University of
St Andrews

School of Economics and Finance Discussion Papers

Non-Monotone Social Learning

Min Zhang

School of Economics and Finance
Online Discussion Paper Series
issn 2055-303X
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

School of Economics and Finance Discussion Paper No. 2008
7 Dec 2020
JEL Classification: D83
Keywords: anti-imitation, information structure, social learning

Non-Monotone Social Learning

Min Zhang*

December 2020

Abstract

We revisit the canonical binary-state model of social learning to investigate the possibility of non-monotone learning: *ceteris paribus*, with some predecessor(s) switching to actions that reveal greater confidence in one state of the world, agents nevertheless become more confident in the other. A necessary and sufficient condition for non-monotone learning is provided in an illustrative setting where agents are either uninformed or partially informed by binary private signals. In a general setting with continuous private signals, we obtain a sufficient condition for non-monotone learning that does not rely on the shape of the underlying information structures apart from a couple of simple boundary requirements. As a result, a social learning process that completes eventually can be well approximated in a way such that non-monotone learning and anti-imitation are bound to happen.

JEL classification: D83

Keywords: Anti-Imitation; Information Structure; Social Learning

*School of Economics and Finance, University of St Andrews; mz47@st-andrews.ac.uk.

1 Introduction

The theoretical literature on social learning has demonstrated that, when rational agents with common preferences act sequentially, they tend to imitate their predecessors and eventually exhibit herd behaviour. In the canonical example of restaurant choice due to Banerjee (1992), agents sequentially choose to eat at one of the two restaurants, A and B, which are equally likely to be the better one, and each agent obtains a binary private signal indicating the better restaurant before making her choice. Assuming all signals have the same precision and agents follow their private signals when indifferent, when one restaurant has been chosen twice more than other, all future agents will ignore their private signals and choose that restaurant.

Rationalizing herd behaviour and imitation is considered as one key contribution of the literature.[1] Yet we are keen to know what else rational social learning may predict. Do rational agents always tend to imitate their predecessors throughout the learning process? When they imitate, do they imitate each and every predecessor? Answers to such questions rely on a thorough understanding of the learning models beyond the well-established long-run predictions. In addition, more studies into the short-run behavioural implications of rational social learning may serve to distinguish rationality from other potential explanations of social learning, both theoretically and empirically.

Let us start by looking into the learning dynamics in the canonical example above. We can see that agents' incentives to imitate their predecessors result from a "monotone" evolution of posterior beliefs. More specifically, each agent forms posterior beliefs that are monotone with respect to her observation, in the sense that she believes one restaurant, say A, is (weakly) more likely to be better than the other when more of her predecessors have chosen A instead of B. When she observes that A has been chosen at least twice more than B, her posterior belief dominates her private information and she then chooses A regardless of her private signal.

Is such monotonicity a general feature of social learning? It seems plausible at first glance. By choosing a restaurant, each agent reveals more confidence in that restaurant than the other, so it seems natural that, if more agents have chosen A instead of B in the past, the history becomes stronger evidence supporting A. However, this is *not* necessarily the case as the inter-

[1] Quoting the preface of Chamley (2004) which provides a comprehensive overview of the literature, "Learning by individuals from the behaviour of others and imitation pervade the social life... herds, fads, bubbles, crashes, and booms are cited as proofs of the irrationality of individuals. However, most of these colourful events will appear in the models of rational agents...".

pretation of each action in the history closely depends on the underlying information structure of the model. In a variety of settings, rational agents could indeed form non-monotone posterior beliefs with respect to their observations, which we refer as "non-monotone learning": *ceteris paribus*, when some predecessor(s) switch to actions revealing greater confidence in one state of the world, agents become more confident in the other state. As a consequence, agents have less incentive to imitate their predecessors, a feature that is absent in the canonical example and perhaps also overlooked in models extended from it.

For instance, in the restaurant-choice problem, certain signal structures could lead agent 3 to believe B is more likely to be the better restaurant when she observes a history (A,B) than when she observes a history (B,B). In other words, fixing agent 2's action, the more agent 1's action reveals confidence in A, the less confident agent 3 becomes in A.[2] To further illustrate the idea, we consider a setting in which agents are heterogeneously informed: some agents are *experts* who have private signals about whether A or B is better with precision $q \in (\frac{1}{2},1)$, while others are *amateurs* who know nothing at all and the fraction of amateurs in the population is $r \in (0,1)$. A necessary and sufficient condition on the signal structure (q,r) for non-monotone learning is provided that requires either r or q to be sufficiently large. To get the intuition, think again about the comparison between history (A,B) and history (B,B). (A,B) reveals one weak piece of evidence against B from the first action A, as it could just be a random choice by an amateur. Meanwhile (A,B) reveals one strong piece of evidence supporting B from the second action B, as it must be an informative choice by an expert since an amateur, lacking private information, would have followed the first action A. On the other hand, (B,B) reveals two weak pieces of evidence supporting B as each action B could come from an amateur. When r (q) is sufficiently large which implies the weak (strong) evidence is fairly insignificant (significant), (A,B) may turn out to be an overall stronger piece of evidence supporting B than (B,B).

Despite the simplicity of the binary setting, this condition that r or q is sufficiently high provides a nice interpretation of the type of information structures that could lead to non-monotone learning. In particular, if learning is conducted in a society where most people barely have any private knowledge but those who do have some knowledges are sufficiently knowledgeable on average, it is bound to be non-monotone. Such information structures are indeed reasonable

[2] Fixing agent 2's action makes the comparison less trivial, since agent 3's posterior belief is always consistent with agent 2's action, *i.e.*, the *Overturning Principle* in Smith and Sørensen (2000).

in many contexts of social learning such as technology adoption or development of medical treatments, where most of us often know little but those technicians or physicians are usually well recognized for their expertise.[3]

We then turn to a general setting with continuous private signals due to Smith and Sørensen (2000). A nice feature of the continuous setting is that it guarantees behavioural difference whenever learning is non-monotone: if one agent forms posterior beliefs that are non-monotone with respect to the action of one predecessor, she would indeed choose one action with higher probability had that predecessor switched to the other action, *i.e.*, she *anti-imitates* that predecessor.[4] We particularly consider continuous signals over compact support that allow for the possibility of positive tails: signal realizations that generate extreme private beliefs have strictly positive unconditional densities. We show that a couple of simple boundary conditions on the underlying signal structure guarantees non-monotone learning, and it is, unlike the condition in the binary setting, independent of the general shape of the signal structure. In particular, such boundary conditions lead to non-monotone belief updates when posterior beliefs are close to 0 or 1, which is guaranteed as long as learning completes eventually. As a result, any information structure that leads to complete learning can be well approximated by an information structure that leads to non-monotone learning. In other words, a social learning process that completes eventually can be well approximated in a way such that non-monotone learning and anti-imitation are bound to happen at some point. On the other hand, in the absence of explicit parametrization of the underlying information structure, we are still in search for a tight condition for non-monotone learning in the general setting with continuous signals.

1.1 Related Literature

The theoretical literature on social learning, since the first papers by Banerjee (1992), Bikhchandani *et al.* (1992), and Smith and Sørensen (2000), has conventionally focused on information aggregation and herd behaviour in the long run, and fewer papers have deliberately investigated the dynamics of learning and its behavioural implications short of the limit. The information structure in the binary-signal setting of this paper is close to that of Callander and Hörner

[3] Learning and information aggregation are indeed more desirable when individuals have limited private knowledge. For example, Conley and Udry (2010) investigate the diffusion of a new agricultural technology in Ghana, and find evidence that farmers adjust their inputs to align with those who were surprisingly successful in previous periods, which indicates the presence of social learning. However the input choices for another crop of known technology indicate an absence of social learning effects.

[4] We argue by Corollary 2.2.1 that in the binary setting non-monotone learning does not necessarily induce such behavioural difference.

(2009), in which agents can only observe the total number of agents having chosen each option though. They show that uninformed agents (*amateurs*) should follow the minority of their predecessors rather than the majority, when informed agents (*experts*) are rare. We differ from them by maintaining the canonical assumption of full observation and providing a precise condition for non-monotone learning under less restrictive information structures.[5]

As for the general setting with continuous signals, several authors have looked into when belief updates are (non-)monotone, *i.e.*, when posterior belief as a function is (non-)monotone in the prior belief. Suen (2004) first gives a sufficient condition for monotone belief updates, and later work by Herrera and Hörner (2012) characterizes a similar set of conditions on the signal distributions. We establish the necessity (but not sufficiency) of non-monotone belief updates for non-monotone learning, and characterize a sufficient condition for non-monotone learning when belief updates are non-monotone as beliefs converge to 0 or 1 eventually, which precisely violates the conditions in Suen (2004) and Herrera and Hörner (2012) but only at the extreme values.[6]

On behavioural implications of social learning, Eyster and Rabin (2014) look at observational structures instead of information structures and question the rationality of imitative behaviour. They provide a necessary and sufficient condition on the observation structures for rational anti-imitation, and the logic lies in the fact that rational agents need to take into account the redundancy of previous actions under those structures. The full observation assumption in the canonical model actually violates their condition, so our findings on non-monotone learning and anti-imitation based on information structures is in a way orthogonal to theirs.[7]

Most empirical and experimental work on social learning follows the main focus in the theoretical literature and justifies imitative behaviour and herding. For example, Moretti (2011) uses box-office data and finds that the sales of movies with positive and negative opening-weekend surprises in demand diverge over time: a movie that experiences larger sales in Week 1 will experience further increasing sales in subsequent weeks. He considers such imitative behaviour by consumers as a result of social learning and further quantifies the effect of social

[5] Callander and Hörner (2009) assume that experts have perfect private information, *i.e.*, $q = 1$. In fact our necessary and sufficient condition can be satisfied solely by $r \geq \frac{\sqrt{5}-1}{2}$ without any requirement on q. Nevertheless the more restrictive information structure was necessary to reduce the computational complexity in their model due to the different observational assumptions.

[6] Recent works by Morris and Yildiz (2019) and Smith *et al.* (2020) consider conditions for (non-)monotone mapping from prior to posterior beliefs in terms of log-likelihood ratios, which are not necessarily compatible with our assumption of positive tails due to the naturally unbounded support.

