A NECESSARY AND SUFFICIENT CONDITION FOR INFORMATION CASCADES

HELIOS HERRERA AND JOHANNES HÖRNER

Abstract. We show that the IHRP (Increasing Hazard Ratio Property) and the IFRP (Increasing Failure Ration Property), known statistical properties of signal distributions, are necessary and sufficient for the absence of information cascades (in finite time) in observational learning.
Keywords: Observational Learning.
JEL codes: D82, D83

1. Introduction

Since the first simple models of informational herding, e.g. Banerjee [1] and Bikhchandani, Hirshleifer and Welch [2] (BHW, henceforth), introduced the notion of information cascades the question of whether these cascades are a general phenomenon, robust to a general information and signal structure, emerged. Information cascades occur if some *finite* history of observed actions leads to a situation in which observational learning completely stops, i.e. from some point on no information can be inferred from observed actions any longer. This happens when a situation occurs in which all agents from some point on want to take the same action regardless of their private information, so nothing at all about their private information is conveyed through the actions they take. While cascades appear to always arise after a finite history of actions in models with discrete (and bounded) signals, e.g. as in the BHW model (see also Chamley [3]), Smith and Sørensen [6] have pointed out that cascades do not seem to occur at least for several examples of continuum distributions of signals. In fact, they show that from some point on there has to be action convergence (herd), but that information is transmitted from these observed actions so learning does not stop: even if the same action is observed from some point on, this

Date: First Version: December 25th 2010, This Version: July 19, 2012.
Herrera: SIPA, Columbia University. Email: hhh2108@columbia.edu. Hörner: Yale University. Email: johannes.horner@yale.edu. This paper supersedes the Cowles Foundation discussion paper No. 1738, "Biased Social Learning ".

action is never a foregone conclusion. In sum, the question of how pervasive or robust cascades are remains yet unanswered.[1]

We characterize when and why cascades may happen in a standard herding model. We show that the presence or absence of cascades is not linked to the continuity or discreteness of the signal structure: cascades can indeed occur after a finite history of actions for continuum signal distributions too. The presence or absence of cascades is linked to properties of the signal distribution already used in the statistical literature, but new to the literature on social learning. These properties, the Increasing Hazard Ratio Property (IHRP henceforth, related to upper cascades) and the Increasing Failure Ratio Property (IFRP, related to lower cascades), are necessary and sufficient for the absence of information cascades. Namely, provided the decision of the first individual depends on his signal, the decision of all later individuals will do as well or, to use the language we adopt in this paper, these properties ensure that the posterior public belief necessarily stays in the learning region provided that the prior belief is in it to begin with. We also show that discrete bounded distributions always violate the IHRP and IFRP properties and hence cascades must indeed happen with the discrete information structure.

We first introduce a basic observational learning setup and the notion of no-cascade region. Then, before giving the general result and its intuition, we show simple illustrative examples of signal distributions for which cascades can or cannot occur.

2. Observational Learning

2.1. **Basic Setup.** Consider a basic herding model with the exact timing and observation structure as the BHW model. Agents arrive sequentially and observe the exact sequence of actions taken by all their predecessors in addition to their private signal. There are two relevant states of the world $\theta \in \{0, 1\}$. Each agent faces a binary choice $a \in \{0, 1\}$ and, without loss of generality, wants to take the action a that is more likely to match the state θ, given all the information available.[2] The payoff of an agent is given by

[1] It should be emphasized that the absence of cascade does not imply that social learning occurs. Whether convergence takes place to limit with information cascade, or whether it convergences asymptotically without cascade, social learning fails in both cases.

[2] Adjusting for the case in which the benefits from actions are unequal is straightforward.

$$u(a=i) := \begin{cases} 1 & \text{if } \theta = i, \\ 0 & \text{otherwise.} \end{cases}$$

Action 1 will be also interpreted as an "investment," and the agent taking it as an "investor." Agents have private information about the state in the form of a signal x extracted out of a compact set $X \subseteq \mathcal{R}$.

To fix ideas, assume without loss of generality that X is either $[0,1]$, $\mathcal{R}_+$ or $\mathcal{R}$, and let $\underline{x} := \inf\{x : x \in X\}$, $\bar{x} := \sup\{x : x \in X\}$. Conditional on state θ, the distribution (c.d.f) is denoted F_θ, and assumed twice differentiable on the interior of X, with density f_θ that is bounded away from zero. (See below for the case of discrete signals, which requires minor adjustments.)

