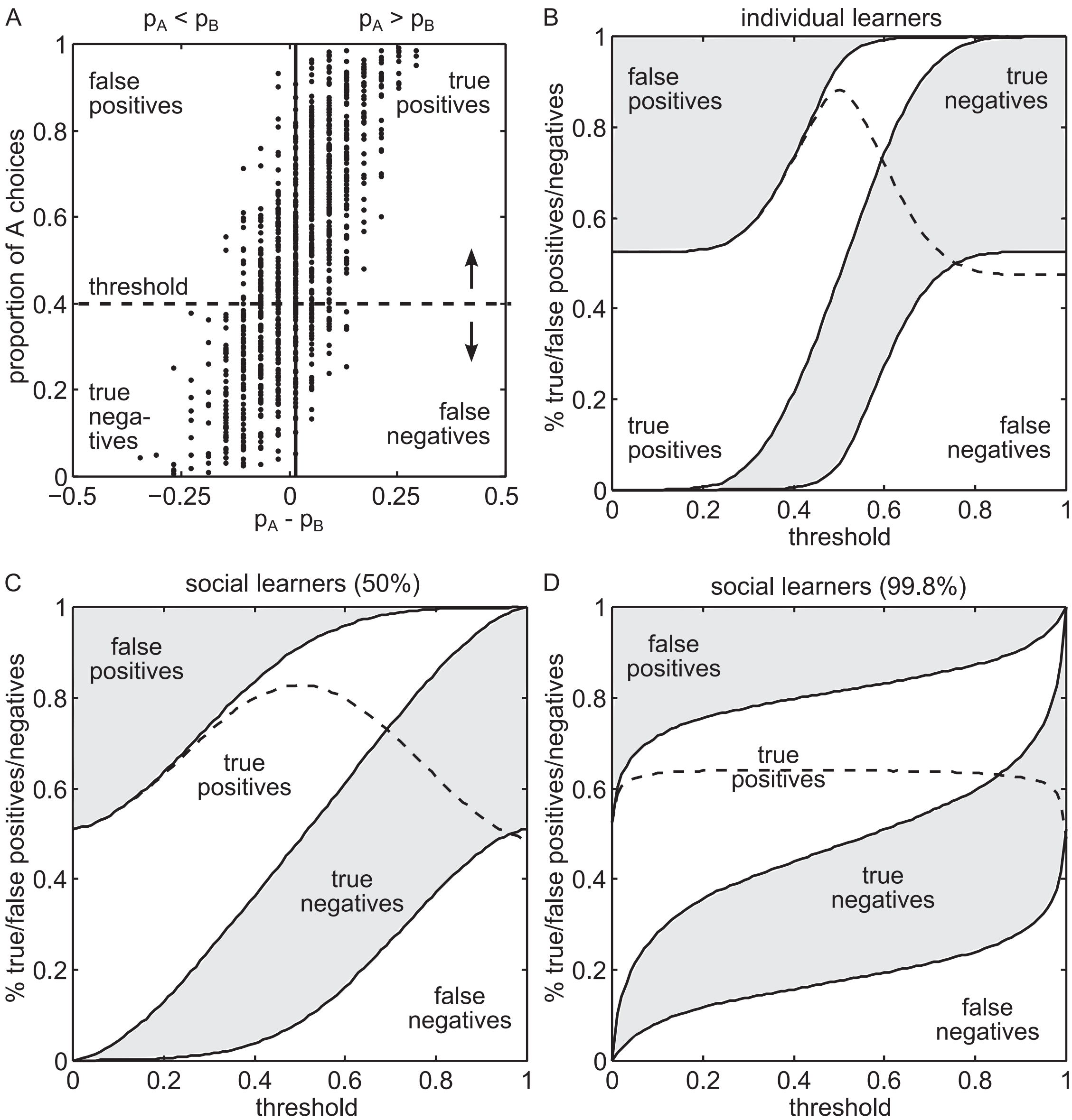

Fig. 12. How well the environmental state can be inferred from behavior.

over *B*. For example, if this threshold is set at 50%, the most obvious choice, the inference would be that *A* is better than *B* if more than half of the population choose *A* and vice versa.

