Adaptive Social Learning*

CHRISTOPH MARCH†

THIS VERSION: February 1, 2016

Abstract

The paper investigates social-learning when the information structure is not commonly known. Individuals repeatedly interact in social-learning settings with distinct information structures. In each round of interaction, they use their experience gained in past rounds to draw inferences from their predecessors’ current decisions. Such adaptation yields rational behavior in the long-run if and only if individuals distinguish social-learning settings and receive rich feedback after each round. Limited feedback may lead individuals to imitate uninformed predecessors. Moreover, adaptation across social-learning settings renders Bayes’ rule payoff-inferior compared to non-Bayesian belief updating rules and suggests that belief-updating rules are likely to be heterogeneous in the population.

KEYWORDS: Informational herding; Adaptation; Analogy-based expectations equilibrium; Non-Bayesian updating.

JEL CLASSIFICATION: C73, D82, D83.

1 Introduction

In many economic settings with observable actions, individuals with limited information about a payoff-relevant state of nature benefit from learning from the actions of other individuals. Such social-learning has been identified among others in financial and microloan markets and with respect to consumption of experience goods like movies or restaurant meals (Cipriani and Guarino, 2014; Zhang and Liu, 2012; Moretti, 2011; Cai et al., 2009).

Beginning with Bikhchandani et al. (1992) and Banerjee (1992), an extensive research program has investigated rational social-learning. This literature assumes that individuals

---

*The paper is based on chapter II of my PhD thesis. I thank in particular Anthony Ziegelmeyer, Christophe Chamley, and Philippe Jehiel for advice, suggestions and support. I also thank Thomas Daske, Christian Feilcke, Florian Inderst, Frédéric Koessler, Dorothea Kübler, Marco Ottaviani, Marco Sahm, Georg Weizsäcker, and seminar audiences at Berlin, Boston, Jena, Munich, Paris and Stony Brook for helpful comments. Finally, I gratefully acknowledge financial support from the ERC (advanced grant GTAPCL) and hospitality of Paris School of Economics during my time as a postdoctoral researcher.

†TUM School of Management, Technische Universität München, Arcisstraße 21, D-80333 Munich, Germany; Email: christoph.march@tum.de.
are rational, form beliefs using Bayes’ rule, and possess common strategic and structural knowledge meaning that Bayesian rationality and the structure of the social-learning setting are commonly known. If the set of actions is discrete, rational social-learning leads individuals to eventually herd on an action. As a consequence, information aggregation fails spectacularly and the economic outcome is inefficient.¹

Despite the many important insights it delivers, rational social-learning also reaches some unsound conclusions. For instance, even in a long herd, individuals never (or only very slowly) become extremely confident in the correctness of the chosen action. Moreover, if an individual with very precise information goes against the herd, her decision overturns it since her successors correctly infer the precise information (Smith and Sørensen, 2000). Many of these conclusions hinge upon the models’ strong underlying assumptions.

Though (commonly known) rationality can be criticized on many grounds, the present paper mainly investigates the implications of assuming that individuals possess structural knowledge. In particular, I question the assumption that the information structure, i.e. the distribution of private information, is commonly known. Common knowledge of the information structure is innocuous only if a “physical” ex ante procedure generates private information (Dekel and Gul, 1997). For example, if the state of nature is the amount of oil in some tract and firms’ private information results from taking soil samples, published experiments provide a thorough understanding of both the prior likelihood of oil and the distribution of samples as a function of the oil in the tract (Hendricks and Kovenock, 1989). On the other hand, fueled by the rapid diffusion of Internet technologies, there are many important economic settings where interaction is anonymous and information sources are wide and dispersed. For example, lenders in online peer-to-peer lending markets successfully use “soft” information – such as written statements of borrowers about the reasons for the loan application – to predict the creditworthiness of borrowers (Iyer et al., 2015; Lin et al., 2013).² Furthermore, some decision settings are characterized by an abundance of diverging information (Heal and Millner, 2014).

In the absence of common knowledge of the information structure, individuals are unable to deduce the informational content of the actions they observe. My main assumption is that repeated interaction enables individuals to acquire an understanding of the information conveyed by observed actions. Specifically, individuals repeatedly interact in social-learning settings and they get feedback about the state of nature and the actions

¹Social-learning is efficient only if individuals choose from a continuum of actions and are rewarded according to the proximity of their action to the most profitable one (Lee, 1993). If the set of actions is discrete and private signals are of bounded strength, rational social-learning quickly stops, and individuals herd on a wrong action with positive probability (Banerjee, 1992; Bikhchandani et al., 1992). With a discrete set of actions and unbounded private signals, the correct action is chosen asymptotically, but information aggregation is often extremely slow and long herds still emerge (Smith and Sørensen, 2000). The results carry over to social-learning settings with general observation structures (Smith and Sørensen, 2008; Acemoglu et al., 2011) and settings with an informationally efficient price process and a sufficiently rich information structure (Avery and Zemsky, 1998; Park and Sabourian, 2011).

²See Oberlechner and Hocking (2004), Kulkarni et al. (2012) for further examples.
of others after each repetition. They use this feedback to assess conditional probabilities of a given history of actions, conditional on each possible state of nature. For a given history and state, the conditional probability is assessed by the relative frequency with which this history occurred across past repetitions in which the given state was realized. I further assume that relative frequencies for each state are combined according to Bayes’ rule with private information, and that individuals pick the action which maximizes their expected payoff in the current repetition. Accordingly, individuals are myopic because they ignore repeated-game considerations.

The paper analyzes the long-run outcome of the above defined adaptive process. Yet, in real-world environments individuals are unlikely to encounter exactly the same strategic setting many times. I therefore assume that repeated interactions take place in several different social-learning settings. I focus on settings which differ only with respect to the information structure.

The first result of the paper establishes sufficient conditions under which long-run behavior mimics rational social learning: First, individuals are able to distinguish settings and second, the state of nature is revealed after each repetition (Proposition 1). The subsequent analysis proves that these assumptions are also necessary. First, I analyze how limited feedback regarding the state of nature affects adaptation. In particular, the state may not be revealed unless a certain action is taken (e.g. a good is bought or an investment is realized). As a consequence, individuals are likely to excessively imitate actions whose payoffs cannot reveal the state (Proposition 2). Such imitation may be based on no information at all. This explains why people are susceptible to following false prophets, joining cults, or relying on anecdotal reasoning. Second, I study the long-run outcome of adaptation across settings. Individuals adapt across settings if they use the feedback from past repetitions regardless of the respective social-learning setting. This may stem from an inability to distinguish settings or a desire to rely on a larger amount of feedback. I show (in Proposition 3) that the long-run outcome of adaptation across settings is captured by an analogy-based expectations equilibrium (Jehiel, 2005). In equilibrium, individuals bundle the decision situations of others into analogy classes and have a correct understanding of average behavior in each class. In the present setting, bundling leads to systematically biased inferences from observed actions. Therefore, long-run behavior does not maximize individuals’ (ex ante) expected payoffs (Proposition 4), and herding may spill-over from one social-learning setting to another. Moreover, adaptation across settings renders Bayes’ rule payoff inferior compared to non-Bayesian belief updating rules: Individuals with the most precise private information benefit from overweighting private information relative to others’ actions. Conversely, individuals with the least precise private information benefit from underweighting private information (Proposition 5).

The results of the paper provide several new insights: First, the results clarify when and why the assumptions underlying rational social-learning are justified. Second, the results suggest that belief updating rules are likely to be heterogeneous in the population.
when individuals are unable to distinguish social-learning settings. This provides an explanation for corresponding empirical findings (see e.g. Palfrey and Wang, 2012). Third, the results straightforwardly lead to a behavioral model of social-learning with heterogeneous belief updating rules which is able to accommodate the experimental regularities on social-learning. Numerous laboratory studies have established that herds emerge later than predicted by rational social-learning and that the length and strength of herds are positively correlated (e.g. Kübler and Weizsäcker, 2004, 2005; Goeree et al., 2007). In particular, individuals become extremely confident even in wrong actions. Though the experiments have stimulated an active behavioral literature (see below), none of the alternative theories organizes well the bulk of the experimental evidence. A behavioral model of social-learning with flexible belief updating rules accommodates the experimental regularities and is also able to capture more recent evidence of a ‘social-confirmation’-bias among participants in a social-learning experiment with a richer information structure (March and Ziegelmeyer, 2015).