Private signals provide valuable information about the state to the agents. The distributions are assumed throughout to satisfy the strict monotone likelihood ratio property (MLRP, henceforth). That is, defining the likelihood ratio of the signal

$$l(x) := \frac{f_1(x)}{f_0(x)},$$

we assume that $l(x)$ is strictly increasing on all the domain X. We let $l(\underline{x}) := \lim_{x \to \underline{x}} l(x) \in \bar{\mathcal{R}}$, $l(\bar{x}) := \lim_{x \to \bar{x}} l(x) \in \bar{\mathcal{R}}$. This guarantees that higher values of the signal lead to higher posterior probabilities that the state is good, for all priors. See Milgrom [5]. Indeed, if p is the public (prior) belief and $L = \frac{p}{1-p}$ is its likelihood ratio, then by Bayes' rule the posterior belief $p(x) = \Pr(\theta|x)$ for an agent with signal x solves

$$\frac{p(x)}{1-p(x)} = l(x) L.$$

Hence, the posterior $p(x)$ is strictly increasing in x. Since agents want to take the action that is more likely to match the state, then agents will take action 0 if $l(x) L < 1$ and action 1 if $l(x) L > 1$. For definiteness, we assume that agents choose randomly if $l(x) L = 1$, though this is an innocuous convention when the distribution of signals is atomless.

2.2. **Cascade Region and Signal Boundedness.** For any signal likelihood ratio $l(x)$, we define the boundaries of the cascade region as

$$\underline{L} := \frac{1}{l(\bar{x})}, \quad \overline{L} := \frac{1}{l(\underline{x})},$$

where $\underline{L}, \overline{L} \in \bar{\mathcal{R}}$.[3] That is, the public belief is in a no-cascade region if the likelihood ratio L is in the interval $\left(\underline{L}, \overline{L}\right)$ and in a cascade region in $[0, \underline{L}) \cup (\overline{L}, +\infty]$.

If the (likelihood ratio of) the public belief L is in the cascade region, then no matter what the private information is, that is, no matter how high or low the private signal $x \in X$ is, actions are a "forgone conclusion" and from that moment on nothing is learned by observing the history of actions: if $L < \underline{L}$ (resp. $L > \bar{L}$) it is optimal to choose action 0 (resp. 1) independently of x –hence the expression cascade.

In terms of actual beliefs, the cascade region boundaries are

$$\underline{p} := \frac{\underline{L}}{1+\underline{L}}, \text{ and } \overline{p} := \frac{\overline{L}}{1+\overline{L}}.$$

We refer to signals as *bounded* if the boundaries $\left(\underline{p}, \overline{p}\right) \subseteq (0,1)$ and *unbounded* if $\underline{p} = 0$ and $\overline{p} = 1$. We shall restrict attention to signals being either bounded or unbounded.

When signals are unbounded, cascades cannot occur, as for any public belief $p \in (0,1)$, signals that are extreme enough to lead the agent to take either action have positive probability.

When signals are bounded, cascades can potentially happen even if we start from a belief inside the cascade region $p \in \left(\underline{p}, \overline{p}\right)$.

2.3. **When can cascades happen for bounded signals?** For any given initial belief $L \in \left(\underline{L}, \overline{L}\right)$, after observing an action $a = 1$, an outside observer deduces that the agent taking that action had a signal above the threshold signal x that solves

$$l(x) L = 1,$$

if such a threshold exists. Hence, by Bayes' rule, the public belief L increases to

$$L^+ := \frac{1 - F_1(x)}{1 - F_0(x)} L.$$

By the same token, after observing an action $a = 0$ the belief L drops to

$$L^- := \frac{F_1(x)}{F_0(x)} L.$$

A cascade can occur if and only if there exists some prior belief $L \in \left(\underline{L}, \overline{L}\right)$ such that $L^+ > \overline{L}$ or $L^- < \underline{L}$.

[3]If $l(\bar{x}) = +\infty$, set $\underline{L} = 0$, and similarly, if $l(\underline{x}) = 0$, set $\bar{L} = +\infty$.

3. Examples of Cascades

3.1. **Discrete Bounded Signals.** The example of BHW shows that, with binary signals of finite precision, after observing two consecutive actions of the same type, the learning process stops. Namely, the public belief becomes so strong in favor of one of the states that no private signal can make the following agent take the contrarian action. Hence, an information cascade occurs.

