DISCUSSION PAPERS
Department of Economics
University of Copenhagen

05-13

Informational Herding and Optimal Experimentation

Lones Smith
Peter Norman Sørensen

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

Informational Herding and Optimal Experimentation[*]

Lones Smith[†]
Department of Economics
University of Michigan

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

August 30, 2005

Abstract

We show that far from capturing a formally new phenomenon, informational herding is really a special case of single-person experimentation — and 'bad herds' the typical failure of complete learning. We then analyze the analogous *team equilibrium*, where individuals maximize the present discounted welfare of posterity. To do so, we generalize Gittins indices to our non-bandit learning problem, and thereby characterize when *contrarian* behaviour arises: (*i*) While herds are still constrained efficient, they arise for a strictly smaller belief set. (*ii*) A log-concave log-likelihood ratio density robustly ensures that individuals should lean more against their myopic preference for an action the more popular it becomes.

[*]We thank Patrick Bolton and three referees in guiding this radical revision, as well as Abhijit Banerjee, Katya Malinova, Meg Meyer, Christopher 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, and Erasmus University 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.norman.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 each may condition his decision both on his endowed private signal about the state of the world, and on all predecessors' decisions. But crucially, he cannot see their private signals.

In this setting, Bikhchandani, Hirshleifer, and Welch (1992) (BHW) and Banerjee (1992) showed that a *herd* eventually arises — eventually, all decision-makers ($\mathcal{DM}$s) make the same choice, possibly unwise. This simple pathological outcome has understandably attracted much attention. Clarifying a claim of BHW, we showed in Smith and Sørensen (2000) (SS) that beliefs converge in the limit upon a *cascade set,* where only one action is taken with probability one. SS proved that cascade sets have interior points — and so bad herds may occur — iff the $\mathcal{DM}$s' private signals are uniformly bounded in strength.

Efficiency and Herding. This paper attempts a definitive welfare analysis of the informational herding model. It has often been expressed that herding is a socially inefficient outcome. We are aware of no general formalization or proof of this intuitive claim for the original herding setting. For instance, Banerjee loosely justifies the 'herding externality' label by showing that it fails to maximize the long run average of payoffs. But this weak statistical claim quite restrictively demands that no weight be placed on current payoffs.[1] Moreover, deducing the efficient solution for this simple payoff objective is problematic: Banerjee's proposed remedy for the externality is simply to exclude early individuals from viewing others' actions at all. We are able to characterize efficient forward-looking behaviour, shedding light on the informational herding externality. We contrast behaviour in the selfish or myopic *herding equilibrium,* and the *team equilibrium* (Radner 1962), where everyone maximizes the present discounted welfare of individuals.

The overarching thesis that we formalize and prove is that it is efficient for individuals to behave in a *contrarian* fashion. By this, we suggestively mean that 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.

First, in the long-run, the belief cascade sets — namely, public beliefs where all further

[1] Vives (1993) introduces a repeated market setting, in which many agents period-by-period come to reveal their original private information through noisy observations. Vives (1997) studies the team problem in this setting, where every agent takes into account that more aggressive trading helps to reveal useful information to the other agents. We adopt here an analogous approach to the herding model.

learning stops — are smaller. For we show that in the team equilibrium, cascade sets monotonically strictly shrink as the discount factor rises. Only in the extreme patience limit are incorrect herds impossible. But we also establish a limit to contrarian behaviour. While incorrect herds might be seen as a market failure, they are in fact constrained-efficient: Cascade sets only vanish as the discount factor tends to one: Even when $\mathcal{DM}$s are very patient, weighting immediate gains very little, informational herds arise in the team equilibrium; with bounded private beliefs, such herds are misguided (ex post) with positive chance. So herding owes to the inability of individuals to signal private information by finitely many actions, and not to their selfishness.

