THE POSITIVE EFFECT OF GARBLING ON SOCIAL LEARNING

ITAI ARIELI, YAKOV BABICHENKO, AND WADE HANN-CARUTHERS

Abstract. This paper explores the dynamics of social learning within the herding model framework. We examine the efficacy of regulatory interventions, specifically focusing on garbling or restricting agents' private information. Two regulator types, adaptive and static, are considered. Surprisingly, both regulators prove equally powerful in maximizing agents' asymptotic welfare. The findings imply that regulators need not actively monitor real-time actions for optimal outcomes; committing to predefined information garbling suffices. We provide a precise characterization of asymptotic welfare in the presence of the regulator, establishing an upper bound for asymptotic welfare in the standard herding model. Finally, we demonstrate that a gap exists between static and adaptive regulators when the regulator aims to maximize the herding probability on one of the actions.

A notable classic example in social learning illustrates that rational agents may fail to aggregate information. Specifically, in the celebrated herding model [6, 5, 13], agents sequentially take actions based on their private information and the history of predecessors' actions. The (hypothetical) aggregation of all the private information of all agents determines the state of the world with probability one. Nevertheless, the agents may herd on inferior action with positive probability. This inefficiency of social learning naturally raises the question: *Can regulation alleviate the inefficiency?*

Different forms of regulations might be considered in the context of social learning (see Section for a discussion of various types of regulations that have been considered in the literature). In this paper, we focus on a regulation that takes a simple form of restricting (i.e., garbling) the private information of agents. Notice that such a non-monetary intervention requires minimalistic assumptions on the abilities of the regulator. In particular, the regulator may have no access to the private information of the agents. Moreover, communication with the agents is not required. Instead, the regulator may restrict the private information of an agent by imposing some filters on the information sources of individuals. At the micro level, the restriction of the information can only damage an individual. However, at the macro level, we will see that a careful choice of garblings might benefit society and cause herding on the superior action with a higher probability.

We consider the canonical herding model: Agents are trying to guess the correct binary state $\theta \in \{0, 1\}$ sequentially and are exposed to private signals that are drawn i.i.d conditional on the state. A regulator who plays the role of an information designer in our model generates a garbling of these private signals. We mainly focus on two types of regulators. The first is the *adaptive regulator* who has access to the public information (i.e., the actions of the agents). The adaptive regulator has the flexibility to design the garbling at time t as a function of the actions up to time t. The second is the *static regulator* who is less powerful than the adaptive one. She commits to the entire sequence of garbling before the interaction starts.

Our main result focuses on a regulator who is trying to maximize the asymptotic welfare of the agents; namely, she is trying to maximize the probability of a herd on the correct action. Surprisingly, we show that both the static and the adaptive regulators are equally powerful. Moreover, we provide a clean formula for the asymptotic welfare of the agents. The existence of such a clean formula for the best social welfare is somewhat surprising given the fact that in the standard model without garbling, no such formula exists.

The implications of this result are twofold. First, it shows that in order to maximize the asymptotic welfare of the society, the regulator does not need to monitor the agents' actions, and it can do as well by committing in advance to a revelation policy. Secondly, the welfare characterization presented in our results establishes a non-trivial upper bound for the asymptotic welfare in the standard herding model without the presence of a regulator.

Our second result considers a regulator with transparent motives that is instead interested in maximizing the probability of the society to asymptotically herd on one of the two actions (say action 1). In contrast to the previous result, we demonstrate a gap between the two monitoring models and show that the regulator cannot achieve the same utility if it is uninformed of the history. A closed-form expression for the optimal guarantees of the adaptive regulator is provided.

Related Literature. The conventional herding model offers profound insights into the dynamics of social learning. [6] and [5] demonstrate that even rational individuals, whose pooled private information is sufficient for identifying the optimal action, may collectively opt for a suboptimal alternative. Specifically, they illustrate the occurrence of information cascades, wherein agents, starting from a specific time period, disregard their private signals and base their decisions solely on the historical choices of others. This may lead to failure of *asymptotic learning* where the probability of selecting the optimal action approaches one with time. In contrast, Smith and Sorensen [13] show that when signals have unbounded precision, then asymptotic learning holds.

Interfering in the social learning process to optimize the social outcome has been studied by [1], [3], [12]. These works focus on the observation structures where agents observe only a partial subset of their predecessors. We take an orthogonal perspective and focus on the private information of the agents. We restrict attention to the observation structure where the agents observe the actions taken by all past agents but we believe that this assumption can be substantially relaxed.

Smith, Sorensen, and Tian [14] study the problem of a social planner who is designing the optimal decision rule for agents to maximize the discounted stream of utilities. Specifically, the social planner has the power to choose a map from private signals to action recommendations, to maximize the discounted sum of the agents' payoffs. They provide comparative statics over the cascading region in the optimum as a function of the discount factor. In contrast to our work, the optimal decision rule ignores informational externalities, and thus agents are not selfish and do not optimize their payoff given their information.

Our work is also related to the intersection of the herding literature with Bayesian persuasion [8]. An example of such integration is found in the related paper by [10], who studies the classic binary-action herding model with a designer that can dynamically send information to the agents in addition to their private signals. As in the second part of our analysis, the sender would like to maximize the probability that the population herds on Action H. The main result derives conditions under which the designer prefers observational learning and under which she prefers to induce a herd from the start. In another paper, [2] considered a herding model with a designer that can choose the agents' private conditional i.i.d. signal. The state space is binary, but the agents and the designer have general utility functions. The designer chooses an information structure to maximize her utility in the asymptotic action that the agents play. The main result identifies necessary and sufficient conditions under which the designer's utility is equal to the maximal possible achieved utility obtained in the Bayesian persuasion model. In contrast with the aforementioned works our designer who is represented by the regulator, cannot generate information and can only garble the agents' given private information.

Finally, as mentioned above, our characterization provides a non-trivial upper bound on the agents' asymptotic welfare in the herding model as a function of their private beliefs distribution. The only known non-trivial lower bound on the asymptotic welfare, which is also valid for a large class of random social networks (see [9]), is obtained by considering the boundary of the support of the private belief regions.

1. Model

The canonical herding model considers an unknown binary state $\theta \in \{0,1\}$. We identify beliefs about θ with real numbers in $[0,1]$ that indicate the probability of $\theta = 1$. The common prior is denoted by $\pi \in [0,1]$. Every agent $t = 1, 2, ...$ observes a private signal about the state. The distribution of posteriors conditional at state $\theta = 0, 1$ is denoted by $F_0, F_1 \in \Delta([0,1])$. The unconditional distribution of posteriors is denoted by $F = \pi F_1 + (1-\pi)F_0$. By the splitting lemma (see [4]) $\mathbb{E}[F] = \pi$ and every F with expectation π can be induced by some signaling structure.

1.1. Garbling.

The regulator is allowed to garble the private signal F of agent t; i.e., to cause agent t to be exposed to a less informative signal in the Blackwell sense. Concretely, the regulator can choose a distribution $G_t \preceq F$ where $\preceq$ denotes the second-order stochastic dominance order. The set of all possible garblings is denoted by $\mathcal{G} = \{G : G \preceq F\}$. The following types of regulators will be considered.

- The *adaptive regulator* commits to a regulation policy that is a mapping $r : H \to \mathcal{G}$, where H is the set of all possible (publically observed) histories of actions; namely, H is simply the set of all finite 0/1 sequences.
- The *static regulator* sets the sequence $r = (G_1, G_2, G_3, ...)$ before the interaction.

The agents are rational and are aware of the regulation policy. In particular, they know how the information of their predecessors was garbled and hence can deduce a posterior from the publicly observed sequence of past actions.

1.2. The Interaction.

All agents have binary actions $a = 0, 1$ and the same utility function, $u(a, \theta) = \mathbb{1}[a = \theta]$. Agents $t = 1, 2, ...$ sequentially observe a private signal $\omega \sim G_t$ (the garbled signal by the regulator), they observe the strategy of the regulator as well as the history of play $(a_1, ..., a_{t-1})$, and take an expected utility maximizing action $a_t \in \{0,1\}$.

Given a garbling policy r (either adaptive or static) we denote by $\text{BNE}(r)$ the set of Bayesian Nash equilibria of the above interaction.[1]

2. Welfare maximizing regulator

For clarity of presentation, we restrict attention to the case where the common prior is $\pi = \frac{1}{2}$. This assumption is not crucial, and our results can be easily generalized for an arbitrary $\pi \in [0,1]$.

In this section, we consider a regulator whose goal is to maximize the asymptotic social welfare; namely, to maximize the probability that the herd will be on the correct action.

Formally, given r, we evaluate the welfare performance for an equilibrium $E \in \text{BNE}(r)$ by $W(E) = \lim_{t\to\infty} \mathbb{P}[a_t = \theta]$.[2] We evaluate the performance of r by the worst-case equilibrium $W(r) = \inf_{E \in \text{BNE}(r)} W(E)$. Our goal is to find a regulation policy r with the best performance. For the classes of adaptive and static regulation policies, we denote by W_{adap} and W_{stat} the expression $\sup_r W(r)$, where the supremum is taken over the adaptive/ static regulation policies correspondingly.

The optimal social welfare that can be induced by garbling (in the adaptive and the static case) can be expressed by the notion of *information width* that is defined below.