It is easy to see that the same logic goes through with a discrete number of signals of finite precision. Namely, an informational cascade eventually occurs after almost every history in a model with a finite number of signals with finite precision (see, for instance, Chamley [3]). We will come back to discrete signals in Section 4.5.

As mentioned, if there are unbounded signals, e.g. discrete signals with arbitrarily high precision, cascades cannot happen.

3.2. **Continuum Bounded Signals.** If instead we consider a model with a continuum of signals (satisfying MLRP), cascades may or may not occur, as the following examples show.

3.2.1. *Symmetric Linear Model: No Cascades.* A simple example is

$$f_1(x;\alpha) = 2\alpha x + (1-\alpha), \qquad f_0(x;\alpha) = 2\alpha(1-x) + (1-\alpha), \qquad x \in [0,1],$$

where $\alpha \in [0,1]$ is a measure of the informativeness of the signal. Indeed, the support of the signal likelihood ratio

$$l(x;\alpha) \in \left[\frac{1-\alpha}{1+\alpha}, \frac{1+\alpha}{1-\alpha}\right] = \left[\underline{L}, \overline{L}\right] \subseteq [0,+\infty)$$

increases in α. In particular,

$$\underline{p} := \frac{1-\alpha}{2}, \qquad \overline{p} := \frac{1+\alpha}{2},$$

so for $\alpha = 0$ any signal is totally uninformative, whereas for $\alpha = 1$ signals are unbounded: the extreme, most informative, signals $x = 0$ and $x = 1$ perfectly reveal the state.

The threshold x solves

$$l(x;\alpha)L = 1 \quad \Longrightarrow \quad x = \frac{1+\alpha - L(1-\alpha)}{2\alpha(1+L)}.$$

Hence, in order not to have cascades for any

$$L \in \left(\underline{L}, \overline{L}\right),$$

it must be that

$$\begin{aligned} L^{-} &= \frac{\alpha x^2 + (1-\alpha)x}{(1+\alpha)x - \alpha x^2} L = \frac{(1-\alpha)(1+L)+2}{(1+\alpha)(1/L+1)+2} > \underline{L}, \\ L^{+} &= \frac{1-(\alpha x^2 + (1-\alpha)x)}{1-((1+\alpha)x - \alpha x^2)} L = \frac{(1+\alpha)(1+L)+2}{(1-\alpha)(1/L+1)+2} < \overline{L}. \end{aligned}$$

Because L^+ and L^- are strictly increasing in L, the minimum and maximum are reached at $\underline{L}$ and $\overline{L}$. But, since we have the following fixed points

$$L^{+}\left(\overline{L}\right) = \overline{L}, \quad L^{-}\left(\underline{L}\right) = \underline{L},$$

then the inequalities above must hold for any interior $L \in \left(\underline{L}, \overline{L}\right)$: cascades cannot occur.

3.2.2. *Cascades with a Continuum of Signals.* Consider now instead:

$$f_0(x) = \frac{3}{2}\left(2(1-x) - (1-x)^2\right), \qquad f_1(x) = 1, \quad x \in [0,1],$$

which satisfies MLRP. For such a signal distribution, the cascade boundaries are

$$\underline{L} = \frac{1}{l(1)} = 0, \quad \overline{L} = \frac{1}{l(0)} = \frac{3}{2}.$$

Note that the distribution functions are

$$F_0(x) = \frac{x(3-x^2)}{2}, \qquad F_1(x) = x.$$

Let us start from an interior belief, i.e. a belief which falls strictly inside the no-cascade region, say

$$L = \frac{4}{3},$$

so that an agent with signal $x = 1/3$ is indifferent between both actions. If the action $a = 1$ is observed, the public belief jumps to L^+, which is in the upper cascade region, as

$$L^{+} = \frac{1-F_1(1/3)}{1-F_0(1/3)} L = \frac{1-\frac{1}{3}}{1-\frac{13}{27}} \frac{4}{3} = \frac{12}{7} > \frac{3}{2} = \overline{L}.$$

Hence, all learning then stops, and the public belief remains "stuck" forever at the value L^+. Hence, cascades may occur with continuous distributions (but bounded signals).