The paper relates to a growing literature on social-learning with bounded rationality. It complements the analysis of Guarino and Jehiel (2013) who assume that individuals only understand the relation between the aggregate distribution of actions and the state of nature. This correct understanding is assumed to emerge from an adaptive process with limited feedback. The two papers therefore characterize the long-run outcome of adaptation under different feedback regimes. The present paper also suggests a re-interpretation of Bohren (2015) who assumes that (i) a fraction of individuals are socially uninformed and decide based only on their private signals, and (ii) other individuals misperceive the exact proportion of uninformed predecessors. Based on the results presented here, Bohren’s (2015) model may be re-interpreted as a social-learning model with (a specific form of) heterogeneous belief updating rules and a non-common prior.3 Finally, the naïve inference model of Eyster and Rabin (2010), in which individuals believe that the action of each previous individual reveals that individual’s private information, can be seen as an alternative model of initial behavior.

A recurrent issue in the literature has been the emergence of extreme confidence in wrong beliefs (Eyster and Rabin, 2010; Guarino and Jehiel, 2013). This confidence cannot emerge in models of rational social-learning, but it has been repeatedly observed in the laboratory and in the field. A social-learning model with belief updating heterogeneity is able to predict extreme, false beliefs alongside a delayed formation of herds, which has been consistently found in the lab as well.

The paper is structured as follows. Section 2 outlines the results of the paper with the help of a simple example. Section 3 introduces the analytical framework, discusses rational social-learning, and formalizes adaptation. Section 4 characterizes the long-run outcome

---

3 See also Bernardo and Welch (2001) and Kariv (2005) both of which introduce individuals who overweight their private information into the standard model of social-learning.
of adaptation under different assumptions regarding individuals’ feedback. Section 5 introduces a model of social learning based on heterogeneous belief updating and discusses its relation to the adaptation results. Section 6 presents extensions of the model towards heterogeneous preferences and an endogeneous timing of decisions. Section 7 concludes. The appendix contains the proofs and some additional analyses.

2 A Simple Example

Consider the following social-learning game with two players: Anna and Bob decide in sequence whether or not to invest. Anna decides first, and Bob decides after having observed Anna’s decision. Payoffs from investing and, respectively, rejecting are identical for both players. The investment payoff $\theta$ takes values 0 and 1 with equal probability, the cost of an investment is $c = 1/2$, and the payoff from rejection is zero. Before deciding, each player $i \in \{A, B\}$ observes a symmetric, binary private signal $s_i \in S = \{0, 1\}$ where $\Pr(s_i = 0 \mid \theta = 0) = \Pr(s_i = 1 \mid \theta = 1) = q_i \in (0.5, 1)$ denotes the signal precision of player $i$. Signals are independent conditional on the state $\theta$.

I will assume throughout that Anna is Bayes-rational which implies that her dominant strategy is to reject if $s_A = 0$ and to invest if $s_A = 1$.

2.1 Rational Social-Learning

In this example, the assumptions of rational social-learning entail that (i) not only Anna but also Bob is Bayes-rational, (ii) Bob knows that Anna is Bayes-rational, i.e. he possesses strategic knowledge, and (iii) both Anna and Bob know the payoff function and the signal precisions $q_A$ and $q_B$, i.e. they possess structural knowledge. Bob’s knowledge implies that he identifies Anna’s decision to invest (reject) with the signal $s_A = 1$ ($s_A = 0$). Accordingly, Bob’s likelihood ratio, i.e. the ratio of probabilities he assigns to the investment payoff being 1 versus 0 given signal $s_B \in \{0, 1\}$ and Anna’s action $x_A \in \{\text{invest}, \text{reject}\}$ is given by

$$
\lambda(s_B, x_A) = \frac{\Pr(\hat{\theta} = 1 \mid s_B, x_A)}{\Pr(\hat{\theta} = 0 \mid s_B, x_A)} = \frac{\Pr(s_B \mid \theta = 1)}{\Pr(s_B \mid \theta = 0)} \cdot \begin{cases} \frac{q_A}{1-q_A} & \text{if Anna invests} \\ \frac{1-q_A}{q_A} & \text{if Anna rejects} \end{cases}.
$$

Bob invests (rejects) if $\lambda(s_B, x_A) > (<) 1$. Therefore, Bob follows his private signal if it confirms Anna’s decision (Anna invests and $s_B = 1$, or Anna rejects and $s_B = 0$). On the other hand, if Anna’s decision contradicts his private signal, Bob follows his private signal if $q_B > q_A$ and imitates Anna’s decision if $q_B < q_A$.\footnote{I omit the case $q_A = q_B$ for which rational social-learning makes no clear-cut prediction.}
2.2 Adaptive Social-Learning

Contrary to rational social-learning, adaptive social-learning assumes that individuals do not possess strategic or structural knowledge, but interact repeatedly. Since Bayes-rational Anna has the dominant strategy to follow her private signal, only Bob’s behavior differs between the two approaches. Concretely, I make the following assumptions: Anna and Bob repeatedly play the social-learning game in rounds \( r = 1, 2, \ldots \). In each round, a new investment payoff is determined randomly and independently from the investment payoff in previous rounds. Players learn the realized investment payoff at the end of a round. This enables Bob to track the relationship between Anna’s decision and the investment payoff, i.e. he keeps counts of how often Anna invests and rejects when the investment payoff is 1 and when it is 0. Finally, both players are Bayes-rational and myopically maximize their expected payoff in the current round.

The long-run outcome of the above defined adaptive process is easily derived: Myopic Anna follows her private signal in each round. Therefore, the relative frequency with which Anna invests when the investment payoff equals 1 (respectively 0) approaches \( q_A \) (respectively \( 1 - q_A \)). As Bob tracks this relative frequency, he eventually infers the same information from Anna’s decision as he could deduce when possessing strategic and structural knowledge. Hence, Bob eventually plays his unique rationalizable strategy.

In summary, the adaptive social-learning outcome coincides with the rational social-learning outcome.

2.3 Adaptive Social-Learning Across Games

Real-world social-learning is likely to take place in a multitude of settings and individuals are unlikely to distinguish those settings in their finest details. I therefore investigate whether adaptive social-learning also leads to rational social-learning when players must simultaneously adapt to multiple games.

I assume the following: Anna and Bob repeatedly play two different social-learning games. The games differ only with respect to the private signal precisions. In game \( k \in \{H, L\} \) the signal precision of player \( i \in \{A, B\} \) is \( q_i^k \in (0.5, 1) \) where \( q_A^L < q_A^H \) without loss of generality. In each round, the game to be played is determined randomly and independently of previous rounds; both games are equally likely. As before, Bob tracks the relation between Anna’s decision and the realized investment payoff across rounds. However, Bob does not distinguish the games, i.e. he keeps single counts of how often Anna invests and rejects when the investment payoff is 1 and when it is 0, and he uses them to learn from Anna’s decision in both games.

As before, Anna follows her private signal in each round and each game. Because each game occurs on average in half of the rounds, the relative frequency with which Anna invests when the investment payoff equals 1 (resp. 0) approaches \( \overline{q}_A = \frac{1}{2} (q_A^L + q_A^H) \) (resp. \( 1 - \overline{q}_A \)) across the two games. Since Bob does not distinguish the two games, this relative
frequency eventually guides his behavior. Accordingly, Bob eventually follows his private signal in game \( k \) if \( q_B^K > q_A^K \), and he eventually imitates Anna’s decision if \( q_B^k < q_A^k \).

Obviously, Bob’s long-run behavior when he adapts across games is not optimal in game \( L \) (resp. \( H \)) if \( q_A^L < q_B^L < q_A^H \) (resp. \( q_A^L < q_B^H < q_A^H \)). Moreover, Bob’s long-run behavior may not maximize his ex ante expected payoff from the randomly selected social-learning game. Consider for instance two games such that signal precisions are low (high) in game \( L \) (resp. \( H \)) since information is scarce (abundant). If Bob’s signal precision is higher than Anna’s in each game \( q_B^k > q_A^k \) for each \( k \), but Anna’s average signal precision is higher than Bob’s signal precision in game \( L \) \( (q_A^L > q_B^L) \), Bob suboptimally imitates Anna’s decision in game \( L \) in the long-run. Equivalently, Bob suboptimally follows his private signal in game \( H \) in the long-run, if \( q_B^k < q_A^k \) for each \( k \) and \( q_B^H > q_A^H \).