[7] Eyster and Rabin (2014) assume continuous action space, hence it is still possible to have anti-imitation in the canonical model with binary actions, even though their precise condition is violated.

learning on movie sales. Cai *et al.* (2009) conduct a field experiment to distinguish imitative social learning from salience, in which they tell diners either the recently popular dishes or the "feature dishes". They find that diners react more strongly to popularity than to salience, suggesting that diners imitate. Our work, on the other hand, will potentially raise the question whether imitative behaviour detected by these papers comes from *rational* learning.[8]

Herding or contrarian behaviour has also been studied in the sequential trading model of financial market introduced by Glosten and Milgrom (1985), and Park and Sabourian (2011), for instance, build on the seminal paper by Avery and Zemsky (1998) and characterize conditions on the underlying information structures that are necessary and sufficient for contrarianism: traders buy (sell) assets after observing histories that reveal bad (good) information about the asset value. However, due to the existence of a competitive market maker who consistently adjusts the bid (ask) prices, these models implicitly impose heterogeneous payoff functions among traders, which makes the underlying mechanism for contrarian behaviour in such models distinct from that for non-monotone learning or anti-imitation in standard social learning models where agents have fully aligned preferences.[9]

The remainder of the paper is structured as follows. Section 2 constitutes the main body of the paper, in which Subsection 2.1 introduces the formal definitions of (non-)monotone learning, Subsection 2.2 discusses the binary setting for illustration, and Subsection 2.3 works on the general setting with continuous signals. Section 3 concludes. All omitted proofs are stacked in the Appendix.

2 A Canonical Model of Social Learning

The model setup is canonical. There is an underlying state of the world, $\theta \in \{A,B\}$, whose realization is unknown to a population of countably infinite agents. Agents hold a common prior of θ, $\Pr(\theta = A) = \Pr(\theta = B) = \frac{1}{2}$. There is an infinite time horizon, $t \in \{1,2,3,\ldots\}$, and at each period t an agent (referred to as agent t hereafter) is chosen to make a once-in-a-life-time binary decision, $a_t \in \{A,B\}$. The history of past actions, $\boldsymbol{h_t} \equiv (a_1,a_2,\ldots,a_t) \in \{A,B\}^t$, is publicly observable

[8] In Cai *et al.* (2009), the popularity of dishes is sorted by the actual number of plates sold in the previous week. So roughly speaking, they are really only looking at how agent 2 reacts to agent 1, and we know that there is no anti-imitation due to the overturning principle.

[9] Similarly, anti-imitation or contrarian behaviour is less surprising in learning models with negative payoff interdependence, such as congestion cost in Eyster *et al.* (2014) or queueing cost in Cripps and Thomas (2019).

for all future agents. Besides, each agent at period t receives a private signal σ_t before making her choice. Conditional on θ, $\{\sigma_t\}_{t=1}^{\infty}$ are independently and identically distributed across t. Agents' preferences are fully aligned, $u(a_t,\theta) = 1_{\{a_t=\theta\}}$, $\forall t \geq 1$.

2.1 Definitions of (Non-)Monotone Learning

Let $p_{t+1} \equiv \Pr(\theta = B|\boldsymbol{h}_t)$ denote the posterior belief of agent $t+1$ after observing history $\boldsymbol{h}_t$ but before acquiring her private signal. p_{t+1} is also referred to as the *public belief* at period $t+1$ since history $\boldsymbol{h}_t$ is publicly observed by all future agents, with $p_1 = \frac{1}{2}$ being the prior. Notice that p_{t+1} is only well defined for equilibrium history $\boldsymbol{h}_t$, hence hereafter we restrict attention to equilibrium histories whenever talking about posterior or public beliefs. [10]

To introduce the (non-)monotone property of learning in question, let us set a simple linear order $\precsim$ on $\{A,B\}$ such that $A \precsim B$ (and of course $B \precsim B$). It then induces a partial order on each Cartesian product $\{A,B\}^t$:

$$\forall t \geq 1,\ \forall \boldsymbol{h}_t, \boldsymbol{h}'_t \in \{A,B\}^t,\ \boldsymbol{h}_t \precsim \boldsymbol{h}'_t \text{ if and only if } a_\tau \precsim a'_\tau \text{ for any } \tau \leq t,$$

and we say $\boldsymbol{h}_t$ is *B-dominated by* $\boldsymbol{h}'_t$ when $\boldsymbol{h}_t \precsim \boldsymbol{h}'_t$.

Definition 1 *Learning is* ***monotone*** *if*

$$\forall t \geq 1,\ \forall \textit{ two equilibrium histories } \boldsymbol{h}_t \textit{ and } \boldsymbol{h}'_t \textit{ such that } \boldsymbol{h}_t \precsim \boldsymbol{h}'_t,\ p_{t+1} \leq p'_{t+1}.$$

Learning is ***non-monotone*** *otherwise, or equivalently,*

$$\exists \textit{ two equilibrium histories } \boldsymbol{h}_t \textit{ and } \boldsymbol{h}'_t \textit{ such that } \boldsymbol{h}_t \precsim \boldsymbol{h}'_t,\ \boldsymbol{h}_t \neq \boldsymbol{h}'_t, \textit{ and } p_{t+1} > p'_{t+1}.$$

The non-monotonicity given by Definition 1 is straightforward: a B-dominated history nevertheless generates greater confidence in state B. We also want to highlight a special case of non-monotone learning, *anti-imitative belief*, which is particularly linked to the anti-imitative behaviour we would like to address later.

[10] As agents are myopic, we do not emphasize a particular notion of equilibrium. Readers can nonetheless assume what we have in mind is the standard *Bayesian Nash Equilibrium* throughout the paper.

Definition 2 *Consider two equilibrium histories* $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$ *that differ only in one action:*

$$\exists \tau \leq t \text{ s.t. } a_\tau = A,\ a'_\tau = B;\ \forall \tau' \leq t \text{ and } \tau' \neq \tau,\ a_{\tau'} = a'_{\tau'}.$$

We say the posterior belief of agent $t+1$ *is* ***anti-imitative*** *of agent* τ *under the pair* $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$ *if* $p_{t+1} > p'_{t+1}$*. We say the posterior belief of agent* $t+1$ *is* ***always anti-imitative*** *of agent* τ *if* $p_{t+1} > p'_{t+1}$ *for every such pair* $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$*.*

Definition 2 for anti-imitative belief is quite intuitive as well: fixing the actions of all the predecessors of agent $t+1$ other than agent τ, agent $t+1$ becomes less confident in state B had agent τ's action switched to reveal more confidence in state B. In other words, it is as if the posterior belief of agent $t+1$ "anti-imitates" agent τ's action.

2.2 A Necessary and Sufficient Condition under Binary Signals

In this subsection, for illustration purpose, we consider the canonical information structure with binary signals as in Bikhchandani *et al.* (1992), with a simple twist of introducing heterogeneity in signal precision, a treatment similar to Callander and Hörner (2009). In particular, agents are of two informational types: either an *expert* who receives a private signal $\sigma_t \in \{A, B\}$ that matches the true state with probability $q \in (\frac{1}{2}, 1)$, or an *amateur* who never receives a private signal. Each agent's type is her private information but the probability of being an amateur, $r \in (0,1)$, is the same across the population and commonly known.

For convenience, in this subsection we work with the posterior log-likelihood ratio, $l_{t+1} \equiv \ln \frac{p_{t+1}}{1-p_{t+1}} = \ln \frac{\Pr(\theta=B|\boldsymbol{h}_t)}{\Pr(\theta=A|\boldsymbol{h}_t)} = \ln \frac{\Pr(\boldsymbol{h}_t|\theta=B)}{\Pr(\boldsymbol{h}_t|\theta=A)}$, instead of the posterior belief p_{t+1} itself. Notice that, since l_{t+1} is a monotone transformation of p_{t+1}, Definition 1 and 2 for *(non-)monotone learning* and *anti-imitative belief*, remain unchanged apart from the notation override from p_{t+1} to l_{t+1}.

To begin the analysis, we first characterize the (equilibrium) dynamics of the public beliefs $\{l_t\}_{t=1}^{\infty}$ by solving the binary Bayesian decision problem of each agent. It is not hard to see that each agent's decision rule is unique up to a tie-breaking rule, and we simply assume that an expert follows her private signal when indifferent, and an amateur chooses randomly between A and B when indifferent.[11]

[11] Tie-breaking assumptions are no longer important for the continuous-signal model in the next subsection.

Lemma 2.2.1 *Let* $M \equiv \ln \frac{r\cdot\frac{1}{2}+(1-r)\cdot q}{r\cdot\frac{1}{2}+(1-r)\cdot(1-q)}$, $L \equiv \ln \frac{r+(1-r)\cdot q}{r+(1-r)\cdot(1-q)}$ *, and* $H \equiv \ln \frac{q}{1-q}$. *Along every equilibrium path* $\boldsymbol{h}_\infty \equiv (a_1, a_2, \ldots, a_t, \ldots)$*, the (public) posterior beliefs evolve in the following way:*

$$l_1 = 0;\ l_2 = \begin{cases} M & \text{if } a_1 = B \\ -M & \text{if } a_1 = A \end{cases};\ \forall t \geq 2,\ l_{t+1} = \begin{cases} l_t + H & \text{if } a_t = B \text{ and } a_{t-1} = A \\ l_t - H & \text{if } a_t = A \text{ and } a_{t-1} = B \\ l_t + L & \text{if } a_t = a_{t-1} = B \text{ and } l_t \leq H \\ l_t - L & \text{if } a_t = a_{t-1} = A \text{ and } l_t \geq -H \\ l_t & \text{otherwise} \end{cases}.$$

Proof. See Appendix. ■

Except for the first action a_1, each action later in the history reveals a strong piece of evidence, a weak piece of evidence, or no evidence at all.[12] If a_{t+1} is different from a_t, it must come from an expert hence it is a strong piece of evidence, which shifts the public belief by $\pm H$. If a_{t+1} is the same as a_t and $|l_t| \leq H$, it is only meaningful if it comes from an expert, which happens with probability $1-r$, hence it is a weak piece of evidence and only shifts the public belief by $\pm L$. If a_{t+1} is the same as a_t but $|l_t| > H$, it reveals no more information as both types will follow the immediate predecessor a_t anyway, hence the public belief remains unchanged thereafter, *i.e.*, an informational cascade.

With the dynamics of posterior beliefs fully characterized by Lemma 2.2.1, we are in the position to provide a necessary and sufficient condition for non-monotone learning in this binary setting. For convenience we present the result using H, M, and L defined in Lemma 2.2.1.