Using this criterion, four outcomes are possible. First, one could conclude that *A* is better than *B* when indeed *A* is better than *B* (true positive); that *A* is better than *B* when instead *B* is better than *A* (false positive); that *A* is worse than *B* when indeed *A* is worse than *B* (true negative); and that *A* is worse than *B* when instead *A* is better than *B* (false negative). These four outcomes, in relation to the chosen threshold, are illustrated in panel A of Fig. 12. In panels B, C, and D, the proportion of true and false positives and negatives as derived from simulations are cumulatively shown. The total proportion of true estimations is shown as the dashed line. Individual learners are shown in panel B. They have a peak of total true estimations of 88.2% for a threshold of 0.5. Social learners at a frequency of 50% are shown in panel C. The peak of total true estimations is close to 0.5 but lower than for individual learners, 82.9%. Social learners at a frequency of 99.8% are shown in panel D. The proportion of true estimations is almost constant for thresholds between 0.05 and 0.95 and very low, hardly exceeding 64%. This shows that it is hard to infer the environmental state from the behavior of market participants when social learning is dominant in the population, which is the case in equilibrium.

Appendix E. Pseudocode

Here is a very high level pseudocode describing how the simulations work. The pseudocode as used for the robustness analysis is shown. The whole code can be found online at https://github.com/BenjaminBossan/coevo. Note that to increase computational speed, the actual code is vectorized, which may make it hard to read.

- Draw parameters t_{max}, r, k_{incr}, p_{incr}, Δp, p_{skill}, and sample size using Latin hypercube sampling
- Initialize population with starting frequencies of the strategies (typically 9000 individual learners, 1000 social learners). Each individual starts with a base fitness/wealth of 10
- Set first p_A and p_B to 0.5 (if not stated altered by environmental parameter Δp)
- Set very first choice of agents to A with probability 50% and else to B
- For generation 1 to 5000:
  - Create a random environment p_A and p_B according to the parameters
  - For time period 1 to t_{max}:
    - For agent 1 to 10,000:
      - Determine if last choice led to success
      - Update propensities for individual learning
      - Increase wealth/fitness by 1 if the agent was successful
      - Decide whether to choose A or B in the next period, depending on the learning strategy of the agent
    - end
  - end
  - Generate the next generation according to evolutionary algorithm, depending on the fitness of the strategies (last generation is replaced), thus changing the frequencies of the learning strategies
  - All 10,000 agents of the new generation start with an initial wealth of 10
  - The first choice of the new generation is set to the last choice made by the parent (vertical transmission)
  - The first p_A and p_B of the next generation is set to the last value of p_A and p_B from the current generation
- end

For other analyses, the code is adapted accordingly. For example, for the behavioral analysis with fixed strategy frequencies, there is no natural selection. The exact working of the learning strategies, which is the most interesting part, is not described here; for this, refer to the main text.