So far, I have assumed that players update beliefs according to Bayes’ rule. A more general adaptive process allows players to also grope for the optimal response to beliefs. I therefore extend the adaptive process by considering flexible belief updating rules where Bob’s generalized likelihood ratio given signal \( s_B \) and Anna’s decision \( x_A \) is given by

\[
\hat{\lambda}(s_B, x_A) = \left( \frac{\Pr(s_B | \theta = 1)}{\Pr(s_B | \theta = 0)} \right)^{\beta_B} \cdot \begin{cases} 
\frac{q_A}{1-q_A} & \text{if Anna invests} \\
\frac{1-q_A}{q_A} & \text{if Anna rejects}
\end{cases}
\]

\( \beta_B > 0 \) denotes Bob’s private information weight. Bob is Bayesian for \( \beta_B = 1 \), overweights private information for \( \beta_B > 1 \), and underweights private information for \( \beta_B < 1 \).

If Bob has a higher signal precision than Anna in each game, adaptation across games may lead him to eventually imitate Anna’s decision in setting \( L \). Bob’s strategy will therefore improve if he overweights his private signal. Equivalently, underweighting his private signal improves Bob’s strategy, if adaptation across games leads him to suboptimally follow his private signal in game \( H \). Figure 1 shows all possible combinations of signal precisions \( (q_A^L, q_A^H, q_B^L, q_B^H) \in \left( \frac{1}{2}, 1 \right)^4 \) for which over- or underweighting his private signal improves Bob’s strategy. Settings with signal qualities in the blue (light blue) area strictly (weakly) favor overweighting of private information, while settings with signal qualities in the orange (light orange) area strictly (weakly) favor underweighting of private information.\(^5\) The result indicates that optimal belief updating rules are likely to be heterogeneous in the population.

2.4 Discussion

There are two main reasons for Bob to adapt across games. First, Bob may be unable to distinguish games since he lacks or neglects relevant information. For instance, Anna’s signal precision may not be known to Bob, and he may fail to recognize that his own and Anna’s signal precision are correlated. Second, Bob’s experience is likely restricted to a

\(^5\) See Proposition A.1 in the appendix for a formal statement and the proof.
finite number of interactions in which case it may be optimal to adapt across games. Appendix B provides numerical results establishing that adaptation across games is optimal for Bob even if it leads to systematically mistaken long-run beliefs as long as the number of repetitions is not too large.

A second remark concerns the feedback Bob receives after each round. Indeed, the investment payoff may not be revealed to Bob unless he invests. Assuming this changes the long-run outcome of the adaptive process. After any finite number of rounds, Bob believes with strictly positive probability that imitating Anna’s decision to reject regardless of his private signal maximizes his expected payoff. At this point however, the investment payoff is no longer revealed to Bob whenever Anna rejects. Bob is therefore unable to ever revise his wrong inference from a rejection by Anna and he imitates this decision in the long-run. Most importantly, this happens with strictly positive probability even if $q_B > q_A$, i.e. if it is optimal for Bob to never imitate Anna’s decision, and if $q_A \approx 0.5$, i.e. if Anna does not possess any valuable information.

3 General Analytic Framework

3.1 A Social-Learning Game

There is a finite sequence of players $t = 1, 2, \ldots, T$ who each choose an action $a_t$ from the set $A = \{0, 1\}$. The sequential order of players is exogenously, randomly determined and players are indexed according to their position. While payoff externalities are absent, player $t$’s payoff from her action depends on the realization of the state of Nature (hence-
forth state) $\tilde{\theta} \in \Theta = \{0, 1\}$.\textsuperscript{6} Ex ante the two states are equally likely. More precisely, player $t$’s payoff for each $t = 1, \ldots, T$ is determined by the vN-M payoff function

$$w_t(a_t, \theta) = \begin{cases} \theta - \frac{1}{2} & \text{if } a_t = 1 \\ 0 & \text{if } a_t = 0 \end{cases}. $$

In the following, action $a = 1$ ($a = 0$) is sometimes referred to as “invest” (“reject”) and the costs of the investment are set equal to $\frac{1}{2}$ merely to simplify the exposition.\textsuperscript{7}

Each player receives a private signal $\tilde{s}_t \in [0, 1]$ about the state. Conditional on the state, signals are independent and identically distributed. When the true state is $\theta$, the signal distribution is given by the cumulative distribution function $G_{\theta}$. $G_0$ and $G_1$ are mutually absolutely continuous and have common support $[b, \bar{b}] \subseteq [0, 1]$. Therefore, a positive, finite Radon-Nikodym derivative exists and satisfies $f(s) = \frac{s - b}{\bar{b} - b}$ (Smith and Sørensen, 2000). To ensure that some signals are informative, I rule out $f = 1$ almost surely. The assumptions imply that $G_0$ first-order stochastically dominates $G_1$ and that $s_t = \Pr(\tilde{\theta} = 1 | s_t)$. I assume that $b > 0$ and $\bar{b} < 1$, i.e. private signals are bounded.

While player $t$ only knows her own private signal, she observes additionally the complete history of previous actions denoted by $h_t = (a_1, \ldots, a_{t-1}) \in H_t = A^{t-1}$ (and $h_1 \equiv \emptyset$). Subsequently, $H = \bigcup_{t=1}^T H_t$ denotes the complete set of histories, and $H_{T+1} = A^T$ denotes the set of final histories with element $h_{T+1} = (a_1, \ldots, a_T)$.

The social-learning game is summarized by the collection $(T, A, \{u_t\}_{t=1}^T, \Theta, (G_\theta)_{\theta \in \Theta})$.

### 3.2 Rational Social-Learning

Rational social-learning relies on four main assumptions. First, players form beliefs about the state of nature by combining all the available information using Bayes’ rule. Second, players maximize their expected utility given these beliefs. I refer to the first two assumptions as Bayesian rationality of players. Third, Bayesian rationality is commonly known; in other words players possess (common) strategic knowledge. Fourth, the social-learning game is commonly known meaning that players also possess (common) structural knowledge.\textsuperscript{8} The distinction between the two forms of common knowledge is important. Indeed, the next section investigates which adaptive process generates long-run outcomes equivalent to rational outcomes of social-learning when players possess strategic knowledge but are deprived from structural knowledge.

Without loss of generality, I focus on behavioral strategies $\sigma_t : [b, \bar{b}] \times H_t \to \Delta(A)$.

---

\textsuperscript{6}Throughout, tilded letters ($\tilde{\theta}$) denote random variables and standard letters ($\theta$) denote realizations.

\textsuperscript{7}Similarly, the state and the action set are binary to simplify the exposition. The results extend to any finite number of actions and states but at significant algebraic cost.

\textsuperscript{8}See Brandenburger (1996) for a similar distinction between strategic and structural uncertainty.
In a slight abuse of notation, $\sigma_t(s_t, h_t)$ denotes player $t$’s probability of investment given her private signal $s_t$ and the history $h_t$. Let $\Sigma_t$ denote the strategy set of player $t$, $\Sigma = \times_{t=1}^T \Sigma_t$ the set of strategy profiles, and $\Sigma_{-t} = \times_{\tau \neq t} \Sigma_{\tau}$ the set of profiles of strategies for all players excluding $t$. The ex-ante expected payoff of player $t$ following strategy $\sigma_t \in \Sigma_t$ for given $\sigma_{-t} \in \Sigma_{-t}$ is given by

$$U_t(\sigma_t, \sigma_{-t}) = \frac{1}{2} \sum_{b \in \Theta} \sum_{h_t \in H_t} Pr(h_t \mid \Hat{\theta} = \theta, \sigma) \int \sum_{a \in A} \sigma_t(a \mid s_t, h_t) u(a, \theta) dG_{\theta}(s_t).$$

(1)

As shown by Tan and Werlang (1988), the assumptions above restrict players to iteratively undominated strategies. Strategy $\sigma_t \in \Sigma_t$ is strictly dominated if there exists $\sigma'_t \in \Sigma_t$ such that $U_t(\sigma'_t, \sigma_{-t}) > U_t(\sigma_t, \sigma_{-t})$ for every $\sigma_{-t} \in \Sigma_{-t}$. A strategy profile $\sigma = (\sigma_t, \ldots, \sigma_T)$ is iteratively undominated if each strategy $\sigma_t$ survives the iterated elimination of strictly dominated strategies. Lemma 1 characterizes all iteratively undominated strategies in terms of sequential best responses to Bayesian consistent beliefs $b_t : [b, \tilde{b}] \times H_t \to [0, 1]$ where $b_t(s_t, h_t) = Pr(\hat{\theta} = 1 \mid \hat{s}_t = s_t, \hat{h}_t = h_t)$.