Proposition 2.2.1 *Learning is non-monotone if and only if*

$$H > 2M + L, \text{ or equivalently, } \frac{r^3}{2r(1-r)^2 + (1-r)^3} > 4q(1-q).$$

In particular, the posterior belief of agent 3 *is always anti-imitative of agent* 1 *when* $H > 2M + L$*, and other cases of non-monotone learning require stronger conditions.*

Proof. See Appendix. ■

[12] The first action a_1 comes from either an informative choice by an expert or a random choice by an amateur, and by Bayes rule it turns out to be a 'mediocre' piece of evidence compared to later actions.

Let us briefly go through the logic behind the anti-imitative belief of agent 3, and leave the discussion over other cases of non-monotone learning to the Appendix. We compare only $\boldsymbol{h}_2 = (A,B)$ and $\boldsymbol{h}'_2 = (B,B)$ for agent 3 as the other comparison between (A,A) and (B,A) is symmetric. History (A,B) reveals one mediocre piece of evidence against state B from the first action A but one strong piece of evidence in favour of state B from the second action B, and by Lemma 2.2.1 the posterior belief after (A,B) can be precisely calculated as $l_3 = -M+H$. On the other hand, history (B,B) reveals one mediocre piece and one weak piece of evidence in favour of B from the first and second action B, respectively, and by Lemma 2.2.1 the posterior belief after (B,B) is $l'_3 = M+L$. To make the posterior belief of agent 3 anti-imitative, we need $l_3 > l'_3$ or simply $H > 2M+L$. It is not hard to see that this inequality can be satisfied by some pair (r,q) such that r or q is sufficiently high: either the weak/mediocre piece of evidence is fairly insignificant due to a large proportion of amateurs in the population, or the strong piece of evidence is sufficiently dominant due to a high precision of experts' signals.[13]

We have shown that, in contrast to the canonical model and what people might have learned from the existing literature, learning is indeed non-monotone in this simple setting with heterogeneous informed agents. We do want to point out that Proposition 2.2.1 does not contradict with the existing results on the long-run behaviour of agents, *e.g.*, an informational cascade occurs eventually or a herd arises eventually.[14] In fact, non-monotone learning does not guarantee anti-imitative behaviour, defined as follows, when private signals are discrete.

Definition 3 *Consider two equilibrium histories $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$ that differ only in one action:*

$$\exists \tau \leq t \text{ s.t. } a_\tau = A,\ a'_\tau = B;\ \forall \tau' \leq t \text{ and } \tau' \neq \tau,\ a_{\tau'} = a'_{\tau'}.$$

*We say agent $t+1$ **anti-imitates** agent τ under the pair $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$ if*

$$\Pr(a_{t+1} = B|\boldsymbol{h}_t) > \Pr(a_{t+1} = B|\boldsymbol{h}'_t).$$

[13] Note that the necessary and sufficient condition in Proposition 2.2.1 can be satisfied solely by $r \geq \frac{\sqrt{5}-1}{2}$ without any requirement on q. Hence, in a way, how many experts there are plays a more important role for non-monotone learning than how expert they are.

[14] See Proposition 1 in Bikhchandani *et al.* (1992) or Theorem 3(a) in Smith and Sørensen (2000).

We say agent $t+1$ ***always anti-imitates*** *agent τ if*

$$\Pr(a_{t+1} = B|\boldsymbol{h}_t) > \Pr(a_{t+1} = B|\boldsymbol{h}'_t) \textit{ for } \text{every } \textit{such pair } (\boldsymbol{h}_t, \boldsymbol{h}'_t).$$

This definition of anti-imitative behaviour is in the similar spirit of *anti-imitation* defined in Eyster and Rabin (2014), except that we use a probabilistic criterion here on account of the binary action space.[15] To appreciate the difference between non-monotone learning, particularly anti-imitative belief, and anti-imitative behaviour, let us reconsider agent 3 after history (A,B) or (B,B). Suppose the posterior belief of agent 3 is already anti-imitative, *i.e.*, $l((A,B)) > l((B,B))$ or $H > 2M + L$. If agent 3 is an amateur, she will choose B after both histories. If agent 3 is an expert, she will follow her private signal after both histories since $l((A,B)) = H - M < H$ and $l((B,B)) = M + L < H$. Therefore, although agent 3's posterior belief is anti-imitative, there is *no* probabilistic difference in her behaviour and she does *not* anti-imitate agent 1.

A corollary is given to highlight the difference between anti-imitative belief and anti-imitative behaviour in the current setting with binary private signals.

Corollary 2.2.1 *The posterior belief of each agent $t \geq 3$ is always anti-imitative of agent 1 if and only if $H > 2M + L$ and $H > 2M + (k^* - 2)L$, where $k^* \equiv \min\{k \in \mathbb{N}^+ | M + kL > H\}$. Nevertheless, at most $(k^* - 3)$ agents anti-imitate agent 1.*

Proof. See Appendix. ■

The first part of the corollary builds on Proposition 2.2.1. Suppose the posterior belief of agent 3 is already anti-imitative, *i.e.*, $l((A,B)) > l((B,B))$. By Lemma 2.2.1, extending the two histories by one same action will shift the posterior belief toward the same direction and by the same amount, $\pm H$ or $\pm L$, therefore the inequality still holds. This logic works for further extensions as well, until the posterior belief after one history, say l_t, grows beyond the precision of private signals, *i.e.*, $|l_t| > \ln \frac{q}{1-q}$. Since then l_t no longer changes and l'_t will catch up until $|l'_t| > \ln \frac{q}{1-q}$ as well. The extra condition precisely accounts for this fact and assures that l'_t will not exceed l_t in the end.

For the second part, since amateurs always follow their immediate predecessors, only the behaviour of experts could be different after two histories that differ in the first action. And

[15] See Definition 5 in Eyster and Rabin (2014) for a comparison. They assume continuous action space as in Lee (1993) and hence define anti-imitation as one agent's action being decreasing in some action(s) she observes.

according to Lemma 2.2.1, that happens only when the posterior belief after one history, say l_t, has exceeded the precision of private signals while the posterior belief after the other, say l'_t, has not. In that situation an expert will follow her immediate predecessor after l_t but follow her private signal after l'_t. However, as we discussed earlier, l'_t will catch up along the history and eventually exceeds the threshold as well. Hence such behavioural difference can only exist for a finite number of future agents, and the upper bound of that number is precisely $(k^* - 3)$. It is worth noting that Corollary 2.2.1 is indeed consistent with the long-run behaviour predicted by the existing literature.

In the next subsection we turn to a setting with continuous private signals, in order to better understand the impact of the underlying information structures on the learning process. Besides, as we will see from Lemma 2.3.1, the continuous model gets rid of the subtle difference between non-monotone learning and anti-imitation and thus allows a clear focus.

2.3 A Sufficient Condition under Continuous Signals

We now consider a setting with continuous private signals similar to the seminal paper by Smith and Sørensen (2000). In particular, we assume that the private signal of each agent t is continuous over a compact support, $\sigma_t \in [\underline{c}, \overline{c}]$. The cumulative distribution function of σ_t conditional on state $\theta \in \{A, B\}\}$, $G_\theta(\cdot)$, is assumed differentiable with density function $g_\theta(\cdot)$ that satisfies the monotone likelihood ratio property: $\frac{g_B(\cdot)}{g_A(\cdot)}$ is strictly increasing over $[\underline{c}, \overline{c}]$. This allows us to identify the private signal σ_t by the private belief of agent t after observing σ_t, $q_t \equiv \Pr(\theta = B|\sigma_t) \in [0, 1]$. In fact we work directly with q_t instead of σ_t as what matters for each agent is the information generated by her private signal rather than the realization of private signal itself. Let $F_\theta(\cdot)$ be the cumulative distribution function of q_t conditional on θ. $F_A(\cdot)$ and $F_B(\cdot)$ capture the underlying information structure, and if both are differentiable with density function $f_A(\cdot)$ and $f_B(\cdot)$ respectively, the unconditional density function of q_t, $f(\cdot) \equiv \frac{1}{2}[f_A(\cdot) + f_B(\cdot)]$, is already sufficient. In particular, Bayesian updating implies a simple “no introspection condition” that $\frac{f_B(x)}{f_A(x)} = \frac{x}{1-x}$ and thus $f_A(x) = 2(1-x)f(x)$ and $f_B(x) = 2xf(x)$.[16] Hence $f(\cdot)$ is referred to as the information structure throughout this subsection, and we impose the following assumptions on it:

Assumption 1 (Full support) $f(\cdot)$ *is strictly positive on* $(0, 1)$.

[16] The private belief structure is commonly used in the literature. Curious readers can look at Appendix *A* in Smith and Sørensen (2000) for the justification of it.

Assumption 2 (Smoothness) *$f(\cdot)$ is continuously twice differentiable on $(0,1)$, and both $\lim_{x\to 0+} f''(x)$ and $\lim_{x\to 1-} f''(x)$ exist.*

Assumption 3 (Positive tails) $f(0) = \lim_{x\to 0+} f(x) > 0$ *and* $f(1) = \lim_{x\to 1-} f(x) > 0$.

The assumption of full support implies that private beliefs are unbounded, which guarantees that no history of actions is off equilibrium, as each agent has strictly positive probability to follow her private signal no matter what the history was. We are therefore free to compare any pair of histories when investigating whether learning is monotone or not. The assumption of smoothness is purely technical, and it simply allows us to conduct higher order Taylor expansions at the boundaries to yield the sufficient condition for non-monotone learning later.[17]

The last assumption of positive tails, on the other hand, is critical for our analysis. Notice that the assumption of full support allows the possibility that $f(0) = 0$ and/or $f(1) = 0$, and in fact such possibility is the default when private signals are over an unbounded support, for which integrability requires the density of extremely high or low realizations to tend to 0.[18] However positive tails may well be admissible when private signals are over a compact support, *e.g.*, $g_A(\sigma) = 2 - 2\sigma$ and $g_B(\sigma) = 2\sigma$ over common support $[0,1]$ generate a uniform information structure, $f(x) \equiv 1$.[19] In a way, our analysis picks up this subtle difference due to the nature of private signals, and highlights the existence of non-monotone learning under the possibility of positive tails that could be overlooked when signals over unbounded support are often considered.[20]

We first claim the equivalence between anti-imitative belief and anti-imitative behaviour, under the current setting with continuous signals. It relieves us from any extra concern about the subtle difference between the two that was present in the binary-signal setting, so that any detection of non-monotone learning now has always a behavioural implication. To see this, notice that any public belief difference as result of non-monotone learning or more specifically anti-imitative belief, will lead an agent to exhibit anti-imitative behaviour as long as her private belief falls in to the gap between the public beliefs generated by the two histories in comparison,

[17] Strictly speaking, we only need twice differentiability around the boundaries, not over the entire support, to conduct the analysis.