Second, in the major result of this paper, we analyze behavior in the short run — that is, at any finite stage before the limit. There we must formalize a notion of contrarianism at the margin: The more the weight of history favours taking an action, the more likely an informed $\mathcal{DM}$ should lean against taking it. This marginal penalty for 'popular' actions is optimal only under a plausible new informational assumption: The unconditional signal distribution must have a log-concave density of the log-likelihood ratio. We believe that this is the first use of log-concavity in the experimentation or herding literatures.

Analyzing behaviour away from the limit in a discrete time dynamic optimization model is well-known to be hard. A famous exception is Gittins' (1979) multi-armed bandit indices, where the optimal decision rule is to choose the action with the highest index. In another key contribution of this paper, we have modified his methodology for our setting. So unlike Gittins' perfect information context, an action's index is not simply its present discounted social value. Rather, individuals' signals are hidden from view, and so social rewards must be translated into private incentives using the marginal social value, to produce the privately estimated social value of each action. Our contrarianism comparative static succeeds by showing that the difference of consecutive indices, which measures the marginal return to a higher posterior threshold, is monotonic in public beliefs. We have also used this new result throughout the analysis as a simplifying tool. For instance, it affords an easy proof of the optimality of belief interval rules.

Experimentation Déjà Vu. This paper was sparked by a simple question about informational herding: Haven't we seen this before? We were piqued by its similarity 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. To us, this had the clear ring of: (i) an action herd occurs, and (ii) with positive chance is misguided.

This paper begins by formally justifying this intuitive link. We prove that informational herding is not a new phenomenon, but a camouflaged context for this old one: single person experimentation, with possible incomplete learning. Our proof respects the herding paradigm quintessence that predecessors' signals be hidden from view. In a nutshell, we replace all $\mathcal{DM}$s by agent machines that automatically map any realized private signals into action choices; the true experimenter then must furnish these automata with optimal history-contingent 'decision rules'. We therefore reinterpret actions in the herding model as the experimenter's stochastic signals, and the $\mathcal{DM}$s' decision rules as his allowed actions. We perform this formal embedding for a general observational learning context. This experimentation embedding is crucial for setting up our planner's problem.

The plan of this paper is as follows. Section 2 describes a general observational learning model, and then re-interprets it as an optimal single-person experimentation model. The herding model and team equilibrium are introduced in section 3, where our action indices are introduced. We then describe optimal strategies using index rules, and prove our short-run contrarian result in section 4. The long-run model is considered in section 5; there we show the cascade sets are non-empty, and that they shrink with the discount factor. A conclusion follows, while many proofs are appendicized.

2 TWO EQUIVALENT LEARNING MODELS

In this section, we first set up a general observational learning model subsuming the herding models. All models in this class are then formally embedded in the experimentation framework. Afterwards, we specialize our findings to the informational herding model.

2.1 The Observational Learning Model

Information. An infinite sequence of *decision-makers* ($\mathcal{DM}$s) $n = 1, 2, \ldots$ acts in that exogenous order. The actions have uncertain payoffs. There is a given common prior belief over the compact space Ω, whose elements ω are the *states of the world.*

The nth $\mathcal{DM}$ observes a partially informative random *private signal* realization σ_n about the state of the world. We may assume WLOG that the private signal observed by a $\mathcal{DM}$ is actually his *private belief* which results from Bayesian updating given σ_n

and the prior;[2] that is, $\sigma_n \in \Sigma$, where $\Sigma = \Delta(\Omega)$ consists of probability measures over Ω. Conditional on the state, the signals are assumed to be i.i.d. across $\mathcal{DM}$s, drawn according to the probability measure $\mu^\omega \in \Delta(\Sigma)$ in state $\omega \in \Omega$.

Bayesian Decision-Making. Everyone chooses from a compact action set A. Action $a \in A$ earns a payoff $u(a, \omega)$ in state $\omega \in \Omega$, the same for all $\mathcal{DM}$s. Before choosing an action, the n'th $\mathcal{DM}$ first observes his private signal/belief and the history consisting of the $n-1$ predecessors' actions.