**Lemma 1.** To any iteratively undominated strategy profile $\sigma$ there exists a belief system $\{b_t^*\}_{t=1}^T$ such that for each $t = 1, \ldots, T$, each $s_t \in [b, \tilde{b}]$, and each $h_t \in H_t$

(i) beliefs are formed according to Bayes’ rule, i.e.

$$b_t^*(s_t, h_t) = \left[1 + \frac{1 - s_t}{s_t} \frac{Pr(h_t \mid \hat{\theta} = 0, \sigma^*)}{Pr(h_t \mid \hat{\theta} = 1, \sigma^*)}\right]^{-1}$$

if $Pr(h_t \mid \hat{\theta} = \theta, \sigma^*) = \prod_{\tau < t} \int_{b}^{\tilde{b}} \sigma_{\tau}^*(a_{\tau} \mid s_{\tau}, h_{\tau}) dG_{\theta}(s_{\tau}) > 0$ for each $\theta \in \Theta$ where $a_{\tau} = h_{\tau}(\tau)$ and $h_{\tau} \subset h_t$ for each $\tau < t$,

(ii) behavioral strategies are sequentially rational, i.e.

$$\sigma_t^*(s_t, h_t) = \begin{cases} 1 & \text{if } b_t^*(s_t, h_t) > \frac{1}{2} \\ 0 & \text{if } b_t^*(s_t, h_t) < \frac{1}{2} \end{cases}.$$ 

The iteratively undominated outcome is almost always unique.\(^\text{10}\)

In the following, $\sigma^*$ denotes the iteratively undominated strategy profile characterized in the lemma. While iterated dominance does not restrict behavior in case of a tie ($b_t^*(s_t, h_t) = 1/2$), ties occur with probability zero almost always. I therefore assume henceforth that the social-learning game does not allow for ties which is why there is

\(^9\)For a given set $M$, $\Delta(M)$ is the set of probability distributions over $M$.

\(^\text{10}\)i.e. the set of parameters of the social-learning game for which there exist multiple iteratively undominated outcomes is a null set.
no need to commit to a specific tie-breaking rule. Absent ties, iterated elimination of dominated strategies yields a unique outcome which is also the unique (Perfect) Bayesian equilibrium outcome.

Rational social-learning entails that players benefit individually from taking into account the information revealed by others’ actions. Collectively, however, rational social-learning is self-defeating (Chamley, 2004b), because the more a player’s decision is influenced by the actions of her predecessors (the more distinct are the probabilities \( \Pr(h_t \mid \theta, \sigma^*) \) for \( \theta \in \Theta \)), the less new information it conveys to her successors (the more signals \( s_t \in S \) yield the same sign of \( b_t^i(s_t, h_t) - 1/2 \)). As a consequence, information aggregation slows down or stops completely which is why the economic outcome is inefficient.

### 3.3 Adaptive Social-Learning

Consider a (finite) family of social-learning games \( (T, A, u, \Theta, (G^k_0, G^k_1)) \), \( k \in \mathcal{K} = \{1, \ldots, K\} \), which differ only in the distribution of private signals, \( G^k_0 \) and \( G^k_1 \), with support \( \big[ \theta^k, \theta^l \big] \). Players are assumed not to know these distributions. They interact repeatedly over the rounds \( r = 1, 2, \ldots \) where in each round one of the \( K \) games is played. Concretely, each round \( r \) begins with the random draws of (i) the game \( k^r \) where \( k^r = k \) with probability \( \pi_k > 0 \), (ii) the state of nature \( \theta^r \) where both states are equally likely in each round, (iii) the sequence of private signals \( \{ s^r_{t} \}_{t=1}^{T} \) where signals are independent across periods and drawn according to \( G^k_{\theta^r} \), and (iv) the order of players such that each player eventually occupies any position in the sequence. Random draws are independent across rounds which guarantees that learning about the state and learning about the structure of the game and the strategies of other players are not confounded. At the end of each round, payoffs are realized.

Adaptation through repeated interaction requires players to receive feedback after each round. I suppose that private signals are never publicly revealed. Accordingly, for a given player \( i \) (in an abuse of notation) the relevant outcome of the game after round \( r \) is given by the tuple \( \omega^i_r = (k^r, \theta^r, t^r_i, h^r_{T+1}, \omega^r_i) \in \Omega \equiv \mathcal{K} \times \Theta \times \{1, \ldots, T\} \times H_{T+1} \) where \( t^r_i \) denotes the position occupied by the player in round \( r \). Following Esponda (2008), I formally capture feedback via the functions \( y_K : \Omega \rightarrow 2^\mathcal{K} \), \( y_\Theta : \Omega \rightarrow 2^\Theta \), and \( y_H : \Omega \rightarrow 2^H \). When the outcome of the game is \( \omega^r \in \Omega \), the player observes that the game \( k^r \) belongs to the set \( y_K(\omega^r) \). Similarly, the player observes that the state \( \theta^r \) belongs to the set \( y_\Theta(\omega^r) \) and that histories \( h^r_t \in y_H(\omega^r) \) where \( t \in \{1, \ldots, T\} \) occurred. Let \( y(\omega^r) = (y_K(\omega^r), y_\Theta(\omega^r), y_H(\omega^r)) \). An adaptation path for round \( r \) is given by \( \zeta^r = (\omega^1, \ldots, \omega^{r-1}) \in \Omega^{r-1} \).

Since players do not know the information structure, they cannot derive the information contained in a history \( h_t \in H_t \), i.e. the (game-specific) probabilities \( \Pr(h_t \mid \theta) \) for each \( \theta \in \Theta \) from their knowledge. Instead, players are assumed to assess this information based

---

111 I assume that a player always remembers her position and I omit the subscript \( i \) henceforth.
on their feedback from past interactions. Formally, for a given game $k$ an assessment for period $t$ is a mapping $\varphi_t^k : \Theta \rightarrow \Delta (H_t)$ assigning to each state $\theta \in \Theta$ a probability distribution $\varphi_t^k (h_t \mid \theta)$ over histories $h_t \in H_t$.

**Definition 1.** Let $\eta > 0$. The adaptive process is given by a sequence of frequencies $\kappa_t^{k,r} : H_t \times \Theta \times \Omega^{-1} \rightarrow \mathbb{N}$, a sequence of assessments $\varphi_t^{k,r} : \Theta \times \Omega^{-1} \rightarrow \Delta (H_t)$, and a sequence of strategic responses $\sigma_t^{k,r} : [b_k, \bar{b}_k] \times H_t \times \Omega^{-1} \rightarrow \Delta (A)$ for each $t = 1, \ldots, T$ and each $k \in \mathcal{K}$, such that

(i) for each $r \geq 1$, $\zeta_r \in \Omega^{-1}$, $k \in \mathcal{K}$, $t = 1, \ldots, T$, $h_t \in H_t$, and $\theta \in \Theta$

$$\kappa_t^{k,r} (h_t, \theta \mid \zeta_r) = \left| \{ 1 \leq r' < r : k \in y_k (\omega^{\rho}) \land h_t \in y_H (\omega^{\rho}) \land \theta \in y_\Theta (\omega^{\rho}) \} \right|, \quad (2)$$

(ii) for each $r \geq 1$, $\zeta_r \in \Omega^{-1}$, $k \in \mathcal{K}$, $t = 2, \ldots, T$, $h_t \in H_t$, and $\theta \in \Theta$

$$\varphi_t^{k,r} (h_t \mid \theta ; \zeta_r) = \frac{\kappa_t^{k,r} (h_t, \theta \mid \zeta_r) + \eta}{\sum_{h_t' \in H_t} \kappa_t^{k,r} (h_t', \theta \mid \zeta_r) + \eta}, \quad (3)$$

(iii) for each $r \geq 1$, $\zeta_r \in \Omega^{-1}$, $k \in \mathcal{K}$, $t = 1, \ldots, T$, $h_t \in H_t$, and $s_t \in [b_k, \bar{b}_k]$

$$\sigma_t^{k,r} (s_t, h_t \mid \zeta_r) = 1 (0) \quad \text{if} \quad \frac{s_t}{1 - s_t} > \left< \frac{\varphi_t^{k,r} (h_t \mid 0; \zeta_r)}{\varphi_t^{k,r} (h_t \mid 1; \zeta_r)} \right>. \quad (4)$$

The adaptive process relies on two main assumptions:12 First, in a given round $r$, period $t$, and game $k$, players form beliefs at history $h_t$ by replacing for each state $\theta \in \Theta$ the (unknown) conditional probability $\Pr (h_t \mid \theta)$ with the relative frequency with which history $h_t$ occurred across relevant past rounds in which the state was $\theta$. The past rounds used to calculate this relative frequency are determined by players’ feedback. For example, players only rely on past rounds in which the same game was played if $y_k (\omega^{\rho}) = \{ k^{r} \}$. In contrast, players rely on past rounds regardless of the respective social-learning game if $y_k (\omega^{\rho}) = \mathcal{K}$. Moreover, some past rounds may not be usable because the realized state was not revealed (i.e. $y_\Theta (\omega^{\rho}) = \emptyset$ for some $\rho < r$). Since the relative frequency is not well defined if no past rounds are usable for a given couple (history,state), I assume that players have arbitrarily small initial weights $\kappa_t^{k,1} (h_t, \theta \mid \emptyset) = \eta > 0$ for each $k$, $t$, $h_t$, and $\theta$ and I focus on the limit as $\eta \rightarrow 0$. In the limit, players attach probability zero to histories not observed previously and they believe that such histories are uninformative.

