Informational Herding and Optimal Experimentation[*]

Lones Smith[†]
Department of Economics
University of Michigan

Peter Norman Sørensen[‡]
Department of Economics
University of Copenhagen

April 16, 2008

Abstract

When privately informed individuals sequentially solve identical finite action choice problems seeing prior choices, herding eventually arises: everyone chooses the same action, ignoring future gains from more information revelation. This paper analyzes the team equilibrium that internalizes this gain using insights from Bayesian experimentation. This equilibrium can be implemented by selfish individuals with a simple rule, and it entails *contrarian* behaviour: (*i*) While informational herding is still constrained efficient, it requires more extreme beliefs. (*ii*) A new log-concavity assumption on signals both precludes cascades, and ensures that individuals should lean more against their myopic preference for actions that become more popular.

[*]We thank the editor and referees, as well as Abhijit Banerjee, Patrick Bolton, Katya Malinova, Meg Meyer, Chris Wallace, and seminar participants at the MIT theory lunch, the Stockholm School of Economics, the Stony Brook Summer Festival on Game Theory (1996), Copenhagen University, the 1997 European Winter Meeting of the Econometric Society (Lisbon), Bristol University, Erasmus University, and the 2005 Workshop on Informational Herding (Copenhagen) for comments on various versions. Smith gratefully acknowledges financial support for this work from NSF grants SBR-9422988 and SBR-9711885, and Sørensen equally thanks the Danish Social Sciences Research Council.

[†]e-mail address: `lones@umich.edu`

[‡]e-mail address: `peter.sorensen@econ.ku.dk`

1 INTRODUCTION

Informational herding has been a subject of much interest for the last fifteen years. The context is seductively simple: An infinite sequence of individuals must decide on an action choice from a finite menu. Everyone has identical preferences and actions, and is endowed private signal about the state of the world. He can additionally condition his choice on see his predecessor's decisions, but not their private signals.

In this setting, Bikhchandani, Hirshleifer, and Welch (1992) and Banerjee (1992) showed that a herd arises — eventually, all decision-makers ($\mathcal{DM}$s) make the same choice, possibly unwise. Smith and Sørensen (2000) identified the signal distributions for which a misguided choice can possibly be made. The belief guided by history alone converges in the limit to a cascade set, and belief convergence implies action convergence by contradiction: Each non-conformist action gives a large belief change. We also found that the cascade set has non-empty interior precisely when the private signals are of uniformly bounded precision. This is equivalent to a positive chance of a herd on an incorrect action.

Here, we undertake a welfare analysis of this informational herding model, formally addressing the intuitive notion that the model's outcome is inefficient. Observational learning involves an informational externality, because everyone's action partially conveys his hidden private signal. The $\mathcal{DM}$s jointly possess enough information to perfectly reveal the true state of the world, yet their sequential individually rational actions do not take into account the value of information to successors. We contrast behaviour in the selfish *herding equilibrium,* and the *team equilibrium* (Radner 1962), where everyone maximizes the sum of discounted expected utilities. In this way, we are able to characterize constrained efficient forward-looking behaviour, shedding light on the informational herding externality.

The constrained efficient team equilibrium can be usefully interpreted as the choice of a social planner who maximizes the discounted sum of expected utilities. The planner sees realized actions but not private signals. He can dictate to each $\mathcal{DM}$ how to translate private signals into actions, and does this to efficiently trade off the direct action-related benefits to the present $\mathcal{DM}$ against the indirect observational learning benefits to later $\mathcal{DM}$s. We show that the planner's optimum can be implemented via a simple mechanism that rewards each $\mathcal{DM}$ solely using his own and his immediate successor's actions.

Our main findings establish the principle that it is constrained efficient for $\mathcal{DM}$s to behave in a *contrarian* fashion. They should skew their choices towards the less popular actions, so that their actions will better reflect their private information. We have found two different manifestations of this principle, one in the short run and one in the long run.

First, the constrained efficient team equilibrium features contrarianism at the margin in the short run: The more the weight of history favours taking a particular action, the more one should lean against taking it. This obtains under a plausible new informational assumption, that the unconditional distribution of the signal log-likelihood ratio has a log-concave density. We think this is the first use of log-concavity in the information literature. As a bonus, we show that this condition precludes jumping into cascades after period one.

Next, with a uniform bound on the informativeness of private signals, the cascade sets exist, and monotonically strictly shrink in the discount factor. They vanish in the perfect patience limit. But with the slightest impatience, informational herds arise in the team equilibrium, and are constrained-efficient. Herding owes not to the selfishness of $\mathcal{DM}$s, but rather to their inability to signal their private information by finitely many actions.

The social planner's problem is formally an optimal experimentation problem, to trade off the current $\mathcal{DM}$'s immediate utility against the gains of future $\mathcal{DM}$s from better information. Adapting and extending familiar tools from the experimentation literature to the herding model greatly advances our analysis. Superficially, it may seem as if the long-run inefficient herding result is just a reincarnation of the well-explored incomplete learning results of the experimentation literature. In fact, the analogy is not so immediate — although our setting has only two states and a finite number of actions, the social planner has at his disposal the continuum of rules mapping private signals into actions.

2 RELATED LITERATURE

Banerjee (1992) introduces the 'herding externality' through an example with an ex-post incorrect herd, after the two first $\mathcal{DM}$s' signals by chance are misleading. Banerjee's proposed remedy for the externality is to exclude some early $\mathcal{DM}$s from viewing others' actions at all, for they then their actions reveal their independent signals.[1] The rationale for this remedy supposes that the welfare loss of the early $\mathcal{DM}$s is negligible when compared with the gains to successors. Our analysis builds on the assumption that in the altruistic team, each $\mathcal{DM}$ attaches a strictly positive weight to his own utility when weighed against the team's utilities. In fact, our team has Banerjee's policy at its disposal, but it is suboptimal. This owes to our contrarianism result, as the team does better when the decision rule responds strictly to changes in the action history.

[1]This idea has been further explored by Sgroi (2002). Closer to our spirit, Doyle (2002) considers the social planner's problem in the endogenous-timing herding model of Chamley and Gale (1994).

In Dow (1991), a consumer first observes a signal realization, but in the next period can only recall the signal's partition interval. In the second and final period, another signal realization is seen, and a choice is made. The consumer's optimal determination of the first-period coarse partitioning of the signal is like our planner's problem. By the Bellman principle, our planner's decisions can be analysed in a two-period setting similar to Dow's.

Vives (1993) explores a model with a fundamentally different sequential structure, and Gaussian information. There, a continuum of privately informed agents act in every period, and then observe a noisy market price statistic summarizing the actions. Reminiscent of herding information externality, the more accurate is the history-based signal about the state, the less actions reveal about private information. Addressing this externality, Vives (1997) studies a team problem in the market setting. He proves an inequality showing that team members choose to reveal more of their private information.[2] Our more elaborate contrarianism comparative statics result here finds that teams shy away *more* from the *more* popular actions.[3] Vives also finds that the optimal long-run Gaussian precision growth is as low as in the selfish model. This may seem analogous to our finding that cascade sets have a non-empty interior in the team setting, but there is no clear logical connection. For instance, Vives' experimentation problem never yields incomplete learning.

The incorrect herding outcome is intuitively related to the familiar failure of complete learning in optimal experimentation. Rothschild's (1974) analysis of the two-armed bandit is a classic example: An impatient monopolist optimally experiments with two possible prices each period, with fixed uncertain purchase chances for each price. Rothschild showed that the monopolist (*i*) eventually settles down on one price, and (*ii*) selects the less profitable price with positive probability. This is analogous to (*i*) an action herd occurs, and (*ii*) with positive chance is ex-post incorrect. However, Easley and Kiefer (1988) prove that with finite state and action spaces, complete learning generically arises. This is puzzling, since the herding outcome arises in a model with two actions and two states.

The formulation of our social planner's problem offers a resolution of this puzzle. Even though each agent chooses from a finite action set, our social planner has no access to private signals, and so cannot dictate the choice among any two actions. Rather, for each history, he chooses a continuously defined rule that maps agents private beliefs into

[2]In a related setting, Medrano and Vives (2001) describe the behaviour of revealing less of the private information as 'contrarianism.' They do not motivate this label, but we find it more natural that contrarian behaviour is deviating more from the public expectation.

[3]Vives always employs the normal learning model, ruling out results like ours on the distributional shape's importance. On the other hand, that model allows the long-run properties of learning to be characterized by the speed with which the precision approaches infinity. Our model offers no analogy.

actions. In the myopic planner case with a zero discount factor, we obtain the original herding model. Hence, we can conclude that the herding outcome is formally to incomplete learning in an experimentation model with a continuous choice space. Seen from this vantage point, it is also not particularly surprising that the incomplete learning result extends from the self-interested model (corresponding to the myopic experimenter) into the domain of our team model (corresponding to somewhat patient experimentation).

The plan of the paper is as follows. The herding model and team equilibrium are introduced in section 3. Section 4 provides a first analysis of the forward-looking signaling problem. Section 5 shows that informational log-concavity precludes cascades. Under this assumption, Section 6 derives the short-run comparative statics result that individuals lean against the conventional wisdom when they care about posterity. Section 7 then shows how cascade sets shrink as individuals grow more patient. Many proofs are appendicized.

3 THE FORWARD-LOOKING HERDING MODEL

A. The Basic Herding Model. An infinite sequence of *decision-makers* ($\mathcal{DM}$s) $n = 1, 2, \ldots$ acts in that exogenous order. The actions have uncertain payoffs. There are two states of the world $\omega \in \{H, L\}$, over which everyone shares a common prior belief. Without loss of generality, we assume a fair prior, with both states equally likely.

The nth $\mathcal{DM}$ observes a partially informative random *private signal* realization σ_n about the state of the world. We may assume without loss of generality that $\sigma_n \in [0, 1]$ is the *private belief* that the state is H, resulting from Bayesian updating given the signal and the prior. In state ω, the signals are i.i.d. across $\mathcal{DM}$s, drawn according to the probability measure μ^ω. To avoid trivialities, some signals are informative, or $\mu^H \neq \mu^L$. Each distribution may contain atoms, but we assume that no signal will perfectly reveal the state of the world, i.e., μ^H and μ^L are mutually absolutely continuous. Let $\text{supp}(\mu)$ denote their common support, i.e. the smallest closed subset of $[0, 1]$ with probability 1. If $\text{supp}(\mu) \subseteq (0, 1)$, then private beliefs are *bounded*; they are *unbounded* if $0, 1 \in \text{supp}(\mu)$ — namely, if arbitrarily strong private beliefs exist.

Given the equi-likely states, the *unconditional distribution of private beliefs* is described by the probability measure $\mu = (\mu^H + \mu^L)/2$. The derivative $d\mu^L/d\mu^H$ of beliefs in the two states is well-defined and finite, by mutual absolute continuity. Bayesian updating implies that $(d\mu^L/d\mu^H)(\sigma) = (1 - \sigma)/\sigma$, so that $d\mu^H/d\mu = 2\sigma$ and $d\mu^L/d\mu = 2(1 - \sigma)$. It suffices to take μ as the primitive signal distribution, and derive the probability measures μ^H, μ^L.

Every $\mathcal{DM}$ chooses from a finite action set $\{1,\ldots,A\}$. Action a earns a payoff $u(a,\omega)$ in state $\omega \in \{H,L\}$, the same for all $\mathcal{DM}$s. We assume that action 1 is best in state L, and action A in state H. No two action payoffs are tied in either state. Before choosing, the n'th $\mathcal{DM}$ first observes σ_n and the history of the $n-1$ predecessors' actions.

Each $\mathcal{DM}$'s Bayes-optimal decision uses the observed action history and his private belief. In equilibrium, each $\mathcal{DM}$ can compute the probability distribution over histories, based on correctly conjectured strategies of all predecessors. The posterior chance of state H given the action history is the *public belief* $\pi \in [0,1]$. Conditioning also on the private signal σ gives the private *posterior belief* $\rho \in [0,1]$. By conditional independence,

$$\rho = r(\pi,\sigma) \equiv \frac{\pi\sigma}{\pi\sigma + (1-\pi)(1-\sigma)}.$$

The $\mathcal{DM}$ picks an action a to maximize his expected payoff $\bar{u}(a,\rho) = (1-\rho)u(a,L) + \rho u(a,H)$.

B. The Altruistic Herding Model. We now consider a different objective, where everyone altruistically aims to maximize a discounted average of payoffs. Individuals are still subject to the informational herding restriction, that they can observe past actions, but not past private signals. Adapting Radner (1962), we call an equilibrium in this revised model a team equilibrium. We underscore however that individuals' preferences are not perfectly aligned as in Radner, since they still weight their own payoffs highest.

In the altruistic game, individual utility depends directly on the choices of others. It is convenient to think of the other $\mathcal{DM}$s as employing *decision rules* — namely, maps x from private beliefs to probability measures over actions (thus allowing for mixed actions). Formally, the rule space X consists of maps $x : [0,1] \to \Delta$, where the simplex Δ denotes probability measures over the finite action set.

A *team equilibrium* is a perfect Bayes-Nash equilibrium of the game where the N'th $\mathcal{DM}$ chooses a rule $x \in X$, and his payoff is the expected discounted sum of his own and future utilities u_n, namely $E[(1-\delta)\sum_{n=N}^{\infty}\delta^{n-N}u_n]$. The discount factor $\delta \in [0,1)$ fixes the trade-off between each $\mathcal{DM}$ and his successors. Posterity matters when $\delta > 0$.

We will also consider an informationally constrained *social optimum.* Here we imagine a social planner who observes the action history, but not the private signals, and dictates to each $\mathcal{DM}$ which rule $x \in X$ to choose. The social planner aims to maximize a welfare criterion, which is again the present value expression $E[(1-\delta)\sum_{n=1}^{\infty}\delta^{n-1}u_n]$.

Whether employed by the members of a team or dictated by the social planner, a *strategy* s_n for the n'th $\mathcal{DM}$ is a map that assigns a rule to each history; $s = (s_1, s_2, \ldots)$

denotes a strategy profile. A *policy* defines a stationary strategy: it is a map $\xi : [0,1] \to X$, where $\xi(\pi)$ is the rule given belief π. If the social optimum exists, then some policy attains the maximum, since this infinite-horizon problem is Markovian in public beliefs π.

A $\mathcal{DM}$'s behavior is interpreted through the conditional chances with which each action is taken. A $\mathcal{DM}$ who applies rule x will take action a with conditional probability $\psi(a,\omega,x) = \int x(\sigma)(a)\mu^{\omega}(d\sigma)$ for $\omega \in \{H, L\}$. This definition extends to the case when π is the public belief, as $\psi(a,\pi,x) = \pi\psi(a,H,x) + (1-\pi)\psi(a,L,x)$. Any observer of the $\mathcal{DM}$'s realized action a updates the public belief to $p(a,\pi,x) = \pi\psi(a,H,x)/\psi(a,\pi,x)$.

4 EQUILIBRIUM VIA GITTINS INDICES

In this section, we extend the idea of the Gittins Index to our multi-agent setting. This permits easy deductions of all the key insights into the forward-looking herding model.

A. The Planner's Problem. The incentives in the team are very similar to those of the social planner, and this yields a version of the one-shot deviation principle.