Each $\mathcal{DM}$'s Bayes-optimal decision uses the observed action history and his own private belief. A $\mathcal{DM}$ can compute the strategies of all predecessors, and can use the common prior to calculate the ex ante probability distribution over histories in either state. Bayes' rule then implies a *public belief* $\pi \in \Sigma$ for any history. A final application of Bayes' rule incorporates the private belief σ to give the private *posterior belief* $\rho \in \Sigma$.

Given the posterior belief ρ, the $\mathcal{DM}$ picks the action $a \in A$ which maximizes his expected payoff $\bar{u}(a, \rho) = \int_\Omega u(a, \omega) d\rho(\omega)$. Let X be the space of *decision rules* — namely, maps x from Σ into $\Delta(A)$, the probability measures $x(\sigma)$ over A. Any rule $x \in X$ induces a distribution over actions for all private beliefs σ.

The Stochastic Process of Beliefs. The distribution of signals σ depends on the state ω, so that the distribution over actions a depends on both the state ω and the decision rule x. There is a density $\psi(\omega, x) \equiv \int x(\sigma)(a)\mu^\omega(d\sigma)$; unconditional on the state, it is $\psi(\pi, x) \equiv \int_\Omega \psi(\omega, x)\pi(d\omega)$. This yields a distribution over next period public beliefs. Thus, $\langle \pi_n \rangle$ follows a Markov process with state-dependent transition chances.

2.2 Informational Herding as Experimentation Déjà Vu

And out of old bookes, in good faithe,
Cometh al this new science that men lere.
— Geoffrey Chaucer (The Assembly of Fowles, line 22)

Our immediate goal is to recast the observational problem outcome as a single person optimization. A first stab brings us to the *forgetful experimenter,* who each period receives a new informative signal, takes an optimal action, and then promptly forgets his signal; the next period, he can reflect only on his action choice. But this is not a model of *Bayes-optimal* experimentation, since it assumes and in fact requires irrational behaviour. How then can an experimenter not observe the private signals, and yet take informative actions?

[2]Hereafter, we therefore often use private belief and private signal interchangeably.

For context, consider McLennan's (1984) sequel to Rothschild (1974). He allowed the monopolist to charge one of a continuum of prices, with two possible linear purchase chance 'demand curves'. McLennan found that the resulting uninformative price when the demand curves crossed may well eventually be chosen by an optimizing monopolist.

Rothschild's and McLennan's models give examples of *potentially confounding actions,* later introduced in EK: Easley and Kiefer (1988). In brief, such actions are optimal for *unfocused* beliefs for which they are invariants (i.e. taking the action leaves the beliefs unchanged). Of particular significance is the proof in EK (on page 1059) that with finite state and action spaces, potentially confounding actions generically do not exist, and thus complete learning must arise.[3] Rothschild and McLennan might be seen as separate anticipations of EK's general insight. Rothschild escapes it by means of a continuous state space, whereas McLennan resorts to a continuous action space. Yet there appears no escape for the herding paradigm, where both flavours of incomplete learning, incorrect limit cascades and confounded learning (see SS), generically arise with two actions and two states. This puzzle suggests the inverse mapping that we now consider.

In recasting our general observational learning model as a single person experimentation problem, we must focus on the myopic experimenter with discount factor 0 (ruling out active experimentation). Steering away from a forgetful experimenter, we shall regard the observational learning story from a new perspective. Consider the nth $\mathcal{DM}$, who uses both the public belief π_n and his private signal σ_n in forming and acting upon his posterior beliefs ρ_n. We may separate these two steps by the conditional independence of π_n and σ_n. Regard Mr. n as: (i) observing π_n, but *not* his private signal; (ii) optimally determining the rule $x \in X$, and submitting it to an agent 'choice' machine; and (iii) letting that machine observe his private signal and take his action $a \in A$ for him. The payoff $u(a, \omega)$ is unobserved, lest that provide an additional signal of the state of the world.