2.1. Width.

As we will see, binary garblings (i.e., garblings that satisfy $|\text{supp}(G_t)| = 2$) play a central role in our model. Among the binary-garblings we would like to focus on the class of the most informative; i.e., the Parto frontier of the binary-garblings with respect to the Blackwell order. Those most informative garblings take a threshold form. For every $q \in [0,1]$ we denote by $F|_q$ the conditional distribution of the low q-quantile of F; i.e., the mass of size q of the lowest values of F. We denote by $F|^{1-q}$ the conditional distribution of the high $(1-q)$-quantile of F; i.e., the mass of size $1-q$ of the highest values of F. The garbling $G(q)$ pools together the mass of $F|_q$ into the posterior $p_-(q) = \mathbb{E}[F|_q]$ and it pools together the mass of $F|^{1-q}$ into the posterior $p_+(q) = \mathbb{E}[F|^{1-q}]$. The class $\{G(q) : q \in (0,1)\}$ is the Pareto frontier.

[1]Multiplicity of equilibria might happen only in the case where some agent t has a posterior exactly $\frac{1}{2}$, in which case both choices $a_t = 0, 1$ can occur in equilibrium.

[2]Note that by the imitation principle, in any equilibrium, $\mathbb{P}[a_{t+1} = \theta] \geq \mathbb{P}[a_t = \theta]$, so this is an increasing bounded sequence which converges to the probability that the agents herd on the correct action.

As usual in social learning, and more generally in aggregation of conditionally independent signals, the log-likelihood ratio plays a central role. We denote by $\mathrm{LL} : (0,1) \to \mathbb{R}$ the log-likelihood ratio transformation $\mathrm{LL}(x) = \log(\frac{x}{1-x})$. For convenience of notations we also define $\mathrm{LL}(0) = -\infty$ and $\mathrm{LL}(1) = \infty$. Its inverse is the logit function $\mathrm{logit} : \mathbb{R} \to (0,1)$ defined by $\mathrm{logit}(x) = \frac{e^x}{1+e^x}$. Again, for notation convenience, we define $\mathrm{logit}(-\infty) = 0$ and $\mathrm{logit}(\infty) = 1$.

The *width* of F is defined as $\mathrm{Wid}(F) = \sup_{q\in(0,1)} \Big(\mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q)) \Big)$; namely, it captures the maximal log-likelihood distance that can be induced from F. The value q that maximizes the width, when it exists, is denoted by q^*.[3] We also denote $p^*_- = p_-(q^*)$ and $p^*_+ = p_+(q^*)$. Perhaps surprisingly, calculating the width is tractable for distributions commonly used in the social learning literature (see the examples in Section 2.3).

2.2. **Main result.** We are now ready to state our main theorem.

Theorem 1. *For every F we have $W_{\mathrm{adap}} = W_{\mathrm{stat}} = \mathrm{logit}(\mathrm{Wid}(F))$.*

The theorem states that the static regulator is as powerful as the adaptive regulator. Both can achieve the social welfare of $\mathrm{logit}(\mathrm{Wid}(F))$ by careful garbling of information. Moreover, even the adaptive regulator cannot exceed the $\mathrm{logit}(\mathrm{Wid}(F))$ bound. Furthermore, our proof is constructive and suggests a simple sequence of garblings (i.e., static garbling) that approaches the social welfare of $\mathrm{Wid}(F)$.

Remark 2. It follows from the proof of Theorem 1 that for every $\epsilon > 0$ there exists $T \in \mathbb{N}$ and two garblings $G_s, G_l \preceq F$ for which the T-periodic sequence of garblings

$$(G_s, ..., G_s, G_l, G_s, ..., G_s, G_l, ...),$$

in which G_s repeats $T-1$ times, guarantees a social welfare of at least $\mathrm{Wid}(F) - \epsilon$. The notation G_s captures the idea that G_s should (typically) reveal *a small* amount of information about the state and is expected to have a small effect on the learning process. The notation G_l captures the idea that G_l should reveal *a large* amount of information about the state and is expected to have a large effect on the learning process. In fact, G_l is precisely the q^*-quantile binary garbling that maximizes the width.

[3]One can show that if signals are bounded, i.e., does not contain $\{0,1\}$ in the support, then the maximum exists.

The exact expression for social welfare (i.e., the probability of herd on the correct action) is a complicated object without any garbling (namely when F is not garbled). Thus, the clean expression of $\operatorname{logit}(\operatorname{Wid}(F))$ is somewhat surprising. Moreover, this result allows us to bound from above the social welfare in the classical social learning setting (without garbling).

Corollary 3. *In the standard herding model in which all agent's signals are distributed according to F, the probability that agents will herd on the correct action is at most* $\operatorname{logit}(\operatorname{Wid}(F))$.

The corollary simply follows from the fact that one option for the regulator is not to garble the information of any agent.[4]

2.3. **Examples.** To illustrate our results, we calculate the width and demonstrate the scope for improvement for three families of distributions: signals that are unbounded from one side, signals distributed ex-ante uniform on an interval, and binary signals.

Example 1: Signals unbounded from one side

Distributions that do not induce arbitrarily strong posteriors, in the sense that their support does not contain 0 or 1, are referred to as *bounded* (see [6], [13]). There are two types of distribution which are not bounded. The first type is the unbounded distribution which was introduced by Smith and Sorensen [13]. Under unbounded distributions, it is known that *asymptotic learning* holds and thus the asymptotic action matches the state with probability 1. Therefore, under unbounded signal, no regulator intervention is needed since the asymptotic welfare is 1 without any intervention.

The second type of distributions which are not bounded are distributions that are unbounded only on one side. Namely, those distributions that contain either 0 or 1, but not both, in the support. For such distributions it holds that for any $\varepsilon > 0$, there is a q such that either $p_-(q) < \varepsilon$ or $p_+(q) > 1 - \varepsilon$. Therefore, since

$$\operatorname{Wid}(F) \geq \operatorname{LL}(p_+(q)) - \operatorname{LL}(p_-(q)) \geq \max\left(|\operatorname{LL}(p_+(q))|, |\operatorname{LL}(p_-(q))|\right),$$

it follows that in either case, $\operatorname{Wid}(F) > \operatorname{LL}(1-\varepsilon) > \log\left(\frac{1}{\varepsilon}\right) - 1$ and thus $\operatorname{Wid}(F) = \infty$. Therefore under regulation, the probability of herding on the correct actions for such distributions can be made arbitrarily close to 1.

[4]In fact, Corollary follows from the first stage in our proof; see Lemma 4.

Interestingly, for every $\epsilon > 0$ there exists a bounded from one side distributions F for which the standard herding model herds on the correct action with a probability of at most $\frac{1}{2}+\epsilon$. For those distributions F the regulator increases the probability of herding from $\frac{1}{2}+\epsilon$, which corresponds to essentially no information aggregation, to approximately 1, which corresponds to the perfect learning of the state.

To construct such an F we consider a distribution with $\mathrm{supp}(F) = [\frac{1}{2} - \varepsilon^2, 1]$ which never cascades (see Appendix I) for some small $\varepsilon > 0$. Since $0 \leq p_t \leq \frac{1}{2} + \varepsilon^2$ and $\mathbb{E}[\frac{1}{2} + \varepsilon^2 - p_t] = \varepsilon^2$, by Markov's inequality $\mathbb{P}[p_t \leq \frac{1}{2} - \varepsilon] \leq \varepsilon$, so

$$\mathbb{E}[\max(p_t, 1 - p_t)] \leq \varepsilon \cdot 1 + (1 - \varepsilon) \cdot (\frac{1}{2} + \varepsilon) = \frac{1}{2} + O(\varepsilon).$$

Since $\mathbb{P}[a_t = \theta] = \mathbb{E}[\max(p_t, 1 - p_t)]$, it follows that the probability of herding on the correct action is $\frac{1}{2} + O(\varepsilon)$. On the other hand, by Theorem 1,

$$W_{\text{adap}} = W_{\text{stat}} = \mathrm{logit}(\mathrm{Wid}(F)) = 1.$$

Example 2: Uniform signals

Suppose $F \sim U([1 - r, r])$ with $r \in (\frac{1}{2}, 1)$. A straightforward calculation gives

$$p_-(q) = \frac{1 - (1 - q)\alpha}{2}$$
$$p_+(q) = \frac{1 + q\alpha}{2}$$

where $\alpha = 2r - 1$, and hence

$$\mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q)) = \log\left(\frac{1 + q\alpha}{1 - q\alpha}\right) + \log\left(\frac{1 + (1 - q)\alpha}{1 - (1 - q)\alpha}\right).$$

This function is quasi-convex and symmetric about $\frac{1}{2}$, so

$$\mathrm{Wid}(F) = \lim_{q \to 1} \mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q)) = \log\left(\frac{1 + \alpha}{1 - \alpha}\right) = \mathrm{LL}(r).$$

By Theorem 1,

$$W_{\text{adap}} = W_{\text{stat}} = \mathrm{logit}(\mathrm{Wid}(F)) = r.$$

Moreover, since $\mathrm{supp}(F) = [1 - r, r]$, the probability of herding on the correct action is at least r, and so must be exactly r. Hence, for such a distribution, the regulator cannot improve the probability of herding on the correct action.