[18] For example, private signals that are, conditional on state A (or B), normally distributed with a low (or high) mean and the same variance satisfy the assumption of full support but fail the assumption of positive tails.

[19] Since for the decision problems in canonical social learning models it suffices to identify the private signals as the private beliefs, it is to some extent without loss of generality to consider only private signals over a compact support (the unit interval in particular) after all.

[20] We could possibly replicate our analysis with the Taylor expansion technique of an even higher order, while relaxing Assumption 3 to $\lim_{x\to 0+} f^{(k)}(x) > 0$ and $\lim_{x\to 1-} f^{(k)}(x) > 0$ for some $k \geq 1$ but tightening Assumption 2 to $f(\cdot)$ being continuously $(k+1)$ times differentiable. This could make certain signal distributions over unbounded support admissible but it offers merely a technical edge over the current analysis without providing extra insights or intuition for the underlying logic of our main result later.

which happens with a strictly positive probability given the assumption of full support.

Lemma 2.3.1 *Let* $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$ *be a pair of histories such that*

$$\exists \tau \leq t \text{ s.t. } a_\tau = A,\ a'_\tau = B;\ \forall \tau' \leq t \text{ and } \tau' \neq \tau,\ a_{\tau'} = a'_{\tau'}.$$

Under Assumption 1, agent $t+1$ *anti-imitates agent* $\tau \leq t$ *under* $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$ *if and only if the posterior belief of agent* $t+1$ *is anti-imitative of agent* τ *under* $(\boldsymbol{h}_t, \boldsymbol{h}'_t)$. *In other words, Definition 2 and 3 are equivalent.*

Proof. See Appendix. ■

Let us start the analysis by solving the Bayesian decision problem of each agent as well as characterizing the dynamics of the public beliefs, $\{p_t\}_{t=1}^{\infty}$.

Lemma 2.3.2 *Agent* t *who forms public belief* p_t *from history and private belief* q_t *from her signal will choose* $a_t = B$ *if and only if* $q_t \geq 1 - p_t$. *Therefore the public belief process* $\{p_t\}_{t=1}^{\infty}$ *in equilibrium is characterized by the following transition functions:*

$$p_{t+1}(p_t, a_t) = \begin{cases} m(p_t) \equiv \frac{p_t F_B(1-p_t)}{p_t F_B(1-p_t) + (1-p_t) F_A(1-p_t)} & \text{if } a_t = A \\ n(p_t) \equiv \frac{p_t [1 - F_B(1-p_t)]}{p_t [1 - F_B(1-p_t)] + (1-p_t)[1 - F_A(1-p_t)]} & \text{if } a_t = B \end{cases}.$$

Proof. See Appendix. ■

As in the binary-signal setting, rational agents here simply compare the public belief generated from the observed history with the private belief generated from their private signals. And the transitions of public beliefs come from the individual decision rules and Bayesian updating as a result. In particular, the transition function, $m(\cdot)$ or $n(\cdot)$, captures how the public beliefs evolve from one period to the next, when the most recent action is A or B.

Next we present an intuitive but important necessary condition for non-monotone learning.

Lemma 2.3.3 *Learning is monotone if the transition functions* $m(\cdot)$ *and* $n(\cdot)$ *characterized in Lemma 2.3.2 are monotone.*

Proof. See Appendix. ■

Quite intuitively, non-monotone learning requires non-monotone transitions of posterior beliefs. To see it clearly, let us think about two different histories $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ such that $\boldsymbol{h}_t \precsim \boldsymbol{h}'_t$. If $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ differ on the last action, *i.e.*, $a_t = A$ and $a'_t = B$, then clearly $p_{t+1} < \frac{1}{2} < p'_{t+1}$ by the overturning principle. If $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ only differ up to some earlier period $\tau < t$, then $p_{\tau+1} < p'_{\tau+1}$, again by the overturning principle. Since after period τ actions are identical between the two histories, agents should update their beliefs using the same transition function from period $\tau+1$ on along the two histories. If both $m(\cdot)$ and $n(\cdot)$ are monotone, then the inequality $p_{s+1} \leq p'_{s+1}$ is preserved for every $s = \tau, \ldots, t$, and hence non-monotone learning, *i.e.*, $p_{t+1} > p'_{t+1}$, is not possible.

Monotone belief updates have been studied in the literature. It can be easily verified that the monotonicity of the transition function $m(\cdot)$ (and $n(\cdot)$), taking into account the no-introspection condition that $\frac{f_B(x)}{f_A(x)} = \frac{x}{1-x}$, is equivalent to the *increasing failure ratio property* (and the *increasing hazard ratio property*) of the signal distributions given by Herrera and Hörner (2012), and is in turn equivalent to the comparable assumptions imposed in Suen (2004). It is worth pointing out that the sufficient condition for non-monotone learning presented later in Proposition 2.3.1 builds solely upon the violation of monotone belief updates around the two extreme values 0 and 1, so an information structure that leads to non-monotone learning may well satisfy those established properties for monotone belief updates in the interior.

On the other hand, we make the following interesting observation about what learning patterns monotone belief updates may generate in our model.

Claim 2.3.1 *Under Assumptions 1 to 3, the learning process never generates moderate posterior beliefs if the transition functions $m(\cdot)$ and $n(\cdot)$ characterized in Lemma 2.3.2 are monotone. In particular, $p_t \notin (\frac{1}{3}, \frac{2}{3})$, $\forall t > 1$.*

Proof. See Apendix. ■

This result depends critically on the assumption of positive tails.[21] In particular, positive tails imply a 'stronger' overturning principle: $p_{t+1} \geq \frac{2}{3}$ if history $\boldsymbol{h}_t$ ends with an action B, or $p_{t+1} \leq \frac{1}{3}$ if $\boldsymbol{h}_t$ ends with an action A. For instance, consider the case where an action B follows a history that reveals a public belief very close to 0. Such action B then reveals a private belief

[21] Note that positive tails does not preclude monotone belief transitions in the first place so Claim 2.3.1 is not empty. For example, the uniform information structure mentioned in page 13, $f(x) \equiv 1$, admits both positive tails and monotone transition functions $m(\cdot)$ and $n(\cdot)$.

very close to 1, and thanks to the positive tails, the informativeness of such extreme private belief is sufficiently strong to not only overturn the public belief, but also overturn it by a considerable margin, *i.e.*, $p_{t+1} \to \frac{2}{3} > \frac{1}{2}$. Given the monotonicity of the transition function $n(\cdot)$, any history that ends with an action B then generates a posterior belief at least $\frac{2}{3}$. Similarly any history that ends with an action A then generates a posterior belief at most $\frac{1}{3}$. We think such preclusion of moderate posterior beliefs contradicts with many empirical applications of social learning and creates extra motivation for studying the possibility of non-monotone belief updates (and thus non-monotone learning) in the model.[22]

One last observation we want to exploit before proceeding to the main result of this subsection is the famous finding by Smith and Sørensen (2000) that unbounded private beliefs, implied by the assumption of full support, guarantees complete learning. We present it without replicating the proof from the original paper.

Lemma 2.3.4 *Under Assumption 1, $p_t \longrightarrow 1_{\{\theta=B\}}$ almost surely as $t \longrightarrow +\infty$.*

Proof. See Theorem $1(b)$ in Smith and Sørensen (2000). ■

Although non-monotone belief updates are necessary for non-monotone learning according to Lemma 2.3.3, they are not sufficient since the decreasing region(s) of a non-monotone $m(\cdot)$ or $n(\cdot)$ may not be reachable in equilibrium with the belief process $\{p_t\}_{t=1}^{\infty}$ being a jumping process due to the discrete action space. Nevertheless, thanks to Lemma 2.3.4 that ensures the two extreme values 0 and 1 are always reachable in the long run, we are now in a position to provide a sufficient condition for non-monotone learning that imposes only a couple of boundary requirements on the information structure around the extreme values.

Proposition 2.3.1 *Under Assumptions 1 to 3, learning is almost surely non-monotone if the information structure $f(\cdot)$ is such that $\lim_{x\to 0+} f'(x) > 3f(0)$ and $\lim_{x\to 1-} f'(x) < -3f(1)$.*

Proof. See Appendix. ■

Despite the tedious algebra behind, one key interim step is to show that, when $f(\cdot)$ diminishes sufficiently fast around 0 and 1 (while keeping positive tails), the transition functions turn

[22] For instance, in medical research physicians are allowed to offer patients randomization to different treatments only if *clinical equipoise* exists, *i.e.*, there is genuine uncertainty in the expert medical community over whether a treatment will be beneficial. In other words, the community of physicians should regard the treatments as (roughly) equally preferable, which implies a fairly moderate posterior belief. See Freedman (1987).

out non-monotone around 0 or 1. In particular, $n(\cdot)$ is decreasing over $(0,\varepsilon)$ for some small ε and $m(\cdot)$ is decreasing over $(1-\varepsilon,1)$. Let us now compare the posterior beliefs after any two histories, $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$, such that $\boldsymbol{h}_t \precsim \boldsymbol{h}'_t$ and $\boldsymbol{h}_t \neq \boldsymbol{h}'_t$. Generically the beliefs are different as well, $p_{t+1} \neq p'_{t+1}$.[23] Complete learning implies that both p_{t+1} and p'_{t+1} will fall in $(0,\varepsilon)$ or $(1-\varepsilon,1)$ when t is sufficiently large. Suppose the true state is A and hence the former is the case. If $p_{t+1} > p'_{t+1}$, learning is non-monotone by definition. If $p_{t+1} < p'_{t+1}$, we can simply extend $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ by an action B, *i.e.*, $\boldsymbol{h}_{t+1} = (\boldsymbol{h}_t, B)$ and $\boldsymbol{h}'_{t+1} = (\boldsymbol{h}'_t, B)$. But now $p_{t+2} = n(p_{t+1}) > n(p'_{t+1}) = p'_{t+2}$, so learning is non-monotone. Notice that when $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ only differ in one action, we effectively get anti-imitation by either agent $t+1$ or agent $t+2$.