The second assumption is that players combine the relative frequencies according to Bayes’ rule with their private signal, and that they rationally respond to the resulting belief. This assumption entails that players are myopic, i.e. they do no engage in strategic

---

12The adaptive process is based on the idea of fictitious play (Brown, 1951).
considerations of repeated play. In section 5, I relax the assumption that players are Bayesian by allowing them to experiment with flexible updating rules.

One feature of the adaptive process deserves special emphasis: While a player assesses the informational content of others’ actions using past experience, she is assumed to know the informational content of her own private signal. The dichotomy relies on an interpretation of social-learning as a process in which informed players (though imperfectly) attempt to learn from others’ decisions. Accordingly, private information is the outcome of an active process of information search and selection of the most credible source. In contrast, observed decisions are driven by unknown information sources and garbled through others’ strategic thinking.

The adaptive process is very specific about how players form and respond to assessments. More general models of adaptation stay agnostic about how players exactly reach their decisions, allow for active experimentation, and focus on the asymptotic properties of the adaptive process (Fudenberg and Levine, 1998). A generalization of the adaptive process in this direction is discussed in Appendix C.

3.4 Convergence

To formalize convergence of the adaptive process I consider the following two metrics on, respectively, the set of assessments and the strategy space:

**Definition 2.** Let $\epsilon > 0$. For each $k \in \mathcal{K}$ and each $t = 1, \ldots, T$

(i) assessments $\varphi^k_t, \hat{\varphi}^k_t : \Theta \to \Delta(H_t)$ are $\epsilon$-close if $|\varphi^k_t(h_t | \theta) - \hat{\varphi}^k_t(h_t | \theta)| < \epsilon$ for each $h_t \in H_t$ and each $\theta \in \Theta$,

(ii) strategy $\sigma^k_t$ plays $\epsilon$-like strategy $\hat{\sigma}^k_t$ at history $h_t \in H_t$ if there exists $B_{\epsilon} \subseteq \left[\begin{array}{c}b^k, \bar{b}^k\end{array}\right]$ such that $G^k_B(\theta) > 1 - \epsilon$ for each $\theta \in \Theta$, and $|\sigma^k_t(s_t, h_t) - \hat{\sigma}^k_t(s_t, h_t)| < \epsilon$ for each $s_t \in B_{\epsilon}$ (Jackson and Kalai, 1997).

Lemma 2 establishes that the two distance functions are consistent meaning that $\epsilon$-closeness of assessments implies $\epsilon$-like play of strategies and vice versa. To state the result, call a strategy profile $\sigma^k$ and a profile of assessments $\{\varphi^k_t\}_{t=1}^T$ corresponding, if (i) $\varphi^k_t(h_t | \theta) = \prod_{t \in T} \int_{b^k}^{\bar{b}^k} \sigma^k_t(a_t | s_t, h_t) dG^k_B(s_t)$ and (ii) $\sigma^k_t(s_t, h_t) = 1(0)$ if $\frac{a_t}{1-a_t} > (<) \frac{\varphi^k_t(h_t | \theta)}{\varphi^k_t(h_t | \bar{\theta})}$ for each $t = 1, \ldots, T$, each $h_t \in H_t$, and each $s_t \in \left[\begin{array}{c}\underline{b}^k, \bar{b}^k\end{array}\right]$.

**Lemma 2.** Fix $k \in \mathcal{K}$ and consider the strategy profiles $\sigma^k, \hat{\sigma}^k$ with corresponding assessments $\{\varphi^k_t\}_{t=1}^T$ and $\{\hat{\varphi}^k_t\}_{t=1}^T$. For each $t = 1, \ldots, T$ and each $\epsilon > 0$ there exists $\delta > 0$ such that

(i) $\sigma^k_t$ plays $\epsilon$-like $\hat{\sigma}^k_t$ at each $h_t \in H_t$ satisfying $\hat{\varphi}^k_t(h_t | \theta) > 0$ for each $\theta \in \Theta$, if $\varphi^k_t$ is $\delta$-close to $\hat{\varphi}^k_t$,
Based on Lemma 2, I will focus on convergence of strategies henceforth. All convergence results are in probabilistic terms with respect to the distribution $\mathbf{P}$ of (infinite) adaptation paths $(\omega^1, \omega^2, \ldots) \in \Omega^\infty$ induced by the objective distributions of the random variables and the rules for the formation of assessments and strategies.

### 4 Long-run Outcomes of Adaptation

In this section I discuss the long-run outcomes of the adaptive process under different feedback regimes. As a benchmark, I first characterize the long-run outcome if players observe the complete sequence of actions and the state at the end of each round, and are able to distinguish games (4.1). Second, I investigate the consequences of constraints on the observation of actions or the state (4.2). Finally, I study the outcome of the adaptive process under the assumption that players do not distinguish games (4.3). While for the sake of clarity I assume mutual knowledge of the social-learning game, I refrain from assuming any higher-order interactive knowledge.

#### 4.1 Adaptation by Game

I first assume that players observe the outcome of the game after each round.

**Proposition 1.** Let $y(\omega^r) = (\{k^r\}, \{\theta^r\}, \{h_t \subset h^r_{T+1}\})$ for each $r \geq 1$ and each $\omega^r \in \Omega$. For each $\epsilon > 0$, each $k \in K$, and each $t = 1, \ldots, T$, the limit strategy $\lim_{r \to \infty} \sigma_{k, r}^{k, *}$ almost surely plays $\epsilon$-like a rationalizable strategy $\sigma_{k, r}^{k, *}$ at all histories satisfying $\varphi_{k, r}^{k, *}(h_t | \theta) > 0$ for each $\theta \in \Theta$ where $\varphi_{k, r}^{k, *}(h_t | \theta)$ are the assessments under rational social-learning.

The proposition establishes that rational social-learning may be justified as the outcome of an adaptive process if players are able to distinguish games, receive ample feedback after each round, and play each game infinitely often. It follows from the dominance-solvability of the game (Milgrom and Roberts, 1991). In a nutshell, since players have a dominant strategy in period 1, period 2-assessments will eventually be correct by the law of large numbers if players distinguish games. Hence, strategies for period 2 (and by induction strategies for later periods) are eventually rationalizable.\(^{13}\)

\(^{13}\)Mutual knowledge of the social learning game is not necessary for the result. In the long-run any possible state and any possible (open subset of) private signal(s) occurs infinitely often such that frequentists may learn their own private signal distribution as long as its support is well-behaved. Moreover, players are able to explore the (finite) action spaces of other players and their payoff function once a small amount of experimentation is assumed.
4.2 Feedback Constraints

Proposition 1 relies on strong assumptions regarding players’ feedback. For instance, though the state must be revealed through the payoffs for at least one of the available actions, it may not be revealed for all of them. If a player rejects an investment opportunity, she may never know what the outcome could have been. To take this possibility into account, I now investigate the impact of feedback constraints on the state. Formally, I assume that \( y_{\omega} (\omega) = \emptyset \) if \( a_{\tau}^r = 0 \) and \( y_{\omega} (\omega) = \{ \theta^r \} \) if \( a_{\tau}^r = 1 \). I maintain the assumption that players get rich feedback on the game and the history.

**Proposition 2.** Assume that for each \( r \geq 1 \) and each \( \omega^r \in \Omega \), \( y_{K} (\omega^r) = k^r \), \( y_{H} (\omega^r) = \{ h_t \in h_T^r \} \), \( y_{\omega} (\omega^r) = \emptyset \) if \( a_{\tau}^r = 0 \), and \( y_{\omega} (\omega^r) = \{ \theta^r \} \) if \( a_{\tau}^r = 1 \). For each \( k \in K \) such that \( G_k^r (1/2) \geq \frac{b_k}{1-b_k} \), there exists \( \varepsilon > 0 \) and \( t > 1 \) such that with strictly positive probability the limit strategy \( \lim_{r \to \infty} \sigma^r \) does not play \( \varepsilon \)-like the rationalizable strategy \( \sigma^r_{k^r} \) at history \( h_t = h_t^0 \equiv (0, \ldots, 0) \) for each \( 0 < \varepsilon < \varepsilon \). Concretely, \( \lim_{r \to \infty} \sigma^r (s_r, h_r) = 0 \) for each \( \tau \geq t \), each \( h_r \equiv h_t^0 \), and each \( s_r \in \left[ b_k, \bar{b_k} \right] \).