Lemma 1 *For any discount factor $\delta < 1$, a social optimum is a team equilibrium.*

Proof: To see that the planner's optimal strategy s is a team equilibrium, suppose that all other $\mathcal{DM}$s use it, but that some $\mathcal{DM}$ n has a strictly better reply $\tilde{x}$ at some history. Then the planner can improve his value at that history by *fully* mimicking this deviation, i.e. by (i) taking $\tilde{x}$ in the first period and (ii) continuing with s as if s_n had been applied at stage n with this history. This contradicts optimality of the policy. $\square$

The social planner's problem is Markovian and thus can be analyzed with dynamic programming tools. Our analysis here follows Aghion, Bolton, Harris, and Jullien (1991) and §9.1–2 of Stokey and Lucas (1989). The planner's value function $v_\delta : \Sigma \mapsto \mathbb{R}$ is $v_\delta(\pi) = \sup_s E[(1-\delta)\sum_{n=1}^{\infty} \delta^{n-1} u_n|\pi]$, where the expectation is over the payoff sequences given the strategy profile s. Recalling the expected payoff function $\bar{u}$, the action chances ψ, and the updated public belief p, the Bellman equation for social optimality is:

$$v_\delta(\pi) = \sup_{x\in X} \left\{ \sum_{a=1}^{A} \psi(a,\pi,x) \left[(1-\delta)\bar{u}(a, p(a,\pi,x)) + \delta v_\delta(p(a,\pi,x))\right] \right\}. \qquad (1)$$

A unique solution v_δ to the Bellman equation exists, and it is convex and continuous in π.

In the classical multi-armed bandit (see Bertsekas (1987), §6.5), an experimenter each period chooses one of n actions, with uncertain independent reward distributions. Gittins

(1979) showed that optimal behaviour is described by index rules: Attach to each action the value of the problem with just that action and the largest possible lump sum retirement reward yielding indifference. Then always choose the action with the highest index.

In our setting, the rewards to the actions are not independent. Nevertheless, the policy employed in a stationary team equilibrium can be described by an analogous index rule: At public belief π and posterior ρ, the $\mathcal{DM}$ chooses the action a with the largest index. This index will include the social payoff as privately estimated by the $\mathcal{DM}$.

Below, we employ the standard notation $\partial g(z)$ for the *subdifferential* of the convex function g at z — i.e., all slopes m that obey $g(z') \geq g(z) + m \cdot (z' - z)$ for all z'.

Proposition 1 (Optimal Behaviour via Index Rules) *Fix a team equilibrium. To each action a and public belief π, there exists a function $\bar{v}_\delta(a, \pi, \rho)$, affine in decision maker n's private posterior belief ρ, such that n's average present value of action a is*

$$w_\delta(a, \pi, \rho) = (1 - \delta)\bar{u}(a, \rho) + \delta \bar{v}_\delta(a, \pi, \rho). \tag{2}$$

In any social optimum, $\bar{v}_\delta(a, \pi, \rho) = v_\delta(p(a)) + m_\delta(a, \pi)(\rho - p(a))$ with $m_\delta(a, \pi) \in \partial v_\delta(p(a))$, where $p(a) \equiv p(a, \pi, s_n(\pi))$ is the public posterior belief induced by action a.

Proof: Action a of individual n at π leads to a subgame where the state-contingent expected discounted future payoff of his successors is $\bar{v}(a, \pi, \omega)$. Then n's expected value of this subgame is $\bar{v}_\delta(a, \pi, \rho) \equiv \rho \bar{v}(a, \pi, H) + (1 - \rho)\bar{v}(a, \pi, L)$. The present value expression (2) follows. In the social optimum, the continuation value is $\bar{v}_\delta(a, \pi, p(a)) = v_\delta(p(a))$. Because the planner can always employ the same subgame strategy starting at an arbitrary public belief p as is optimal at $p(a)$, we have $\bar{v}_\delta(a, \pi, p(a)) \leq v_\delta(p(a))$. Thus, the slope of this affine function necessarily lies in the subdifferential $\partial v_\delta(p(a))$. □

That the planner can always ensure himself a payoff function tangent to the value function by simply not adjusting his policy was critical to this proof. This idea also implies convexity of the value function (Lemma 2 of Fusselman and Mirman (1993)).

B. Belief Interval Rules. Since the expected payoff of each action is affine in the posterior belief ρ, a selfish $\mathcal{DM}$ will use a belief interval rule: Action a is optimally taken for beliefs in some subinterval of $[0, 1]$. The team takes into account that actions can also be used to signal information to successors. To this end, a team member may adopt an action that is not optimal at any of his possible posterior beliefs. Yet, the team will use interval rules. Intuitively, the interval structure is not only myopically best, but it also ensures the greatest information value, by producing the riskiest posterior belief distribution.

Corollary 1 (Private Belief Intervals) *Fix any team equilibrium strategy s. For any public belief π, the rule $s_n(\pi)$ is described by a dissection of $[0,1]$ into closed intervals $I_a(\pi) \equiv I_a$, generically overlapping at endpoints only, with action a optimal iff $\sigma \in I_a$.*

Proof: By Proposition 1, the value of each action depends affinely on the posterior $r(\pi, \sigma)$. Thus, an action is optimal for $r(\pi, \sigma)$ in an interval. Since r is increasing in its second argument, an action is optimal on an interval for σ. □

By Lemma 1, the planner's solution too is described by the interval rules of Corollary 1. So the search for optimal rules can be narrowed down to a compact set. An optimal rule then exists (eg. Aghion, Bolton, Harris, and Jullien (1991), Theorem 4.1), and so by Lemma 1, a team equilibrium also exists. The appendix completes this formal argument:

Corollary 2 (Existence) *A social planner's* policy $\xi^\delta : [0,1] \to X$ *and team equilibrium both exist. Further, the correspondence $\pi \mapsto \xi^\delta(\pi)$ is upper hemi-continuous in π.*

Not only are actions optimally chosen for private beliefs inside an interval, but the resulting public beliefs settle down in intervals — so called *cascade sets.* This folk result in herding easily follows for our forward-looking model.

Corollary 3 (Long Run Public Beliefs) *In a social optimum, any action a is taken for all private beliefs σ iff the public beliefs π lie inside the cascade set $J_a(\delta)$, where*
(i) $J_a(\delta) \subset [0,1]$ is empty, a point, or a closed interval.
(ii) $0 \in J_1(\delta)$ and $1 \in J_A(\delta)$ for any $\delta \in [0,1)$, and $\cup_{a=1}^{A} J_a(\delta) \neq [0,1]$.
(iii) When $\delta_1 \geq \delta_2$, for any a, $J_a(\delta_1) \subseteq J_a(\delta_2)$.
(iv) With unbounded private beliefs, only the extreme sets $J_1(\delta), J_A(\delta)$ are nonempty.
(v) If the private beliefs are bounded, then $J_1(\delta) = [0, \underline{\pi}(\delta)]$ and $J_A(\delta) = [\bar{\pi}(\delta), 1]$, where $0 < \underline{\pi}(\delta) < \bar{\pi}(\delta) < 1$. For large enough $\delta < 1$, all cascade sets disappear except for $J_1(\delta)$ and $J_A(\delta)$, while $\lim_{\delta \to 1} J_1(\delta) = \{0\}$ and $\lim_{\delta \to 1} J_A(\delta) = \{1\}$.

Proof of (i)–(iii): For part (*i*), the certain choice of a achieves optimality iff $v_\delta(\pi) = \bar{u}(a, \pi)$. As $\bar{u}(a, \pi)$ is affine in π, and v_δ is weakly convex, this equality holds on a closed interval $J_a(\delta)$. For (*ii*), action 1 is myopically strictly optimal when $\pi = 0$. Since it updates to continuation belief $\pi = 0$ for any rule, it is also dynamically optimal for any discount factor $\delta \in [0,1)$. A similar proof holds for $\pi = 1$. If $\cup_{a=1}^{A} J_a(\delta) = [0,1]$, then $v_\delta(\pi) = \max_a \bar{u}(a, \pi)$ is piecewise linear, and information has no value, a contradiction at any kink of $v_\delta(\pi)$. Part (*iii*) follows from the observation that the value function is weakly increasing in δ. Its full proof, as well as proofs of (*iv*) and (*v*), are in the Appendix. □

Figure 1 illustrates that the value function is affine on the cascade sets.

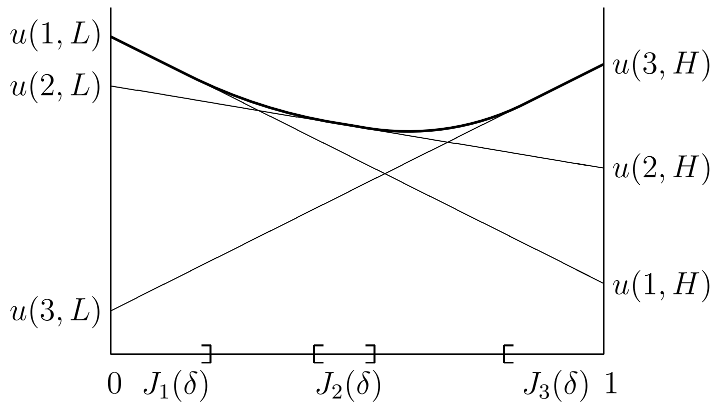

Figure 1: **Payoffs and Public Beliefs.** Graph of $v_\delta(\pi)$ with three actions.

C. Dominated Actions are Socially Valuable. The herding paradigm explores the implications of using a finite action screen mesh on a continuous belief. If individuals could more precisely convey their beliefs, then the welfare would intuitively rise. This translates into the result that in the herding model, more information can be transmitted with more actions. This effect is so strong that the actions can even be myopically dominated.[4]

Corollary 4 *Suppose that beliefs are bounded, and that posterity matters ($\delta > 0$). Suppose that action A dominates $A-1$, as $u(A,\omega) = u(A-1,\omega)+\varepsilon$. For small enough $\varepsilon > 0$, the social planner induces action $A-1$ with positive probability in an open set of public beliefs.*

Proof: Suppose that the planner never uses action $A-1$. The value function v is affine on the cascade set $J_A = [\bar{\pi}, 1]$, but not on any interval extending J_A to the left — i.e. it is strictly convex at $\bar{\pi}$. At belief $\bar{\pi}$, it is optimal to choose A. Consider the alternative rule x that maps the private beliefs below 1/2 into $A-1$, and others into A. This rule implies $p(A-1, \bar{\pi}, x) < \bar{\pi} < p(A, \bar{\pi}, x)$. Since the value function is strictly convex at $\bar{\pi}$, the expected continuation value exceeds $v(\bar{\pi})$ by some $\eta > 0$. Since the policy change implies a myopic loss less than ε, the modified policy beats the optimal policy when $\delta\eta > (1-\delta)\varepsilon$. □

Since we have seen that the team might well employ a dominated action, it is no longer even clear how it should order the action intervals (as a function of the public belief). To illustrate why the team might swap the natural order, consider a two-action example. Suppose that myopically action 2 is better than action 1 in state H, but worse in state L. To communicate accurately the lowest private beliefs, the team may employ the reverse-order policy that takes action 1 for all but very low beliefs. Swapping actions to respect

[4]Our model thus illustrates the benefits of costly communication inside the sequential decision making framework. Shiller (1995) argues that limits to communication are essential for herding.

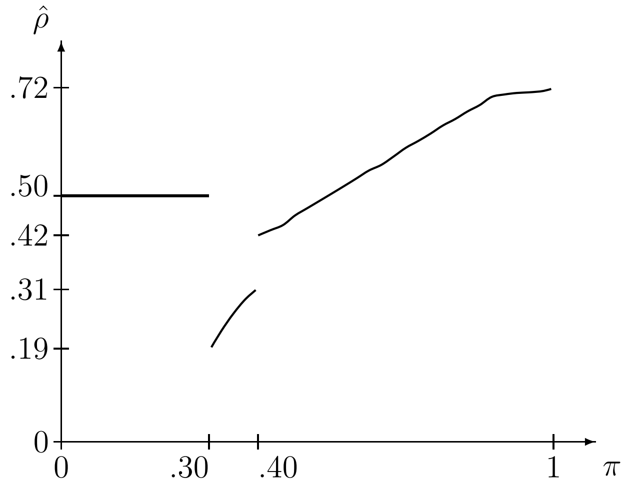

Figure 2: **Action ordering.** The unconditional private belief density is $f(\sigma) = 7^8\sigma^6/4^7$ over the range $(0, 4/7)$, with mean belief $1/2$. Action $a = 1, 2$ has payoff a in state H and $3 - a$ in state L. The myopic ($\delta = 0$) basin for action 1 is $(0, 3/7)$. For $\delta = .95$, the graph shows the individual's posterior belief $\hat{\rho}(\pi) = r(\pi, \hat{\sigma}(\pi))$ at the private belief threshold $\hat{\sigma}(\pi)$. The basin for action 1 now shrinks to $[0, .3)$. For $\pi \in (.3, .4) \subset (0, 3/7)$, reversing the action order is optimal: the myopically optimal action 1 is taken at high beliefs, while the myopically worse action 2 is taken at the less likely low private beliefs.

the myopic order might lead to a loss when the public belief is already low, as it raises the chance that the inefficient action 2 is taken. Figure 2 portrays this possibility based on numerical simulations of the optimal policy in an explicit example. In this example, the last $\mathcal{DM}$ using his own information may have taken action 2 and yet push the public belief into the cascade set for action 1. A team herd on some action (say, action 1) need not be a string of $\mathcal{DM}$s copying the choice of the last $\mathcal{DM}$ who used his private information.

D. Implementation of the Social Optimum. The team problem assumes that every $\mathcal{DM}$ altruistically cares about posterity. If the $\mathcal{DM}$s are really selfish, can the planner implement the optimal solution through a scheme of transfers? The planner cannot observe the private signals, so transfers may depend only on the observed action history.

When the planner's policy prescribes for a $\mathcal{DM}$ an interval rule that does not swap the myopic interval order, it suffices to reward the $\mathcal{DM}$ on the basis of his own actions. The planner can move the selfish $\mathcal{DM}$'s threshold between each pair of actions up (down) by taxing (subsidizing) the action to be taken above the threshold. But transfers based on the $\mathcal{DM}$'s own action can never reverse the myopic ordering of actions, and therefore are not sufficient if the optimal action ordering differs from the optimal action ordering.[5]

To implement the optimal policy, the planner can instead modify each selfish $\mathcal{DM}$'s payoff function from $\bar{u}$ to the Gittins index w_δ. This requires a change in the slope of payoffs in the private belief. So the planner needs a transfer that depends on the never observed true state ω. The successor's action provides the instrument, as we now see.

Proposition 2 (Implementation) *A social optimum can be implemented by transfers that depend only on the public belief π, a $\mathcal{DM}$'s own action, and his successor's action.*

Proof: Fix public belief π, and a social optimum. We design a transfer depending on a $\mathcal{DM}$'s own action a and his successor's action b, so that the $\mathcal{DM}$ selfishly chooses as if the payoff function is $w_\delta(a, \pi, \rho)$ of (2). We give here the proof in the main case when neither the present nor the next $\mathcal{DM}$ are in a cascade set; the other case is in the appendix.