Thus, the observational learning model corresponds to a single-person experimentation model where: The state space is Ω. At stage n, the experimenter $\mathcal{EX}$ chooses an action (the rule) $x \in X$. Given this choice, a random observable statistic $a \in A$ is realized with chance $\psi(\omega, x)$ in state ω. Finally, the period's payoff $u(a, \omega)$ is realized but not observed.[4] When private beliefs σ have distribution μ^ω in state ω, and when $\mathcal{EX}$ fully disounts future payoffs, then he chooses the same optimal decision rule x described in section 2, resulting in

[3]Eg: payoffs in a one-armed bandit, with a potentially confounding safe arm, are not generic in $\mathbb{R}^2$.

[4]This model doesn't strictly fit into the EK mold, where stage payoffs depend only on the action and the observed signal, but (unlike here) not on the parameter $\omega \in \Omega$. This is the structure of Aghion, Bolton, Harris, and Jullien (1991), who admit unobserved payoffs. Alternatively, we could posit that $\mathcal{EX}$ has fair insurance, and only sees/earns his expected payoff each period and not his random realized payoff.

Observational Learning Model	Impatient Experimenter Model
State: $\omega \in \Omega$	State: $\omega \in \Omega$
Public Belief after nth $\mathcal{DM}$: π_n	Belief after n observations: π_n
Optimal decision rule: $x \in X$	Optimal action: $x \in X$
Private signal/belief of nth $\mathcal{DM}$: σ_n	Randomness in the nth experiment: σ_n
Action taken by each $\mathcal{DM}$: $a \in A$	Observable signals: $a \in A$
Density over actions: $\psi(\omega, x)$	Density over observables: $\psi(\omega, x)$
Payoffs: private information	Payoffs: unobserved

Table 1: **Embedding.** This table displays how our infinite person observational learning model fits into the impatient single person experimentation model.

action $a \in A$ with chance $\psi(\omega, x)$. Table 1 summarizes this embedding of the observational learning model into a myopic experimentation model.

Notice how this addresses both lead puzzles. First, the experimenter never knows the private beliefs σ, and thus does not forget them. Second, incomplete learning (bad herds) are entirely consistent with EK's generic finding of complete learning for models with finite action and state spaces. Simply put, actions do not map to actions but to signals when one rewrites the observational learning model as an experimentation model. The true action space for $\mathcal{EX}$ is the infinite space X of decision rules.[5]

3 FORWARD-LOOKING INFORMATIONAL HERDING

3.1 Two Forward-Looking Informational Herding Models

We now specialize to the simpler informational herding framework, following SS, with two states and exactly $A < \infty$ actions. 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.

This observational learning paradigm involves an informational externality, because taking an action partially conveys one's hidden private signal. The individuals jointly possess enough information to perfectly reveal the true state of the world, yet given the sequential structure, do not. SS prove that with bounded private beliefs, the resulting

[5]Such an embedding is well-known and obvious for rational expectations pricing models, since the price is publicly observed, and an inverse mapping is not required.

SS considered two modifications of the informational herding paradigm. With i.i.d. 'crazy' preference types, one adds an exogenous chance of a noisy signal (i.e. random action). With individuals randomly drawn from one of T different preference types, let the $\mathcal{EX}$ choose a T-vector of optimal decision rules from X^T with (only) the choice machine seeing the task and private belief, and choosing the action a.

herding equilibrium is sometimes *ex post* inefficient. But is this *ex ante* inefficient, subject to the sequential structure, and hidden signal assumption? To address this question, we consider alternative objectives, where everyone is altruistic, and aims to maximize an discounted average of payoffs. 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.

Assume a state space $\Omega = \{H, L\}$, with both states equi-likely ex ante (prior chance 1/2). Private belief σ is the chance of state H, so that $\Sigma = [0,1]$. To avoid trivialities, μ^H and μ^L are not (a.s.) identical, so that some signals are informative. Each distribution may contain atoms, but to ensure that no signal will perfectly reveal the state of the world, we insist that μ^H and μ^L be mutually absolutely continuous. Let $\text{supp}(\mu)$ denote their common support. If $\text{supp}(\mu) \subseteq (0,1)$, then private beliefs are *bounded*; they are *unbounded* if $\text{co}(\text{supp}(\mu)) = [0,1]$ — 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. The 'no introspection' property of SS asserts $(d\mu^L/d\mu^H)(\sigma) = (1-\sigma)/\sigma$, so that $d\mu^H/d\mu = 2\sigma$ and $d\mu^L/d\mu = 2(1-\sigma)$. We can thus take μ to be the primitive distribution of the private signal, from which the state-dependent probability measures μ^H and μ^L are derived.