The proposition shows that adaptive social-learning may differ from rational social-learning if feedback on the state is conditional on a player’s decision. In particular, players may in the long-run imitate the rejections of their predecessors although they would not do so if they knew the true informational content of previous decisions. More broadly, herds on an action that does not reveal the state through its payoffs are more likely in the long-run of adaptive social-learning than under rational social-learning, and more likely to be wrong.

The rationale behind the result is simple: After any finite number of rounds, players are with strictly positive probability convinced that the evidence conveyed by a sequence of rejections is sufficiently strong to swamp any private signal since private signals are bounded. Players therefore imitate their predecessors and are no longer able to revise their wrong assessments because of the constrained feedback. Formally, the limit outcome is a self-confirming equilibrium (see e.g. Dekel et al., 2004). The condition \( G_k^r (1/2) / G_k^r (1/2) < 1 \) guarantees that it is not optimal to imitate the first rejection in game \( k \). The result is constrained to histories containing only rejections since players switch positions in the social-learning game. Thus, players who invest before the herd starts would be able to revise their assessments. The result could be strengthened by assuming, additionally, that players receive no feedback about subsequent choices.

Proposition 2 continues to hold if the first players in the sequence are uninformed (and uninformed players move early for ulterior motives) as long as this is not mutually known. The proposition therefore helps explain why players sometimes follow others based on no

---

14 Another possibility is that a player does not get feedback on subsequent choices. Yet, since a player’s position in the sequence is randomly determined in each round, each player acts in period \( T \) infinitely often and observes the complete history of actions. Therefore, the result of Proposition 1 is robust with respect to limited feedback on the history.
apparent reason. For example, in recent years a growing number of parents refuse to
vaccinate their children for fear of harmful side effects despite the extensive statistical
evidence to the contrary. One reason seems to be a strong reliance on anecdotal evidence
(see e.g. Moran et al., 2015). In the model considered here, players are forced to rely on
anecdotal reasoning since feedback constraints prevent them from collecting representative
statistical information.

4.3 Adaptation Across Multiple Social Learning Games

In the field, individuals rarely encounter exactly the same strategic situation a large
number of times. Therefore, they are likely to extrapolate experience across games they
deeom similar (see e.g. Fudenberg, 2006). In this subsection, I investigate this idea by
assuming that players receive a coarse feedback $y^r_K(\omega^r) = K$ about the game.

There are several reasons why players adapt across games. First, absent common
knowledge of the information structure, players may be unable to distinguish games.
Grimm and Mengel (2012) show that experimental subjects extrapolate especially in
complex environments where information to distinguish games is scarce. Though a player
could assess the informational content of others’ actions conditional on her own private
signal,$^{15}$ she may fail to account for the correlation between her private signal and the
actions of others (Esponda, 2008), or ignore the importance of this factor altogether (see
e.g. Ross, 1977). Second, players may hold a simplified representation of the world due
to limited cognitive abilities (see e.g. Samuelson, 2001; Mengel, 2012). Third, bundling
experiences can be optimal when experience with a collection of games is scarce because
the larger amount of available data overcompensates the loss in predictive accuracy (Al-
Najjar and Pai, 2014; Mohlin, 2014). Appendix B presents simulation results for simple
social-learning games which establish this principle for the setup of this paper.

The first result of the subsection shows that adaptation across games converges to
an analogy-based expectations equilibrium (Jehiel, 2005; Jehiel and Ettinger, 2010, ABEE
henceforth). In a general ABEE, each player partitions the decision nodes of other players
into analogy classes and has a correct understanding of average behavior in each class.
In the long-run ABEE considered here, players use a separate analogy class for each
history-state-pair $(ht, \theta)$ and their expectation about the (average) distribution of actions
at history $ht$ if the state is $\theta$ is correct. However, each analogy-class bundles all decision
situations with the same history and state regardless of the game $k$ and players’ private
signals. I therefore refer to the partition as the information-anonymous analogy partition.
The ABEE additionally assumes that players combine their analogy-based expectations
with their private signal using Bayes’ rule and best respond to the resulting beliefs.$^{16}$

---

$^{15}$This is not possible if games differ with respect to preferences; see Section 6.1.

$^{16}$See appendix A for a formal definition and a characterization of the ABEE. The assumption of
Bayesian updating is relaxed in section 5.
Proposition 3. Let \( y(\omega^r) = (K, \{\theta^r\}, \{h_t \in h_{t+1}^r\}) \) for each \( r \geq 1 \) and each \( \omega^r \in \Omega \). For each \( \epsilon > 0 \), each \( k \in K \), and each \( t = 1, \ldots, T \), the limit strategy \( \lim_{r \to \infty} \sigma_{t}^{k,r} \) almost surely plays \( \epsilon \)-like an ABEE strategy \( \sigma_{t}^{k,A} \) at all histories satisfying \( \varphi_t(h_t | \theta) > 0 \) for each \( \theta \in \Theta \).

As Proposition 1, Proposition 3 rests upon the dominance-solvability of the social-learning game. The existence of a dominant strategy in period 1 implies that players eventually correctly assess the average choice probabilities conditional on the state, where the average is taken across games. Accordingly, by best responding to assessments, players eventually play the ABEE strategy in period 2, and in later periods by induction.

Comparing analogy-based social-learning with rational social-learning yields two interesting results. First, while the ABEE strategies coincide with the rationalizable strategies in period 1, players will in general not make correct inferences in a given game in later periods. The ABEE strategies may therefore be suboptimal in a given game. Second, ABEE strategies do not discriminate between games. More precisely, \( \sigma_{t}^{k,A}(s_t, h_t) = \sigma_{t}^{k,A}(s_t, h_t) \) for any two games \( k, \ell \in K \), any period \( t \) and history \( h_t \in H_t \), and any signal \( s_t \in [b_k, b_k] \cap [b_{\ell}, b_{\ell}] \). Among all strategies \( \sigma_t : \cup_{K} [b_{k}, b_{k}] \times H_t \to \Delta(A) \) which do not discriminate between games in this sense, the ABEE strategy need not maximize the \textit{ex ante expected payoff} in the randomly selected social learning game given by \( U_t(\sigma_t, \{\sigma_{t-1}^{k,A}\}_{k \in K}) = \sum_{k \in K} \pi_k U_t^k(\sigma_t, \sigma_{t-1}^{k,A}) \) where \( U_t^k(\sigma_t, \sigma_{t-1}) \) is the expected payoff in game \( k \) given in (1). The following proposition formalizes these two results. Following Smith and Sørensen (2000), I call a property \textit{generic} if it holds for an open and dense subset of parameters.

Proposition 4. Generically,

1. there exists \( k \in K \), \( t > 1 \), \( h_t \in H_t \) and \( \bar{\epsilon} > 0 \) such that the ABEE strategy \( \sigma_{t}^{k,A} \) does not play \( \epsilon \)-like a rationalizable strategy \( \sigma_{t}^{k,*} \) at history \( h_t \) for each \( 0 < \epsilon < \bar{\epsilon} \).

2. there exists \( t > 1 \) and \( \tilde{\sigma}_t : \cup_{k} [b_{k}, b_{k}] \times H_t \to \Delta(A) \) such that

\[
U_t(\tilde{\sigma}_t, \{\sigma_{t-1}^{k,A}\}_{k \in K}) > U_t(\sigma_{t}^{k,A}, \{\sigma_{t-1}^{k,A}\}_{k \in K}) .
\]

Proposition 4 shows that long-run behavior is not optimal when players do not distinguish games while adapting. Accordingly, opportunities to improve upon long-run behavior exist. Section 5 explores one such opportunity by assuming that players use more flexible belief updating rules than Bayes’ rule. The proposition also provides an explanation for unjustified contagion or spill-over between games. If a herd forms after history \( h_t \) in game \( k \), players may conclude that following others at \( h_t \) is also optimal in game \( \ell \neq k \) even if this is not warranted by fundamentals. This problem may be exacerbated when allowing for limited feedback on the state as in Proposition 2.