We transform the current $\mathcal{DM}$'s selfish payoffs to those given by the Gittins index, by designing transfers as follows. For any action a of the current $\mathcal{DM}$, let the successor take action b for his lowest private beliefs with respective conditional chances $\psi(b, H)$ and $\psi(b, L)$. Then $\psi(b, H) \neq \psi(b, L)$ since b is taken for the lowest beliefs, but is not always taken. Give the $\mathcal{DM}$ the transfer $t(a, b)$ when the successor takes b, and $t(a, \neg b)$ otherwise. In state ω, the selfish $\mathcal{DM}$'s expected transfer from action a is $\psi(b, \omega)t(a, b) + (1 - \psi(b, \omega))t(a, \neg b)$. Because $\psi(b, H) \neq \psi(b, L)$, there exist transfers $t(a, b), t(a, \neg b)$ such that $\psi(b, \omega)t(a, b) + (1 - \psi(b, \omega))t(a, \neg b) = \delta\bar{v}_\delta(a, \pi, \omega)/(1 - \delta)$ for both states $\omega = H, L$. The expected transfer with private belief ρ is then $\delta\bar{v}_\delta(a, \pi, \rho)/(1 - \delta)$ as desired. □

5 CASCADES ARE INFORMATIONALLY RARE

Informational herding papers are often said to belong to the "cascades literature". We now argue that models that yield cascades violate a natural robustness property that most common continuous signal distributions obey. While we need this result later, we flesh it out in this section since it is of independent interest, and turns on a novel extension of the log-concavity notion beyond random variables, to informative signals.

Assume $A = 2$ actions for simplicity. By Corollary 1, a rule is described by an action ordering and a *private signal threshold* $\hat{\sigma}$, separating the private belief intervals for the two actions. At public belief π, the private and posterior beliefs σ and ρ are one-to-one

[5] Bru and Vives (2002) consider the properties of an incentive compatible mechanism that cannot implement the optimum of Vives (1997).

by $\rho = r(\pi, \sigma)$. A rule is thus equivalently defined by its action ordering and the posterior belief threshold $\hat{\rho} \equiv r(\pi, \hat{\sigma})$.

Under myopia ($\delta = 0$), the equilibrium rule lets $\hat{\rho}$ equal the unique belief at which the $\mathcal{DM}$ is indifferent among the two actions. Under the more general assumption that $\hat{\rho}$ is constant in π with lower (higher) private posteriors mapped into action ℓ (h), we will derive that the continuation public beliefs[6] $p(\ell, \pi, \hat{\rho})$ and $p(h, \pi, \hat{\rho})$ strictly rise in π.

Action ℓ (or h) is taken whenever such a private posterior falls below (or above) a threshold. The public posterior is an average of unobserved private posteriors that lead to the action. We need such truncated averages to be monotone. In the spirit of Burdett (1996), a log-concavity assumption is the missing ingredient — here for signals rather than random variables. Since Bayes' rule is additive in the log-likelihood ratios of signals and prior beliefs, the distribution of the log-likelihood ratio is a natural primitive of our model.

Let f be the *unconditional private belief density* for μ. By this, we mean that the state densities can be written as $f^H(\sigma) = \sigma f(\sigma)$ and $f^L(\sigma) = (1-\sigma)f(\sigma)$. Associate to each the private signal σ the log-likelihood ratio $\Lambda = \Lambda(\sigma) = \log(\sigma/(1-\sigma))$, with inverse $\sigma = \sigma(\Lambda)$. By a change of variables, its density is $\phi(\Lambda) \equiv f(\sigma(\Lambda))\sigma'(\Lambda) = f(\sigma(\Lambda))e^{\Lambda}/(1+e^{\Lambda})^2$.

A-1 *The log-likelihood ratio density $\phi(\Lambda)$ exists, and is strictly log-concave.*

We showed in Smith and Sorensen (2000) (our "bounded beliefs example") that a cascade is possible in finite time only if the posterior public belief after an action is not monotone in the prior belief. Conversely, strict monotonicity implies that cascades do not arise in finite time. In particular, the discrete signal examples of cascades in the myopic model of Bikhchandani, Hirshleifer, and Welch (1992) all violate assumption A-1.

Proposition 3 (Ruling Out Cascades) *Assume A-1 and keep constant a rule's action ordering and posterior belief threshold $\hat{\rho}$. Then the continuation public beliefs $p(\ell, \pi, \hat{\rho})$ and $p(h, \pi, \hat{\rho})$ strictly rise in π. Thus, no cascade can start that does not initially obtain.*

Proof: Let $\hat{\sigma}(\pi)$ solve $\hat{\rho} = r(\pi, \hat{\sigma})$. Since an increase in $p(\ell, \pi, \hat{\rho})$ is equivalent to an increase in $p(\ell, \pi, \hat{\rho})/[1 - p(\ell, \pi, \hat{\rho})]$, we need only show the latter rises in $\hat{\rho}$. Now,

$$\frac{p(\ell, \pi, \hat{\rho})}{1 - p(\ell, \pi, \hat{\rho})} = \left(\frac{\pi}{1-\pi}\right) \frac{\int_0^{\hat{\sigma}(\pi)} f^H(\sigma) d\sigma}{\int_0^{\hat{\sigma}(\pi)} f^L(\sigma) d\sigma} = \left(\frac{\pi}{1-\pi}\right) \frac{\int_0^{\hat{\sigma}(\pi)} \sigma\, f(\sigma) d\sigma}{\int_0^{\hat{\sigma}(\pi)} (1-\sigma)\, f(\sigma) d\sigma} \qquad (3)$$

given densities $f^H(\sigma) = 2\sigma f(\sigma)$ and $f^L(\sigma) = 2(1-\sigma)f(\sigma)$ for the measures μ^H and μ^L.

[6] As the policy rule x is represented by the posterior threshold $\hat{\rho}$, we replace $p(a, \pi, x)$ with $p(a, \pi, \hat{\rho})$ and $\psi(a, \pi, x)$ with $\psi(a, \pi, \hat{\rho})$.

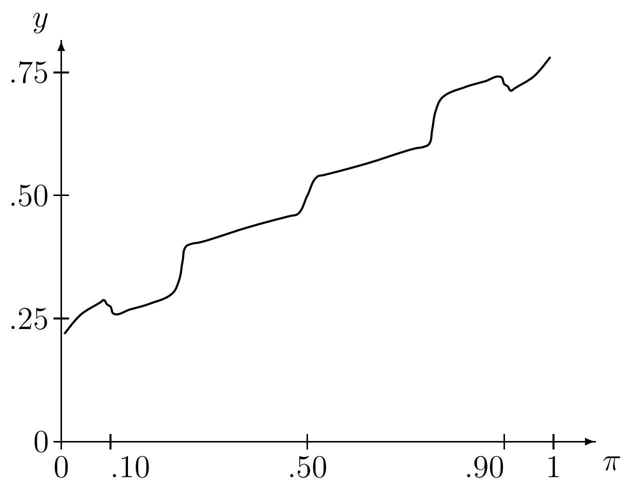

Figure 3: **Non-monotonicity.** The unconditional private belief density is symmetric, unbounded and piecewise constant, being 1/3 on the three intervals $(0, .2)$, $(.25, .75)$, and $(.8, 1)$, and being 7 on the two intervals $[.2, .25]$ and $[.75, 8]$. It violates log-concavity. Action $a = 1, 2$ has payoff a in state H and $3 - a$ in state L. When $\delta = .85$, the graph shows the threshold private posterior belief $\hat{\rho}(\pi)$ against the public belief π. When $\pi \approx .9$, then $\hat{\rho}(\pi) \approx .72$, corresponding to a private belief threshold $\hat{\sigma} \approx .25$ where the density is high. Increasing π, passing over the high plateau of the belief distribution, the continuation belief is decreasing. To mitigate this fall, $\hat{\rho}(\pi)$ is locally decreasing.

Define the function
$$\gamma(\lambda, \pi) = \frac{\pi(1-\pi)\phi(\Lambda(\sigma(\lambda)))}{(\pi + \lambda(1-\pi))\lambda}$$

Changing densities via the posterior likelihood ratio $\lambda = \pi\sigma/[(1-\pi)(1-\sigma)]$, we prove in the Appendix (Claim 4) that $(1-\pi)(1-\sigma)f(\sigma)d\sigma = \gamma(\lambda, \pi)d\lambda$ and $\pi\sigma f(\sigma)d\sigma = \lambda\gamma(\lambda, \pi)d\lambda$.

This change of variables allows us to rewrite (3) as:

$$\frac{p(\ell, \pi, \hat{\rho})}{1 - p(\ell, \pi, \hat{\rho})} = \frac{\int_0^{\hat{\rho}/(1-\hat{\rho})} \lambda\gamma(\lambda, \pi)\, d\lambda}{\int_0^{\hat{\rho}/(1-\hat{\rho})} \gamma(\lambda, \pi)\, d\lambda} = E\left[\lambda \middle| 0 \le \lambda \le \frac{\hat{\rho}}{1-\hat{\rho}}\right].$$

Since $\gamma(\lambda, \pi)$ obeys the monotone likelihood ratio property by Claim 4, (3) rises in π.

A similar proof works for h, given $p(h, \pi, \hat{\rho})/(1 - p(h, \pi, \hat{\rho})) = E[\lambda | \lambda \ge \hat{\rho}/(1-\hat{\rho})]$. □

Assumption A-1 is violated by standard atomic signal distributions, but all well-known continuous distributions are log-concave (see Marshall and Olkin (1979), §18.B.2.d). Figure 3 illustrates how the failure of the log-concavity assumption A-2 can undermine both Proposition 3 and the comparative statics result to come, Proposition 4.

6 CONTRARIANISM IS GENERALLY OPTIMAL

6.1 A Two Period Example

While our theory obtains for infinite horizon models, we illustrate it in a closed-form two period example that captures the essence of the informational herding externality.[7] A professor and a student share a common prior π, and observe conditionally iid signals σ with state-dependent densities $f^H(\sigma) = 2\sigma$ and $f^L(\sigma) = 2(1-\sigma)$. The unconditional density is $f(\sigma) = 1$ over $(0,1)$. (The induced unconditional density over log-likelihood ratios is log-concave.) The professor sees a signal, and takes an action; his student observes his action, but not his signal. The professor selflessly takes action $a \in \{\ell, h\}$ to maximize his student's expected payoff, where $u(h,H) = u(\ell,L) = 1, u(\ell,H) = u(h,L) = 0$.

If the student starts with a continuation public belief p, then she takes action h when her signal $\sigma \geq 1-p$. Given the conditional signal densities, her value function is $V_S(p) = (1-p)\int_0^{1-p} 2(1-\sigma)d\sigma + p\int_{1-p}^1 \sigma d\sigma = 1-p+p^2$. The professor obviously employs a threshold rule $\hat{\sigma} = \hat{\sigma}(\pi)$: He chooses ℓ if his signal $\sigma < \hat{\sigma}$, and h if $\sigma \geq \hat{\sigma}$. His problem is to maximize $V(\pi) = E_\pi V_S(p)$, where $p = p(a, \pi, \hat{\sigma})$ is his student's public belief (denoted p_a later on). Since $\pi = E[p|\pi]$, we get $V(\pi) = E[V_S(p)|\pi] = E(1-p+p^2|\pi) = 1-\pi+\pi^2+E[(p-\pi)^2|\pi]$. Then the professor's optimal value exceeds the myopic student value $V_S(\pi) = 1-\pi+\pi^2$ by the variance of beliefs, which is the value of information. Then

$$E[(p-\pi)^2|\pi] = \frac{\pi - p_\ell}{p_h - p_\ell}(p_h - \pi)^2 + \frac{p_h - \pi}{p_h - p_\ell}(\pi - p_\ell)^2 = (p_h - \pi)(\pi - p_\ell) \qquad (4)$$

by the martingale property of beliefs, where the two continuation public beliefs are $p_\ell = [\pi\hat{\sigma}^2]/[\pi\hat{\sigma}^2 + (1-\pi)(2\hat{\sigma} - \hat{\sigma}^2)]$ and $p_h = [\pi(1-\hat{\sigma}^2)]/[\pi(1-\hat{\sigma}^2) + (1-\pi)(1-2\hat{\sigma}+\hat{\sigma}^2)]$. Maximizing (4), we find the optimal threshold $\hat{\sigma}(\pi) = (\pi - 1 + \sqrt{\pi - \pi^2})/(2\pi - 1)$ if $\pi \neq 1/2$, with limit $\hat{\sigma}(1/2) = 1/2$. Then the threshold private posterior belief is $\hat{\rho}(\pi) = r(\pi, \hat{\sigma}(\pi)) = [\pi - \sqrt{\pi - \pi^2}]/[2\pi - 1]$. Thus $\hat{\rho}(\pi)$ is increasing in π, illustrating Proposition 4 below. We observe that the possible signals conveyed by the professor's action are not ordered by sufficiency, and so Blackwell's Theorem does not allow us to compare the value of any two signals. But intuitively, the professor tries to better communicate the state of the world by erring on the side of a more equally weighted private signal afforded by the informationally optimal private belief threshold $\hat{\sigma} = 1/2$ away from the myopic threshold $\hat{\sigma} = 1 - \pi$.

[7] We could bring finite horizon models under the umbrella of this theory, but dealing with finite and infinite horizon cases greatly complicates the notation and the dynamic programming proof.

6.2 The Infinite Horizon Model with Two Actions

Assume here $A = 2$ actions. By Corollary 3, both actions are taken with a positive chance for non-cascade public beliefs π, i.e. in the open interval $M = [0,1] \setminus (J_1(\delta) \cup J_2(\delta))$. We will derive a local comparative statics result, valid near a fixed $\pi^* \in M$, provided:

A-2 *The optimal rule $x(\pi)$ is unique for π in an open neighbourhood $\mathcal{N}(\pi^*) \subseteq M$ of π^*.*

Assumption A-2 is weak, generically valid on an open dense subset of beliefs $\pi^* \in M$. We have, however, not identified any sufficient condition for a global comparative statics result. The example of Figure 2 satisfies A-1 but violates A-2 at certain isolated $\pi^* \in M$.

Recall that in the myopic case when $\delta = 0$, the optimal posterior threshold $\hat{\rho}(\pi)$ is *constant* in π. We provide conditions under which the best team policy at $\delta > 0$ encourages *contrarian* behaviour relative to the myopic policy. (We fix $\delta > 0$, and so suppress the δ super and subscripts.) Specifically we prove that $\hat{\rho}(\pi)$ *increases* in π. So for higher public beliefs π, the $\mathcal{DM}$ is discouraged from choosing the increasingly popular action. This experimentation benefits successors, since a less likely action moves public beliefs more.