We now give a necessary and sufficient condition for information cascades to occur. Given that cascades cannot occur with unbounded signals, we assume throughout that signals are bounded.

4. General Results

4.1. **Increasing Hazard Ratio Property (IHRP).** Define the *hazard ratio* at the signal x as the ratio of the *hazard functions*[4] $\frac{f_\theta}{1-F_\theta}$ conditional on the good and the bad state, that is,

$$H(x) := \frac{1-F_0(x)}{1-F_1(x)} l(x).$$

The (strict) *increasing hazard ratio property* (IHRP) holds if this mapping is strictly increasing.

This property was first introduced in the statistical literature by Kalashnikov and Rachev [4]. When ranking the life x of two different populations θ, say $\theta = 0$ and $\theta = 1$, IHRP is referred to as the "ageing faster property" which probably originates from the fact that the hazard function $f_\theta / (1 - F_\theta)$ is the instantaneous probability of death for an agent in population θ. Hence, IHRP means that the probability of death for an agent in population $\theta = 1$ increases faster as x increases than the probability of death for an agent in population $\theta = 0$.

4.2. **Updating Monotonicity.** The relevance of IHRP for an observational learning model comes from the fact that the posterior likelihood ratio L^+ after an $a = 1$ action (good news, henceforth) is given by

$$L^+ = \frac{1-F_1(x)}{1-F_0(x)} L,$$

where x is the cut-off signal and L is the prior likelihood ratio. Since the cut-off signal solves $L = 1/l(x)$, then the posterior likelihood ratio is equal to

$$L^+ = 1/H(x).$$

To put it differently, IHRP states that this posterior likelihood ratio L^+ is decreasing in the cut-off signal, or alternatively, since the cut-off signal x is decreasing in the prior likelihood ratio L (under MLRP), that the posterior likelihood ratio L^+ is increasing in the prior likelihood ratio L.

To be a little more formal, let $L^+ : \mathcal{R} \to \mathcal{R}$ denote this function mapping the prior likelihood ratio L into the posterior likelihood ratio after good news L^+:

$$L^+(L) = \frac{1}{H(l^{-1}(1/L))}.$$

[4]The hazard function is often also called *hazard rate*: we avoid the term hazard rate to avoid confusion with the term *hazard ratio*.

and let us say that the signals satisfy *updating monotonicity after good news* (UMG) if L^+ is strictly increasing. Given our discussion, the following is immediate.

Proposition 1. *Under MLRP, IHRP is necessary and sufficient for UMG.*

A corresponding property can be defined for the event in which an agent takes the action $a = 0$ (bad news). Define the *failure ratio* K by

$$K(x) = \frac{F_0(x)}{F_1(x)} l(x),$$

for all $x \in X$. The (strict) *increasing failure ratio property* (IFRP) holds if this mapping is strictly increasing. Observe that the posterior likelihood ratio, conditional on such an event, is given by the function $L^- : \mathcal{R} \to \mathcal{R}$ defined as

$$L^-(L) = \frac{1}{K(l^{-1}(1/L))}.$$

It follows from IFRP and MLRP that this posterior likelihood ratio after bad news is an increasing function of the prior likelihood ratio L. Let us say that the signals satisfy *updating monotonicity after bad news* (UMB) if L^- is strictly increasing. Then it immediately follows that, under MLRP, IFRP is necessary and sufficient for UMB.

4.3. **Updating Monotonicity and Cascades.** As it turns out, these concepts characterize the existence of cascades. Take without loss of generality $X = [0, 1]$. If the public likelihood ratio L lies in $\left(\underline{L}, \overline{L}\right)$, as defined previously, an agent's strategy will depend on his private signal: he will take action $a = 1$ if and only if his signal is above some cut-off in $(0, 1)$. (Recall that we are focusing on the case of bounded signals, so that $\left(\underline{L}, \overline{L}\right)$ is a proper subset of $(0, +\infty)$, that is $\left(\underline{p}, \overline{p}\right)$ is a proper subset of $(0, 1)$.) On the other hand, if this ratio ever exits the interval $\left[\underline{L}, \overline{L}\right]$, a cascade starts and observational learning stops: all agents take the same action independently of their signal.