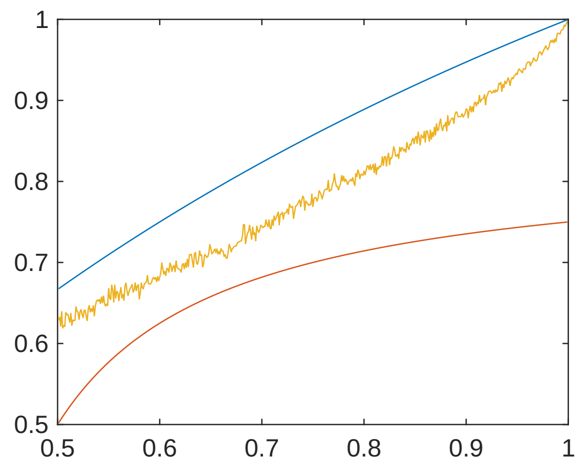

FIGURE 1. Correct herd probability for binary signals with support $\{\frac{1}{3}, x\}$

Example 3: Binary signals

Let $\mathrm{supp}(F) = \{r_-, r_+\}$ for $r_- < \frac{1}{2} < r_+$ denote the two points in the support. Since F is a binary garbling of itself,

$$\mathrm{Wid}(F) \geq \mathrm{LL}(r_+) - \mathrm{LL}(r_-).$$

Moreover, for any $q \in (0,1)$, $p_-(q) \geq r_-$ and $p_+(q) \leq r_+$, so

$$\mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q)) \leq \mathrm{LL}(r_+) - \mathrm{LL}(r_-).$$

It follows that

$$\mathrm{Wid}(F) = \mathrm{LL}(r_+) - \mathrm{LL}(r_-).$$

By Theorem 1,

$$W_{\mathrm{adap}} = W_{\mathrm{stat}} = \mathrm{logit}(\mathrm{Wid}(F)) = \frac{r_+(1-r_-)}{r_+(1-r_-) + r_-(1-r_+)}.$$

There is no known closed-form expression for the probability of herding on the correct action for binary signals in the standard herding model (without garbling). A well-known lower bound is given by

$$\frac{(\frac{1}{2} - r_-)r_+ + (r_+ - \frac{1}{2})(1-r_-)}{r_+ - r_-}.$$

In Figure 1, we plot W_{adap} (top curve), the lower bound (bottom curve), and an approximation for the probability of herding on the correct action using simulations, with $r_- = \frac{1}{3}$ and r_+ ranging from $\frac{1}{2}$ to 1.

2.4. **Proof Idea of Theorem 1.** A central notion for analyzing social learning dynamics is the *public belief* martingale $(p_t)_{t=0}^{\infty}$. The public belief captures the belief about the state of an outside observer who is not exposed to any information beyond the actions of the agents. Initially the public belief is the prior $p_0 = \pi$ ($p_0 = \frac{1}{2}$ with our simplifying assumption that the prior is $\pi = \frac{1}{2}$). In every period t the public belief has two possible updates conditional on the last action being 0 or 1. By the law of iterated expectation, one can see that

$$a_t = 0 \Rightarrow p_t \leq \frac{1}{2} \text{ and } a_t = 1 \Rightarrow p_t \geq \frac{1}{2}. \tag{1}$$

Namely, the last played action indicates whether the public belief belongs to $[0, \frac{1}{2}]$ or to $[\frac{1}{2}, 1]$.

Denote by $\underline{p} = \min\{x \in \operatorname{supp} F\}$ and $\overline{p} = \max\{x \in \operatorname{supp} F\}$ the two most extreme signals of F. Notice that if $\mathrm{LL}(p_t) < -LL(\overline{p})$ then there is a cascade on the action $a = 0$ because no signal is sufficiently informative to switch the belief of an individual above $\frac{1}{2}$. Similarly if $\mathrm{LL}(p_t) > -LL(\overline{p})$ then there is a cascade on the action $a = 1$ because no signal is sufficiently informative to switch the belief of an individual below $\frac{1}{2}$. In other words, all states $x \notin [\mathrm{logit}(-LL(\overline{p})), \mathrm{logit}(-LL(\underline{p}))]$ are absorbing for the martingale of public belief. The segment of non-absorbing states is denoted by $I = [\mathrm{logit}(-LL(\overline{p})), \mathrm{logit}(-LL(\underline{p}))]$.

The public belief martingale also captures the probability of herding on the correct action. Denote by p_∞ the limit distribution of the martingale (p_t). The probability of a correct herding is given by

$$W = \int_0^{\frac{1}{2}} (1-x) dp_\infty + \int_{\frac{1}{2}}^1 x dp_\infty. \tag{2}$$

Namely, if in the limit we are absorbed in the public belief $x > \frac{1}{2}$, then there is herding on $a = 1$ and, indeed the probability of this herding being correct is x; similarly for the first expression of Equation (2).

All the above observations hold also for social learning with garbling. The only difference is that the private signals of the agents are drawn according to $G_t \preceq F$.

Intuitively, Equation (2) together with the fact that the non-absorbing region is I indicates that in order to increase W we would like to achieve the following phenomenon: in the last step before a cascade, the public belief "jumps maximally out of I". So the question is: *how far from I can the public belief jump?* The answer to this question turns out to be the width.

Lemma 4. *For every sequence of garblings $(G_t)_t$ we have $|\mathrm{LL}(x)| \leq \mathrm{Wid}(F)$ for every $x \in \operatorname{supp} p_\infty$.*

Namely, $\mathrm{LL}(x) \leq \mathrm{Wid}(F)$ means that the public belief cannot jump to the right of $\mathrm{logit}(\mathrm{Wid}(F))$, and $LL(x) \geq -\mathrm{Wid}(F)$ means that the public belief cannot jump to the left of $\mathrm{logit}(-\mathrm{Wid}(F))$.

Proof of Lemma 4. Observe that the private signals drawn from $G_t \preceq F$ that yield the actions $a_t = 0, 1$ have a threshold structure: all signals (of G_t) below $\mathrm{logit}(-\mathrm{LL}(p_t))$ yield the action $a_t = 0$, and all signals above $\mathrm{logit}(-\mathrm{LL}(p_t))$ yield the action $a_t = 1$. Thus the public signals corresponding to actions $a_t = 0, 1$ are precisely $p_-(q), p_+(q)$, where q is the probability of action $a_t = 0$ given the history of actions up to time t.

Now, observe that

$$\mathrm{LL}(p_t) = \begin{cases} \mathrm{LL}(p_{t-1}) + \mathrm{LL}(p_-(q)), & \text{if } a_t = 0 \\ \mathrm{LL}(p_{t-1}) + \mathrm{LL}(p_+(q)), & \text{if } a_t = 1 \end{cases}$$

If $q = 0$ or $q = 1$, then a_t conveys no information, so $p_t = p_{t-1}$. If $0 < q < 1$, then

$$\mathrm{LL}(p_{t-1}) + \mathrm{LL}(p_-(q)) \leq 0 \leq \mathrm{LL}(p_{t-1}) + \mathrm{LL}(p_+(q)),$$

so if $a_t = 0$, then

$$0 \geq \mathrm{LL}(p_t) = \mathrm{LL}(p_{t-1}) + \mathrm{LL}(p_-(q)) \geq -[\mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q))],$$

and if $a_t = 1$, then

$$0 \leq \mathrm{LL}(p_t) = \mathrm{LL}(p_{t-1}) + \mathrm{LL}(p_+(q)) \leq \mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q)),$$

and in either case,

$$|\,\mathrm{LL}(p_t)| \leq \mathrm{LL}(p_+(q)) - \mathrm{LL}(p_-(q)) \leq \mathrm{Wid}(G_t) \leq \mathrm{Wid}(F).$$

Putting these observations together, it follows that

$$|\,\mathrm{LL}(p_t)| \leq \max\left(|\,\mathrm{LL}(p_{t-1})|, \mathrm{Wid}(F)\right)$$

almost surely. Since $|\,\mathrm{LL}(p_0)| = 0 \leq \mathrm{Wid}(F)$, it follows by induction that $|\,\mathrm{LL}(p_t)| \leq \mathrm{Wid}(F)$ for all t almost surely, and the claim follows. $\square$

The proof of Lemma 4 instructs us also how to reach the bound of $\mathrm{Wid}(F)$: Ideally, for a cascade from the right side of I we want the public belief to reach $\mathrm{LL}(p_{t-1}) = -\mathrm{LL}(p_-^*) - \epsilon$ at time $t-1$ for some small $\epsilon > 0$. At the time t, we would like to garble the information via the q^*-quantile threshold garbling which results in a public belief with $\mathrm{LL}(p_t) \approx \mathrm{Wid}(F) - \epsilon$ if the action is $a_t = 1$. Otherwise, if the action is $a_t = 0$, the resulting public belief will be approximately $\mathrm{LL}(p_t) \approx -\epsilon$ and will remain inside the interval I.

For a cascade from the left side of I we want a symmetric picture: the public belief should reach $\mathrm{LL}(p_{t-1}) = -\mathrm{LL}(p^*_+) + \epsilon$. At the time t, we again use the q^*-quantile threshold garbling which potentially will bring us to a public belief of $\mathrm{LL}(p_t) \approx -\mathrm{Wid}(F) + \epsilon$. Otherwise, if the action would be $a_t = 1$ the public belief will be approximately $\mathrm{LL}(p_t) \approx \epsilon$ and will remain inside the interval I. Figure 2 summarizes the desirable plan of the regulator on the behavior of the public belief.

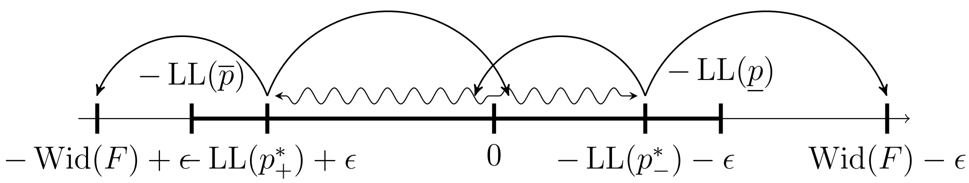

FIGURE 2. The Figure demonstrates the desirable plan of the regulator in the log-likelihood space. Wavy arrows capture a slow dynamic movement of the public belief towards one of the points $-\mathrm{LL}(p^*_+) + \epsilon$ or $-\mathrm{LL}(p^*_-) - \epsilon$. Arc arrows capture the possible jumps of the public belief that occurs in a single period of the q^*-quanite garbling. The bold line corresponds to the interval I (i.e., the non-cascade region of the public belief).