Proposition 4 (Contrarianism) *Assume $\pi^* \in M$ and $\delta > 0$. Under A-1 and A-2, contrarian behaviour is encouraged: the threshold $\hat{\rho}(\pi)$ strictly increases in $\pi \in \mathcal{N}(\pi^*)$.*

We derive this result by carefully analyzing the planner's first order condition. In light of (1), and our assumed action order, the maximand of the Bellman equation is[8]

$$B(\pi, \hat{\rho}) = \psi(\pi, \hat{\rho})[(1-\delta)\bar{u}_\ell(p_\ell) + \delta v(p_\ell)] + (1 - \psi(\pi, \hat{\rho}))[(1-\delta)\bar{u}_h(p_h) + \delta v(p_h)] \qquad (5)$$

where $\bar{u}_a(\hat{\rho}) = \bar{u}(a, \hat{\rho})$, $\psi(\pi, \hat{\rho}) = \psi(\ell, \pi, \hat{\rho})$, and $p_\ell = p(\ell, \pi, \hat{\rho}) < \hat{\rho} < p_h = p(h, \pi, \hat{\rho})$. Differentiability of the value function in problems of learning is an open problem.[9] We now offer a proof of differentiability where we need to apply the first order condition.

Step 1 (First Order Condition) *If the posterior threshold $\hat{\rho}$ solves $\max_{\hat{\rho}} B(\pi, \hat{\rho})$, then the value function v is differentiable at the continuation belief p_a after action $a = \ell, h$. Also,*

$$B_{\hat{\rho}}(\pi, \hat{\rho}) = \frac{\partial \psi}{\partial \hat{\rho}} \{(1-\delta)\bar{u}_\ell(\hat{\rho}) + \delta[v(p_\ell) + v'(p_\ell)(\hat{\rho} - p_\ell)] - (1-\delta)\bar{u}_h(\hat{\rho}) - \delta[v(p_h) + v'(p_h)(\hat{\rho} - p_h)]\}. \quad (6)$$

[8] As before, replace $p(a, \pi, x)$ with $p(a, \pi, \hat{\rho})$ and $\psi(a, \pi, x)$ with $\psi(a, \pi, \hat{\rho})$, and denote the action taken at low beliefs as ℓ and the action taken at high beliefs as h. The action ordering cannot switch in the open neighbourhood $\mathcal{N}(\pi^*)$, for local uniqueness and a uhc rule correspondence precludes policy jumps.

[9] We thank Rabah Amir, David Easley, Andrew McLennan, Paul Milgrom, Len Mirman, Yaw Nyarko, and Ed Schlee for discussions about the differentiability of the value function in experimentation problems. Amir (1996) establishes differentiability at all continuation states in a deterministic problem.

Proof: If we assume first that v is differentiable at p_ℓ and p_h, then $B_{\hat\rho}(\pi,\hat\rho)$ equals

$$\frac{\partial\psi}{\partial\hat\rho}\left((1-\delta)\bar u_\ell(p_\ell)+\delta v(p_\ell)\right)+\psi\left((1-\delta)\frac{\partial\bar u_\ell(p_\ell)}{\partial p_\ell}\frac{\partial p_\ell}{\partial\hat\rho}+\delta v'(p_\ell)\frac{\partial p_\ell}{\partial\hat\rho}\right)$$
$$-\frac{\partial\psi}{\partial\hat\rho}\left((1-\delta)\bar u_h(p_h)+\delta v(p_h)\right)+(1-\psi)\left((1-\delta)\frac{\partial\bar u_h(p_h)}{\partial p_h}\frac{\partial p_h}{\partial\hat\rho}+\delta v'(p_h)\frac{\partial p_h}{\partial\hat\rho}\right).$$

Recalling that $\bar u_\ell$ and $\bar u_h$ are affine functions, and using Lemma 2 below, this produces (6).

If the convex v is not differentiable at p_ℓ or p_h or both, then its right derivative strictly exceeds its left — i.e. a 'convex kink'. Since p_ℓ and p_h are increasing functions of $\hat\rho$, and since $\partial\psi/\partial\hat\rho>0$ and $p_\ell<\hat\rho<p_h$, (6) applies to one-sided derivatives — in other words, $B_{\hat\rho-}(\pi,\hat\rho)<B_{\hat\rho+}(\pi,\hat\rho)$. But optimality of $\hat\rho$ implies that the left derivative is nonnegative, and the right derivative nonpositive. The convex kink cannot then obtain. □

Lemma 2 *The threshold $\hat\rho$ separating posterior beliefs following actions, and the chance ψ of action h, jointly obey: $(\hat\rho-p_\ell)\partial\psi/\partial\hat\rho=\psi\partial p_\ell/\partial\hat\rho$ and $(p_h-\hat\rho)\partial\psi/\partial\hat\rho=(1-\psi)\partial p_h/\partial\hat\rho$.*

The lemma, proved in the Appendix, admits a simple intuition. Observe that $\hat\rho$ and p_ℓ are the *marginal* and *average* private posteriors leading to ℓ, while ψ is the chance of ℓ. Think of these as a firm's marginal cost MC, average cost AC, and quantity Q. Then $(\hat\rho-p_\ell)\partial\psi/\partial\hat\rho=\psi\partial p_\ell/\partial\hat\rho$ is the producer theory result that $(MC-AC)/Q=AC'(Q)$.

The first order condition in Step 1 is related to our index result of Proposition 1. Since $\partial\psi/\partial\hat\rho>0$, the first order condition says that the indices coincide at knife-edge beliefs: $w_\delta(\ell,\pi,\hat\rho)=w_\delta(h,\pi,\hat\rho)$. This could also be rewritten as the equality at belief $\hat\rho$ of a myopic gain and a dynamic loss from a marginal shift in the threshold belief (see Figure 4):

$$(1-\delta)\left[\bar u_\ell(\hat\rho)-\bar u_h(\hat\rho)\right]=\delta\left[(v(p_h)+v'(p_h)(\hat\rho-p_h))-(v(p_\ell)+v'(p_\ell)(\hat\rho-p_\ell))\right]. \quad (7)$$

In other words, the optimal threshold lies at the intersection of the two tangents.[10]

The next two more technical Lemmas are proved in the Appendix.

Lemma 3 (Tangents to a Convex Function) *Fix $0\le z_1<z_2<z_3<z_4<z_5\le 1$. Let $\tau_j:\mathbb{R}\to\mathbb{R}$ be tangent functions to v at z_j for $j=1,2,4,5$. Then $\tau_1(z_3)\le\tau_2(z_3)$ with strict inequality unless v is affine between z_1 and z_2, and $\tau_5(z_3)\le\tau_4(z_3)$ with strict inequality unless v is affine between z_4 and z_5.*

[10]Our result generalizes Dow's (1991) Proposition 2, which relies on the perfect patience as well as a particularly simple second-period value function. We note in passing that Dow's Example 3 illustrates the multiplicity of optimal solutions which can arise in this class of problems.

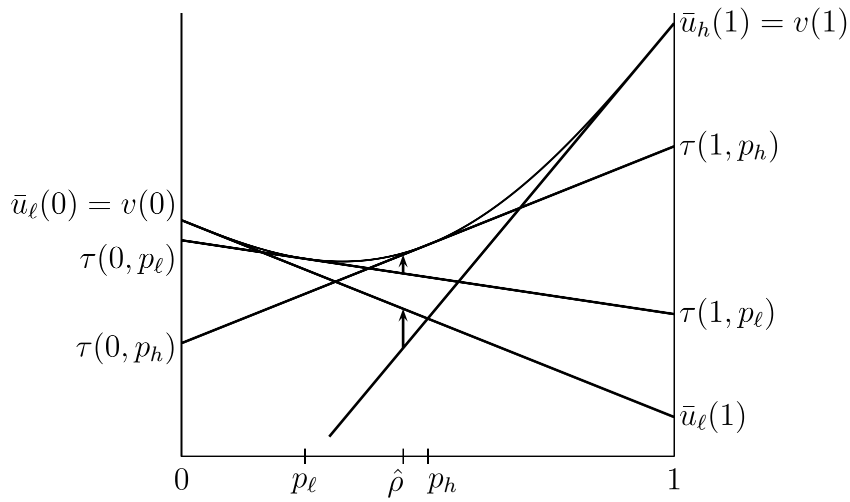

Figure 4: **First order condition.** The convex curve is the value function v in the beliefs. It intersects the myopic value functions at 0 and 1. Given the threshold posterior belief $\hat{\rho}$, we depict the tangents τ to v at the posteriors p_ℓ, p_h. The vertical arrows over $\hat{\rho}$ indicate the myopic loss and dynamic gain from a marginal change in $\hat{\rho}$. The first order condition states that the lengths of these arrows stand in proportion $\delta/(1-\delta)$.

Lemma 4 (Strict Convexity) *The value function v is strictly convex on M.*

By Proposition 3, the public posteriors p_ℓ, p_h increase with the public prior belief π. By the two lemmas, the RHS of (7) falls in p_ℓ and p_h, and so in π, since the value function is convex (Figure 4). This yields a sufficient condition for an increasing optimizer $\hat{\rho}(\pi)$.

Step 2 (Single-Crossing Property) *For all beliefs $\pi, \pi' \in \mathcal{N}(\pi^*)$ with $\pi < \pi'$, the function $B(\pi, \hat{\rho})$ obeys the single crossing property $B_{\hat{\rho}-}(\pi', \hat{\rho}(\pi)) > 0$, and so $\hat{\rho}(\pi') > \hat{\rho}(\pi)$.*

Proof: First, rewrite $B_{\hat{\rho}-}(\pi, \hat{\rho}) = \Upsilon(\hat{\rho}, p(\ell, \pi, \hat{\rho}), p(h, \pi, \hat{\rho}))\partial\psi(\pi, \hat{\rho})/\partial\hat{\rho}$, where

$$\Upsilon(\hat{\rho}, p_\ell, p_h) \equiv (1-\delta)[\bar{u}_\ell(\hat{\rho}) - \bar{u}_h(\hat{\rho})] + \delta[\tau(\hat{\rho}, p_\ell) - \tau(\hat{\rho}, p_h)] \tag{8}$$

and where $\tau(\hat{\rho}, p) \equiv v(p) + v'(p-)(\hat{\rho} - p)$ is the affine function left-tangent to v at p.

Since $B_{\hat{\rho}-}(\pi, \hat{\rho}(\pi)) = 0$, we have $B_{\hat{\rho}-}(\pi', \hat{\rho}(\pi)) > 0$ iff

$$\Upsilon(\hat{\rho}(\pi), p(\ell, \pi', \hat{\rho}(\pi)), p(h, \pi', \hat{\rho}(\pi))) > \Upsilon(\hat{\rho}(\pi), p(\ell, \pi, \hat{\rho}(\pi)), p(h, \pi, \hat{\rho}(\pi))).$$

Since only the continuation beliefs change, this is equivalent to

$$\tau(\hat{\rho}(\pi), p(\ell, \pi', \hat{\rho}(\pi))) - \tau(\hat{\rho}(\pi), p(h, \pi', \hat{\rho}(\pi))) > \tau(\hat{\rho}(\pi), p(\ell, \pi, \hat{\rho}(\pi))) - \tau(\hat{\rho}(\pi), p(h, \pi, \hat{\rho}(\pi))).$$

When $\pi < \pi'$, we have $p(\ell,\pi,\hat\rho(\pi)) < p(\ell,\pi',\hat\rho(\pi)) < \hat\rho(\pi) < p(h,\pi,\hat\rho(\pi)) < p(h,\pi',\hat\rho(\pi))$ by Proposition 3. Lemmas 3 and 4 imply $\tau(\hat\rho(\pi), p(h,\pi,\hat\rho(\pi))) \geq \tau(\hat\rho(\pi), p(h,\pi',\hat\rho(\pi)))$ and $\tau(\hat\rho(\pi), p(\ell,\pi',\hat\rho(\pi))) \geq \tau(\hat\rho(\pi), p(\ell,\pi,\hat\rho(\pi)))$. Since some continuation belief p_a lies inside M by Claim 3 (Appendix), at least one inequality is strict. The inequality follows.

Next, we prove that $\hat\rho(\pi') > \hat\rho(\pi)$. First, if $\hat\rho(\pi') = \hat\rho(\pi)$, then $B_{\hat\rho-}(\pi', \hat\rho(\pi')) = B_{\hat\rho-}(\pi', \hat\rho(\pi)) > 0$, contradicting the optimality of $\hat\rho(\pi')$. Next, if $\hat\rho(\pi') < \hat\rho(\pi)$, then by a variant of the Fundamental Theorem of Calculus for convex functions (see Claim 5),

$$B(\pi', \hat\rho(\pi)) - B(\pi', \hat\rho(\pi')) = \textstyle\int_{\hat\rho(\pi')}^{\hat\rho(\pi)} B_{\hat\rho-}(\pi', z)dz.$$

Since the optimizer $\hat\rho(\pi)$ is unique, the u.h.c. map $\pi \mapsto \hat\rho(\pi)$ is also continuous. By the Intermediate Value Theorem, for all beliefs $z \in (\hat\rho(\pi'), \hat\rho(\pi))$, there then exists $\pi'' \in (\pi, \pi')$ with $\hat\rho(\pi'') = z$. Thus, the above integrand $B_{\hat\rho-}(\pi', z) = B_{\hat\rho-}(\pi', \hat\rho(\pi'')) > 0$, and so $B(\pi', \hat\rho(\pi)) > B(\pi', \hat\rho(\pi'))$. This contradicts optimality of $\hat\rho(\pi')$. □

6.3 Efficient Contrarianism with Multiple Actions

The proof of Proposition 4 avoids considering the cross partial derivative of the Bellman function in beliefs and thresholds. Our attempts to extend this method directly to the case $A > 2$ has fallen short. We proceed with a more traditional approach to a comparative statics exercise, using the implicit function theorem to analyse the first order condition. This requires slightly stronger assumptions, that we first state.

Step 1 can be extended to multiple actions, to show that the value function is differentiable at the continuation beliefs. We make a slightly stronger assumption:

A-3 *The value function is twice continuously differentiable at the continuation beliefs.*

The definition of the non-cascade region M extends to this case. Even on M, some actions may be taken with probability zero (i.e. *inactive*). But in generic problems, the following assumption will hold at almost every π^* (valid in the two-period example for all $\pi^* \in M$):

A-2′ *A unique rule $x(\pi)$ employing the same set of $\alpha > 2$ active actions is optimal in an open neighbourhood $\mathcal{N}(\pi^*) \subseteq M$ of π^*.*

Given the revised assumption A-2′, we excise (and ignore) all inactive actions. Denote the posterior belief thresholds by $\hat\rho = (\hat\rho_1, \ldots, \hat\rho_{\alpha-1})$, with $\hat\rho_1 < \cdots < \hat\rho_{\alpha-1}$.