Observe now that, by definition of $\overline{L}$, and of the mapping L^+, $\overline{L}$ is a fixed-point of this mapping, as $\frac{1-F_1(0)}{1-F_0(0)}\overline{L} = \overline{L}$. Similarly, $\underline{L}$ is a fixed-point of the mapping L^-. Therefore, updating monotonicity after good news guarantees that, if the prior likelihood ratio starts below $\overline{L}$, the posterior remains below it. Similarly, UMB guarantees that the posterior likelihood ratio is above $\underline{L}$ if the prior is. See Figure 1.

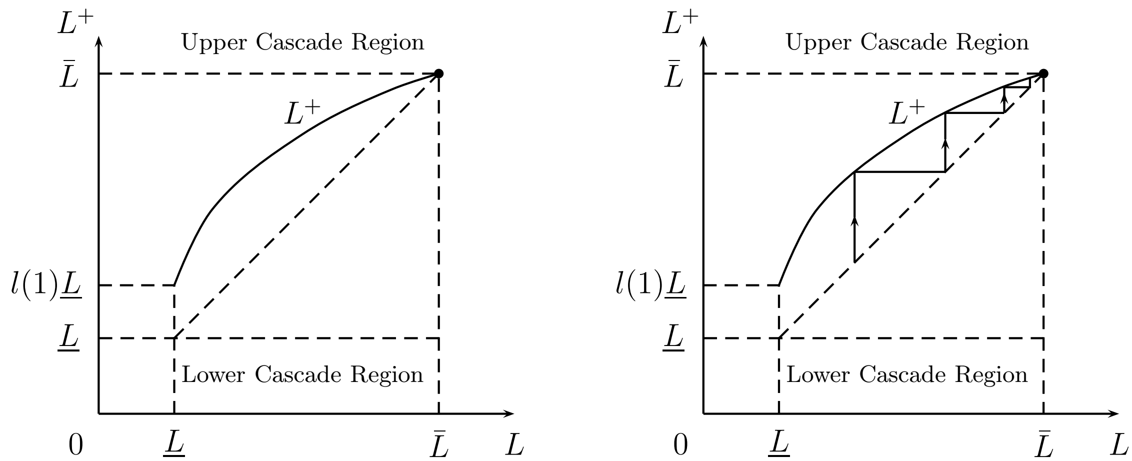

FIGURE 1. The map L^+ and its relationship to updating.

As a result, along with MLRP, IHRP (and IFRP) is a sufficient condition guaranteeing that, provided that the first agent uses a strategy that depends on his private signal, then all later agents will: if the game does not start with a cascade, a cascade cannot occur. Conversely, it is a necessary condition, in the sense that, if either condition is violated, it is possible to find a public belief p and a cost of investment c (i.e., a threshold $\bar{L}$) for which a cascade occurs with positive probability, i.e. one of the two actions will update the belief into the cascade region: simply consider an interval in which the map L^+ is decreasing, choose the prior log-likelihood belief as well as the threshold $\bar{L}$ in this interval, with the prior below the threshold, of course. One investment suffices to start a cascade. We summarize this discussion in our main Proposition.

Proposition 2. *Assume MLRP. If IHRP and IFRP holds, and the prior public belief is in* $(\underline{p}, \overline{p})$, *a cascade cannot occur. Conversely, if either IHRP or IFRP does not hold, there exists (an open set of values for) a public belief and private signal such that a cascade occurs.*

4.4. **Intuition for IHRP Failure and Cascades.** To get some intuition for these results, recall that $H(x)$ is the ratio of $h_1(x)/h_0(x)$, where

$$h_\theta(x) := \frac{f_\theta(x)}{1 - F_\theta(x)}$$

is the hazard function in state θ, namely the density at the cut-off x, conditional on an action $a = 1$. If the public belief increases, the threshold x decreases, and therefore, according to IHRP,

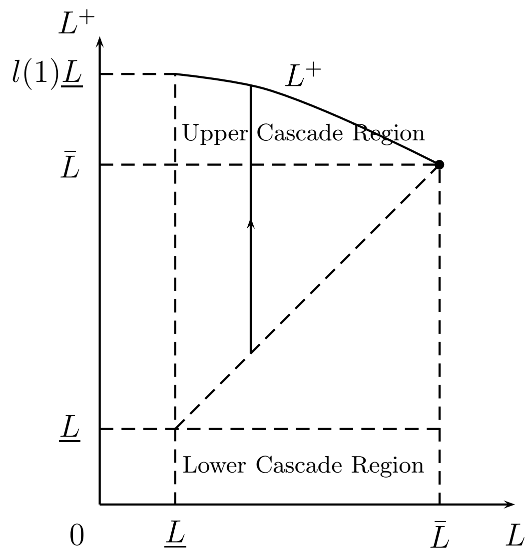

FIGURE 2. One investment triggers a cascade.