An important (and somewhat surprising) observation is that the garbling that should be applied at both points, near $\mathrm{logit}(-\mathrm{LL}(p^*_+))$ and near $\mathrm{logit}(-\mathrm{LL}(p^*_-))$, is *the same* garbling.

If we follow these instructions, eventually we will reach a cascade either ϵ-close to $\mathrm{logit}(\mathrm{Wid}(F))$ or ϵ-close to $\mathrm{logit}(-\mathrm{Wid}(F))$. To follow these instructions we have two challenges: (a) to bring the public belief near $\mathrm{logit}(-\mathrm{LL}(p^*_+))$ or near $\mathrm{logit}(-\mathrm{LL}(p^*_-))$ and (b) to identify the time when we are there. Notice that challenge (b) is actually a challenge when we consider a static regulator who should in advance determine the sequence of garblings and cannot track the location of the public belief.

We show that challenges (a) and (b) can be accomplished even by the static regulator. To achieve these goals we design a garbling such that if it is applied repeatedly then the public belief (almost) continuously moves to one of the points $\mathrm{logit}(-\mathrm{LL}(p^*_+))$ or $\mathrm{logit}(-\mathrm{LL}(p^*_-))$ with the guarantee of not crossing them. Moreover, the time on which the public belief reaches these points can be accurately predicted and hence is addressed too.

3. A regulator with transparent motives

We now consider a regulator whose goal is to maximize the probability of herding on a particular action, say $a = 1$. In this case, given a policy r, we take the regulator's utility for an equilibrium $E \in \mathrm{BNE}(r)$ to be $U(E) = \lim_{t\to\infty} \mathbb{P}[a_t = 1]$, and we again evaluate r according to the worst-case equilibrium $U(r) = \inf_{E\in\mathrm{BNE}(r)} U(E)$. As in the previous section, we denote by U_{adap} and U_{stat} the supremum of $U(r)$ taken over adaptive and static regulation policies, respectively.

For this problem, the prior plays a more important role. In particular, for $\pi > \frac{1}{2}$ the solution is trivial. The regulator can prevent any information, and the agents herd on action $a = 1$ with probability 1. For $\pi = \frac{1}{2}$, the regulator can design an initial garbling G_1 such that after step $t = 1$ the public belief will be $p_t > \frac{1}{2}$ with probability arbitrary close to 1 and prevent any further information in latter steps, and hence again can ensure a utility that approaches 1. Conversely, for $\pi \le \mathrm{logit}(-\mathrm{LL}(\bar{p}))$, no matter what policy the regulator chooses, it is an equilibrium for all agents to take action $a = 0$, in which case the probability that the agents to herd on action $a = 1$ is 0. Hence, our main result restricts attention to priors in the interval $\pi \in (\mathrm{logit}(-\mathrm{LL}(\bar{p})), \frac{1}{2})$.

Using Lemma 4 and the martingale condition, it is easy to establish an upper bound on the utility of any policy:

$$\mathbb{P}(p_\infty \ge \frac{1}{2}) = \frac{\pi - \mathbb{E}(p_\infty \,|\, p_\infty < \frac{1}{2})}{\mathbb{E}(p_\infty \,|\, p_\infty \ge \frac{1}{2}) - \mathbb{E}(p_\infty \,|\, p_\infty < \frac{1}{2})} \le \frac{\pi - \mathrm{logit}(-\mathrm{Wid}(F))}{\frac{1}{2} - \mathrm{logit}(-\mathrm{Wid}(F))}.$$

Our result shows that for a binary-supported F, the adaptive regulator can achieve this bound but the static regulator cannot.

Theorem 5. *If F is binary-supported and* $\mathrm{logit}(-\mathrm{LL}(\bar{p})) < \pi < \frac{1}{2}$, *then*

$$U_{\mathrm{stat}} < U_{\mathrm{adap}} = \frac{\pi - \mathrm{logit}(-\mathrm{Wid}(F))}{\frac{1}{2} - \mathrm{logit}(-\mathrm{Wid}(F))}.$$

This shows that if the incentives of the regulator are different from the agents' incentives, then a gap may exist between the static and the adaptive regulator.

Remark 6. While Theorem 5 is stated for binary signals, a similar statement holds for general signals. Unlike in Theorem 5, the upper bound depends on π; however, for any distribution and prior, $U_{\mathrm{stat}} < U_{\mathrm{adap}}$.

Proof idea of Theorem 5. In Theorem 1, the goal of the regulator was to attract as much as possible the distribution p_∞ to the boundaries (0 and 1). This is formally reflected by the objective in Equation (2). Here the objective is to maximize its mass on $[\frac{1}{2}, 1]$ (namely to maximize $\mathbb{P}(p_\infty \geq \frac{1}{2})$).

To reach this objective, the regulator wants to cause p_∞ to be supported on two points: a and $\frac{1}{2}$ with a being as low as possible. Lemma 4 tells us that the lowest value for a would be $a = \operatorname{logit}(-\operatorname{Wid}(F))$. We show that it is attainable by an adaptive regulator, but not by the static regulator.

The adaptive regulator can cause $\operatorname{supp}(p_\infty) \subset [a, a+\epsilon)] \cup [\frac{1}{2}, \frac{1}{2}+\epsilon]$ by modifying the ideas of Theorem 1. For every $p_t \in (\operatorname{logit}(-\operatorname{LL}(p^*_+)), \frac{1}{2})$ we design a garbling that causes the public belief either to move towards $\operatorname{logit}(-\operatorname{LL}(p^*_+))$ but not to cross it or to jump to $\frac{1}{2}+\epsilon$. We apply these garblings sequentially as far as we move towards $\operatorname{logit}(-\operatorname{LL}(p^*_+))$. If we jump to $\frac{1}{2}+\epsilon$ we stop revealing any information. After finitely many steps towards $\operatorname{logit}(-\operatorname{LL}(p^*_+))$ we are very close to it. Then, to make a "big jump" to either $a+\epsilon$ or to $\frac{1}{2}+\epsilon$ we make a single iteration of the q^*-quantile threshold garbling.

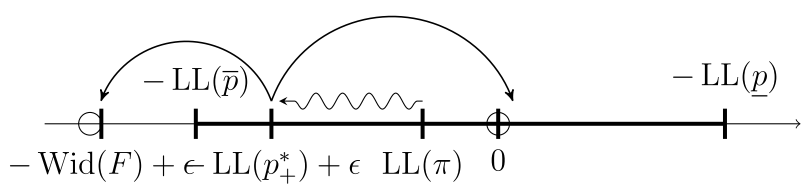

FIGURE 3. The Figure illustrates the desirable plan of the regulator in the log-likelihood space. The wavy arrow captures a slow dynamic movement of the public belief towards the point $-\operatorname{LL}(p^*_+)+\epsilon$ (or a jump to ϵ). Arc arrows capture the possible jumps of the public belief that occurs in a single period of the q^*-quanite garbling. The bold line corresponds to the interval I (i.e., the non-cascade region of the public belief). The circles indicate the two ideal points in $\operatorname{supp}(p_\infty)$ for the regulator that are reached (up to ϵ) by the plan.

The static regulator cannot follow these instructions. The commitment in advance to apply the q^*-quantile threshold garbling at some future time t can be harmful: for example, if at time t the public belief is at $[\frac{1}{2}, \frac{1}{2}+\epsilon]$, then a play of action 1 will cause the public belief being far from $\frac{1}{2}$. On the other hand, if she commits never to use the q^*-quantile threshold garbling (or a similar garbling with a similar effect), then she is unable to make the desirable big jump to $\operatorname{logit}(-\operatorname{Wid}(F))$.

4. Conclusion

The impact of regulation on society can be significant. In our work, we examine a regulator in the standard herding model that can garble agents' private information. We have considered two types of regulators: a dynamic regulator that can garble private information as a function of public information and a static regulator that has to pre-commit to its garbling strategy. Another weaker type of regulator that we didn't consider is the fixed regulator that pre-commits to a garbling strategy that is identical for all agents. In this regard, one may show that for signals that are bounded from one side (see Example 1 in Section 2.3) our main Theorem 1 can be strengthened, and there exists a (nearly) optimal fixed garbling policy.

A fundamental assumption in the standard herding model that we have used is that agents observe all past predecessors' actions when they take their actions. A natural question is whether our results are still valid under partial observability of the agents. We believe that our results carry forward to a large class of deterministic observation structures. One such class is obtained where the agents observe their k direct predecessors.

The classic result of Smith and Sorensen [13] demonstrates that when signals are unbounded, the asymptotic probability of taking the correct action approaches one with time. More recently, Rosenberg and Vieille [11] show that for a large class of unbounded signals, learning can be very slow, and the expected time for the first correct action is infinite.[5] In light of this striking negative result, a natural question to ask is whether regulation of the kind we study here can improve the speed of learning.

[5]More on the asymptotic speed of learning can be found in Hann-Caruthers, Martynov, and Tamuz [7].

References

[1] Daron Acemoglu, Munther A Dahleh, Ilan Lobel, and Asuman Ozdaglar. Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236, 2011.

[2] Itai Arieli, Ronen Gradwohl, and Rann Smorodinsky. Herd design. *American Economic Review: Insights*, 5(4):460–76, 2023.

[3] Itai Arieli, Fedor Sandomirskiy, and Rann Smorodinsky. On social networks that support learning. *Available at SSRN 3728165*, 2021.

[4] Robert J Aumann, Michael Maschler, and Richard E Stearns. *Repeated games with incomplete information.* MIT press, 1995.