References

Alchian, A., 1950. Uncertainty, evolution, and economic theory. J. Polit. Econ. 58 (3), 211–221.
Anderson, L., Holt, C., 1997. Information cascades in the laboratory. Am. Econ. Rev. 87 (5), 847–862.
Apesteguia, J., Huck, S., Oechssler, J., 2007. Imitation-theory and experimental evidence. J. Econ. Theory 136 (1), 217–235.
Banerjee, A., 1992. A simple model of herd behavior. Q. J. Econ. 107 (3), 797–817.
Barberis, N., Shleifer, A., Vishny, R., 1998. A model of investor sentiment. J. Financ. Econ. 49 (3), 307–343.
Baum, W., 1979. Matching, undermatching, and overmatching in studies of choice. J. Exp. Anal. Behav. 32 (2), 269.
Benmelech, E., Dlugosz, J., 2009. The credit rating crisis. NBER Macroecon. Annu. 2009, 161–207.
Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. J. Polit. Econ. 100 (5), 992–1026.
Birch, S., Vauthier, S., Bloom, P., 2008. Three- and four-year-olds spontaneously use others' past performance to guide their learning. Cognition 107 (3), 1018–1034.
Bonabeau, E., 2002. Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. U. S. A. 99 (Suppl 3), 7280.
Borda, J., 1781. Mémoire sur les élections au scrutin. Histoire de l'académie royale des sciences 2, 657–665.
Bossan, B., 2013. The evolution of social learning (Ph.D. thesis). Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät I http://www.nbn-resolving.de/urn:nbn:de:kobv:11-100213702
Boyd, R., Richerson, P., 1982. Cultural transmission and the evolution of cooperative behavior. Hum. Ecol. 10 (3), 325–351.
Boyd, R., Richerson, P., 1985. Culture and the Evolutionary Process. University of Chicago Press.
Boyd, R., Richerson, P., 1995. Why does culture increase human adaptability? Ethol. Sociobiol. 16 (2), 125–143.
Carswell, J., 1960. The South Sea Bubble. Cresset Press.
Darwin, C., 1859. On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. D. Appleton, New York.
Economist, 2011. The foolishness of crowds. The Economist 71, 1463.
Enquist, M., Eriksson, K., Ghirlanda, S., 2007. Critical social learning: a solution to Rogers's paradox of nonadaptive culture. Am. Anthropol. 109 (4), 727–734.
Fama, E., 1991. Efficient capital markets: II. J. Finance, 1575–1617.
Feldman, M., Aoki, K., Kumm, J., 1996. Individual versus social learning: evolutionary analysis in a fluctuating environment. Anthropol. Sci. 104, 209–232.
Fisher, R., 1930. The Genetical Theory of Natural Selection. Clarendon Press, Oxford.
Friedman, M., 1953. Essays in Positive Economics. University of Chicago Press.
Friedman, M., 1957. A Theory of the Consumption. Princeton University Press, Princeton.
Gintis, H., 2000. Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Behavior. Princeton University Press.
Giraldeau, L., Valone, T., Templeton, J., 2002. Potential disadvantages of using socially acquired information. Philos. Trans. R. Soc. Lond. B: Biol. Sci. 357 (1427), 1559.
Grossman, S., Stiglitz, J., 1976. Information and competitive price systems. Am. Econ. Rev. 66 (2), 246–253.
Grossman, S., Stiglitz, J., 1980. On the impossibility of informationally efficient markets. Am. Econ. Rev. 70 (3), 393–408.
Hammerstein, P., Hagen, E., 2005. The second wave of evolutionary economics in biology. Trends Ecol. Evol. 20 (11), 604–609.
Harding, R., Fullerton, S., Griffiths, R., Bond, J., Cox, M., Schneider, J., Moulin, D., Clegg, J., 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. Am. J. Hum. Genet. 60 (4), 772.