The Team Equilibrium. We first suppose that every $\mathcal{DM}$ is altruistic, but subject to the informational herding restriction. A *strategy* s_n for the n'th $\mathcal{DM}$ is a map from history to X, assigning the rule $x_n \in X$ for each history; $s = (s_1, s_2, \ldots)$ denotes a strategy profile. A *team equilibrium* is a Bayes-Nash equilibrium of the game where the nth $\mathcal{DM}$ maximizes the average present welfare of posterity, themselves included $E[(1-\delta)\sum_{k=0}^{\infty} \delta^k u_{n+k} | \pi_n]$.

The Social Optimum. We next reinterpret $\mathcal{EX}$'s problem as that of an informationally constrained *social optimum*: Maximize the average present value of posterity's welfare $E[(1-\delta)\sum_{n=1}^{\infty} \delta^{n-1} u_n]$. Here, the realized payoff sequence is $\langle u_n \rangle$ for $\mathcal{DM}$'s $n = 1, 2, \ldots$. So the $\mathcal{EX}$'s objectives are perfectly aligned with a hypothetical social planner. To respect the herding restriction, we assume that the planner neither knows the state nor can observe the individuals' private signals, but can both observe and induce any actions taken.

A *policy* is the Markovian restriction of a strategy: a map $\xi : [0,1] \to X$, where $\xi(\pi)$ is the rule given belief π. If the social optimum exists, then there exists a policy which attains the maximum, since the problem is Markovian in public beliefs π, via Bayes rule.

The Second Welfare Theorem. Inspired by Vives (1997), we now briefly recast an insight of Radner (1962) for our framework.

Lemma 1 *For any discount factor $\delta < 1$, any social optimum is a team equilibrium. Further, the most efficient team equilibrium is described by a Markov policy.*

Proof: As the planner can use any team equilibrium strategy, the second claim follows from the first. So fix the current $\mathcal{DM}$ and public belief π. To see that the planner's optimum s is a team equilibrium, assume that successors use it, but that some $\mathcal{DM}$ n has a strictly better reply $\hat{x}$. Then the planner can improve his value at π by *fully* mimicking this deviation, i.e. by (i) taking $\hat{x}$ in the first period and (ii) continuing with s as if the first period history had been generated by $s_n(\pi)$. This contradicts optimality of the policy. $\square$

3.2 Optimal Behaviour via Index Rules

Dynamic Programming. Our analysis here follows Aghion, Bolton, Harris, and Jullien (1991) and §9.1–2 of Stokey and Lucas (1989). The value function $v_\delta(\cdot) : \Sigma \mapsto \mathbb{R}$ for the planning problem with discount factor δ 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. Recall that $\bar{u}(a, \pi)$ denotes the expected payoff from action a at belief π. Define $\psi(\omega, x) = \int x(\sigma)(a)\mu^\omega(d\sigma)$ for $\omega = L, H$, and put $\psi(\pi, x) = \pi\psi(H, x) + (1-\pi)\psi(L, x)$, in a suggestive abuse of notation. If $q(a, \pi, x) = \pi\psi(H, x)/\psi(\pi, x)$ is the Bayes-updated *posterior belief* from π when action a is observed and rule x is applied, we may formulate the Bellman equation:

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

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

Generalized Gittins Indices. We now consider the classical multi-armed bandit (see Bertsekas (1987), §6.5). A patient experimenter each period must choose one of n actions, with uncertain independent reward distributions. He must then trade-off the informational and myopic payoffs associated with each action. Gittins (1979) showed that optimal behaviour in that model can be described by simple 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, each period, choose the action with the highest index.