The boundary conditions in Proposition 2.3.1 can be easily satisfied by either simple density functions over $[0,1]$ or common distributions (over unbounded support) truncated to $[0,1]$:[24]

1. *Linear Density*: $f(x) = (4-8s)\left|x-\frac{1}{2}\right| + 2s$, $\forall s \in (\frac{5}{7},1)$;

2. *Quadratic Density*: $f(x) = s(x-\frac{1}{2})^2 + (1-\frac{s}{12})$, $\forall s \in (-6,-2)$;

3. *(Truncated) Double Exponential Distribution*: $f(x) = \frac{\exp(-\frac{|x-\frac{1}{2}|}{s})}{2s(1-\exp(-\frac{1}{2s}))}$, $\forall s \in (0,\frac{1}{3})$;

4. *(Truncated) Normal Distribution*: $f(x) = \frac{s}{2\int_0^{s/2} e^{-t^2} dt} \cdot \exp[-s^2(x-\frac{1}{2})^2]$, $\forall s > 2\sqrt{3}$;

5. *(Truncated) Cauchy Distribution*: $f(x) = \frac{s}{2\arctan(\frac{s}{2})} \cdot \frac{1}{1+s^2(x-\frac{1}{2})^2}$, $\forall s > 2\sqrt{3}$.

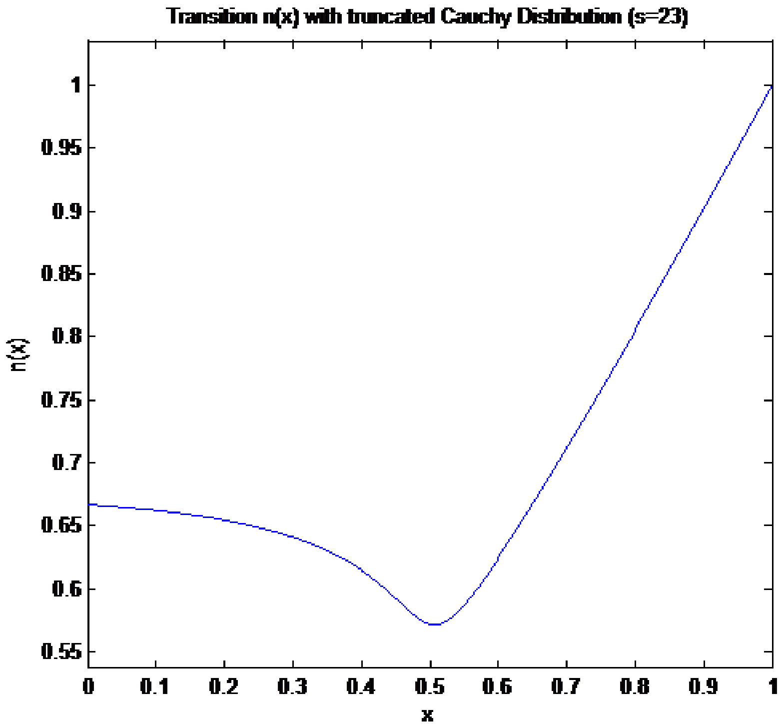
Notice that, although Proposition 2.3.1 in principle only induces non-monotone belief updates around the extreme values, it does not necessarily mean that non-monotone learning and anti-imitation can only happen in the long run. In Figure 1 we plot the transition function $n(\cdot)$ under the example of (truncated) Cauchy distribution with a sufficiently high parameter $s = 23$, which indeed has a very large decreasing region over $(0,\frac{1}{2})$. In this case learning is non-monotone as long as we are comparing two histories with the same last two actions that are alternating, *e.g.*, (A,A,B) versus (B,A,B), (A,A,A,B) versus (A,B,A,B), etc.

In fact, Proposition 2.3.1 does not require any particular shape of the underlying information structure $f(\cdot)$ at the interior of its support. One nice implication is that we can construct information structures satisfying Proposition 2.3.1 that is sufficiently close to any information

[23] For a generic density function $g(\cdot)$, we can always construct two different histories long enough that generate different posterior beliefs.

[24] For convenience we impose $f(x) = f(1-x)$ on all the examples so that the two boundary conditions in Proposition 2.3.1 are satisfied simultaneously. As a result some examples may violate differentiability at $x = \frac{1}{2}$, which does not affect the boundary conditions at all.

Figure 1: An example of large decreasing region of the transition function(s)



structure of interest, which is summarized as the corollary below. In other words, a social learning process that completes eventually can be well approximated in a way such that non-monotone learning and anti-imitation are bound to happen at some point.

Corollary 2.3.1 *For any information structure $f(\cdot)$ that satisfies Assumptions 1 to 3, there exists a sequence of information structures $\{f_k(\cdot)\}_{k=1}^{\infty}$ such that*

1. for any $x \in [0,1]$, $\lim_{k\to\infty} f_k(x) = f(x)$;[25]

2. for any k sufficiently large, learning is almost surely non-monotone under $f_k(\cdot)$.

Proof. See Appendix. ■

3 Conclusion

In this paper we reconsider the standard social-learning model where agents act myopically and share common preferences. We study how underlying information structures affect the

[25] $\lim_{k\to\infty} f_k(x) = f(x)$ refers to point-wise convergence of the probability density functions (and thus the cumulative distribution functions), so readers can interpret it as convergence in distribution, *i.e.*, weak convergence, of the random variables (the private beliefs) behind.

evolution of posterior beliefs as well as the behaviour of rational agents. We show that learning is not always monotone: rational agents often form posterior beliefs that are non-monotone with respect to the actions they observe. As a result, alongside the long-run herd behaviour that has been well established by the existing literature, it is rational for agents to anti-imitate some of the predecessors during the learning process.

We first look into a simple setting with binary private signals and two informational types of agents, *i.e.*, agents are either uninformed or informed by a private signal with certain precision. Learning is non-monotone when the probability of being uninformed is sufficiently large or the private signals are sufficiently precise. In particular, under such information structures, the third agent as well as each agent after her always form posterior beliefs that are non-monotone with respect to the action of the first agent. Consequently, some of them anti-imitate the first agent: *ceteris paribus*, they are more likely to choose one action when the first agent has switched to the opposite, even in the absence of any strategic concern or preference heterogeneity.

Next we consider a canonical setting with continuous private signals, particularly when the signals have compact support which allows the possibility of positive tails. We establish a link between (non-)monotone learning and (non-)monotone belief updates that have been studied in the literature. A sufficient condition for non-monotone learning is obtained, which does not depend on the general shape of the underlying information structure apart from a couple of boundary requirements. This suggests that the information structures behind many social learning processes of interest may indeed be very close to those that induce non-monotone learning and anti-imitative behaviour by rational agents. We are still in search for a tight condition for non-monotone learning, and due to the lack of parametrization in the continuous setting, it remains technically challenging and further work out of this paper is most welcome.

We treat this paper as an interesting contribution to the literature of social learning, where the majority of existing work has been devoted to studying the efficiency of information aggregation and the long-run behaviour of agents. Our work hopefully points out that, during the learning process, rational agents may indeed act quite differently from what we might expect based on the asymptotic outcomes we already know. It is clearly important as well to understand the behavioural implications of rational social learning short of the limit, a lot of which are still puzzling or at least cannot be simply deduced from the existing results.

References

AVERY, C., AND P. ZEMSKY (1998): "Multidimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, 88(4), 724–748.

BANERJEE, A. V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107(3), 797–817.

BIKHCHANDANI, S., D. HIRSHLEIFER, AND I. WELCH (1992): "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades," *Journal of Political Economy*, 100(5), 992–1026.

CAI, H., Y. CHEN, AND H. FANG (2009): "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, 99(3), 864–882.

CALLANDER, S., AND J. HÖRNER (2009): "The Wisdom of the Minority," *Journal of Economic Theory*, 144(4), 1421–1439.

CHAMLEY, C. P. (2004): *Rational Herds: Economic Models of Social Learning*, Cambridge University Press.

CONLEY, T. G., AND C. R. UDRY (2010): "Learning about a New Technology: Pineapple in Ghana," *American Economic Review*, 100(1), 35–69.

CRIPPS, M., AND C. THOMAS (2019): "Strategic Experimentation in Queues," *Theoretical Economics*, 14(2), 647-708.

EYSTER, E., AND M. RABIN (2014): "Extensive Imitation is Irrational and Harmful," *Quarterly Journal of Economics*, 129(4), 1861–1898.

EYSTER, E., A. GALEOTTI, N. KARTIK, AND M. RABIN (2014): "Congested Observational Learning," *Games and Economic Behavior*, 87, 519–538.

FREEDMAN, B. (1987): "Equipoise and The Ethics of Clinical Research," *The New England Journal of Medicine*, 317(3):141–145.

GLOSTEN, L., AND P. MILGROM (1985): "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders," *Journal of Financial Economics*, 14(1), 71–100.

LEE, I. H. (1993): "On the Convergence of Informational Cascades," *Journal of Economic Theory*, 61(2), 395–411.

HERRERA, H., AND J. HÖRNER (2012): "A Necessary and Sufficient Condition for Informational Cascades," working paper.

MORETTI, E. (2011): "Social Learning and Peer Effects in Consumption: Evidence from Movie Sales," *Review of Economics Studies*, 78(1), 356–393.

MORRIS, S., AND M. YILDIZ (2019): "Crises: Equilibrium Shifts and Large Shocks," *American Economic Review*, 109(8), 2823–2854.

PARK, A., AND H. SABOURIAN (2011): "Herding and Contrarian Behavior in Financial Markets," *Econometrica*, 79(4), 973–1026.

SUEN, W. (2004): "The Self-Perpetuation of Biased Beliefs, " *Economic Journal*, 114(495), 377–396.

SMITH, L., AND P. SØRENSEN (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68(2), 371–398.

SMITH, L., P. SØRENSEN, AND J. TIAN (2020): "Informational Herding, Optimal Experimentation, and Contrarianism," *Review of Economic Studies*, forthcoming.