We extend the definition of the Bellman maximand to the many-actions setting:

$$B(\pi,\hat{\rho}) \equiv \sum_{a=1}^{\alpha} \psi(a,\pi,\hat{\rho})\left[(1-\delta)\bar{u}(a,p(a,\pi,\hat{\rho})) + \delta v(p(a,\pi,\hat{\rho}))\right] \tag{9}$$

The final weak assumption, generically satisfied, states that the necessary second order condition for optimization holds strictly (which also held in the two-period example):

A-4 *The matrix of second derivatives $B_{\hat{\rho}\hat{\rho}}(\pi,\hat{\rho}(\pi))$ is negative definite for $\pi \in \mathcal{N}(\pi^*)$.*

Proposition 5 *Assume $\pi^* \in M$ and $\delta > 0$. Given assumptions A-1, A-2′, A-3, and A-4, contrarian behaviour is encouraged — the threshold vector $\hat{\rho}(\pi)$ strictly rises in $\pi \in \mathcal{N}(\pi^*)$.*

Proof: Shorten $p_a = p(a,\pi,\hat{\rho})$ and $p_{a+1} = p(a+1,\pi,\hat{\rho})$, and suppress arguments of (9) with subscripts. Let τ_a be the linear function tangent to v at p_a. Differentiating (9) as in (8):

$$B_{\hat{\rho}_a}(\pi,\hat{\rho}) \equiv \frac{\partial\psi(a,\pi,\hat{\rho})}{\partial\hat{\rho}_a}\left[(1-\delta)\bar{u}_a(\hat{\rho}_a) + \delta\tau_a(\hat{\rho}_a) - (1-\delta)\bar{u}_{a+1}(\hat{\rho}_a) - \delta\tau_{a+1}(\hat{\rho}_a))\right] \tag{10}$$

Now, a marginal change in the threshold $\hat{\rho}$ affects $B_{\hat{\rho}_a}$ only through $\hat{\rho}_a$, $\hat{\rho}_{a-1}$, and $\hat{\rho}_{a+1}$ (via p_a and p_{a+1}). Greater $\hat{\rho}_{a-1}$ pushes up p_a, and since $p_a < \hat{\rho}_a$, the tangent to the convex value function $\tau_a(\hat{\rho}_a)$ will increase. In turn, $B_{\hat{\rho}_a\hat{\rho}_{a-1}} \geq 0$ from (10). By Claim 8 in the Appendix, the inverse matrix $-B^{-1}_{\hat{\rho}\hat{\rho}}(\pi,\hat{\rho}(\pi))$ of a supermodular and negative definite (A-4) matrix has all non-negative elements, and is strictly positive on the diagonal. By the implicit function theorem, the first order equation $B_{\hat{\rho}}(\pi,\hat{\rho}(\pi)) = 0$ yields $\hat{\rho}_\pi(\pi) = -B^{-1}_{\hat{\rho}\hat{\rho}}(\pi,\hat{\rho}(\pi))B_{\pi\hat{\rho}}(\pi,\hat{\rho}(\pi))$. On the other hand, all elements of the vector $B_{\pi\hat{\rho}}(\pi,\hat{\rho}(\pi))$ are strictly positive — by our proof of Step 2 in §6.2. So $\hat{\rho}_\pi(\pi)$ has strictly positive entries. □

7 CASCADE SETS SHRINK WITH PATIENCE

7.1 Long Run Behaviour in the Optimal Solution

Turning to the main focus of the informational herding literature, we now explore what happens in the long run. Before developing our more novel result here, we first address the constrained efficiency of herding. Public beliefs converge by the martingale convergence theorem, and cannot be completely wrong in the limit:

Lemma 5 *The belief process $\langle\pi_n\rangle$ is a martingale unconditional on the state, converging a.s. to some limiting random variable π_∞. The limit π_∞ is concentrated on $(0,1]$ in state H.*

Smith and Sørensen (2000) prove that the public belief process converges upon the cascade set, where learning no longer active occurs. In this setting, we can strengthen this result:

Lemma 6 (Convergence of Beliefs) *Consider a solution of the planner's problem.*
(i) The limit belief π_∞ has support in the cascade sets $J_1(\delta) \cup \cdots \cup J_A(\delta)$. In particular, π_∞ is concentrated on the truth for unbounded private beliefs.
(ii) With bounded private beliefs, learning is incomplete for any $\delta \in [0, 1)$: That is, unless $\pi_0 \in J_A(\delta)$, there is positive probability in state H that $\pi_\infty \notin J_A(\delta)$.
(iii) With bounded private beliefs, the chance of complete learning ($\pi_\infty \in J_A(\delta)$ in state H) tends to 1 as $\delta \uparrow 1$.

Proof: For (i), observe that at least two actions occur with positive chance for any belief π not in any cascade set. By the interval structure, the highest such action is more likely in state H, and the lowest in state L. So the next period's belief differs from π with positive probability. Intuitively, or by the characterization result for Markov-martingale processes in Appendix B of Smith and Sørensen (2000), π cannot lie in the support of π_∞.

For parts (ii) and (iii), note that the likelihood ratio $(1 - \pi_n)/\pi_n$ is a martingale conditional on state H. By Corollary 3 it is bounded when beliefs are bounded, and hence the mean of its limit equals its prior value $(1 - \pi_0)/\pi_0$. Since $\lim_{\delta \to 1} \underline{\pi}(\delta) = 0$, the weight that π_∞ places on $J_1(\delta)$ in state H must vanish as $\delta \to 1$. □

A herd obtains on action a at stage N if individuals $n = N, N + 1, N + 2, \ldots$ choose action a. While a cascade implies a herd, the converse is false. To show that herds arise, we extend the logic of Smith and Sørensen (2000), by first generalizing the Overturning Principle to the forward-looking case. Claim 11 in the Appendix proves that for beliefs π near $J_a(\delta)$, actions other than a push the updated public belief far from its current value. So convergence of beliefs implies convergence of actions — a limit cascade implies a herd.

Since beliefs must converge to cascade sets (Lemma 6), herding is constrained best:

Proposition 6 (Herding is Constrained Efficient) *Consider any planner's solution:*
(i) A herd eventually starts.
(ii) With unbounded private beliefs, the herd is on the ex post optimal action.
(iii) With bounded private beliefs, unless $\pi_0 \in J_A(\delta)$, there is positive chance that the herd is on an ex post sub-optimal action.
(iv) The chance of an incorrect herd with bounded private beliefs vanishes as $\delta \uparrow 1$.

The proof is in the appendix. It should come as no surprise that complete learning happens with unbounded beliefs, for this occurs even with selfish individuals (i.e. myopic learning).

More interesting is that the planner optimally incurs the risk of an ever-lasting incorrect herd. We see that herding is quite a robust property of the observational learning paradigm.

7.2 Monotonic Cascade Sets

We finally establish our second major result comparative statics result: With bounded beliefs, the cascade sets strictly shrink as the patience rises. This is the long-run analogue of our short-run comparative static Proposition 4, since it means that for any positive discount factor, some individuals should in the limit behave in a contrarian fashion, violating the cascade. However, here no other assumptions are needed. Instead we exploit the fact that the planner is indifferent about learning at the edge of a cascade set, and thus strictly prefers to learn if he is slightly more patient.

Proposition 7 *Assume bounded beliefs. Non-empty cascade sets strictly shrink when δ rises: For all actions a, if $\delta_2 > \delta_1$ and $J_a(\delta_1) \neq \emptyset$, then $J_a(\delta_2) \subset J_a(\delta_1)$.*

The proof in the Appendix exploits two key results, each of sufficient independent interest that we summarize them in the text (but appendicize their proofs).

Lemma 7 (Strict Value Monotonicity) *The value function increases strictly with δ outside the cascade sets: for $\delta_2 > \delta_1$, $v_{\delta_2}(\pi) > v_{\delta_1}(\pi)$ for all $\pi \notin J_1(\delta_2) \cup \cdots \cup J_A(\delta_2)$.*

Intuitively, in the experimenting region, a more patient planner enjoys a higher value.

Two rules x and $\tilde{x}$ are equivalent if they induce actions with the same chances, i.e. $\psi(a, \omega, x) = \psi(a, \omega, \tilde{x})$ for all actions a and states $\omega \in \{H, L\}$. As needed in Step 1 in §6.2, we now characterize the differentiability of the value function at the edge of cascade sets.

Lemma 8 (Differentiability) *Let $\pi^* \in (0, 1)$ be an endpoint of cascade set $J_a(\delta)$. If all rules optimal at π^* are equivalent, then v_δ is differentiable at the belief π^*.*

Since cascade sets *do* change with the discount factor, we can now explain why we proved in Lemma 1 that the social optimum is an equilibrium, but not the converse. For there exists a suboptimal team equilibrium when $\delta > 0$. Indeed, herding on action a is a team equilibrium when $\pi \in J_a(0) \setminus J_a(\delta)$. If every successor considers $\mathcal{DM}$ n's action uninformative, the best n can do is to maximize the current payoff. Since $\pi \in J_a(0)$, it is individually optimal to take action a. Yet, since $\pi \notin J_a(\delta)$, this rule is suboptimal.

A APPENDIX: OMITTED PROOFS

Proof of Corollary 2. Given the interval structure of Corollary 1 and the possibility of mixing, at any belief π the planner chooses the chances $\psi(a,\pi,\xi^\delta)$ of each action, and their order. Represent this by the product of an $A \times A$ permutation matrix and the vector $(\psi(a,\pi,\xi^\delta), a = 1,\ldots,A)$. This choice set is compact in $\mathbb{R}^{A^2}$. Since the objective function in the Bellman equation (1) is continuous in this choice matrix and in π, the claim now follows from the Theorem of the Maximum (e.g. Theorem I.B.3 of Hildenbrand (1974)). Finally, the map from these matrices to rules is continuous. □

Remaining Proof of Corollary 3. Let $\tilde{u}(\pi) = \max_a \bar{u}(a,\pi)$ be the payoff frontier.[11]

First (*iii*): By Claims 1 and 2 below, when $\delta_1 \geq \delta_2$, $v_{\delta_1}(\pi) \geq v_{\delta_2}(\pi) \geq \tilde{u}(\pi) \geq \bar{u}(a,\pi)$ for all π. For $\pi \in J_a(\delta_1)$, $v_{\delta_1}(\pi) = \bar{u}(a,\pi)$ and thus $v_{\delta_2}(\pi) = \bar{u}(a,\pi)$. Hence, $\pi \in J_a(\delta_2)$.

(*iv*): Smith and Sørensen (2000) establish for the selfish model that all $J_a(0)$ are empty, except for $J_1(0) = \{0\}$ and $J_A(0) = \{1\}$. Now apply (*ii*) and (*iii*).

(*v*): We first prove that for sufficiently low beliefs it is optimal to let the rule x induce 1; the argument for high beliefs is similar. Action 1 is optimal at belief $\pi = 0$, and there is no tie, so 1 is the optimal selfish choice for beliefs $\pi \leq \pi'$, for some $\pi' > 0$. In particular, $\bar{u}(1,\pi) > \bar{u}(a,\pi) + \eta$ for all $a \neq 1$ for some $\eta > 0$, and for all beliefs π in the interval $[0, \pi'/2]$. No action can reveal a stronger private signal than any $\sigma \in \operatorname{supp}(\mu) \subseteq [\underline{\sigma}, \bar{\sigma}] \subset (0,1)$. So any initial belief π is updated to at most $\bar{p}(\pi) = \pi\bar{\sigma}/[\pi\bar{\sigma} + (1-\pi)(1-\bar{\sigma})]$. For π small enough, $\bar{p}(\pi) \in [0, \pi'/2]$ and $\bar{p}(\pi) - \pi$ is arbitrarily small. By continuity of v_δ, $v_\delta(\bar{p}(\pi)) - v_\delta(\pi)$ is less than $\eta(1-\delta)/\delta$ for small enough π. By the Bellman equation (1), any action $a \neq 1$ is strictly suboptimal for such small beliefs.

We finally prove the last part of (*v*). Consider first any action $a \notin \{1, A\}$, and suppose δ is such that the cascade set $J_a(\delta)$ is non-empty. Let the rule x take $a-1$ and a for the private belief σ in the respective intervals $I_{a-1} = [0, \hat{\sigma}]$ and $I_a = [\hat{\sigma}, 1]$, where $1 > \mu([\hat{\sigma}, 1]) > 0$. Updating with the optimistic news that $\sigma \in I_a$ leads to an upward revision of the public belief: There exists $\varepsilon > 0$ such that $p(a,\pi,x) - \pi \geq \varepsilon$ for all $\pi \in J_a(\delta) \subset (0,1)$. Denoting by π'' the upper bound of $J_a(\delta)$, write $[\pi', \pi''] = [\pi'' - \varepsilon/2, \pi''] \cap J_a(\delta)$. Since the convex function v_δ strictly exceeds the affine $\bar{u}(a,\cdot)$ outside $J_a(\delta)$, and since $v_\delta(\pi) = \bar{u}(a,\pi)$ inside $J_a(\delta)$, there exists $\eta > 0$ so small that $\psi(a,\pi,x)v_\delta(p(a,\pi,x)) + \psi(a-1,\pi,x)v_\delta(p(a-1,\pi,x)) >$

[11] This differs from $v_0(\pi) \equiv \sup_x \sum_a \psi(a,\pi,x)\bar{u}(a,p(a,\pi,x))$ which allows the myopic individual to observe one signal σ before obtaining the ex post value $\tilde{u}(r(\pi,\sigma))$. In the example of section 6.1, $\tilde{u}(\pi) = \max\langle\pi, 1-\pi\rangle$ and $v_0(\pi) = 1 - \pi + \pi^2$.

$v_\delta(\pi)+\eta$ for all $\pi \in [\pi', \pi'']$. We prove that the interval $[\pi', \pi'']$ is excised from $J_a(\delta')$ once $\delta' > \delta$ is sufficiently large. If this were not true, then $v_{\delta'}(\pi') = \bar{u}(a, \pi')$, and there is an expected gain of at least η in the continuation value of the the Bellman equation (1) by switching from the cascade rule to rule x. For δ' sufficiently large, this continuation gain dominates any first-period loss, proving sub-optimality of the cascade rule at π', and hence in $[\pi', \pi'']$. By iterating this procedure a finite number of times, each time excising length $\varepsilon/2$ from interval $J_a(\delta)$, we see that $J_a(\delta)$ evaporates for large enough δ.

If $a = 1$ or A, apply this procedure repeatedly: $J_a(\delta) \cap I$ vanishes for δ near 1. □

The proof used two claims. The Bellman operator T_δ is given by $T_\delta v$ equal to the RHS of (1). Note that for $v \geq v'$ we have $T_\delta v \geq T_\delta v'$. As is standard, T_δ is a contraction, and v_δ is its unique fixed point in the space of bounded, continuous, weakly convex functions.

Claim 1 *The sequence $\{T_\delta^n \tilde{u}\}$ consists of pointwise increasing weakly convex functions that converge to v_δ. The value v_δ weakly exceeds $\tilde{u}$, and strictly so outside the cascade sets.*

Proof: To maximize $\sum_{a=1}^{A} \psi(a, \pi, x)\,[(1-\delta)\bar{u}_a(p(a, \pi, x)) + \delta \tilde{u}(p(a, \pi, x))]$ over x for given π, one rule $\tilde{x}$ almost surely chooses the myopically optimal action. Then $p(\tilde{x}(\sigma), \pi, \tilde{x}) = \pi$ a.s., resulting in value $\tilde{u}(\pi)$. Optimizing over all $x \in X$, $T_\delta \tilde{u}(\pi) \geq \tilde{u}(\pi)$ for all π. By induction, $T_\delta^n \tilde{u} \geq T_\delta^{n-1}\tilde{u}$, yielding a pointwise increasing sequence converging to the fixed point $v_\delta \geq \tilde{u}$. Finally, when π is outside the cascade sets, by definition it is *not* optimal to almost surely induce one action. So, $v_\delta(\pi) > \tilde{u}(\pi)$ $\forall \delta \in [0, 1)$ and $\forall \pi \notin \cup_{a=1}^{A} J_a(\delta)$. □

Claim 2 *When $\delta_1 \geq \delta_2$, $v_{\delta_1}(\pi) \geq v_{\delta_2}(\pi)$ for all π.*

Proof: Clearly, $\sum_{a=1}^{A} \psi(a, \pi, x)\bar{u}_a(p(a, \pi, x)) \leq \sum_{a=1}^{A} \psi(a, \pi, x) v(p(a, \pi, x))$ for any x and any function $v \geq \tilde{u}$. If $\delta_1 > \delta_2$, then $T_{\delta_1}\tilde{u} \geq T_{\delta_2}\tilde{u}$, since more weight is placed on the larger component of the RHS of (1). Because one possible policy under δ_1 is to choose the ξ optimal under δ_2, we have $T_{\delta_1}^n \tilde{u} \geq T_{\delta_2}^n \tilde{u}$. Let $n \to \infty$ and apply Claim 1. □

Remaining Proof of Proposition 2. If the optimal rule at π prescribes to choose action a with probability one, then $\pi \in J_a(\delta) \subset J_a(0)$, and it is myopically optimal to choose a. Hence, no transfer is needed in this case.