[5] Abhijit V Banerjee. A simple model of herd behavior. *Quarterly Journal of Economics*, 107(3):797–817, 1992.

[6] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, 100(5):992–1026, 1992.

[7] Wade Hann-Caruthers, Vadim V Martynov, and Omer Tamuz. The speed of sequential asymptotic learning. *Journal of Economic Theory*, 173:383–409, 2018.

[8] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.

[9] Ilan Lobel and Evan Sadler. Information diffusion in networks through social learning. *Theoretical Economics*, 10(3):807–851, 2015.

[10] Caio Lorecchio. Persuading crowds. *UB Economics–Working Papers, 2022, E22/434*, 2022.

[11] Dinah Rosenberg and Nicolas Vieille. On the efficiency of social learning. *Econometrica*, 87(6):2141–2168, 2019.

[12] Daniel Sgroi. Optimizing information in the herd: Guinea pigs, profits, and welfare. *Games and Economic Behavior*, 39(1):137–166, 2002.

[13] Lones Smith and Peter Sørensen. Pathological outcomes of observational learning. *Econometrica*, 68(2):371–398, 2000.

[14] Lones Smith, Peter Norman Sørensen, and Jianrong Tian. Informational herding, optimal experimentation, and contrarianism. *The Review of Economic Studies*, 88(5):2527–2554, 2021.

Appendix A. Appendix I: Sequential learning model dynamics

A.1. **Notation.** Denote by $h_t = (a_1, \dots, a_t)$ the history of actions up to time t. Given a policy r and an equilibrium $E \in \mathrm{BNE}(r)$, we denote by $p_t := \mathbb{P}[\theta = 1 \,|\, h_t]$ the public belief at time t and by $\ell_t = \mathrm{LL}(p_t)$ the log-likelihood ratio of the public belief. Observe that

$$\begin{aligned}\ell_{t+1} &= \ell_t + \log\left(\frac{\mathbb{P}[a_{t+1} \,|\, h_t, \theta = 1]}{\mathbb{P}[a_{t+1} \,|\, h_t, \theta = 0]}\right)\\ &= \ell_t + \mathrm{LL}(\mathbb{E}[s_{t+1} \,|\, s_{t+1} \in A_{t+1}^{a_{t+1}}]),\end{aligned}$$

where A_{t+1}^a is the set of signals for $a_{t+1} = a$. Observe that if $\mathrm{LL}(s) < -\ell_t$ then $s \in A_{t+1}^0$ and $s \notin A_{t+1}^1$. Similarly, if $\mathrm{LL}(s) > -\ell_t$, then $s \notin A_{t+1}^0$ and $s \in A_{t+1}^1$.

Define

$$\begin{aligned}E_-^G(x) &:= p_-(G(\mathrm{logit}(-x)))\\ E_+^G(x) &:= p_+(G(\mathrm{logit}(-x))).\end{aligned}$$

and

$$\begin{aligned}d_-^G(x) &:= x + \mathrm{LL}(E_-^G(x))\\ d_+^G(x) &:= x + \mathrm{LL}(E_+^G(x)).\end{aligned}$$

It follows that if $a_{t+1} = 0$, then

$$\sup_{-x < -\ell_t} d_-^{G_{t+1}}(x) \le \ell_{t+1} \le d_-^{G_{t+1}}(\ell_t),$$

and if $a_{t+1} = 1$, then

$$d_+^{G_{t+1}}(\ell_t) \le \ell_{t+1} \le \inf_{-x > -\ell_t} d_+^{G_{t+1}}(x).$$

In particular, if $\mathrm{logit}(-\ell_t)$ is not an atom of G_{t+1}, then

$$a_{t+1} = 0 \implies \ell_{t+1} = \ell_t + d_-^{G_{t+1}}(\ell_t)$$

and

$$a_{t+1} = 1 \implies \ell_{t+1} = \ell_t + d_+^{G_{t+1}}(\ell_t).$$

We will denote $\underline{p} = \inf(\mathrm{supp}(F))$ and $\overline{p} = \sup(\mathrm{supp}(F))$.

A.2. **Signals that never cascade.** Throughout this subsection, we assume that $G_t = F$ for all t. Denote $\mathcal{I} = (-\operatorname{LL}(\overline{p}), -\operatorname{LL}(\underline{p}))$. We say that F *never cascades* if $d_-^F(\mathcal{I}), d_+^F(\mathcal{I}) \subseteq \mathcal{I}$. Observe that if F is nonatomic and $G_t = F$ for all t, then in any equilibrium, $\ell_t \in \mathcal{I}$ for all t; in particular, for any history h_{t-1}, $0 < \mathbb{P}[a_t = 1 \mid h_{t-1}] < 1$. As the following result demonstrates, for any interval I containing $\frac{1}{2}$, there is a distribution that never cascades whose support is I.

Theorem 7. *For any closed interval $I \subseteq [0,1]$ with $\inf(I) < \frac{1}{2} < \sup(I)$, there is a nonatomic distribution F that never cascades with $\operatorname{supp}(F) = I$.*

If F never cascades, then a standard argument shows that in any equilibrium, $\operatorname{supp}(\ell_\infty) = \{-\operatorname{LL}(\overline{p}), -\operatorname{LL}(\underline{p})\}$. The following result shows that the time required for ℓ_t to be close in distribution to ℓ_∞ can be uniformly bounded. For $\varepsilon > 0$, define

$$\mathcal{B}_\varepsilon = \{x \in \mathcal{I} : |-\operatorname{LL}(\overline{p}) - x| < \varepsilon \text{ or } |-\operatorname{LL}(\underline{p}) - x| < \varepsilon\}.$$

Proposition 8. *If F never cascades, then for any $\varepsilon > 0$, there is a T such that in any equilibrium, if $\ell_t \in [-\varepsilon, \varepsilon]$ for some history h_t,*

$$\mathbb{P}[\ell_{t+T} \in \mathcal{B}_\varepsilon \mid h_t] \geq 1 - \varepsilon.$$

For the proofs of Theorem 7 and Proposition 8, see Appendix III.

Appendix B. Appendix II: Proofs of main results

Proof of Theorem 1. By Lemma 4, $W_{\text{stat}} \leq \operatorname{logit}(\operatorname{Wid}(F))$. Hence it is sufficient to prove that $W_{\text{stat}} \geq \operatorname{logit}(\operatorname{Wid}(F))$. Moreover, if F is not binary, then

$$\operatorname{logit}(\operatorname{Wid}(F)) = \operatorname{logit}(\sup_{F' \in \mathcal{G}_{\text{bin}}}(\operatorname{Wid}(F'))) = \sup_{F' \in \mathcal{G}_{\text{bin}}}(\operatorname{logit}(\operatorname{Wid}(F'))),$$

where $\mathcal{G}_{\text{bin}}$ is the set of binary $F' \preceq F$. Since $W_{\text{stat}}(F) \geq W_{\text{stat}}(F')$ whenever $F' \preceq F$, it follows that it is sufficient to prove the claim when F is binary.

So assume F is binary. Define

$$\rho := e^{\min(\operatorname{LL}(\underline{p}), -\operatorname{LL}(\overline{p}))} - e^{-\operatorname{Wid}(F)}.$$

Observe that

$$\begin{aligned}
\mathbb{P}[s_t = \overline{p} \,|\, \theta = 0] &= e^{-\mathrm{LL}(\overline{p})} \cdot \mathbb{P}[s_t = \overline{p} \,|\, \theta = 1] \\
&= e^{-\mathrm{LL}(\overline{p})} \cdot (1 - \mathbb{P}[s_t = \underline{p} \,|\, \theta = 1]) \\
&= e^{-\mathrm{LL}(\overline{p})} \cdot (1 - e^{\mathrm{LL}(\underline{p})} \cdot \mathbb{P}[s_t = \underline{p} \,|\, \theta = 0]) \\
&\geq e^{-\mathrm{LL}(\overline{p})} \cdot (1 - e^{\mathrm{LL}(\underline{p})}) \\
&\geq \rho.
\end{aligned}$$

A similar calculation shows that $\mathbb{P}[s_t = \underline{p} \,|\, \theta = 1] \geq \rho$. Hence, for $s \in \{\underline{p}, \overline{p}\}$ and $t \in \{0, 1\}$, $\mathbb{P}[s_t = s \,|\, \theta = t] \geq \rho$.

Fix $\eta > 0$, and choose N and $\varepsilon < \eta$ such that $(1-\rho)^N < \frac{\eta}{2}$ and $(1-\varepsilon)(1-N\varepsilon) > 1 - \frac{\eta}{2}$.

By Theorem 7 and Proposition 8, there is a $G \preceq F$ and a T such that if $G_{nT+1} = \cdots = G_{(n+1)T-1} = G$ then $\ell_{nT} \in \mathcal{I} \implies \ell_{(n+1)T-1} \in \mathcal{I}$ and

$$\mathbb{P}[\ell_{(n+1)T-1} \in \mathcal{B}_\varepsilon \,|\, \ell_{nT} \in [-\varepsilon, \varepsilon]] \geq 1 - \varepsilon.$$

Let r be the policy where $G_{nT} = F$ and $G_{nT+k} = G$ for $n \geq 1$ and $1 \leq k \leq T-1$.