Appendix

Proof of Lemma 2.2.1. The first action a_1 is either a random choice by the tie-breaking assumption if agent 1 is an amateur or the same as σ_1 if she is an expert. Hence by Bayes rule,

$$l_2 = \ln \frac{\Pr(a_1|\theta = B)}{\Pr(a_1|\theta = A)} = \begin{cases} \ln \frac{r\cdot\frac{1}{2}+(1-r)\cdot q}{r\cdot\frac{1}{2}+(1-r)\cdot(1-q)} = M & \text{if } a_1 = B \\ \ln \frac{r\cdot\frac{1}{2}+(1-r)\cdot(1-q)}{r\cdot\frac{1}{2}+(1-r)\cdot q} = -M & \text{if } a_1 = A \end{cases}.$$

For all $t \geq 2$, $a_{t+1} \neq a_t$ implies agent $t+1$ must be an expert with $\sigma_{t+1} = a_{t+1}$. Hence by Bayes rule,

$$\begin{aligned} l_{t+1} &= \ln \frac{\Pr(\boldsymbol{h}_t = (a_1, a_2, \ldots, a_t, a_{t+1} \neq a_t)|\theta = B)}{\Pr(\boldsymbol{h}_t = (a_1, a_2, \ldots, a_t, a_{t+1} \neq a_t)|\theta = A)} \\ &= l_t + \ln \frac{\Pr(a_{t+1}|\theta = B)}{\Pr(a_{t+1}|\theta = A)} = \begin{cases} l_t + \ln \frac{q}{1-q} = l_t + H & \text{if } a_{t+1} = B \\ l_t + \ln \frac{1-q}{q} = l_t - H & \text{if } a_{t+1} = A \end{cases}. \end{aligned}$$

If $a_{t+1} = a_t$ and $|l_t| \leq \ln \frac{q}{1-q} = H$, we know that agent $t+1$ is either an amateur or an expert with $\sigma_{t+1} = a_{t+1}$. Hence by Bayes rule,

$$\begin{aligned} l_{t+1} &= \ln \frac{\Pr(\boldsymbol{h}_t = (a_1, a_2, \ldots, a_t, a_{t+1} = a_t)|\theta = B)}{\Pr(\boldsymbol{h}_t = (a_1, a_2, \ldots, a_t, a_{t+1} = a_t)|\theta = A)} \\ &= l_t + \ln \frac{\Pr(a_{t+1} = a_t|\theta = B, |l_t| \leq H)}{\Pr(a_{t+1} = a_t|\theta = A, |l_t| \leq H)} = \begin{cases} l_t + \ln \frac{r+(1-r)q}{r+(1-r)(1-q)} = l_t + L & \text{if } a_{t+1} = B \\ l_t + \ln \frac{r+(1-r)(1-q)}{r+(1-r)q} = l_t - L & \text{if } a_{t+1} = A \end{cases}. \end{aligned}$$

If $a_{t+1} = a_t$ but $|l_t| > \ln \frac{q}{1-q} = H$, agent $t+1$ always follows agent t regardless of her type and her private signal, hence the public belief remains unchanged. ■

Proof of Proposition 2.2.1. We prove the proposition in several steps.

Step 1: *The posterior belief of agent* 3 *is anti-imitative of agent* 1 *if and only if* $H > 2M + L$.

To see this, let us consider agent 3's posterior belief after observing $\boldsymbol{h}_2 = (A, B)$ or $\boldsymbol{h}'_2 = (B, B)$:

$$l_3 = -M + H \text{ and } l'_3 = M + L \Longrightarrow l_3 > l'_3 \text{ iff } H > 2M + L.$$

Due to the symmetric information structure of the private signals, the other comparison between

$\widetilde{\boldsymbol{h}}_2 = (A,A)$ and $\widetilde{\boldsymbol{h}}_2' = (B,A)$ will yield the same inequality: [26]

$$\widetilde{l}_3 = -M - L \text{ and } \widetilde{l}_3' = M - H \Longrightarrow \widetilde{l}_3 > \widetilde{l}_3' \text{ iff } H > 2M + L.$$

Hence the posterior belief of agent 3 is (always) anti-imitative of agent 1 if and only if $H > 2M + L$.

Step 2: *The posterior belief of agent* $t > 3$ *is anti-imitative of agent* 1 *only if* $H > 2M + L$.

Let us start with agent 4. By symmetry, we only focus on two comparisons: $\boldsymbol{h}_3 = (A,B,B)$ with $\boldsymbol{h}_3' = (B,B,B)$, $\widetilde{\boldsymbol{h}}_3 = (A,B,A)$ with $\widetilde{\boldsymbol{h}}_3' = (B,B,A)$.

$$\begin{aligned} & l_4 = l_3 + L,\ l_4' = l_3' + L \\ \Longrightarrow\quad & l_4 > l_4' \text{ iff } l_3 > l_3' \text{ iff } H > 2M + L; \\ & \widetilde{l}_4 = l_3 - H,\ \widetilde{l}_4' = l_3' - H \\ \Longrightarrow\quad & \widetilde{l}_4 > \widetilde{l}_4' \text{ iff } l_3 > l_3' \text{ iff } H > 2M + L. \end{aligned}$$

Hence $H > 2M + L$ is necessary (and sufficient) for agent 4's posterior belief to be anti-imitative of agent 1. In fact we can see that, as long as each new action added to the histories updates the posterior beliefs in exactly the same way, we do not need extra conditions for anti-imitative beliefs of future agents. However the updates will differ when one posterior belief has absolute value bigger than H but not the other, and we do need extra conditions to account for that.

Consider the earliest such instance: $\boldsymbol{h}_4 = (A,B,B,B)$ with $\boldsymbol{h}_4' = (B,B,B,B)$. The posterior belief after $\boldsymbol{h}_4$ is $l_5 = -M + H + 2L = H + (2L - M) > H$ since it is easy to verify that $L < M < 2L$. Therefore the only equilibrium history extended from $\boldsymbol{h}_4$ is $\boldsymbol{h}_t = (A,B,B,B,\dots,B)$ for $t \geq 4$ and the only valid equilibrium history extended from $\boldsymbol{h}_4'$ to compare is $\boldsymbol{h}_t' = (B,B,B,B,\dots,B)$ for $t \geq 4$. By Lemma 2.2.1 it is not difficult to calculate the corresponding l_{t+1} and l_{t+1}':

$$l_{t+1} = l_5 = H + (2L - M),\ \forall t \geq 4;$$
$$l_{t+1}' = \begin{cases} M + (t-1)L & \text{if } 4 \leq t \leq k^* + 1 \\ M + k^* L & \text{if } t > k^* + 1 \end{cases},$$

[26] We will use the symmetric structure of the private signals to simplify certain proofs later as well, if possible.

where $k^* \equiv \min\{k \in \mathbb{N}^+ | M + kL > H\}$.[27] Clearly, $l_{t+1} > l'_{t+1}$ for every $t \geq 4$ iff

$$H + (2L - M) > M + k^* L \Longleftrightarrow H > 2M + (k^* - 2)L.$$

Note that $H > 2M + L$ and $M > L$ implies that $k^* \geq 3$, so this extra condition is not redundant.

Moreover, the same extra condition will be yielded if we compare two histories where the actions herd on A rather than B eventually. For example, let us consider $\widetilde{\boldsymbol{h}}_4 = (A,A,A,A)$ and $\widetilde{\boldsymbol{h}}'_4 = (B,A,A,A)$. By symmetry $\widetilde{l}'_5 = -l_5 < -H$, hence the only equilibrium history extended from $\widetilde{\boldsymbol{h}}'_4$ is $\widetilde{\boldsymbol{h}}'_t = (B,A,A,A,\ldots,A)$ for $t \geq 4$ and the only valid equilibrium history extended from $\widetilde{\boldsymbol{h}}_4$ to compare is $\widetilde{\boldsymbol{h}}_t = (A,A,A,A,\ldots,A)$ for $t \geq 4$. By symmetry again,

$$\widetilde{l}'_{t+1} = -l_{t+1} = -H - (2L - M),\ \forall t \geq 4;$$
$$\widetilde{l}_{t+1} = -l'_{t+1} = \begin{cases} -M - (t-1)L & \text{if } 4 \leq t \leq k^* + 1 \\ -M - k^* L & \text{if } t > k^* + 1 \end{cases},$$

where k^* is the same as before. Clearly, $\widetilde{l}_{t+1} > \widetilde{l}'_{t+1}$ for every $t \geq 4$ iff

$$-M - k^* L > -H - (2L - M) \Longleftrightarrow H > 2M + (k^* - 2)L.$$

For a general pair of equilibrium histories that differ only in the first action, agents might flip between A and B multiple times before a herd starts eventually. However, we do not need other extra conditions to take care of them because, as argued earlier, all those action update the posterior beliefs in exactly the same way. Therefore, the only condition required here, in addition to $H > 2M + L$, is $H > 2M + (k^* - 2)L$.

Step 3: *$l_{t+1} \leq l'_{t+1}$ whenever $\boldsymbol{h}_t \precsim \boldsymbol{h}'_t$ and $a_1 = a'_1$.* [28]

Take two histories $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ such that $\boldsymbol{h}_t \precsim \boldsymbol{h}'_t$ and $a_1 = a'_1$. Denote $\tau^o \geq 2$ as the first period where $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ differ and $\tau^1 < t$ as the last period where they differ. Clearly $l_{\tau^o} = l'_{\tau^o}$ as the two histories are identical up to period $\tau^o - 1$, and without loss of generality let us assume that $l_{\tau^o} = l'_{\tau^o} = y \geq 0$.

[27] It takes $(k^* - 3)$ periods for l'_{t+1} to grow until beyond H.

[28] In the current setting with binary signals, learning cannot be non-monotone if two histories starting with the same action are in comparison. This is perhaps not a desired consequence of the model setup, however the majority of the analysis can be reproduced as one can think of the comparison between two histories whose first few actions coincide as having a potentially biased prior $l_1 \neq 0$ and then $a_1 \neq a'_1$. Such modelling limitation also disappears as we move to the later setting with continuous signals where initial actions are no longer critical.

Now we take a look at the following part of the two histories, from period τ^o to period τ^1+1:

$$(a_{\tau^o},\dots,a_{\tau^1},a_{\tau^1+1}) = (A,a_{\tau^o+1},\dots,a_{\tau^1-1},A,a_{\tau^1+1}),(a'_{\tau^o},\dots,a'_{\tau^1},a'_{\tau^1+1}) = (B,a'_{\tau^o+1},\dots,a'_{\tau^1-1},B,a'_{\tau^1+1}).$$

Note that a_{τ^1+1} and a'_{τ^1+1} are the same and there are strictly less B in $(a_{\tau^o},\dots,a_{\tau^1})$ than in $(a'_{\tau^o},\dots,a'_{\tau^1})$. Denote p as the number of A's in $(a_{\tau^o},\dots,a_{\tau^1})$ and q as the number of B's in $(a_{\tau^o},\dots,a_{\tau^1})$, as well as p' and q' accordingly in $(a'_{\tau^o},\dots,a'_{\tau^1})$. Clearly $p>p'$ and $q<q'$.