so does $H(x)$. That is, as the public belief increases, and conditional on $a = 1$, it becomes less likely that the investor was the least optimistic (among agents who would invest) in the good state relative to the bad state. To see why this is consistent with updating monotonicity, suppose that it were the case that, to the contrary, the least optimistic investor was conditionally more likely in the good state than in the bad state. Because of MLRP, an investment would still be good news, but, since higher signals are better news, this would mitigate the good news. The likelihood ratio might not increase much for higher values of L, and this might violate the monotonicity of the mapping L^+.

While it is not clear *a priori* whether it is more reasonable to assume that the hazard ratio is increasing or decreasing (or neither), note that, in case it is decreasing, the failure of updating monotonicity implies that a single decision to invest will trigger an investment cascade. That is, the dichotomy is rather extreme: if the first agent uses a strategy that depends on his signal, all later agents will do so if the hazard ratio is increasing (assuming IFRP), while a single investment will trigger a cascade if the hazard ratio is decreasing. The latter phenomenon characteristic of the *decreasing hazard ratio property* (DHRP) can be most easily understood with the help of the following figure. In Figure 2, one single positive action $a = 1$ always starts a positive cascade.

Note that DHRP is inconsistent with both unbounded signals and MLRP, but it is consistent with MLRP and signals that are not unbounded, as our previous DHRP example illustrated. Namely,

$$f_0 = \frac{3}{2}\left(2\left(1-x\right)-\left(1-x\right)^2\right), \quad f_1 = 1$$

has DHRP over all the domain, because the hazard ratio is

$$H\left(x\right) = \frac{x+2}{3x+3},$$

which is decreasing. So, for any initial prior belief an observed good news action leads always to a positive cascade.

4.5. **IHRP with Discrete Signals.** That IHRP (and IFRP) are the necessary and sufficient conditions for (the absence of) cascades appears to be new. The early literature (for instance, BHW) established that cascades occur when signals are discrete, while later contributions (in particular, Smith and Sørensen [7]) showed that this result does not necessarily hold with continuous signals. Indeed, it is easy to see that IHRP and IFRP are necessarily violated in the case of discrete signals.

Take a discrete signal model where the signals are $x \in \{x_1, x_2, ..., x_k, ..., x_n\}$. Then UMG is violated on some points of the domain of the public belief. Indeed, by Bayes' rule the posterior likelihood ratio L^+ after good news

$$L^+ = \frac{\sum\limits_{j\geq k} \mathbb{P}_1[x_j]}{\sum_{j\geq k} \mathbb{P}_0[x_j]} L,$$

decreases discontinuously in L at every threshold indifference point x_k. Namely, for every x_k there is a value of $L = \widetilde{L}$ for which

$$\frac{\mathbb{P}_1[x_k]}{\mathbb{P}_0[x_k]}\widetilde{L} = 1$$

and the function $L^+\left(L\right)$ is discontinuously decreasing at every $L = \widetilde{L}$.

The hazard ratio $H\left(x\right)$, defined as

$$H\left(x\right) := \frac{\mathbb{P}_1[x]}{\sum_{j\geq k} \mathbb{P}_1[x_j]} \Big/ \frac{\mathbb{P}_0[x]}{\sum\limits_{j\geq k} \mathbb{P}_0[x_j]}, \qquad x_{k-1} < x \leq x_k$$

is also decreasing at every threshold indifference point x_k. (On the other hand, the hazard ratio is increasing over each interval defined by these thresholds, which goes to show that there are natural examples in which H is not monotone.)