Denote

$$\begin{aligned}
X_n &:= \{|\ell_{nT}| < \varepsilon\} \\
Y_n &:= \{\varepsilon < |\ell_{nT}| < \mathrm{Wid}(F) - \varepsilon\} \\
Z_n &:= \{|\ell_{nT}| \geq \mathrm{Wid}(F) - \varepsilon\}.
\end{aligned}$$

Observe that since $X_n \implies \ell_{n(T+1)-1} \in \mathcal{I}$,

$$\mathbb{P}[Y_{n+1} \,|\, X_n] \leq \mathbb{P}[\ell_{n(T+1)-1} \in \mathcal{I} \setminus \mathcal{B}_\varepsilon \,|\, X_n] = 1 - \mathbb{P}[\ell_{n(T+1)-1} \in \mathcal{B}_\varepsilon \,|\, X_n] \leq \varepsilon,$$

so

$$\mathbb{P}[Y_{n+1}] \leq \mathbb{P}[Y_n] + \varepsilon \cdot \mathbb{P}[X_n] \leq \mathbb{P}[Y_n] + \varepsilon,$$

and it follows that for all n, $\mathbb{P}[Y_n] \leq n\varepsilon$. Thus,

$$\begin{aligned}
\mathbb{P}[Z_{n+1}] &\geq \mathbb{P}[Z_n] + \mathbb{P}[Z_{n+1} \,|\, X_n] \cdot \mathbb{P}[X_n] \\
&\geq \mathbb{P}[Z_n] + (1-\varepsilon)\rho(1 - n\varepsilon - \mathbb{P}[Z_n]) \\
&= (1 - (1-\varepsilon)\rho)\mathbb{P}[Z_n] + \rho(1-\varepsilon)(1-n\varepsilon) \\
&\geq (1-\rho)\mathbb{P}[Z_n] + \rho(1 - \frac{\eta}{2}).
\end{aligned}$$

Moreover, if $\mathbb{P}[Z_n] \geq 1 - \frac{\eta}{2} - (1-\rho)^n$, then

$$\begin{aligned}\mathbb{P}[Z_{n+1}] &\geq (1-\rho)[1 - \frac{\eta}{2} - (1-\rho)^n] + \rho(1 - \frac{\eta}{2}) \\ &= 1 - \frac{\eta}{2} - (1-\rho)^{n+1}.\end{aligned}$$

Since $\mathbb{P}[Z_0] = 0 \geq 1 - \frac{\eta}{2} - (1-\rho)^0$, it follows by induction that

$$\mathbb{P}[Z_N] \geq 1 - \frac{\eta}{2} - (1-\rho)^N \geq 1 - \eta.$$

Hence,

$$\begin{aligned}\mathbb{P}[|\ell_\infty| \geq \mathrm{Wid}(F) - \eta] &\geq \mathbb{P}[|\ell_\infty| \geq \mathrm{Wid}(F) - \varepsilon] \\ &\geq \mathbb{P}[|\ell_{NT}| \geq \mathrm{Wid}(F) - \varepsilon] \\ &\geq 1 - \eta.\end{aligned}$$

It follows that $W_{\text{stat}} \geq (1-\eta)\,\mathrm{logit}(\mathrm{Wid}(F) - \eta)$. Since this holds for all $\eta > 0$, $W_{\text{stat}} \geq \mathrm{logit}(\mathrm{Wid}(F))$.

□

B.1. **Proof of Theorem 5.** We will prove Theorem 5 in two parts. Denote

$$\bar{U} = \frac{\pi - \mathrm{logit}(-\mathrm{Wid}(F))}{\frac{1}{2} - \mathrm{logit}(-\mathrm{Wid}(F))}.$$

As shown above,

$$U_{\text{stat}} \leq U_{\text{adap}} \leq \bar{U}.$$

We begin by showing that for the adaptive regulator, this upper bound is achieved.

Proposition 9. *If F is binary-supported and* $\mathrm{logit}(-\mathrm{LL}(\bar{p})) < \pi < \frac{1}{2}$, *then*

$$U_{\text{adap}} \geq \bar{U}.$$

Proof. Fix $\varepsilon > 0$, and define the policy r_ε as follows. If ℓ_t is above 0 or below $-\mathrm{LL}(a)$, send an uninformative signal. If $-\mathrm{LL}(\overline{p}) + \varepsilon < \ell_t \leq 0$, send the signal G with $\mathrm{supp}(G) = \{\mathrm{logit}(-\varepsilon), \mathrm{logit}(-\ell_t + \varepsilon)\}$. Otherwise, send the signal F.

First, note that for all sufficiently small ε, this policy is well-defined, since $\underline{p} \leq \mathrm{logit}(-\varepsilon) < \frac{1}{2}$, and for $-\mathrm{LL}(\overline{p}) + \varepsilon < \ell_t \leq 0$, $\frac{1}{2} < \mathrm{logit}(-\ell_t + \varepsilon) \leq \overline{p}$, so $G \preceq F$.

Now, observe that under r_ε, p_∞ is supported in $[0, \mathrm{logit}(-\mathrm{Wid}(F)+\varepsilon)] \cup (\frac{1}{2}, \varepsilon]$. It follows from the martingale condition that

$$\pi = \mathbb{E}[p_\infty] \leq \mathbb{P}[p_\infty < \frac{1}{2}] \cdot \mathrm{logit}(-\mathrm{Wid}(F)+\varepsilon) + \mathbb{P}[p_\infty > \frac{1}{2}] \cdot \mathrm{logit}(\varepsilon)$$

so

$$U(r_\varepsilon) = \mathbb{P}[p_\infty > \frac{1}{2}] \geq \frac{\pi - \mathrm{logit}(-\mathrm{Wid}(F)+\varepsilon)}{\mathrm{logit}(\varepsilon) - \mathrm{logit}(-\mathrm{Wid}(F)+\varepsilon)}.$$

Hence,

$$U_{\mathrm{adap}}(\pi) \geq \sup_\varepsilon U(r_\varepsilon) \geq \bar{U}.$$

□

We next show that the first inequality above is strict; that is, the adaptive regulator can always do strictly better than the static regulator.

Proposition 10. *If F is binary-supported and* $\mathrm{logit}(-\mathrm{LL}(\bar{p})) < \pi < \frac{1}{2}$, *then*

$$U_{\mathrm{stat}} < \bar{U}.$$

The proof of Proposition 10 is somewhat more involved than the proof of Proposition 9. For ease of exposition, we split the argument into several Lemmas (see proofs below).

Proof of Proposition 10. Fix a policy r. First, in order for $U(r)$ to be close $\bar{U}$, most of the mass of p_∞ must be concentrated near $\mathrm{logit}(-\mathrm{Wid}(F))$ and $\frac{1}{2}$. Denote

$$\begin{aligned}
\delta_0 &= \mathbb{E}[p_\infty \,|\, p_\infty < \frac{1}{2}] - \mathrm{logit}(-\mathrm{Wid}(F)) \\
\delta_1 &= \mathbb{E}[p_\infty \,|\, p_\infty \geq \frac{1}{2}] - \frac{1}{2} \\
C &= \min\left(\frac{1}{2} - \pi, \pi - \mathrm{logit}(-\mathrm{Wid}(F))\right).
\end{aligned}$$

Lemma 11.

$$U(r) \leq \bar{U} - C \cdot \max(\delta_0, \delta_1).$$

So it is sufficient to show that δ_0 and δ_1 cannot both be close to 0. As the following Lemma demonstrates, in order for δ_0 to be close to 0, there must be some time t where G_t is close to F. Say that $G \preceq F$ is an ε-approximation of F if $|G_t(\underline{p}+\varepsilon) - F(\underline{p})| < \sqrt{\varepsilon}$ and $G(\overline{p}-\varepsilon) - G(\underline{p}+\varepsilon) < \sqrt{\varepsilon}$.

Lemma 12. *For every $\varepsilon > 0$ sufficiently small, there is a $\delta > 0$ such that if $\delta_0 < \delta$ then G_t is an ε-approximation of F for some t.*

On the other hand, observe that

$$\delta_1 = \mathbb{E}[p_\infty - \frac{1}{2} \,|\, p_\infty \geq \frac{1}{2}] \geq \mathbb{E}[f(p_\infty)],$$

where $f(x) = \mathbb{1}(x \geq \frac{1}{2}) \cdot (x - \frac{1}{2})$. Since f is convex and p_t is a martingale, $\mathbb{E}[f(p_t)]$ is increasing in t. In particular,

$$\delta_1 \geq \sup_t \mathbb{E}[f(p_t)].$$

Now, for any t, π', and $\kappa > \frac{1}{2}$,

$$\mathbb{E}[f(p_{t+1})] \geq (\kappa - \frac{1}{2}) \cdot \mathbb{P}[p_{t+1} \geq \kappa, p_t \geq \pi', a_{t+1} = 1].$$

Hence, we have the following.

Lemma 13. *Suppose for that for some $\varepsilon > 0$, $\kappa > \frac{1}{2}$, and π' there exists a $D > 0$ such that if G_{t+1} is an ε-approximation of F then*

$$\mathbb{P}[p_{t+1} \geq \kappa, p_t \geq \pi', a_{t+1} = 1] \geq D.$$

Then there is a $\underline{\delta} > 0$ such that $\max(\delta_0, \delta_1) \geq \underline{\delta}$.