Suppose $a_{\tau^1+1}=a'_{\tau^1+1}=B$. By simple calculation we can get

$$l_{\tau^1+2} = y+(q-p+1)L < y+(q'-p'+1)L = l'_{\tau^1+2}.$$

Suppose $a_{\tau^1+1}=a'_{\tau^1+1}=A$. By similar calculation we get

$$l_{\tau^1+2} = y+(q-p)L-H < y+(q'-p')L-H = l'_{\tau^1+2}.$$

If $t>\tau^1+1$, we continue to have $l_{t+1}<l'_{t+1}$ as the actions after period τ^1+1 are identical between the two histories.

Note that this result implies that only anti-imitative beliefs with respect to agent 1 are possible, so step 1 and 2 have covered all the possible cases of anti-imitative beliefs. We cover the remaining cases of non-monotone learning in the next step.

Step 4: *To get $l_{t+1}>l'_{t+1}$ for $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ such that $\boldsymbol{h}_t \precsim \boldsymbol{h}'_t$ and they differ in at least 2 actions, a strictly stronger condition than $H>2M+L$ is required.*

By step 3 we know that, to get $l_{t+1}>l'_{t+1}$, a_1 and a'_1 must be different, *i.e.*, $a_1=A$ and $a'_1=B$. Denote $\tau^2>1$ as the last period where h_t and h'_t differ. Denote p as the number of A's in $(a_2,\dots,a_{\tau^2})$ and q as the number of B's in $(a_2,\dots,a_{\tau^2})$, as well as p' and q' accordingly in $(a'_2,\dots,a'_{\tau^2})$. Clearly $p>p'$ and $q<q'$.

Without loss of generality let us assume $a_{\tau^2+1}=a'_{\tau^2+1}=B$. By simple calculation we have

$$\begin{aligned} & l_{\tau^2+2} = -M+H+(q-p)L;\ l'_{\tau^2+2} = M+(q'-p'+1)L \\ \Longrightarrow\quad & l_{\tau^2+2} > l'_{\tau^2+2} \text{ if and only if } H > 2M+[2(p-p')+1]L \geq 2M+3L. \end{aligned}$$

The inequality above is clearly stronger than $H>2M+L$. If $t>\tau^2+1$, we continue to have

$l_{t+1} > l'_{t+1}$ provided that a herd has not started in either history, or we need *extra* conditions to take care of possible herds like we did in step 2. Hence we always need a strictly stronger condition than $H > 2M + L$.

Finally we rewrite the inequality $H > 2M + L$ in terms of r and q:

$$\begin{aligned} H > 2M+L \quad &\Longleftrightarrow \quad \ln\frac{q}{1-q} > 2\ln\frac{r\cdot\frac{1}{2}+(1-r)\cdot q}{r\cdot\frac{1}{2}+(1-r)\cdot(1-q)} + \ln\frac{r+(1-r)\cdot q}{r+(1-r)\cdot(1-q)} \\ &\Longleftrightarrow \quad \frac{q}{1-q} > [\frac{r\cdot\frac{1}{2}+(1-r)\cdot q}{r\cdot\frac{1}{2}+(1-r)\cdot(1-q)}]^2 \cdot \frac{r+(1-r)\cdot q}{r+(1-r)\cdot(1-q)} \\ &\Longleftrightarrow \quad \frac{r^3}{2r(1-r)^2+(1-r)^3} > 4q(1-q). \end{aligned}$$

Note that this condition can indeed be satisfied by some admissible pair (r,q), for instance,

$$\forall r \in [\frac{\sqrt{5}-1}{2}, 1),\ \forall q \in (\frac{1}{2}, 1),\ \frac{r^3}{2r(1-r)^2+(1-r)^3} \geq 1 > 4q(1-q).$$

■

Proof of Corollary 2.2.1. The first part was already proved in step 2 of the proof of Proposition 2.2.1. For the second part, clearly agent 2 will not anti-imitate agent 1 so let us consider two equilibrium histories with length at least 2 that differ only in the first action. Without loss of generality let us assume the last action of both histories is B, *i.e.*, $\boldsymbol{h}_t = (A, \ldots, B), \boldsymbol{h}'_t = (B, \ldots, B)$, $t \geq 2$.

If agent $t+1$ is an amateur, she always follows her immediate predecessor and chooses B, so an amateur cannot anti-imitate agent 1. If agent $t+1$ is an expert, the only way to induce anti-imitative behaviour of her is that $l_{t+1} > H$ and $l'_{t+1} \leq H$, in which case agent $t+1$ always chooses B after $\boldsymbol{h}_t$ but still follows her private signal σ_{t+1} after l'_{t+1}. Then unconditional on her informational type, the probability of agent $t+1$ choosing B is,

$$\begin{aligned} &\Pr(a_{t+1} = B|\boldsymbol{h}_t) = r\cdot 1 + (1-r)\cdot 1 = 1; \\ &\Pr(a_{t+1} = B|\boldsymbol{h}'_t) = r\cdot 1 + (1-r)\cdot \Pr(\sigma_{t+1} = B|\boldsymbol{h}'_t) = \frac{1+r}{2} < 1. \end{aligned}$$

However, l_{t+1} never changes after period $t+1$ while l'_{t+1} will gradually increase until it exceeds H according to Lemma 2.2.1, and there is no more behavioural difference for either type from then on. The number of periods it takes for l'_{t+1} to increase until beyond H has been calculated in the proof of Proposition 2.2.1, which is at most $k^* - 3$. ■

Proof of Lemma 2.3.1. Note that each agent t chooses $a_t = B$ if and only if $q_t \geq 1 - p_t$. This is because, due to the conditional independence between $\boldsymbol{h}_{t-1}$ and σ_t,

$$\frac{\Pr(\theta = B|\boldsymbol{h}_{t-1}, \sigma_t)}{\Pr(\theta = A|\boldsymbol{h}_{t-1}, \sigma_t)} = \frac{\Pr(\theta = B|\boldsymbol{h}_{t-1})\Pr(\theta = B|\sigma_t)}{\Pr(\theta = A|\boldsymbol{h}_{t-1})\Pr(\theta = A|\sigma_t)} = \frac{p_t q_t}{(1-p_t)(1-q_t)}$$

$$\Longrightarrow \quad \Pr(\theta = B|\boldsymbol{h}_{t-1}, \sigma_t) \geq 0.5 \text{ iff } \frac{p_t q_t}{(1-p_t)(1-q_t)} \geq 1 \text{ iff } q_t \geq 1 - p_t.$$

Therefore Definition 2 and 3 are clearly equivalent:

$$\begin{aligned}
& \Pr(a_{t+1} = B|\boldsymbol{h}_t) - \Pr(a_{t+1} = B|\boldsymbol{h}'_t) > 0 \\
\Longleftrightarrow \quad & \Pr(q_{t+1} \geq 1 - p_{t+1}) - \Pr(q_{t+1} \geq 1 - p'_{t+1}) > 0 \\
\Longleftrightarrow \quad & \Pr(q_{t+1} \in [1 - p_{t+1}, 1 - p'_{t+1})) > 0 \\
\Longleftrightarrow \quad & p_{t+1} > p'_{t+1}.
\end{aligned}$$

■

Proof of Lemma 2.3.2. The transition functions can be derived according to the decision rule verified in the proof of Lemma 2.3.1. When $a_t = A$,

$$\begin{aligned}
p_{t+1} &\equiv \Pr(\theta = B|\boldsymbol{h}_t) = \Pr(\theta = B|\boldsymbol{h}_{t-1}, a_t = A) \\
&= \frac{\Pr(a_t = A|\theta = B, \boldsymbol{h}_{t-1})\Pr(\theta = B|\boldsymbol{h}_{t-1})}{\Pr(a_t = A|\theta = B, \boldsymbol{h}_{t-1})\Pr(\theta = B|\boldsymbol{h}_{t-1}) + \Pr(a_t = A|\theta = A, \boldsymbol{h}_{t-1})\Pr(\theta = A|\boldsymbol{h}_{t-1})} \\
&= \frac{p_t F_B(1-p_t)}{p_t F_B(1-p_t) + (1-p_t)F_A(1-p_t)}.
\end{aligned}$$

When $a_t = B$,

$$\begin{aligned}
p_{t+1} &\equiv \Pr(\theta = B|\boldsymbol{h}_t) = \Pr(\theta = B|\boldsymbol{h}_{t-1}, a_t = B) \\
&= \frac{\Pr(a_t = B|\theta = B, \boldsymbol{h}_{t-1})\Pr(\theta = B|\boldsymbol{h}_{t-1})}{\Pr(a_t = B|\theta = B, \boldsymbol{h}_{t-1})\Pr(\theta = B|\boldsymbol{h}_{t-1}) + \Pr(a_t = B|\theta = A, \boldsymbol{h}_{t-1})\Pr(\theta = A|\boldsymbol{h}_{t-1})} \\
&= \frac{p_t[1 - F_B(1-p_t)]}{p_t[1 - F_B(1-p_t)] + (1-p_t)[1 - F_A(1-p_t)]}.
\end{aligned}$$

■

Proof of Lemma 2.3.3. Suppose that $m(\cdot)$ and $n(\cdot)$ are both monotone, *i.e.*,

$$\forall 0 \leq x \leq x' \leq 1,\ m(x') \geq m(x) \text{ and } n(x') \geq n(x).$$

Let $\mathscr{F}$ be the set of all finite-order compositions of $m(\cdot)$ and $n(\cdot)$, *i.e.*,

$$\mathscr{F} = \cup_{i\geq 1}\mathscr{F}_i, \text{ with } \mathscr{F}_1 = \{m,n\} \text{ and } \mathscr{F}_{i+1} = \{m\circ f, n\circ f | \forall f \in \mathscr{F}_i\} text.$$

Obviously all functions in $\mathscr{F}$ are monotone: $\forall f \in \mathscr{F}$, $\forall 0 \leq x \leq x' \leq 1$, $f(x') \geq f(x)$.