It is finally instructive to compare the approach to related literature. First, the information-anonymous analogy partition is finer than the payoff-relevant analogy parti-
tion considered by Guarino and Jehiel (2013) since players distinguish behavior at different histories. On the other hand, it is distinct from the private information analogy partition (Jehiel and Koessler, 2008). Second, players might adapt across games that differ in other aspects such as payoffs. In the present game, the simple structure of payoffs, the absence of payoff externalities, and the realization of payoffs at the end of each round facilitate the identification of a player’s own payoff function. An extension of the model where games also differ with respect to the distribution of preferences yields similar results and is investigated in section 6.1.

5 Heterogeneous Belief Updating

Thus far, I have assumed that players form beliefs according to Bayes’ rule. A more general adaptive process allows players to also adjust their responses to beliefs. As in Section 2, I consider flexible belief updating rules where player $t$’s belief in game $k$ given history $h_t$, signal $s_t$, and assessments $\varphi^k_t (h_t \mid \theta)$ for each $\theta \in \Theta$ is given by

$$b^k_t (s_t, h_t) = \left[ 1 + \left( \frac{1-s_t}{s_t} \right)^{\beta_t} \frac{\varphi^k_t (h_t \mid 0)}{\varphi^k_t (h_t \mid 1)} \right].$$ (5)

$\beta_t$ is player $t$’s private information weight. A player is Bayesian, if $\beta_t = 1$, overweights private information if $\beta_t > 1$, and underweights private information if $\beta_t < 1$ (see e.g. Grether, 1980; Hung and Plott, 2001; Palfrey and Wang, 2012). In the following, I assume that players adjust their private information weight alongside their assessments. Concretely, in the spirit of reinforcement learning, players pick the weight which yields a higher expected payoff. I focus on the long-run. Accordingly, the selected long-run private information weight must maximize the ex ante expected payoff given long-run assessments.

The simple example of Section 2 suggests that the long-run belief updating rule depends on the family of social-learning games. I explore this possibility in an extension of the model where the precision of private signals has a player-specific component. Formally, in game $k$ player $i$ draws signals from $\{0, 1\}$ according to probabilities $\Pr(\tilde{s}^k_i = 1 \mid \tilde{\theta} = 1) = \Pr(\tilde{s}^k_i = 0 \mid \tilde{\theta} = 0) = q^k_i$ where

$$q^k_i = \frac{1}{2} + \frac{1}{2} \exp (\bar{\rho}^k_i + \nu_i).$$

$\bar{\rho}^k_i$ determines the average signal precision in game $k$ and is a measure for the ease with which information can be collected in this game. By contrast, $\nu_i$ is a player-specific signal.

17 Guarino and Jehiel (2013) do not consider multiple games.

18 The coarsest common refinement is given by the analogy partition $\tilde{A}_i = \{ \alpha (h_t, \theta, s_i) \}$ for which player $i$ assesses others’ behavior separately for each history, state, and realization of his private signal.

19 See e.g. Steiner and Stewart (2008); Mengel (2012); Grimm and Mengel (2012).
precision component which measures the ease with which player \( i \) collects information.\(^{20}\) Player \( i \) is better informed the larger is \( \nu_i \). Let \( V \) denote the cumulative distribution function of player-specific signal precision components.

I study the long-run outcome of adaptation across games when diversely informed players are allowed to combine public and private information differently. The limit outcome is an extended ABEE given by the mapping \( \beta^A: \mathbb{R} \to \mathbb{R}_+ \) such that (i) player \( i \) with quality component \( \nu_i \) invests (rejects) in period \( t \) of game \( k \) given history \( h_t \in H_t \) and signal \( s_t^k \in [b^k, \bar{b}^k] \) if
\[
\left( \frac{s_t^k}{1 - s_t^k} \right) ^{\beta^A(\nu_i)} > (>) \frac{\tilde{\varphi}_t(h_t | 0)}{\tilde{\varphi}_t(h_t | 1)},
\]
and (ii) assessments \( \tilde{\varphi}_t(h_t | \theta) = \sum_k \pi_k \tilde{\varphi}_t^k(h_t | \theta) \) take into account the distribution of player-specific signal precision components and the associated updating rules via
\[
\tilde{\varphi}_t^k(h_t | \theta) = \prod_{\tau < t} \int \sum_{s_{\tau}} \Pr_k(s_{\tau} | \theta, \nu_{\tau}) \sigma_{\tau}^{k, \beta^A} (a_{\tau} | s_{\tau}, h_{\tau}) \ dV(\nu_{\tau})
\]
where \( \Pr_k(s_{\tau} | \theta, \nu_{\tau}) \) denotes the probability that the player in period \( \tau \) who has signal precision component \( \nu_{\tau} \) observes signal \( s_{\tau} \) in game \( k \) when the state is \( \theta \), and \( \sigma_{\tau}^{k, \beta^A} \) are the strategies determined by (6). Proposition 5 characterizes the equilibrium mapping \( \beta^A \) which maximizes the ex ante expected payoff for each \( \nu \in \text{supp}(V) \).

**Proposition 5.** There exist thresholds \( \nu, \overline{\nu} \in \mathbb{R} \) such that the extended ABEE \( \beta^A \) that maximizes the ex ante expected payoff for each \( \nu \in \text{supp}(V) \) satisfies \( \beta^A(\nu_t) \geq 1 \) for each \( \nu_t > \overline{\nu} \) and \( \beta^A(\nu_t) \leq 1 \) for each \( \nu_t < \nu \).

Proposition 5 states that players with a high (low) signal precision benefit from over-weighting (underweighting) private information. Figure 2 illustrates the result for two equally likely social-learning games \((K = 2, \pi_1 = \pi_2 = \frac{1}{2})\), and player-specific signal precisions, \( q^k \), \( k = 1, 2 \), distributed uniformly around average qualities \( q^k \), where \( q^k \in \{0.6, 0.65, \ldots, 0.9\} \), and \( q^k \in \{q^k - 0.09, q^k - 0.08, \ldots, q^k + 0.09\} \). In the figure, different colors denote equilibrium values \( \beta^A(\nu_t) \) for different pairs \((q^1, q^2)\).\(^{21}\) As one can see, especially players whose signal precision is slightly larger (smaller) than the average signal precision deviate the most from Bayesian updating in equilibrium, since they are most likely to suffer from distorted assessments.

It is noteworthy that exactly what constitutes a high or low precision of private information is context-dependent. This may be illustrated with the help of a simple story:

\(^{20}\) I abuse notation and index players with \( i \) to emphasize that the second component of signal precision is player-specific rather than period-specific.

\(^{21}\) The equilibria have been calculated numerically. The results and the code are available from the author upon request.
University students usually have diverse educational backgrounds. Some students attend specific university-preparatory schools. In such an environment, a student may experience that imitating others is a successful strategy even if she considers herself well informed. This student may therefore develop a tendency to underweight her private information. In contrast, other students come from more rural areas where schools have large catchment areas and a correspondingly diverse student body. A top-of-the-class student from such a school may learn that imitation is rarely beneficial even if his own information is weak. He may therefore come to overweight his private information. Though adaptive, these tendencies to apply non-Bayesian updating rules show up as biases at the university.

The proposition suggests that heterogeneous updating rules will evolve. It therefore provides a theoretical background for corresponding evidence (e.g. El-Gamal and Grether, 1995; Delavande, 2008; Palfrey and Wang, 2012). Moreover, it motivates a straightforward, behavioral extension of the standard model of social-learning in which players are assumed to have private information weights drawn from a given distribution. The extension is straightforward, since the equilibrium of the extended game in which the distribution of private information weights is commonly known is equivalent to the Bayesian equilibrium of a standard game with a distorted distribution of private information (March and Ziegelmeyer, 2009, Lemma 3.2). On the other hand, the new, behavioral model of social-learning better captures the established experimental regularities under the assumption that a sufficiently large proportion of the population overweights private information. In this case, compared to the standard model cascades emerge later (since overweighters herd later), beliefs become more extreme (since each action stems from an overweighter with positive probability), and therefore the length and strength of cascades are correlated (since more and more moderate overweighters enter the herd). Moreover, heterogeneous
updating rules improve economic efficiency. Indeed, if the support of the distribution of private information weights is unbounded, social-learning is complete even with private signals of bounded strength (March and Ziegelmeyer, 2009, Corollary 3.4). The behavioral model also accommodates new findings such as the coexistence of the ‘overweighting-of-private-information’ bias with a ‘social-confirmation’ bias (March and Ziegelmeyer, 2015). Finally, its adaptive foundations distinguish the model from other suggested explanations such as the level-k model or models based on a limited depth of reasoning.