Let

$$\kappa = \operatorname{logit}(\frac{1}{2} \min(\mathrm{LL}(\pi') + \mathrm{LL}(\overline{p}), -\mathrm{LL}(\underline{p})))$$

and

$$\pi' = \frac{\underline{p} + \pi}{2}.$$

Observe that if G_{t+1} is an ε-approximation of F, then

$$\begin{aligned} |1 - G_{t+1}(\overline{p} - \varepsilon) - (1 - F(\underline{p}))| &\leq |G_{t+1}(\overline{p} - \varepsilon) - G_{t+1}(\underline{p} + \varepsilon)| \\ &\quad + |G_{t+1}(\underline{p} + \varepsilon) - F(\underline{p})| \\ &< 2\sqrt{\varepsilon}, \end{aligned}$$

so

$$\mathbb{P}[s_{t+1} > \overline{p} - \varepsilon] = 1 - G_{t+1}(\overline{p} - \varepsilon) > 1 - F(\underline{p}) - 2\sqrt{\varepsilon}.$$

Now, if $a_{t+1} = 1$, then

$$\ell_{t+1} \geq d_+^{G_{t+1}}(\ell_t) = \ell_t + \mathrm{LL}(\mathbb{E}[s_{t+1} \,|\, s_{t+1} \geq \operatorname{logit}(-\ell_t)]).$$

If $p_t \geq \kappa$ and $a_{t+1} = 1$, then $p_{t+1} \geq p_t \geq \kappa$. For ε sufficiently small, if $\pi' \leq p_t \leq \kappa$, then

$$\begin{aligned}\mathbb{E}[s_{t+1} \,|\, s_{t+1} \geq \operatorname{logit}(-\ell_t)] &\geq \mathbb{E}[s_{t+1} \,|\, s_{t+1} \geq \underline{p} + \varepsilon] \\ &\geq \frac{1 - F(\underline{p}) - 2\sqrt{\varepsilon}}{1 - F(\underline{p}) - \sqrt{\varepsilon}} \cdot (\overline{p} - \varepsilon) \\ &\geq \operatorname{logit}(\frac{1}{2}(-\operatorname{LL}(\pi') + \operatorname{LL}(\overline{p}))),\end{aligned}$$

so

$$\ell_{t+1} \geq \operatorname{LL}(\pi') + \frac{1}{2}(-\operatorname{LL}(\pi') + \operatorname{LL}(\overline{p})) \geq \operatorname{LL}(\kappa).$$

Thus, for ε sufficiently small, $p_t \geq \pi'$ and $a_{t+1} = 1$ implies $p_{t+1} \geq \kappa$, so

$$\begin{aligned}\mathbb{P}[p_{t+1} \geq \kappa, p_t \geq \pi', a_{t+1} = 1] &= \mathbb{P}[p_t \geq \pi', a_{t+1} = 1] \\ &\geq \mathbb{P}[s_{t+1} > \overline{p} - \varepsilon \,|\, p_t \geq \pi'] \cdot \mathbb{P}[p_t \geq \pi'].\end{aligned}$$

Now,

$$\begin{aligned}\mathbb{P}[s_{t+1} > \overline{p} - \varepsilon \,|\, p_t \geq \pi'] &\geq \mathbb{P}[s_{t+1} > \overline{p} - \varepsilon \,|\, \theta = 0] \\ &\geq 2(1 - \overline{p})\mathbb{P}[s_{t+1} > \overline{p} - \varepsilon] \\ &> 2(1 - \overline{p})(1 - F(\underline{p}) - 2\sqrt{\varepsilon})\end{aligned}$$

and by the martingale condition

$$\begin{aligned}\mathbb{P}[p_t \geq \pi'] &= \frac{\pi - \mathbb{E}[p_t \,|\, p_t < \pi']}{\mathbb{E}[p_t \,|\, p_t \geq \pi'] - \mathbb{E}[p_t \,|\, p_t < \pi']} \\ &\geq \frac{\pi - \pi'}{\operatorname{logit}(\operatorname{Wid}(F)) - \pi'} \\ &\geq \pi - \pi' \\ &= \frac{1}{2}(\pi - \underline{p}).\end{aligned}$$

Hence, for ε sufficiently small,

$$\mathbb{P}[p_{t+1} \geq \kappa, p_t \geq \pi', a_{t+1} = 1] \geq \frac{1}{2}(1 - \overline{p})(1 - F(\underline{p}))(\pi - \underline{p}).$$

$\square$

Proof of Lemma 11. Observe that

$$U(r) \leq \mathbb{P}[p_\infty \geq \frac{1}{2}] = \frac{\pi - (\delta_0 + \operatorname{logit}(-\operatorname{Wid}(F))}{(\delta_1 + \frac{1}{2}) - (\delta_0 + \operatorname{logit}(-\operatorname{Wid}(F))}$$

so

$$\begin{aligned}
\bar{U} - U(r) &\geq \frac{\pi - \operatorname{logit}(-\operatorname{Wid}(F))}{\frac{1}{2} - \operatorname{logit}(-\operatorname{Wid}(F))} - \frac{\pi - (\delta_0 + \operatorname{logit}(-\operatorname{Wid}(F))}{(\delta_1 + \frac{1}{2}) - (\delta_0 + \operatorname{logit}(-\operatorname{Wid}(F))} \\
&\geq (\pi - \operatorname{logit}(-\operatorname{Wid}(F)) \cdot (\frac{1}{2} - \operatorname{logit}(-\operatorname{Wid}(F)) + \delta_1 - \delta_0) \\
&- (\pi - \operatorname{logit}(-\operatorname{Wid}(F)) - \delta_0) \cdot (\frac{1}{2} - \operatorname{logit}(-\operatorname{Wid}(F))) \\
&= (\pi - \operatorname{logit}(-\operatorname{Wid}(F)) \cdot (\delta_1 - \delta_0) + \delta_0 \cdot (\frac{1}{2} - \operatorname{logit}(-\operatorname{Wid}(F))) \\
&= (\pi - \operatorname{logit}(-\operatorname{Wid}(F)) \cdot \delta_1 + (\frac{1}{2} - \pi) \cdot \delta_0 \\
&\geq C \cdot (\delta_1 + \delta_0) \\
&\geq C \cdot \max(\delta_0, \delta_1).
\end{aligned}$$

□

To prove Lemma 12, we will use the following intermediary Lemma.

Lemma 14. *For all sufficiently small $\eta > 0$, if $\mathbb{E}[s_{t+1} \,|\, s_{t+1} \leq \overline{p} - \eta] < \underline{p} + \eta$, then G_{t+1} is an ε-approximation of F.*

Proof. Let $\alpha = G_{t+1}(\underline{p} + \eta)$ and $\beta = G_{t+1}(\overline{p} - \eta)$. By Markov's inequality,

$$\begin{aligned}
\mathbb{P}[s_{t+1} > \underline{p} + \sqrt{\eta} \,|\, s_{t+1} \leq \overline{p} - \eta] &= \mathbb{P}[s_{t+1} - \underline{p} > \sqrt{\eta} \,|\, s_{t+1} \leq \overline{p} - \eta] \\
&\leq \frac{\mathbb{E}[s_{t+1} - \underline{p} \,|\, s_{t+1} \leq \overline{p} - \eta]}{\sqrt{\eta}} \\
&< \sqrt{\eta},
\end{aligned}$$

so $\beta - \alpha < \sqrt{\eta}$. Hence,

$$\begin{aligned}
\mathbb{E}[s_{t+1}] &\geq \underline{p} \cdot \alpha + (\overline{p} - \eta) \cdot (1 - \beta) \\
&= [\underline{p} \cdot \alpha + \overline{p} \cdot (1 - \alpha)] + \overline{p} \cdot (\alpha - \beta) - \eta \cdot (1 - \beta) \\
&\geq [\underline{p} \cdot \alpha + \overline{p} \cdot (1 - \alpha)] - 2\sqrt{\eta}
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}[s_{t+1}] &\leq (\underline{p} + \eta) \cdot \alpha + (\overline{p} - \eta) \cdot (\beta - \alpha) + \overline{p} \cdot (1 - \beta) \\
&= [\underline{p} \cdot \alpha + \overline{p} \cdot (1 - \alpha)] + \eta \cdot \alpha - \eta \cdot (\beta - \alpha) \\
&\leq [\underline{p} \cdot \alpha + \overline{p} \cdot (1 - \alpha)] + 2\sqrt{\eta},
\end{aligned}$$

and since $\mathbb{E}[s_{t+1}] = \frac{1}{2}$, it follows that

$$|\underline{p} \cdot \alpha + \overline{p} \cdot (1 - \alpha) - \frac{1}{2}| \leq 2\sqrt{\eta}.$$

Moreover, $\frac{1}{2} = \underline{p} \cdot F(\underline{p}) + \overline{p} \cdot (1 - F(\underline{p}))$, so

$$|(\overline{p} - \underline{p}) \cdot (\alpha - F(\underline{p}))| \leq 2\sqrt{\eta},$$

and hence $|\alpha - F(\underline{p})| \leq \frac{2}{\overline{p}-\underline{p}} \cdot \sqrt{\eta}$.