To understand this, suppose without loss of generality that there are three possible signals. Suppose that for a given public belief L observing the action $a = 1$ (good news) transmits the information that the agent had a high signal, whereas observing the action $a = 0$ (bad news) transmits the information that the agent had either a low signal or a medium signal. Now assume that L increases. At some point, L will reach a threshold level $\widetilde{L}$ so that that observing the good news action instead transmits the information that the agent had a high signal or also a medium signal. Namely, given this higher prior, also the medium signal is good enough to make the agent take the $a = 1$ action. However, now the information transmitted by the good news action is all of a sudden not as good news as knowing as before that the agent must have had just a high signal. Hence the updating to L^+ drops discontinuously at $\widetilde{L}$. In other words, for some $\varepsilon > 0$, updating monotonicity is violated: $L^+\left(\widetilde{L} - \varepsilon\right) > L^+\left(\widetilde{L}\right)$. Hence, observing good news at $\widetilde{L} - \varepsilon$ may lead to a cascade. The latter will definitely always occur if $\widetilde{L}$ is the highest threshold at which this discontinuity occurs, that is the threshold $\widetilde{L} = \overline{L}$ at which the set of signals consistent with action $a = 1$ goes from being all but the lowest possible signal to all private signals. Note that the cascade is driven precisely by the discontinuous drop of L^+ at $\widetilde{L}$.

4.6. **Relation between MLRP and IHRP or DHRP.** Of course, IHRP neither implies, nor is implied by MLRP. While MLRP states that higher signals are better news, IHRP states that this remains true, conditional on truncations i.e. substituting $f_\theta(x)$ with $h_\theta(x)$. Therefore, these two stochastic orders are related, as IHRP implies MLRP if in addition $h_\theta(x)$ decreases in the state θ.

For θ belonging to any ordered state space, the properties of monotone likelihood ratio property (MLRP), the *decreasing hazard function property* (DHFP) and the increasing hazard ratio property (IHRP) can be stated as

$$\begin{aligned}
\text{MLRP} \quad &\Longleftrightarrow \quad \frac{\partial}{\partial x}\ln\left(f_\theta\left(x\right)\right) \text{ is increasing in } \theta,\\
\text{DHFP} \quad &\Longleftrightarrow \quad \frac{\partial}{\partial x}\ln\left(1-F_\theta\left(x\right)\right) \text{ is increasing in } \theta,\\
\text{IHRP} \quad &\Longleftrightarrow \quad \frac{\partial}{\partial x}\ln\left(\frac{f_\theta\left(x\right)}{1-F_\theta\left(x\right)}\right) \text{ is increasing in } \theta.
\end{aligned}$$

Proposition 3. *The following holds:*

$$\begin{aligned}
&\textit{IHRP+DHFP} \quad \textit{implies} \quad \textit{MLRP},\\
&\textit{MLRP} \quad \textit{implies} \quad \textit{DHFP}.
\end{aligned}$$

Proof. The first implication is trivial as, for $\bar{\theta} > \underline{\theta}$,

$$\frac{\partial}{\partial x}\ln f_{\bar{\theta}}\left(x\right) - \frac{\partial}{\partial x}\ln f_{\underline{\theta}}\left(x\right) > \frac{\partial}{\partial x}\ln\left(1-F_{\bar{\theta}}\left(x\right)\right) - \frac{\partial}{\partial x}\ln\left(1-F_{\underline{\theta}}\left(x\right)\right) > 0.$$

For the second implication, we need to show that, for $\bar{\theta} > \underline{\theta}$, $x \in X$,

$$\frac{f_{\bar{\theta}}\left(x\right)}{\int_x^{\bar{x}} f_{\bar{\theta}}\left(z\right)dz} < \frac{f_{\underline{\theta}}\left(x\right)}{\int_x^{\bar{x}} f_{\underline{\theta}}\left(z\right)dz},$$

which is equivalent to

$$\begin{aligned}
\int_x^{\bar{x}}\left(f_{\underline{\theta}}\left(x\right)f_{\bar{\theta}}\left(z\right) - f_{\bar{\theta}}\left(x\right)f_{\underline{\theta}}\left(z\right)\right)dz \quad &> \quad 0,\\
\int_x^{\bar{x}} f_{\underline{\theta}}\left(x\right)f_{\underline{\theta}}\left(z\right)\left(l_{\bar{\theta},\underline{\theta}}\left(z\right) - l_{\bar{\theta},\underline{\theta}}\left(x\right)\right)dz \quad &> \quad 0,
\end{aligned}$$

where $l_{\bar{\theta},\underline{\theta}} = f_{\bar{\theta}}/f_{\underline{\theta}}$ and the last inequality is implied by MLRP. □

Also, because Smith and Sørensen [6] show that monotonicity of the mapping L^+ is implied by the assumption that the private belief log-likelihood ratio is log-concave, while IHRP is necessary and sufficient, it follows indirectly that the log-concavity assumption is stronger than IHRP (and arguably more complicated to verify). On the implications of log-concavity, see also Smith and Sørensen [7].