The remaining case is where the current $\mathcal{DM}$ is not in a herd, but some of the $\mathcal{DM}$'s actions lead to a herd starting from the successor. By Claim 3 below, if all continuation beliefs land in cascade sets, then the signal distribution must be binary, and at each of the two signal realizations the $\mathcal{DM}$ must start the herd by taking the action that the successors will forever copy. In this binary signal case, the dynamically optimal rule is again myopically optimal, and no transfer is needed. Otherwise, some but not all actions lead to a herd starting from the successor. Again, Claim 3 shows that all such actions lead into the same cascade set, say $J_{b'}(\delta)$. In this case, we let the transfer to these herd-inducing actions be zero, and modify the state-dependent payoffs to all other actions a' as $w_\delta(a',\pi,\rho)/(1-\delta) - \delta\bar{u}(b',\rho)/(1-\delta)$. This is done precisely as above, using transfers that depend on one of the $\mathcal{DM}$'s next actions b following a'. To conclude, we show that the modified payoff functions lead the selfish $\mathcal{DM}$ to choose in accordance with the planner's rule. If a and a' are two actions not leading to a cascade set, then the modified payoffs satisfy $w_\delta(a,\pi,\rho)/(1-\delta)-\delta\bar{u}(b',\rho)/(1-\delta) \geq w_\delta(a',\pi,\rho)/(1-\delta)-\delta\bar{u}(b',\rho)/(1-\delta)$ precisely when $w_\delta(a,\pi,\rho) \geq w_\delta(a',\pi,\rho)$ as desired. If a leads to a cascade set, and a' not, then the modified payoffs satisfy $\bar{u}(a,\rho) \geq w_\delta(a',\pi,\rho)/(1-\delta) - \delta\bar{u}(b',\rho)/(1-\delta)$ precisely when $w_\delta(a,\pi,\rho) = (1-\delta)\bar{u}(a,\rho) + \delta\bar{u}(b',\rho) \geq w_\delta(a',\pi,\rho)$ as desired. Finally, if both a and a' lead to the cascade set, $\bar{u}(a,\rho) \geq \bar{u}(a',\rho)$ if and only if $w_\delta(a,\pi,\rho) \geq w_\delta(a',\pi,\rho)$. □

The proof used the following claim. Let us call the private signal distribution *2S (Two Signals)* if its support contains only two isolated points.

Claim 3 (Unreachable Cascade Sets) *If 2S fails, then for any π not in any δ-cascade set: (★) at most one δ-cascade set is reached by the continuation beliefs following actions at π, and there is positive chance of an action with continuation belief not in any δ-cascade set. If 2S holds, then (★) obtains for all non-cascade beliefs π except possibly at no more than $A-1$ points, each the unique belief between any pair of nonempty myopic cascade sets $J_{a'}(0)$ and $J_{a''}(0)$ from which both cascade sets can be reached. In that case, a' is chosen when the continuation belief is in $J_{a'}(0)$, and a'' is chosen when it is in $J_{a''}(0)$.*

Proof: With unbounded beliefs and no fully revealing private signals, the continuation belief never lies in a cascade set; so assume bounded beliefs. Let $\underline{\sigma} = \min \operatorname{supp}(F)$ and $\bar{\sigma} = \max \operatorname{supp}(F)$ be the most extreme private beliefs in the support. For any π and any action a'', Bayesian updating implies that $r(\pi,\underline{\sigma}) \leq p(a,\pi,x) \leq r(\pi,\bar{\sigma})$. Consider any two nonempty myopic cascade sets $J_{a'}(0) < J_{a''}(0)$, and let $\pi' = \max J_{a'}(0)$ and $\pi'' = \min J_{a''}(0)$. By myopic optimality, $r(\pi',\bar{\sigma}) \leq r(\pi'',\underline{\sigma})$. Hence, cascade set $J_{a''}(0)$ is unreachable from any $\pi \leq \pi'$, and $J_{a'}(0)$ is unreachable from any $\pi \geq \pi''$. Consider now $\pi \in (\pi',\pi'')$. If all possible actions at π led into a cascade set, then $r(\pi,\underline{\sigma}) \leq \pi'$ and $r(\pi,\bar{\sigma}) \geq \pi''$. But these inequalities can only hold with equality since

$$r(r(\pi,\bar{\sigma}),\underline{\sigma}) \geq r(\pi'',\underline{\sigma}) \geq r(\pi',\bar{\sigma}) \geq r(r(\pi,\underline{\sigma}),\bar{\sigma}) = r(r(\pi,\bar{\sigma}),\underline{\sigma})$$

as Bayes-updating commutes. So, between $J_{a'}(0)$ and $J_{a''}(0)$ there exists at most one point $\tilde{\pi}$ which can satisfy both equations; moreover, such a point exists iff *2S* holds. Indeed, given *2S,* we may simply choose $\tilde{\pi}$ to solve $r(\tilde{\pi},\bar{\sigma}) = \pi''$, while if *2S* fails, then with positive chance, a nonextreme signal is realized, and the posterior p is not in a cascade set. With $\delta > 0$ we have weakly smaller cascade sets by Corollary 3.

Finally, assume *2S.* Consider any $\tilde{\pi}$ with reachable cascade sets $J_{a'}(\delta)$ and $J_{a''}(\delta)$. Then the rule $\tilde{x}$ mapping $\underline{\sigma}$ into a' (low signal to π') and $\bar{\sigma}$ into a'' (high signal to π'') is optimal. By convexity, $v_\delta(\pi)$ is at most the average of $v_\delta(\pi'')$ and $v_\delta(\pi')$ (weights given by transition chances), and $\tilde{x}$ achieves this average. Moreover, by myopic optimality of choosing a' (resp. a'') when the posterior belief is in $J_{a'}(0)$ (resp. $J_{a''}(0)$), this is the only optimal rule. □

Claim Needed for the Proof of Proposition 3.

Claim 4 $(1-\pi)(1-\sigma)f(\sigma)d\sigma = \gamma(\lambda,\pi)d\lambda$ *and* γ *satisfies the MLRP.*

Proof: Since $\phi(\Lambda) \equiv f(\sigma(\Lambda))\sigma'(\Lambda)$:

$$(1-\pi)(1-\sigma)f(\sigma)d\sigma = (1-\pi)(1-\sigma)\phi(\Lambda(\sigma(\lambda)))\frac{d\Lambda}{d\sigma}\frac{d\sigma}{d\lambda}d\lambda = \frac{\pi(1-\pi)\phi(\Lambda(\sigma(\lambda)))}{(\pi+\lambda(1-\pi))\lambda}d\lambda.$$

Further, the MLRP means $\partial^2 \log(\gamma(\lambda,\pi))/(\partial\lambda\partial\pi) > 0$. Indeed, this inequality holds iff the following has a positive cross partial:

$$\log\frac{\phi(\Lambda(\sigma(\lambda)))}{(\pi+\lambda(1-\pi))} = \log\phi(\Lambda(\sigma(\lambda))) - \log(\pi+\lambda(1-\pi)).$$

The second term is clear. As $\Lambda(\sigma(\lambda)) = \log(\sigma(\lambda)/(1-\sigma(\lambda))) = \log((1-\pi)\lambda/\pi)$, we have

$$\frac{\partial}{\partial\lambda}\log\phi(\Lambda(\sigma(\lambda))) = \frac{\phi'(\Lambda(\sigma(\lambda))}{\phi(\Lambda(\sigma(\lambda)))}\frac{\partial}{\partial\lambda}\Lambda(\sigma(\lambda)) = \frac{\phi'(\log((1-\pi)\lambda/\pi))}{\lambda\phi(\log((1-\pi)\lambda/\pi))}$$

which is increasing in π since $\log((1-\pi)\lambda/\pi)$ is strictly decreasing in π and since ϕ'/ϕ is a decreasing function, by the log-concavity assumption. □

Proof of Lemma 2. We prove the first relation — the other is symmetric. Put $\psi = \pi\psi^H + (1-\pi)\psi^L$, where $\psi^\omega = \mu^\omega([0,\hat{\sigma}])$ for the private belief threshold $\hat{\sigma} = r(1-\pi,\hat{\rho})$. Recalling that $d\mu^H/d\mu = 2\sigma$ and $d\mu^L/d\mu = 2(1-\sigma)$, differentiation of ψ yields $\partial\psi/\partial\hat{\rho} =$

$2f(\hat{\sigma})[\pi\hat{\sigma} + (1-\pi)(1-\hat{\sigma})](\partial\hat{\sigma}/\partial\hat{\rho})$. Differentiating the quotient $p_\ell = \pi\psi^H/\psi$, we find

$$\psi\frac{\partial p_\ell}{\partial\hat{\rho}} = \psi\frac{\partial}{\partial\hat{\sigma}}\left(\frac{\pi\psi^H}{\pi\psi^H + (1-\pi)\psi^L}\right)\frac{\partial\hat{\sigma}}{\partial\hat{\rho}} = \frac{\pi(1-\pi)[\hat{\sigma}\psi^L - (1-\hat{\sigma})\psi^H]f(\hat{\sigma})}{\psi}\frac{\partial\hat{\sigma}}{\partial\hat{\rho}}.$$

Finally, the relation obtains once we see that

$$(\hat{\rho} - p_\ell)\frac{\partial\psi}{\partial\hat{\rho}} = 2\left(\frac{\pi\hat{\sigma}}{\pi\hat{\sigma} + (1-\pi)(1-\hat{\sigma})} - \frac{\pi\psi^H}{\psi}\right)f(\hat{\sigma})[\pi\hat{\sigma} + (1-\pi)(1-\hat{\sigma})]\frac{\partial\hat{\sigma}}{\partial\hat{\rho}}. \qquad \square$$

Proof of Lemma 3. When v is affine on $[z_1, z_2]$, tangents τ_1 and τ_2 coincide, and so $\tau_1(z_3) = \tau_2(z_3)$. Otherwise, the tangents differ, with τ_2 steeper than τ_1. Since v is convex, the tangent at z_1 lies below v at z_2, and so $\tau_1(z_2) \leq v(z_2) = \tau_2(z_2)$. Also, $\tau_2(z_3) - \tau_2(z_2) > \tau_1(z_3) - \tau_1(z_2)$. Altogether, $\tau_1(z_3) < \tau_2(z_3)$. The other case is similar. □

Claim for the Proof of Proposition 4. The main step of the proof uses a claim:

Claim 5 (Differentiability) *The function B has a well-defined left derivative $B_{\hat{\rho}-}(\pi, \hat{\rho})$ obeying $B(\pi, \hat{\rho}) - B(\pi, \hat{\rho}') = \int_{\hat{\rho}'}^{\hat{\rho}} B_{\hat{\rho}-}(\pi, z)\,dz$ for $\hat{\rho}, \hat{\rho}' \in (0, 1)$.*

Proof: The only issue are the kinks (if any) of v. Being convex, it is left-differentiable, and obeys $v(p(a, \pi, \hat{\rho})) - v(p(a, \pi, \hat{\rho}')) = \int_{p(a,\pi,\hat{\rho}')}^{p(a,\pi,\hat{\rho})} v'(z-)\,dz$, by Rockafellar's Corollary 24.2.1. □

Proof of Lemma 4. Contrary to strict convexity, assume that v is affine on some $[z_1, z_2] \subseteq M$. By Claim 6 below, the same contingent strategy, and so first-period rule x, is optimal throughout $[z_1, z_2]$. Bayes-updating after any observation a continuously and monotonically maps the prior belief interval $[z_1, z_2]$ into a posterior belief image interval $[p(a, z_1, x), p(a, z_2, x)]$. At this point, the optimal contingent strategy for all the future is still constant, so that — by Claim 6 — the value function is affine on this image interval.

By Claim 7 below, at any $\pi \in [z_1, z_2]$, every solution to $\max_{\hat{\rho}\in[0,1]} \tilde{B}(\pi, \hat{\rho})$ is an optimal posterior threshold. The problem $\max_{\hat{\rho}\in[0,1]} \tilde{B}(\pi, \hat{\rho})$ can be solved easily ex post: At any posterior belief ρ, take the action a with the greatest $\tilde{w}_a(\rho)$, i.e. the same action on one side of the intersection point $\hat{\rho}$ of the affine functions $\tilde{w}_\ell$ and $\tilde{w}_h$. (Since $\pi \in M$, both actions must be taken, precluding $\hat{\rho} = 0, 1$.) Since the tangent functions at the continuation beliefs, and thus $\tilde{w}_a$, do not depend on π, the optimal solution is invariant to $\pi \in [z_1, z_2]$.

The constant rule x implies a constant private belief threshold $\hat{\sigma}$. But the relation $\hat{\rho} = r(\pi, \hat{\sigma})$ uniquely determines π, and so cannot be satisfied on all of $[z_1, z_2]$. Contradiction.□

The proof uses the following two claims.

Claim 6 *The value function is affine on an interval* $[z_1, z_2]$ *iff the same contingent strategy is optimal starting from anywhere in that interval.*

Proof: The optimal contingent strategy for all the future, starting from any belief $\pi \in [z_1, z_2]$, yields some state-dependent values $\bar{v}^H$ and $\bar{v}^L$. The line joining the points $(0, \bar{v}^L)$ and $(1, \bar{v}^H)$ is the tangent to v at π. As in the proof of Proposition 1, the same contingent strategy is optimal starting anywhere in $[z_1, z_2]$. Conversely, holding constant the contingent strategy for all the future merely adjusts the chances of $\bar{v}^H$ and $\bar{v}^L$. □

For the next claim, let $\hat{\rho}^* \in [0, 1]$ be an optimal posterior threshold at belief $\pi \in [0, 1]$, inducing continuation beliefs p_a after seeing action $a = \ell, h$. Define the affine function $\tilde{w}_a(\hat{\rho}) = (1 - \delta)\bar{u}_a(\hat{\rho}) + \delta\tau_a(\hat{\rho})$ for action $a = \ell, h$, where τ_a is the tangent to v at the continuation belief p_a after action $a = \ell, h$.