6 Extensions

6.1 Preferences

I have assumed so far that players share the same utility function. More general social-learning games allow for heterogeneous preferences and assume that the distribution of preference types is commonly known (Smith and Sørensen, 2000). Obviously, this common knowledge assumption is equally questionable, and it is likely that players adapt across games which differ in the underlying distribution of preferences. I illustrate that similar results obtain under this assumption with the help of the simple example of Section 2. Accordingly, let Anna and Bob interact repeatedly in games \( k \in \{H, L\} \) which are equally likely. To highlight the role of preferences, I assume that \( q^H_A = q^L_A = q \) for each \( i \in \{A, B\} \).

Assume first that Anna occasionally does not care about the investment payoff and invests or rejects regardless of her private signal. Concretely, each time game \( k \) is played Anna picks her action regardless of her private signal with probability \( \gamma^k_A \) where \( \gamma^k_A < \gamma^H_A \) without loss of generality. I assume that Anna invests or rejects with equal probability in this case. Conversely, in any given repetition of game \( k \) Anna follows her private signal with probability \( 1 - \gamma^k_A \). The assumptions imply that the relative frequency with which Anna invests (rejects) when the investment payoff is 1 (0) approaches \( \hat{q}^k_A = \frac{1}{2} \gamma^k_A + (1 - \gamma^k_A) q_A \). If Bob distinguishes games, he eventually imitates Anna’s decision (follows his private signal) if \( q^H_B < (>) \hat{q}^k_A \). In contrast, if Bob adapts across games he eventually imitates Anna’s decision (follows his private signal) if \( q^H_B < (>) \hat{q}^k_A \) where \( \hat{q}^k_A = \frac{\gamma^H_A + \gamma^L_A}{2} \). This strategy is not optimal for Bob in game \( L \) (\( H \)) if \( \hat{q}^L_A < \hat{q}^H_A \) and Bob benefits from underweighting (overweighting) his private signal.

An even stronger result obtains if Anna is assumed to have opposed preferences in the two games. Assume for instance that Anna’s payoff function satisfies \( u^H_A(1, \theta) = \theta - \frac{1}{2} \), \( u^H_A(1, \theta) = \frac{1}{2} - \theta \), and \( u^H_A(0, \theta) = 0 \) for each \( k \in \{H, L\} \) and each \( \theta \in \Theta \). It follows that Anna invests (rejects) given signal \( s_A = 1 \) and rejects (invests) given signal \( s_A = 0 \) in game \( L \) (\( H \)). Accordingly, the relative frequency with which Anna invests (rejects) across rounds where the investment payoff is 1 (0) approaches \( q_A \) in game \( L \), \( 1 - q_A \) in game \( H \), and \( \frac{1}{2} \) across games. Therefore, Bob eventually follows his private signal in each game when adapting across games. In contrast, it would be optimal for Bob to imitate Anna’s
decision in game $L$ and to anti-imitate Anna’s decision in game $H$ if $q_B < q_A$.\footnote{Bob anti-imitates Anna’s decision if he chooses $a_B \neq a_A$ regardless of his private signal.}

It is straightforward to extend the results to the general social-learning game. Preference heterogeneity is therefore another hindrance to successful adaptation. Indeed, since players cannot use their private signal realizations to distinguish games, adaptation across games which differ with respect to preferences may be unavoidable. Accordingly, preference heterogeneity strengthens the results.

### 6.2 Endogeneous Timing of Decisions

The analysis has focused on settings where the timing of decisions is given exogenously. With an endogeneous timing of decisions, players must foresee future social-learning opportunities in addition to interpreting the observed history of actions (see e.g. Chamley, 2004a). As shown below, the adaptive process need not converge to a Perfect Bayesian equilibrium (PBE) when feedback constraints prevail.

Following Chamley (2004a), I extend the social-learning game as follows: First, each player $i = 1, \ldots, N$ has the option to make one irreversible investment in one of the periods $t = 1, 2, \ldots$. Second, the payoff from investing in period $t$ is given by $\delta^{t-1} \cdot (\theta - c)$ where $0 < \delta < 1$ is the discount factor and the cost of investment satisfies $\underline{b} \leq c < \bar{b}$. In any PBE of this game, there exists a threshold $s_i^*$ such that players with private signals $s > s_i^*$ ($s < s_i^*$) invest in period 1 (delay investment for at least one period). Moreover, there exists at least one PBE with $c < s_i^* < \bar{b}$ if $\underline{b} < c < \bar{b}$, there exists a PBE such that $s_i^* = \bar{b}$ if $c = \bar{b}$, and there may be multiple equilibria in both cases (Chamley, 2004a, Theorem 1).

Assume that the distribution of private information and the equilibrium strategies are not commonly known. Accordingly, players must acquire the information necessary to assess the option value of delay through repeated play of the social-learning game. I focus on the case of a single game. Let $\omega^r_i \equiv (\theta^r, t^r_i, h^r)$ denote the outcome of the social-learning game in repetition $r$ where $h^r = (x^r_1, \ldots, x^r_N)$ is the history of the number of investments $x^r_i$ in each period $t = 1, \ldots, N$ and $t^r_i = \infty$ if the player never invested (the game lasts at most $N$ periods). Let $y_H(\omega^r)$ and $y_\Theta(\omega^r)$ denote the feedback on respectively the history and the state of nature given the outcome $\omega^r$. Below, I characterize the long-run outcome of the adaptive process assuming that (i) $y_H(\omega^r_i) = \{h_i \subset h^r : t < t^r_i\}$, i.e. player $i$ who invests in period $t$ in a given round does not receive feedback about the number of investments in period $t$ or later, and (ii) $y_\Theta(\omega^r_i) = \theta^r$ if $t^r_i < \infty$ and $y_\Theta(\omega^r_i) = \emptyset$ if $t^r_i = \infty$, i.e. players only receive feedback about the realized state when they invest.

**Proposition 6.** Let $y_H(\omega^r_i) = \{h_i \subset h^r : t < t^r_i\}$, $y_\Theta(\omega^r_i) = \theta^r$ if $t^r_i < \infty$, and $y_\Theta(\omega^r_i) = \emptyset$ if $t^r_i = \infty$. In the long-run of the adaptive process almost surely,

1. all players invest in period 1 if $c = \bar{b}$.
2. players with $s_i > c$ invest in period 1 and players with $s_i < c$ never invest if $\underline{b} < c < \bar{b}$.
The proposition entails that players fail to assess the option value of delay. The reason is simple (a formal proof is omitted): A priori, delay has no value and players invest (in period 1) if and only if it is profitable, i.e. if $s_i > c$. Feedback constraints prevent all players from acquiring information since players who invest in period 1 receive no feedback about the number of investments in period 1 and players who never invest are not informed about the state of nature.

7 Conclusion

The paper scrutinizes models of rational social-learning through the lens of adaptation. Adaptation generates rational behavior in the long-run if and only if individuals are able to distinguish social learning settings, receive ample feedback, and their experience with each setting grows without bounds. Limited opportunities for adaptation lead to mistaken inferences from others’ actions and render Bayes’ rule payoff inferior compared to non-Bayesian belief updating rules.

The paper offers some new directions for experimental and theoretical research on social-learning. First, to the best of my knowledge the existing experimental studies have considered laboratory settings in line with standard economic models of social-learning. These experimental settings were the obvious candidates for testing economic models of social-learning, but they have a limited ecological validity. If subjects perceive the laboratory setting as artificial, deviations from rational behavior might not come as a surprise and do not constitute conclusive evidence against rational social-learning. Novel economic experiments should test the rational view of social-learning in settings resembling those encountered by human participants in the field with structural uncertainty and sufficient opportunities to adapt.

Second, the paper shows that a thorough investigation of the assumptions underlying rational social-learning straightforwardly leads to a behavioral model of social-learning with heterogeneous belief updating rules. This behavioral model of social-learning complements recent attempts which have relaxed the assumptions of rational social-learning by arguing for greater psychological realism. Indeed, some of the psychologically more plausible assumptions would disappear when adaptation is considered. More generally, the acknowledgment that individuals face limited opportunities for adaptation is likely to deliver fruitful insights in other economic domains.

---

23 For an experimental study to possess ecological validity, the methods and the setting of the study must approximate the real-life situation that is under investigation. Ecological validity is independent from external validity which relates to the ability of a study’s results to generalize.

24 For example, Hertwig et al. (2004) find that “decisions from experience and decisions from description can lead to dramatically different choice behavior”.

25 For instance individuals which are naïve in the sense of Eyster and Rabin (2010) should be surprised by the clustering of decisions which does not tally with independently distributed private information.
References