Proposition 4. *A necessary condition for DHRP for all $x \in X$ is $l\left(\bar{x}\right) = +\infty$ and $l\left(\underline{x}\right) > 0$.*

Proof. With MLRP, there has to be an interior "neutral news" signal $x_{NN} \in X$, i.e. a signal that solves

$$l(x_{NN}) = 1.$$

By l'Hôpital's rule,

$$l(\bar{x}) < +\infty \quad \Longrightarrow \quad H(\bar{x}) = 1,$$

and

$$H(\underline{x}) = l(\underline{x}) < 1.$$

So in order to have DHRP, it must be that $H(\underline{x}) > H(\bar{x}) \geq 0$. This implies $l(\bar{x}) = +\infty$ and $l(\underline{x}) > 0$. □

Call DFRP the (strict) Decreasing Failure Ratio Property.

Corollary 5. *DHRP and DFRP cannot simultaneously hold. Also, if either DHRP or DFRP holds, the distribution cannot be symmetric, i.e. it cannot be that, for all* $x \in [0, \bar{x} - \underline{x}]$, $f_1(\underline{x} + x) = f_0(\bar{x} - x)$.

Proof. DHRP implies $l(\bar{x}) = +\infty$ and $l(\underline{x}) > 0$ and DFRP implies $l(\bar{x}) < +\infty$ and $l(\underline{x}) = 0$. Evidently we cannot have symmetry when $l(\bar{x}) = +\infty$ and $l(\underline{x}) > 0$. □

4.7. **Commonly Used Distribution Families.** Most but not all distributions commonly used in economics satisfy IHRP. One exception is the Pareto distribution, which satisfies DHRP, but violates MLRP.

4.7.1. *Exponential.* The exponential distribution,

$$f_\theta(x) = \frac{e^{-\frac{x}{\theta}}}{\theta}, \quad x > 0,$$

satisfies strict MLRP. Because its hazard function is constant,

$$\frac{f_\theta(x)}{1 - F_\theta(x)} = \frac{1}{\theta},$$

the hazard ratio is constant too. Hence the exponential distribution is a knife edge case, the posterior after a good news truncation is constant regardless of the prior, so we have a cascade: from any prior any good news puts the posterior at the upper boundary of the cascade region.

4.7.2. *Distributions that satisfy IHRP strictly.* As is readily verified, these include the normal distribution $\mathcal{N}(\theta, \sigma^2)$; the power distributions $F_\theta(x) = x^\theta$, with $\theta > 1$, $x \in [0,1]$; the gamma distribution $f_\theta(x) = e^{-x}x^{\theta-1}/\Gamma(\theta)$ $(\theta \geq 1,\ x \geq 0)$; the chi-distribution $f_\theta(x) = e^{-x^2/2}x^{\theta-1}/\left(2^{(\theta-2)/2}\Gamma(\theta)\right)$ $(\theta \geq 1,\ x \geq 0)$, as well as the chi-squared distribution, $f_\theta(x) = e^{-x/2}x^{(\theta-2)/2}/\left(2^{\theta/2}\Gamma(\theta/2)\right)$ $(\theta \geq 2,\ x \geq 0)$. All these distributions satisfy MLRP.

References

[1] Banerjee, A. (1992). "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, **107**, 797–817.

[2] Bikhchandani, S., Hirshleifer D. and I. Welch (1992). "A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades," *Journal of Political Economy*, **100**, 992–1026.

[3] Chamley C. (2003). "Rational Herds: Economic Models of Social Learning," *Princeton University Press.*

[4] Kalashnikov, V. V. and S. T. Rachev (1986). "Characterization of queuing models and their stability," in *Probability Theory and Mathematical Statistics*, ed. Yu K. Prohorov et al., **2**, 37–53, VNU Science Press, Amsterdam.

[5] Milgrom, P. R. (1981). "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, **12**, 380–391.

[6] Smith, L. and P. Sørensen (2001). "Pathological Outcomes of Observational Learning," *Econometrica*, **68**, 371–398.

[7] Smith, L. and P. Sørensen (2008). "Informational Herding and Optimal Experimentation," mimeo, University of Michigan.