Claim 7 (An Auxiliary Affine Optimization) *The posterior belief threshold* $\hat{\rho} \in [0, 1]$ *is an optimal posterior threshold belief at* π *if* $\hat{\rho}$ *maximizes*

$$\tilde{B}(\pi, \hat{\rho}) = \psi(\pi, \hat{\rho})\tilde{w}_\ell(\hat{\rho}) + (1 - \psi(\pi, \hat{\rho}))\tilde{w}_h(\hat{\rho}).$$

Also, $\hat{\rho}^*$ *solves* $\max_{\hat{\rho} \in [0,1]} \tilde{B}(\pi, \hat{\rho})$.

Proof: Since v is convex, the tangent functions are weakly below v. Then all $\hat{\rho}$ satisfy $B(\pi, \hat{\rho}) \geq \tilde{B}(\pi, \hat{\rho})$. Since $\tau_a(p_a) = v(p_a)$, we also have $\tilde{B}(\pi, \hat{\rho}^*) = B(\pi, \hat{\rho}^*)$. Thus, if $\tilde{B}(\pi, \hat{\rho}) \geq (>) \tilde{B}(\pi, \hat{\rho}^*)$ then $B(\pi, \hat{\rho}) \geq \tilde{B}(\pi, \hat{\rho}) \geq (>) \tilde{B}(\pi, \hat{\rho}^*) = B(\pi, \hat{\rho}^*)$. □

Matrix Result Needed in Proof of Proposition 5.

Claim 8 *Let* B *be a negative definite real matrix with non-negative off-diagonal elements. Then all entries in* B^{-1} *are nonpositive, and strictly negative on the diagonal.*

Proof: As the inverse B^{-1} of a negative definite matrix exists and is negative definite, the diagonal elements of B^{-1} are negative. That $B^{-1} \leq 0$ can be concluded from Theorems 2′ and 4 of McKenzie (1960). To keep the presentation self-contained, we offer here an alternative proof. For any vector $\beta \geq 0$, consider the function $W(c, d) = c'Bc + d\beta'c$ where $d \in \mathbb{R}$. Note that $W_c(c, d) = 2c'B + d\beta'$. Then W is supermodular in c since B has all non-negative off-diagonal elements. Also, W has increasing differences in (c, d) since $W_{cd} = \beta' \geq 0$. Since B has full rank, the first order condition $2Bc = -d\beta$ has a unique solution $c^*(d)$. Since $W_{cc} = B$ is negative definite, $c^*(d)$ is the unique maximizer of W. The monotone comparative statics result of Topkis implies that $c^*(d)$ is weakly increasing.

Application of the implicit function theorem to $2Bc + d\beta = 0$ gives $c_d^*(d) = -(1/2)B^{-1}\beta$. Since $c^*(d)$ is increasing, $B^{-1}\beta \leq 0$. Since $\beta \geq 0$ is arbitrary, we conclude that $B^{-1} \leq 0$. □

Remaining Proof of Lemma 7. By Claim 1 above, $v_{\delta_1}(\pi) > \tilde{u}(\pi)$ for π outside the δ_1-cascade sets. Fix π outside the δ_2-cascade sets. If π lies in a δ_1-cascade set we're done, as $v_{\delta_1}(\pi) = \tilde{u}(\pi) < v_{\delta_2}(\pi)$. Suppose π lies outside the δ_1-cascade sets.

Assume first that π satisfies (★) of Claim 3 for δ_1 (and thus also for δ_2). Then at least one action a is taken with positive chance inducing a belief $p(\pi, \xi^{\delta_1}(\pi), a)$ not in a δ_1-cascade set. Thus, $v_{\delta_1}(p(\pi, \xi^{\delta_1}(\pi), a)) > \tilde{u}(p(\pi, \xi^{\delta_1}(\pi), a))$. Since $\delta_2 > \delta_1$,

$$v_{\delta_1}(\pi) = (T_{\delta_1} v_{\delta_1})(\pi) < (T_{\delta_2} v_{\delta_1})(\pi) \leq (T_{\delta_2} v_{\delta_2})(\pi) = v_{\delta_2}(\pi). \tag{11}$$

Next assume that some $\tilde{\pi}$ between $J_{a'}(\delta_1)$ and $J_{a''}(\delta_1)$ fails (★) for δ_1. If (11) holds at $\tilde{\pi}$, we are done. Assume not. Claim 3 noted that between consecutive cascade sets such $\tilde{\pi}$ must be unique, and that it implied *2S*. In that case, (11) holds in a punctured neighbourhood $(\pi', \tilde{\pi}) \cup (\tilde{\pi}, \pi'')$ of $\tilde{\pi}$, where $\pi' = \max J_{a'}(\delta_1)$ and $\pi'' = \min J_{a''}(\delta_1)$. Also, from the last paragraph of Claim 3's proof, v_{δ_1} was everywhere an affine function on $[\pi', \pi'']$, which in turn, is a supporting tangent line to the convex function v_{δ_2} at $\tilde{\pi}$ (see Claim 2). As it touches v_{δ_2} at $\tilde{\pi}$ only, $v_{\delta_2}(\pi') > v_{\delta_1}(\pi')$ and $v_{\delta_2}(\pi'') > v_{\delta_1}(\pi'')$.

To find a lower bound to $v_{\delta_2}(\tilde{\pi})$, apply rule $\tilde{x}$ from Claim 3's proof at the belief $\tilde{\pi}$. Since $\tilde{x}$ maps $\underline{\sigma}$ into $\pi' \in J_{a'}(\delta_1)$ and $\bar{\sigma}$ into $\pi'' \in J_{a''}(\delta_1)$, it yields myopic first-period values $\bar{u}_{a'}(\pi') = v_{\delta_1}(\pi')$ and $\bar{u}_{a''}(\pi'') = v_{\delta_1}(\pi'')$, and continuation values $v_{\delta_2}(\pi')$ and $v_{\delta_2}(\pi'')$. By the right hand side of (1), this mixture is worth strictly more than $v_{\delta_1}(\tilde{\pi})$:

$$\begin{aligned} v_{\delta_1}(\tilde{\pi}) &= \psi(a', \tilde{\pi}, \tilde{x}) v_{\delta_1}(\pi') + \psi(a'', \tilde{\pi}, \tilde{x}) v_{\delta_1}(\pi'') \\ &< \psi(a', \tilde{\pi}, \tilde{x}) \left[(1-\delta_2) v_{\delta_1}(\pi') + \delta_2 v_{\delta_2}(\pi')\right] + \psi(a'', \tilde{\pi}, \tilde{x}) \left[(1-\delta_2) v_{\delta_1}(\pi'') + \delta_2 v_{\delta_2}(\pi'')\right]. \end{aligned}$$

Given this contradiction, (11) must hold at $\tilde{\pi}$. □

Proof of Lemma 8. Starting at any π, an optimal strategy yields state-contingent values $\bar{v}^L(\pi)$ and $\bar{v}^H(\pi)$. The affine function which is $\bar{v}^L(\pi)$ at 0 and $\bar{v}^H(\pi)$ at 1 is then tangent to the value function at π.

Assume now that the value function v is not differentiable at $\tilde{\pi}$, and that all optimal rules at $\tilde{\pi}$ are equivalent. We show this leads to a contradiction. Since it is optimal to take a forever at $\tilde{\pi}$, one tangent to v at $\tilde{\pi}$ is the affine function $\bar{u}(a, \rho)$. Either the left or the right derivative of v at $\tilde{\pi}$ must differ from the slope of $\bar{u}(a, \rho)$, say the left derivative (when the left derivative differs, necessarily $a > 1$). Denote the left tangent function by τ, then $\tau(0) > \bar{u}(a, 0) = u(a, L)$ and $\tau(1) < \bar{u}(a, 1) = u(a, H)$. Since $u(1, L) \geq \tau(0) > u(a, L)$, a unique $k > 0$ exists satisfying $\tau(0) = ku(1, L) + (1 - k)u(a, L)$.

As v is convex, it is differentiable almost everywhere. When v is differentiable at π, the tangent function is uniquely determined, being $\bar{v}^L(\pi)$ at 0 and $\bar{v}^H(\pi)$ at 1. Since v is convex, $\bar{v}^L(\pi) \geq \tau(0)$ and $\bar{v}^H(\pi) \leq \tau(1)$ when $\pi < \tilde{\pi}$.

Now choose N so large and ε so small that $k/2 \geq 1 - (1 - \delta^N)(1 - \varepsilon)$. Note that 1 is strictly the best action in state L. Then by Claim 9 below, for all $\pi < \tilde{\pi}$ close enough to $\tilde{\pi}$, the optimal state-contingent values $\bar{v}^L(\pi)$ is at most

$$\begin{aligned}
\bar{v}^L(\pi) &\leq (1 - \delta^N)(1 - \varepsilon)u(a, L) + [1 - (1 - \delta^N)(1 - \varepsilon)]u(1, L) \\
&\leq (1 - k/2)u(a, L) + (k/2)u(1, L) \\
&< (1 - k)u(a, L) + ku(1, L) = \tau(0) \leq \bar{v}^L(\pi)
\end{aligned}$$

since $u(1, L) > u(a, L)$, as noted above. Contradiction. □

The proof of Lemma 8 used the following claim.

Claim 9 $\forall N \in \mathbb{N}\ \forall \varepsilon > 0\ \exists \eta > 0$ *so that if* $|\pi - \tilde{\pi}| < \eta$ *then under any optimal strategy from* π*, action* a *is taken for the first* N *periods with chance at least* $1 - \varepsilon$ *in states* H, L.

Proof: Fix $\varepsilon < 1/3$. By Claim 10, for π_n near $\tilde{\pi}$, action a occurs with chance at least $1 - \varepsilon$ in each state starting from π_n. If a occurs, then π_{n+1} obeys $|\pi_{n+1} - \pi_n| \leq 4\tilde{\pi}(1 - \tilde{\pi})\eta$, by Bayes rule. So $|\pi_{n+1} - \pi_n|$ can be chosen arbitrarily small when a occurs, for π_n near $\tilde{\pi}$.

Choose the initial π so close to $\tilde{\pi}$ that if a occurs for the next N periods, the posterior stays so close to $\tilde{\pi}$ that a occurs with conditional chance at least $(1 - \varepsilon)^{1/N}$ each period. □

Claim 10 $\forall \varepsilon > 0\ \exists \eta > 0$ *such that when* $|\pi - \tilde{\pi}| < \eta$*, any optimal rule at* π *induces action* a *with chance at least* $1 - \varepsilon$ *in states* H, L.

Proof: Since $\tilde{\pi} \in J_a(\delta)$, one optimal rule at $\tilde{\pi}$ induces a with chance one. By the equivalence assumption, this property is shared by all rules optimal at $\tilde{\pi}$. Next, if $\psi(a, \pi, x) = \pi\psi(a, H, x) + (1 - \pi)\psi(a, L, x)$ is near 1, so are both $\psi(a, H, x)$ and $\psi(a, L, x)$ for π near $\tilde{\pi}$. The claim follows from the upper hemi-continuity of Corollary 2. □

Proof of Proposition 6 We first cite the extended (conditional) Second Borel-Cantelli Lemma in Corollary 5.29 of Breiman (1968): Let $Y_1, Y_2, \ldots$ be any stochastic process, and $D_n \in \mathcal{F}(Y_1, \ldots, Y_n)$, the induced sigma-field. Then almost surely

$$\{D_n \text{ infinitely often (i.o.)}\} = \{\sum_{n=1}^{\infty} P(D_{n+1}|Y_n,\ldots,Y_1) = \infty\}.$$

Now, fix an optimal policy ξ^δ. Choose $\varepsilon > 0$ to satisfy Claim 11 below for all actions $1,2,\ldots,A$. For fixed a, define events $E_n = \{\pi_n$ is ε-close to $J_a(\delta)\}$, $F_n = \{\psi(a,\pi_n,\xi^\delta(\pi_n)) < 1-\varepsilon\}$, and $G_{n+1} = \{|\pi_{n+1}-\pi_n| > \varepsilon\}$. If $E_n \cap F_n$ is true, then scenario (*ii*) in Claim 11 obtains, and so $P(G_{n+1}|\pi_n) \geq \varepsilon/A$. Then $\sum_{n=1}^{\infty} P(G_{n+1}|\pi_1,\ldots,\pi_n) = \infty$ conditional on $E_n \cap F_n$ i.o. By the above Borel-Cantelli Lemma, almost surely G_n obtains i.o. conditional on $E_n \cap F_n$ i.o. But since $\langle \pi_n \rangle$ almost surely converges by Lemma 5, G_n i.o. is a zero chance event. By implication, $E_n \cap F_n$ i.o. has probability zero.

Consider the event C that $\langle \pi_n \rangle$ has a limit in $J_a(\delta)$ and $E_n \cap F_n$ occurs only finitely often. By definition, C implies that eventually $E_n \setminus (F_n \cup G_{n+1})$. But $E_n \setminus F_n$ implies that every action $a' \neq a$ leads to G_{n+1}, by Claim 11 (*i*). Action a is then eventually taken on C. Sum over all a to get a chance one event, by Corollary 3 and Lemma 5. □

We used a claim that generalizes the Overturning Principle of the herding model to the forward-looking model. When π is near $J_a(\delta)$, action a should occur with high chance. More precisely, any other actions distinctly shift beliefs, *or* there was a non-negligible probability of observing some other action which would distinctly shift beliefs.

Claim 11 (Overturning Principle) *For $\delta \in [0,1)$, assume $J_a(\delta) \neq \emptyset$. Then there exists $\varepsilon > 0$ and an ε-neighbourhood $K \supset J_a(\delta)$, such that $\forall \pi \in K \cap (0,1)$, either:*
(i) $\psi(a,\pi,\xi^\delta(\pi)) \geq 1-\varepsilon$, and $|p(a',\pi,\xi^\delta(\pi)) - \pi| > \varepsilon$ for all $a' \neq a$ that occur; or
(ii) $\psi(a,\pi,\xi^\delta(\pi)) < 1-\varepsilon$, and $\psi(a',\pi,\xi^\delta(\pi)) \geq \varepsilon/A$, $|p(a',\pi,\xi^\delta(\pi)) - \pi| > \varepsilon$ for some a'.

Proof: Choose $\eta > 0$ small enough such that for any π sufficiently close to $J_a(\delta)$, we have $\psi(a',\pi,\xi^\delta(\pi)) < 1-\eta$ for any $a' \neq a$. If such η does not exist, since the optimal rule correspondence is u.h.c., almost surely taking action a' is optimal at some $\tilde{\pi} \in J_a(\delta)$. This is impossible, as a' incurs a strict myopic loss, and captures no information gain.

First, assume bounded private beliefs. By (*iv*) of Corollary 3, for π close enough to 0 or 1, the *only* optimal rule is to stop learning. Thus, we need only consider π in some closed subinterval I of $(0,1)$. Recall $\underline{\sigma} = \min \operatorname{supp}(F)$ and $\bar{\sigma} = \max \operatorname{supp}(F)$. By the existence of informative beliefs, $\underline{\sigma} < 1/2 < \bar{\sigma}$. Let $\varepsilon > 0$ be the minimum of η, $\mu^H([\underline{\sigma},(2\underline{\sigma}+1)/4])$, and $\mu^L([(2\bar{\sigma}+1)/4,\bar{\sigma}])$ (notice that $(2\sigma+1)/4$ is the midpoint between σ and $1/2$).