Hayek, F.A., 1979. Law, Legislation and Liberty: The Political Order of a Free People. University of Chicago Press.
Henrich, J., Boyd, R., 1998. The evolution of conformist transmission and the emergence of between-group differences. Evol. Hum. Behav. 19 (4), 215–241.
Henrich, J., Gil-White, F., 2001. The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. Evol. Hum. Behav. 22 (3), 165–196.
Herrnstein, R., 1961. Relative and absolute strength of response as a function of frequency of reinforcement. J. Exp. Anal. Behav. 4 (3), 267.
Huck, S., Normann, H., Oechssler, J., 1999. Learning in Cournot oligopoly – an experiment. Econ. J. 109 (454), 80–95.
Kaelbling, L., Littman, M., Moore, A., 1996. Reinforcement learning: a survey. J. Artif. Intell. Res. 4, 237–285.
Kameda, T., Nakanishi, D., 2002. Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment: an evolutionary simulation and an experiment with human subjects. Evol. Hum. Behav. 23 (5), 373–393.
Kendal, J., Giraldeau, L., Laland, K., 2009. The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. J. Theor. Biol. 260 (2), 210–219.
Keynes, J., 1936. The General Theory of Employment, Interest, and Money. Harcourt (reprinted 1965).
Klucharev, V., Hytoenen, K., Rijpkema, M., Smidts, A., Fernández, G., 2009. Reinforcement learning signal predicts social conformity. Neuron 61 (1), 140–151.
Knight, F., 1921. Risk, Uncertainty, and Profit. Houghton Mifflin Company.
Li, D., 2000. On default correlation: a copula function approach. J. Fixed Income 9 (4), 43–54.
McElreath, R., Bell, A., Efferson, C., Lubell, M., Richerson, P., Waring, T., 2008. Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. Philos. Trans. R. Soc. B: Biol. Sci. 363 (1509), 3515.
Mesoudi, A., Whiten, A., Laland, K., 2006. Towards a unified science of cultural evolution. Behav. Brain Sci. 29 (4), 329–346.
Miller, G.A., 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. 63 (2), 81.
Ohta, T., 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Syst. 23, 263–286.
Rendell, L., Fogarty, L., Laland, K., 2010. Rogers' paradox recast and resolved: population structure and the evolution of social learning strategies. Evolution 64 (2), 534–548.
Richerson, P., Boyd, R., 2005. Not by Genes Alone: How Culture Transformed Human Evolution. University of Chicago Press.
Robbins, H., 1952. Some aspects of the sequential design of experiments. Bull. Am. Math. Soc. 58 (5), 527–535.
Rogers, A., 1988. Does biology constrain culture? Am. Anthropol. 90 (4), 819–831.
Salmon, F., 2009. Recipe for disaster: the formula that killed wall street. Wired Mag. 17 (3), 17-03.
Schelling, T., 1978. Micromotives and macrobehavior. Norton, New York.
Schlag, K., 1999. Which one should I imitate? J. Math. Econ. 31 (4), 493–522.
Shiller, R., 1981. Do stock prices move too much to be justified by subsequent changes in dividends? Am. Econ. Rev. 71 (3), 421–436.
Shiller, R., 1990. Market volatility and investor behavior. Am. Econ. Rev. 80 (2), 58–62.
Shiller, R., 2003. From efficient markets theory to behavioral finance. J. Econ. Perspect. 17 (1), 83–104.
Shiller, R., Fischer, S., Friedman, B., 1984. Stock prices and social dynamics. Brook. Pap. Econ. Act. 1984 (2), 457–510.
Shiller, R., Pound, J., 1989. Survey evidence on diffusion of interest and information among investors. J. Econ. Behav. Org. 12 (1), 47–66.
Spann, M., Skiera, B., 2009. Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. J. Forecast. 28 (1), 55–72.
Stewart, J.B., 1991. Den of Thieves. Touchstone.
Takahata, N., 1993. Allelic genealogy and human evolution. Mol. Biol. Evol. 10 (1), 2–22.
Tett, G., 2009. Fool's gold: how unrestrained greed corrupted a dream, shattered global markets and unleashed a catastrophe. Little Brown GBR.
The Financial Crisis Inquiry Commission, 2011. The Financial Crisis Inquiry Report. http://www.fcic.law.stanford.edu
Vega-Redondo, F., 1997. The evolution of Walrasian behavior. Econometrica 65 (2), 375–384.
Vriend, N., 2002. Was Hayek an ACE? South. Econ. J. 68 (4), 811–840.
Vulkan, N., 2000. An economist's perspective on probability matching. J. Econ. Surv. 14 (1), 101–118.
Wakano, J., Aoki, K., Feldman, M., 2004. Evolution of social learning: a mathematical analysis. Theor. Popul. Biol. 66 (3), 249–258.
Weibull, J., 1995. Evolutionary Game Theory. The MIT press.
Wolfers, J., Zitzewitz, E., 2004. Prediction markets. J. Econ. Perspect. 18 (2), 127–141.
Wright, S., 1931. Evolution in Mendelian populations. Genetics 16 (2), 97.
Young, H., 1975. Social choice scoring functions. SIAM J. Appl. Math., 824–838.
Ziegelmeyer, A., Koessler, F., Bracht, J., Winter, E., 2010. Fragility of information cascades: an experimental study using elicited beliefs. Exp. Econ. 13 (2), 121–145.