Now, observe that $G_{t+1}(\overline{p} - \varepsilon) - G_{t+1}(\underline{p} + \varepsilon) \leq \beta - \alpha < \sqrt{\eta}$ and

$$\alpha \leq G_{t+1}(\underline{p} + \varepsilon) \leq \beta$$

so

$$\alpha - F(\underline{p}) \leq G_{t+1}(\underline{p} + \varepsilon) - F(\underline{p}) \leq \beta - F(\underline{p}).$$

Since $|\alpha - F(\underline{p})| \leq \frac{2}{\overline{p}-\underline{p}} \cdot \sqrt{\eta}$ and $|\beta - F(\underline{p})| \leq |\alpha - F(\underline{p})| + |\beta - \alpha| \leq \left(\frac{2}{\overline{p}-\underline{p}} + 1\right) \cdot \sqrt{\eta}$, it follows that for η sufficiently small, $G_{t+1}(\overline{p} - \varepsilon) - G_{t+1}(\underline{p} + \varepsilon) < \varepsilon$ and $|G_{t+1}(\underline{p} + \varepsilon) - F(\underline{p})| < \varepsilon$. □

Proof of Lemma 12. Suppose $x \geq -\operatorname{LL}(\overline{p})$, $y \leq -\operatorname{Wid}(F) + \gamma$, and for some t, $\mathbb{P}[\ell_t = x] > 0$ and $\mathbb{P}[\ell_{t+1} = y \mid \ell_t = x] > 0$. Let $\Delta\ell = y - x$. Observe that $\Delta\ell \leq \operatorname{LL}(up) + \gamma$, and since $\Delta\ell \geq \operatorname{LL}(\underline{p})$, $x \leq -\operatorname{LL}(\overline{p}) + \gamma$. Hence,

$$\mathbb{E}[s_{t+1} \mid s_{t+1} \leq -\operatorname{LL}(x)] = \operatorname{logit}(\Delta\ell) \leq \operatorname{logit}(\operatorname{LL}(\underline{p}) + \gamma).$$

Fix $\eta > 0$. For γ sufficiently small, it follows that

$$\mathbb{E}[s_{t+1} \mid s_{t+1} \leq \overline{p} - \eta] \leq \operatorname{logit}(\operatorname{LL}(\underline{p}) + \gamma) < \underline{p} + \eta.$$

Now, observe that if δ_0 is sufficiently small, then there must be a t such that $\mathbb{P}[\ell_t \leq -\operatorname{Wid}(F) + \gamma] = 0$ and $\mathbb{P}[\ell_{t+1} \leq -\operatorname{Wid}(F) + \gamma] > 0$. It follows that for any $\eta > 0$, if δ_0 is sufficiently small, then for some t,

$$\mathbb{E}[s_{t+1} \mid s_{t+1} \leq \overline{p} - \eta] < \eta,$$

and the result then follows. □

Appendix C. Appendix III: Proofs for signals that never cascade

Proof of Theorem 7. Let $\underline{p} = \inf(I)$ and $\overline{p} = \sup(I)$, and define

$$F_n(t) := \begin{cases} 0, & t \leq \underline{p} \\ c_n^-(t - \underline{p})^n, & \underline{p} < t \leq \frac{1}{2} \\ 1 - c_n^+(\overline{p} - t)^n, & \frac{1}{2} < t < \overline{p} \\ 1, & t \geq \overline{p} \end{cases}$$

where $c_n^- = \frac{\overline{p}-\frac{1}{2}}{\overline{p}-\underline{p}} \cdot \frac{1}{\left(\frac{1}{2}-\underline{p}\right)^n}$ and $c_n^+ = \frac{\frac{1}{2}-\underline{p}}{\overline{p}-\underline{p}} \cdot \frac{1}{\left(\overline{p}-\frac{1}{2}\right)^n}$.

A straightforward calculation shows that $\mathbb{E}[F_n] = \frac{1}{2}$. We will show that $d_+^{F_n}(\mathcal{I}) \subseteq \mathcal{I}$ for all sufficiently large n. A similar argument shows that $d_-^{F_n}(\mathcal{I}) \subseteq \mathcal{I}$ for all sufficiently large n, and the claim then follows.

Let $x \in \mathcal{I}$. If $x < 0$,

$$\begin{aligned} E_+^{F_n}(x) &= \mathbb{E}_{X\sim F_n}[X \mid X \geq \operatorname{logit}(-x)] \\ &= \frac{n \operatorname{logit}(-x) + \overline{p}}{n+1} \end{aligned}$$

so

$$\begin{aligned} d_+^{F_n}(x) &= x + \log\left(\frac{n \operatorname{logit}(-x) + \overline{p}}{n(1-\operatorname{logit}(-x)) + (1-\overline{p})}\right) \\ &= x + \mathrm{LL}(\operatorname{logit}(-x)) + \log\left(\frac{1 + \frac{\overline{p}}{n \operatorname{logit}(-x)}}{1 + \frac{1-\overline{p}}{n(1-\operatorname{logit}(-x))}}\right) \\ &\leq \log\left(1 + \frac{\overline{p}}{n \operatorname{logit}(-x)}\right) \\ &\leq \frac{2\overline{p}}{n} \end{aligned}$$

where the last inequality follows from the fact that $\log(1+y) \leq y$ and $\operatorname{logit}(-x) \geq \frac{1}{2}$. Hence, for $n > \frac{2\overline{p}}{-\mathrm{LL}(\underline{p})}$,

$$-\mathrm{LL}(\overline{p}) < x < d_+(x) < -\mathrm{LL}(\underline{p}),$$

so $d_+^{F_n}(x) \in \mathcal{I}$.

If $x > 0$,

$$\begin{aligned} E_+^{F_n}(x) &= \mathbb{E}_{X\sim F_n}[X \mid X \geq \operatorname{logit}(-x)] \\ &= \frac{\mathbb{E}[F_n] - F_n(\operatorname{logit}(-x)) \cdot \mathbb{E}_{X\sim F_n}[X \mid X < \operatorname{logit}(-x)]}{1 - F_n(\operatorname{logit}(-x))} \\ &= \frac{1}{2} + \left(\frac{1}{2} - E_-^{F_n}(x)\right) \cdot \frac{F_n(\operatorname{logit}(-x))}{1 - F_n(\operatorname{logit}(-x))}. \end{aligned}$$

Observe that $F_n(\operatorname{logit}(-x))$ is decreasing in n and $E_-^{F_n}(x)$ is increasing in n, so $E_+^{F_n}(x)$ and hence $d_+^{F_n}(x)$ is decreasing in n. Moreover, since $F_n(\operatorname{logit}(-x)) \to 0$ as $n \to \infty$, it follows that $E_+^{F_n}(x) \to \frac{1}{2}$ as $n \to \infty$, and thus

$$\lim_{n\to\infty} d_+^{F_n}(x) = x.$$

Define

$$A_n = \{x \in [0, -\operatorname{LL}(\underline{p}) \,:\, d_+^{F_n}(x) < -\operatorname{LL}(\underline{p})\}.$$

Then A_n, is relatively open, $A_n \subset A_m$ whenever $m > n$ and $\cup A_n = [0, -\operatorname{LL}(r_-))$.

Now,

$$(d_+^{F_n})'(x) = 1 + \frac{(E_+^{F_n})'(x)}{E_+^{F_n}(x)(1 - E_+^{F_n}(x))}.$$

Since

$$(E_+^{F_n})'(x) = \left(\frac{\frac{1}{2} - E_-^{F_n}(x)}{1 - F_n(\operatorname{logit}(-x))}\right)' \cdot F_n(\operatorname{logit}(-x)) + \left(\frac{\frac{1}{2} - E_-^{F_n}(x)}{1 - F_n(\operatorname{logit}(-x))}\right) \cdot F_n(\operatorname{logit}(-x))',$$

$(E_+^{F_n})'(-\operatorname{LL}(\underline{p})) = 0$ for $n \geq 2$, so $(d_+^{F_n})'(-\operatorname{LL}(\underline{p})) = 1$ for $n \geq 2$, and by continuity, it follows that for all x sufficiently close to $-\operatorname{LL}(\underline{p})$, $(d_+^{F_n})'(x) > 0$. Moreover, $\lim_{x \to -\operatorname{LL}(\underline{p})} d_+^{F_n}(x) = -\operatorname{LL}(\underline{p})$. Hence, $d_+^{F_n}(x) < -\operatorname{LL}(\underline{p})$ for all x sufficiently close to $-\operatorname{LL}(\underline{p})$.

It follows that there is an $\varepsilon > 0$ such that $(-\operatorname{LL}(\underline{p}) - \varepsilon, -\operatorname{LL}(\underline{p})) \subseteq A_2 \subset A_n$ for all $n \geq 2$. Moreover, by compactness, $[0, -\operatorname{LL}(\underline{p}) - \varepsilon] \subseteq A_n$ for all n sufficiently large. Thus, $[0, -\operatorname{LL}(r_-)) \subseteq A_n$ for all n sufficiently large, and hence $d_+^{F_n}(\mathcal{I}) \subseteq \mathcal{I}$ for all n sufficiently large. $\square$

Proof of Proposition 8. We show equivalently there is a T such that if the prior π is such that $\operatorname{LL}(\pi) \in [-\varepsilon, \varepsilon]$, then in any equilibrium

$$\mathbb{P}[\ell_T \in \mathcal{B}_\varepsilon] \geq 1 - \varepsilon.$$

To begin, observe that since F is nonatomic, the distribution of p_t is the same in any equilibrium. Let

$$g(z) = \min\left(|\operatorname{logit}(-\operatorname{LL}(\underline{p})) - z|, |\operatorname{logit}(-\operatorname{LL}(\overline{p})) - z|\right).$$

Observe that since g is concave and p_t is a martingale, $\mathbb{E}[g(p_t)]$ is decreasing, and since F never cascades, $g(p_t) \to 0$ almost surely, so $\lim_{t\to\infty} \mathbb{E}[g(p_t)] = 0$.

Let $\eta \in (0, \varepsilon)$ such that for $x \in \mathcal{I}$, $g(\operatorname{logit}(x)) < \eta$ implies $x \in \mathcal{B}_\varepsilon$. Observe that if $\mathbb{E}[g(p_t)] < \eta^2$, then by Markov's inequality,

$$\mathbb{P}[g(p_t) \geq \eta] \leq \frac{\eta^2}{\eta} = \eta < \varepsilon,$$

so

$$\mathbb{P}[\ell_t \in \mathcal{B}_\varepsilon] \geq \mathbb{P}[g(p_t) < \eta] = 1 - \mathbb{P}[g(p_t) \geq \eta] > 1 - \varepsilon.$$

Now, observe that $\mathbb{E}[g(p_t)]$ is continuous in π, so for every π there is a $T(\pi)$ such that $\mathbb{E}[g(p_t)] < \eta^2$ for all $t \geq T(\pi)$ and π' in a neighborhood of π. Since $[-\varepsilon, \varepsilon]$ is compact, the result then follows. $\square$