Assume $\psi(a,\pi,\xi^\delta(\pi)) \geq 1-\varepsilon$ for some $\pi \in I$. By Corollary 1, any action $a' \neq a$ is a.s. only taken for beliefs within either $[\underline{\sigma},(2\underline{\sigma}+1)/4]$ or $[(2\bar{\sigma}+1)/4,\bar{\sigma}]$. Any such a' implies case (*i*) (selecting, if necessary, ε even smaller).

If instead $\psi(a,\pi,\xi^\delta(\pi)) < 1-\varepsilon$, then each action is taken with chance less than $1-\varepsilon$. By construction of ε, different actions are taken at the two extreme private beliefs (by the interval structure of the optimal rule). At least one of the A actions occurs with chance at least ε/A, does not include private beliefs near 1/2, and therefore moves beliefs by at least ε (selecting, if necessary, ε even smaller), as claimed in case (ii).

Next consider unbounded private beliefs. Let the absolute slope of the value function v have upper bound κ. Since no two payoffs are tied at 0, there exists a small $\zeta > 0$ such that the myopic action payoffs $\bar{u}(a,\rho)$ maintain the same ranking, and the difference $|\bar{u}(\tilde{a},\rho) - \bar{u}(\check{a},\rho)|$ exceeds $\kappa\zeta$ for all $\tilde{a} \neq \check{a}$, for all $\rho \in [0,\zeta]$.

Assume that π is near the cascade set $\{0\}$ — the other case, $\{1\}$, is similar. Then only one a'' can have low continuation belief $p(a'',\pi,\xi^\delta(\pi)) \in [0,\zeta]$. If not, consider the altered policy that redirects private beliefs from two such actions into the myopically higher of the two. This yields a first-period payoff gain of more than $\kappa\zeta$, and a future value loss of at most $\kappa\zeta$ (for p remains in $[0,\zeta]$). So the altered policy is a strict improvement.

Assume $\psi(1,\pi,\xi^\delta(\pi)) \geq 1-\varepsilon$. Then $p(1,\pi,\xi^\delta(\pi)) \leq \pi/(1-\varepsilon) \leq \zeta$, for small enough π and ε. As only action $a''=1$ has continuation belief in $[0,\zeta]$, case (i) is satisfied.

Finally, assume $\psi(1,\pi,\xi^\delta(\pi)) < 1-\varepsilon$. Then $\psi(a'',\pi,\xi^\delta(\pi)) < 1-\varepsilon$. Otherwise, $a'' \neq 1$ and a myopic gain of at least $(1-\varepsilon)\zeta - \varepsilon U$ obtains from swapping the private beliefs for 1 and a'', without any change in future value (here U denotes the maximal possible myopic payoff difference). Thus there is a gain if ε is small enough: contradiction. Since $\psi(a'',\pi,\xi^\delta(\pi)) < 1-\varepsilon$ there must exist some other action taken with chance at least ε/A yielding continuation belief outside $[0,\zeta]$. Thus, case (ii) holds. □

Proof of Proposition 7. Since $J_a(\delta_2) \subseteq J_a(\delta_1)$ by Corollary 3, and $J_a(\delta) = \{\pi | v_\delta(\pi) - \bar{u}_a(\pi) = 0\}$ is closed by continuity, we need only prove $\breve{\pi} \equiv \min J_a(\delta_1) \notin J_a(\delta_2)$.

Case 1. Assume that at public belief $\breve{\pi}$ and with discount factor δ_1, some optimal rule $\tilde{x}$ does not a.s. take action a. Instead, with positive chance, $\tilde{x}$ takes some action a' producing a posterior $p(a',\breve{\pi},\tilde{x})$ not in any δ_1-cascade set. So from Lemma 7, $v_{\delta_2}(p(a',\breve{\pi},\tilde{x})) > v_{\delta_1}(p(a',\breve{\pi},\tilde{x})) \geq \bar{u}_{a''}(p(a',\breve{\pi},\tilde{x}))$ for all a'', and as we can always employ the rule $\tilde{x}$ with the discount factor δ_2,

$$\begin{aligned} v_{\delta_2}(\breve{\pi}) &\geq \textstyle\sum_{a''=1}^{A} \psi(a'',\breve{\pi},\tilde{x})\,[(1-\delta_2)\bar{u}_{a''}(p(a'',\breve{\pi},\tilde{x})) + \delta_2 v_{\delta_2}(p(a'',\breve{\pi},\tilde{x}))] \\ &> \textstyle\sum_{a''=1}^{A} \psi(a'',\breve{\pi},\tilde{x})\,[(1-\delta_1)\bar{u}_{a''}(p(a'',\breve{\pi},\tilde{x})) + \delta_1 v_{\delta_1}(p(a'',\breve{\pi},\tilde{x}))] = v_{\delta_1}(\breve{\pi}). \end{aligned}$$

Consequently, we have $v_{\delta_2}(\breve{\pi}) > v_{\delta_1}(\breve{\pi}) = \bar{u}_a(\breve{\pi})$ and so $\breve{\pi} \notin J_a(\delta_2)$.

CASE 2. Next suppose that the optimal rule at $\breve{\pi}$ with discount factor δ_1 is unique. Then the partial derivative $v'_{\delta_1}(\breve{\pi})$ exists by Lemma 8. By the convexity of the value function, any selection from the subdifferential $\partial v_{\delta_1}(\pi)$ converges to $v'_{\delta_1}(\breve{\pi})$ as π increases to $\breve{\pi}$. Since the optimal rule correspondence is upper hemicontinuous by the Maximum Theorem, and uniquely valued at $\breve{\pi}$, the posterior belief $p(a', \pi, \xi^{\delta_1}(\pi))$ is continuous in π at $\breve{\pi}$ for any optimal selection ξ^{δ_1} and any action a'.

Recall $\underline{\sigma} = \min \operatorname{supp}(\mu)$. As the optimal rule at $\breve{\pi}$ almost surely prescribes action a, we let $p(a, \breve{\pi}, \xi^{\delta_1}(\breve{\pi})) = \breve{\pi}$, and $p(a'', \breve{\pi}, \xi^{\delta_1}(\breve{\pi})) = r(\breve{\pi}, \underline{\sigma}) \equiv \breve{\rho}$ for $a'' \neq a$. By their definition, $w_{\delta_1}(a, \pi, \rho)$ and $w_{\delta_1}(a'', \pi, \rho)$ are then jointly continuous in (π, ρ) at $(\breve{\pi}, \breve{\rho})$. [In the expression for w_{δ_1}, $m_{\delta_1}(a'', \rho)$ lies between the slopes of $\bar{u}_1$ and $\bar{u}_A$, and is multiplied by a function that is continuous and vanishing at $(\breve{\pi}, \breve{\rho})$, given $p(a'', \breve{\pi}, \xi^{\delta_1}(\breve{\pi})) = \breve{\rho}$.] Also, $w_{\delta_1}(a, \breve{\pi}, \breve{\rho}) \geq w_{\delta_1}(a'', \breve{\pi}, \breve{\rho})$ since $\breve{\pi}$ lies in the cascade set $J_a(\delta_1)$, while $w_{\delta_1}(a, \pi, \breve{\rho}) < w_{\delta_1}(a'', \pi, \breve{\rho})$ for $\pi < \breve{\pi}$, since $\breve{\pi}$ is the endpoint of $J_a(\delta_1)$. So $w_{\delta_1}(a, \breve{\pi}, \breve{\rho}) = w_{\delta_1}(a'', \breve{\pi}, \breve{\rho})$ by continuity. This equality can be rewritten in a very useful form:

$$\bar{u}_a(\breve{\rho}) - \bar{u}_{a''}(\breve{\rho}) = \delta_1 [\bar{u}_a(\breve{\rho}) - \bar{u}_{a''}(\breve{\rho}) + v_{\delta_1}(\breve{\rho}) - v_{\delta_1}(\breve{\pi}) - m_{\delta_1}(a, \breve{\pi})(\breve{\rho} - \breve{\pi})] . \tag{12}$$

Moreover, from Proposition 1, $m_{\delta_1}(a, \breve{\pi})$ is the slope of $\bar{u}_a$, because the function $\bar{v}_\delta(a, \breve{\pi}, \rho) = v_{\delta_1}(\breve{\pi}) + m_{\delta_1}(a, \breve{\pi})(\rho - \breve{\pi})$ evaluates the prospect of taking action a forever.

We prove $w_{\delta_2}(a, \breve{\pi}, \breve{\rho}) < w_{\delta_2}(a'', \breve{\pi}, \breve{\rho})$, and so conclude $\breve{\pi} \notin J_a(\delta_2)$. If not, assume $w_{\delta_2}(a, \breve{\pi}, \breve{\rho}) \geq w_{\delta_2}(a'', \breve{\pi}, \breve{\rho})$, i.e. $\breve{\pi} = \min J_a(\delta_2)$. Subtracting $w_{\delta_1}(a, \breve{\pi}, \breve{\rho}) \geq w_{\delta_1}(a'', \breve{\pi}, \breve{\rho})$, we have the contradiction:

$$\begin{aligned}
0 &\geq [w_{\delta_2}(a, \breve{\pi}, \breve{\rho}) - w_{\delta_2}(a'', \breve{\pi}, \breve{\rho})] - [w_{\delta_1}(a, \breve{\pi}, \breve{\rho}) - w_{\delta_1}(a'', \breve{\pi}, \breve{\rho})] \\
&= (\delta_2 - \delta_1)[\bar{u}_a(\breve{\rho}) - \bar{u}_{a''}(\breve{\rho}) - v_{\delta_1}(\breve{\pi})] + \delta_2 v_{\delta_2}(\breve{\rho}) - \delta_1 v_{\delta_1}(\breve{\rho}) - \delta_2 m_{\delta_2}(a, \breve{\pi})(\breve{\rho} - \breve{\pi}) + \delta_1 m_{\delta_1}(a, \breve{\pi})(\breve{\rho} - \breve{\pi}) \\
&> (\delta_2 - \delta_1)[\bar{u}_a(\breve{\rho}) - \bar{u}_{a''}(\breve{\rho}) - v_{\delta_1}(\breve{\pi})] + \delta_2 v_{\delta_1}(\breve{\rho}) - \delta_1 v_{\delta_1}(\breve{\rho}) - \delta_2 m_{\delta_1}(a, \breve{\pi})(\breve{\rho} - \breve{\pi}) + \delta_1 m_{\delta_1}(a, \breve{\pi})(\breve{\rho} - \breve{\pi}) \\
&= (\delta_2 - \delta_1)\left[\bar{u}_a(\breve{\rho}) - \bar{u}_{a''}(\breve{\rho}) - v_{\delta_1}(\breve{\pi}) + v_{\delta_1}(\breve{\rho}) - m_{\delta_1}(a, \breve{\pi})(\breve{\rho} - \breve{\pi})\right] \\
&= (\delta_2 - \delta_1)\left[\bar{u}_a(\breve{\rho}) - \bar{u}_{a''}(\breve{\rho})\right] / \delta_1 \geq 0.
\end{aligned}$$

Indeed, when $\breve{\pi} \in J_a(\delta_2)$, one optimal policy ξ^{δ_2} induces a almost surely at belief $\breve{\pi}$, so that $p(a'', \breve{\pi}, \xi^{\delta_1}(\breve{\pi})) = p(a'', \breve{\pi}, \xi^{\delta_2}(\breve{\pi})) = \breve{\rho}$. The first equality then follows from (2) for each index, and $v_{\delta_2}(\breve{\pi}) = v_{\delta_1}(\breve{\pi})$ when $\breve{\pi} \in J_a(\delta_1) \cap J_a(\delta_2)$. The second exploits $m_{\delta_1}(a, \breve{\pi}) = m_{\delta_2}(a, \breve{\pi})$ (true as both are the slope of $\bar{u}_a$), and $v_{\delta_2}(\breve{\rho}) > v_{\delta_1}(\breve{\rho})$, as given by Lemma 7. The final inequality follows since $\breve{\pi} \in J_a(\delta_1) \subseteq J_a(0)$, so that a is myopically optimal at $\breve{\rho}$.

References

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991): "Optimal Learning by Experimentation," *Review of Economic Studies*, 58, 621–654.

Amir, R. (1996): "Sensitivity Analysis of Multisector Optimal Economic Dynamics," *Journal of Mathematical Economics*, 25, 123–141.

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

Bertsekas, D. (1987): *Dynamic Programming: Deterministic and Stochastic Models*. Prentice Hall, Englewood Cliffs, N.J.

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

Breiman, L. (1968): *Probability*. Addison-Wesley, Reading, Mass.

Bru, L., and X. Vives (2002): "Informational Externalities, Herding and Incentives," *Journal of Institutional and Theoretical Economics*, 158, 91–105.

Burdett, K. (1996): "Truncated Means and Variances," *Economics Letters*, 52, 263–267.

Chamley, C., and D. Gale (1994): "Information Revelation and Strategic Delay in a Model of Investment," *Econometrica*, 62, 1065–1085.

Dow, J. (1991): "Search Decisions with Limited Memory," *Review of Economic Studies*, 58, 1–14.

Doyle, M. (2002): "Informational Externalities, Strategic Delay, and the Search for Optimal Policy," ISU Economics Working Paper.

Easley, D., and N. Kiefer (1988): "Controlling a Stochastic Process with Unknown Parameters," *Econometrica*, 56, 1045–1064.

Fusselman, J. M., and L. J. Mirman (1993): "Experimental Consumption for a General Class of Disturbance Densities," in *General Equilibrium, Growth, and Trade II*, ed. by J. G. et al., pp. 367–392. Academic Press, San Diego.

Gittins, J. C. (1979): "Bandit Processes and Dynamical Allocation Indices," *Journal of the Royal Statistical Society, Series B*, 41, 148–177.

Hildenbrand, W. (1974): *Core and Equilibria of a Large Economy*. Princeton University Press, Princeton.

MARSHALL, A. W., AND I. OLKIN (1979): *Inequalities: Theory of Majorization and Its Applications*. Academic Press, San Diego.

MCKENZIE, L. (1960): "Matrices with Dominant Diagonals and Economic Theory," in *Mathematical Methods in the Social Sciences, 1959*, ed. by A. et al. Stanford University Press, Stanford.

MEDRANO, L. A., AND X. VIVES (2001): "Strategic Behavior and Price Discovery," *Rand Journal of Economics*, 32, 221–248.

RADNER, R. (1962): "Team Decision Problems," *Annals of Mathematical Statistics*, 33, 857–881.

ROTHSCHILD, M. (1974): "A Two-Armed Bandit Theory of Market Pricing," *Journal of Economic Theory*, 9, 185–202.

SGROI, D. (2002): "Optimizing Information in the Herd: Guinea Pigs, Profits, and Welfare," *Games and Economic Behavior*, 39, 137–166.

SHILLER, R. J. (1995): "Conversation, Information, and Herd Behavior," *American Economic Review*, 85, 181–185.

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

STOKEY, N. L., AND R. E. LUCAS (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Mass.

VIVES, X. (1993): "How Fast do Rational Agents Learn?," *Review of Economic Studies*, 60, 329–347.

——— (1997): "Learning from Others: A Welfare Analysis," *Games and Economic Behavior*, 20, 177–200.