Now take any two histories $\boldsymbol{h}_t$ and $\boldsymbol{h}_t'$ such that $\boldsymbol{h}_t \precsim \boldsymbol{h}_t'$ and $\boldsymbol{h}_t \neq \boldsymbol{h}_t'$. Let $\overline{\tau} \equiv \max\{\tau \leq t | a_\tau = A$ and $a'_\tau = B\}$. If $\overline{\tau} = t$, $p_{t+1} < \frac{1}{2} < p'_{t+1}$ according to the overturning principle. If $\overline{\tau} < t$,

$$p_{\overline{\tau}+1} < \frac{1}{2} < p'_{\overline{\tau}+1} \text{ by again the overturning principle;}$$
$$a_\tau = a'_\tau,\ \forall \overline{\tau} < \tau \leq t \Longrightarrow \exists f \in \mathscr{F} \text{ s.t. } p_{t+1} = f(p_{\overline{\tau}+1}),\ p'_{t+1} = f(p'_{\overline{\tau}+1}).$$

Since f is monotone, $p_{t+1} = f(p_{\overline{\tau}+1}) \leq f(p'_{\overline{\tau}+1}) = p'_{t+1}$. Therefore learning is always monotone. ∎

Proof of Claim 2.3.1. We first show that $\lim_{x\to 0+} n(x) = \frac{2}{3}$. According to the characterization of the transition function $n(\cdot)$ in Lemma 2.3.2,

$$\lim_{x\to 0+} \frac{n(x)}{1-n(x)} = \lim_{x\to 0+} \frac{x[1-F_B(1-x)]}{(1-x)[1-F_A(1-x)]}.$$

Given that $f_\theta(\cdot)$ and thus $F_\theta(\cdot)$ are continuously twice differentiable, Taylor expansion yields

$$1-F_A(1-x) = 1-F_A(1)+f_A(1)x - \frac{\lim_{x\to 0+} f'_A(1-x)}{2}x^2 + O(x^3) = f(1)x^2 + O(x^3);$$
$$1-F_B(1-x) = 1-F_B(1)+f_B(1)x + O(x^2) = 2f(1)x + O(x^2).$$

Hence $\lim_{x\to 0+}\frac{n(x)}{1-n(x)} = \lim_{x\to 0+}\frac{2f(1)x^2+O(x^3)}{f(1)x^2+O(x^3)} = 2 \Longrightarrow \lim_{x\to 0+} n(x) = \frac{2}{3}$. Similar derivation yields $\lim_{x\to 1-} m(x) = \frac{1}{3}$.

On the other hand, according to the characterization of $n(\cdot)$ and $m(\cdot)$ in Lemma 2.3.2, it is obvious that $\lim_{x\to 0+} m(x) = 0$ and $\lim_{x\to 1-} n(x) = 1$. Therefore, if $n(\cdot)$ and $m(\cdot)$ are monotone, $n(x) \in [\frac{2}{3}, 1]$ and $m(x) \in [0, \frac{1}{3}]$ for any $x \in [0,1]$. Notice that any posterior belief p_t must be within the image of $n(\cdot)$ or $m(\cdot)$, so $p_t \notin (\frac{1}{3}, \frac{2}{3})$, $\forall t > 1$. ∎

Proof of Proposition 2.3.1. We first show that the boundary conditions stated in the proposition guarantee particularly decreasing transitions of the public beliefs at the boundaries, *i.e.*,

$$\lim_{x\to 0+} f'(x) > 3f(0) \implies \lim_{x\to 1-} \frac{dm(x)}{dx} < 0;$$
$$\lim_{x\to 1-} f'(x) < -3f(1) \implies \lim_{x\to 0+} \frac{dn(x)}{dx} < 0;$$

Notice that $sgn[\frac{dm(x)}{dx}] = sgn[\frac{d(\frac{m(x)}{1-m(x)})}{dx}]$, and $\frac{d(\frac{m(x)}{1-m(x)})}{dx}$ can be derived according to Lemma 2.3.2:

$$\begin{aligned}
\lim_{x\to 1-} \frac{d(\frac{m(x)}{1-m(x)})}{dx} &= \lim_{x\to 1-} \frac{xF_B(1-x)}{(1-x)F_A(1-x)}[\frac{1}{x(1-x)} - \frac{f_B(1-x)}{F_B(1-x)} + \frac{f_A(1-x)}{F_A(1-x)}] \\
&= \lim_{z\to 0+} \frac{F_A(z)F_B(z) + z(1-z)[f_A(z)F_B(z) - F_A(z)f_B(z)]}{z^2F_A^2(z)} \text{ (with } z \equiv 1-x) \\
&= \lim_{z\to 0+} \frac{F_A(z)F_B(z) + 2z(1-z)f(z)[(1-z)F_B(z) - zF_A(z)]}{z^2F_A^2(z)}.
\end{aligned}$$

To simplify the expressions later, let $R \equiv f(0) > 0$ and $W \equiv \lim_{x\to 0+} f'(x)$. Given the assumption of twice differentiability of $f_\theta(\cdot)$, Taylor expansion yields that

$$\begin{aligned}
f(z) &= f(0) + \lim_{z\to 0+} f'(z)z + O(z^2) = R + Wz + O(z^2); \\
F_A(z) &= 0 + f_A(0)z + \frac{\lim_{z\to 0+} f_A'(z)}{2}z^2 + O(z^3) = 2Rz + (W-R)z^2 + O(z^3); \\
F_B(z) &= 0 + f_B(0)z + \frac{\lim_{z\to 0+} f_B'(z)}{2}z^2 + \frac{\lim_{z\to 0+} f_B''(z)}{6}z^3 + O(z^4) = Rz^2 + \frac{2}{3}Wz^3 + O(z^4).
\end{aligned}$$

Substitute these into the expression for $\frac{d(\frac{m(x)}{1-m(x)})}{dx}$,

$$\begin{aligned}
\lim_{x\to 1-} \frac{d(\frac{m(x)}{1-m(x)})}{dx} &= \lim_{z\to 0+} \frac{F_A(z)F_B(z) + 2z(1-z)f(z)[(1-z)F_B(z) - zF_A(z)]}{z^2F_A^2(z)} \\
&= \lim_{z\to 0+} \frac{(R^2 - \frac{1}{3}RW)z^4 + O(z^5)}{4R^2z^4 + O(z^5)} \\
&= \frac{3R-W}{12R} < 0 \text{ since } \frac{W}{R} > 3 \text{ and } R > 0.
\end{aligned}$$

Therefore $\lim_{x\to 1-} \frac{dm(x)}{dx} < 0$ when $\lim_{x\to 0+} f'(x) > 3f(0)$, and by similar algebra it can be verified that $\lim_{x\to 0+} \frac{dn(x)}{dx} < 0$ when $\lim_{x\to 1-} f'(x) < -3f(1)$.

Moreover, by continuous differentiability of the transition functions $m(\cdot)$ and $n(\cdot)$,

$$\exists \varepsilon > 0 \text{ s.t. } \forall x \in (0,\varepsilon),\ \frac{dn(x)}{dx} < 0;\ \forall x \in (1-\varepsilon, 1),\ \frac{dm(x)}{dx} < 0.$$

Suppose $\theta = A$. Consider two histories, $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$, such that

$$\text{i. } \boldsymbol{h}_t \precsim \boldsymbol{h}'_t \text{ and } \boldsymbol{h}_t \neq \boldsymbol{h}'_t; \text{ ii. } p_{t+1} \neq p'_{t+1} \text{ and } \{p_{t+1}, p'_{t+1}\} \subset (0, \varepsilon).$$

Such histories exist almost surely for t sufficient large, thanks to Lemma 2.3.4. If $p_{t+1} > p'_{t+1}$, learning is non-monotone by definition. If $p_{t+1} < p'_{t+1}$, we can extend $\boldsymbol{h}_t$ and $\boldsymbol{h}'_t$ by an action B, *i.e.*, $\boldsymbol{h}_{t+1} = (\boldsymbol{h}_t, B)$ and $\boldsymbol{h}'_{t+1} = (\boldsymbol{h}'_t, B)$. Now we have $p_{t+2} = n(p_{t+1}) > n(p'_{t+1}) = p'_{t+2}$, so learning is again non-monotone. The case for $\theta = B$ can be similarly proved by using the decreasing transition $m(\cdot)$ and extending the pair of histories in comparison by an action A if necessary. ■

Proof of Corollary 2.3.1. For instance, construct

$$f_k(x) \equiv f(x) - \frac{1}{k} + \frac{1}{2k} b_k(x) + \frac{1}{2k} b_k(1-x),$$

where $k \geq \max\{\frac{1}{\inf_{x\in(0,1)} f(x)}, 1\}$ and $b_k(x) \equiv \frac{x(1-x)^k}{B(2,1+k)}$ is the probability density function of a Beta distribution. Clearly, $\lim_{k\to\infty} f_k(x) = f(x)$ for any $x \in [0,1]$, and $f_k(\cdot)$ satisfies Assumption 1 to 3 since $f(\cdot)$ does. Then according to Proposition 2.3.1, it suffices to verify that, for k sufficiently large, $\lim_{x\to 0+} \frac{df_k(x)}{dx} > 3f_k(0)$ and $\lim_{x\to 1-} \frac{df_k(x)}{dx} < -3f_k(1)$. Note that, by Stirling's approximation of the Beta function, $\lim_{x\to 0+} \frac{db_k(x)}{dx} = O(k^2)$ and $\lim_{x\to 1-} \frac{db_k(x)}{dx} = -O(k^2)$ as $k \to \infty$. Hence, for k sufficiently large,

$$\begin{aligned}
\lim_{x\to 0+} \frac{d[b_k(x) + b_k(1-x)]}{dx} &> 2k[3f(0) - \lim_{x\to 0+} f'(x)] - 6; \\
\lim_{x\to 1-} \frac{d[b_k(x) + b_k(1-x)]}{dx} &< -2k[3f(1) + \lim_{x\to 1-} f'(x)] + 6.
\end{aligned}$$

Therefore we have

$$\begin{aligned}
\lim_{x\to 0+} \frac{df_k(x)}{dx} &= \lim_{x\to 0+} f'(x) + \frac{1}{2k} \lim_{x\to 0+} \frac{d[b_k(x) + b_k(1-x)]}{dx} \\
&> \lim_{x\to 0+} f'(x) + 3f(0) - \lim_{x\to 0+} f'(x) - \frac{3}{k} = 3f_k(0); \\
\lim_{x\to 1-} \frac{df_k(x)}{dx} &= \lim_{x\to 1-} f'(x) + \frac{1}{2k} \lim_{x\to 1-} \frac{d[b_k(x) + b_k(1-x)]}{dx} \\
&< \lim_{x\to 1-} f'(x) - 3f(1) - \lim_{x\to 1-} f'(x) + \frac{3}{k} = -3f_k(1).
\end{aligned}$$

■