We now argue that the policy employed in a team equilibrium has a likewise appealing form: For a given 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., the set of all slopes m that obey $g(\hat{z}) \geq g(z) + m \cdot (\hat{z} - z)$ for all $\hat{z}$.

Proposition 1 (The Index Rule) *Fix a team equilibrium s, and a decision maker n who has arrived at prior belief π. To each action $a \in A$, there exists a function $\bar{v}(a, \pi, \rho)$, affine in n's private posterior belief ρ, such that n's average present value of action a is*

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

where $q(a) \equiv q(a, \pi, s_n(\pi))$ is the public posterior belief induced by action a. Moreover, in the social optimum, $\bar{v}(a, \pi, \rho) = v_\delta(q(a)) + m(a, \pi)(\rho - q(a))$ where $m(a, \pi) \in \partial v_\delta(q(a))$.

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

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)).

3.3 Optimal Action Belief Interval Rules

SS shows that the myopic $\mathcal{DM}$ uses a belief interval rule: Action a is taken for beliefs in a subinterval of $[0, 1]$, and for generic payoffs, these intervals overlap only at endpoints. This is also true with patience, as Lemma 2 proves.

Lemma 2 *Fix any team equilibrium strategy s. For any public belief π, the rule $s_n(\pi)$ is almost surely described by a dissection of $[0, 1]$ into closed intervals $I_a(\pi) \equiv \bar{I}_a$, generically overlapping at endpoints only, such that action a is 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. □

Intuitively, the interval structure is not only myopically best, but it also ensures the greatest information value, by producing the riskiest posterior belief distribution.

By Lemma 1, the planner's solution too is described by the interval rules of Lemma 2. 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 also establishes part (b) below:

Lemma 3 (a) *A social planner's* policy $\xi^\delta : [0,1] \to X$ *and team equilibrium both exist.*
(b) *The correspondence* $\pi \mapsto \xi^\delta(\pi)$ *is upper hemi-continuous in* π.

Communicating private information by a finite mesh size is a problem that is not without history. Sobel (1953) investigated an interval structure in a simple statistical decision problem. More recently, Dow (1991) has also studied the nature of such a coarse information process in a two period search model. A consumer first observes one signal $\sigma_1 \in \mathbb{R}$, but can remember in the future only whether $\sigma_1 \in E_i$, for an endogenous coarse partition $E_1, \ldots, E_n$ of $\mathbb{R}$. In the second period, the consumer observes $\sigma_2 \in \mathbb{R}$, and then chooses an action $a \in \{1, 2\}$ to maximize the expected payoff given E_i and σ_2. Since no decision is made in the first period, this corresponds to $\delta = 1$ for us. Like us, Dow shows that the $\mathcal{DM}$ communicates to his future selves with connected intervals of signals.[6]

We first characterize the *cascade sets*, the belief regions where learning stops.

Lemma 4 (a) *An optimal policy almost surely induces action a iff $\pi \in J_a(\delta)$, where $J_a(\delta) \subset [0,1]$ is empty, a point, or an interval.*
(b) $0 \in J_1(\delta)$ *and* $1 \in J_A(\delta)$ *for any* $\delta \in [0,1)$, *and* $\cup_{a=1}^{A} J_a(\delta) \neq [0,1]$.

Proof: It is optimal to induce any action a almost surely iff $v_\delta(\pi) = \bar{u}(a, \pi)$. As $\bar{u}(a, \pi)$ is affine in π, and $v_\delta(\cdot)$ is weakly convex, this equality holds on a closed interval $J_a(\delta)$. Also, 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 $\pi = 1$. Finally, 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. This is impossible at any kink of $v_\delta(\pi)$. $\square$

Interval Ordering. Lemma 2 does not say that intervals are arranged in the myopic order. For while swapping the interval order preserves the informational content of the actions, it might entail a myopic loss. Eg., let action 2 be slightly better than action 1

[6]Chernoff (1952) interprets hypothesis acceptance or rejection as a bi-partition of the available information, and he studies the optimal such partition. Again, when $\delta < 1$ our problem is more complex.

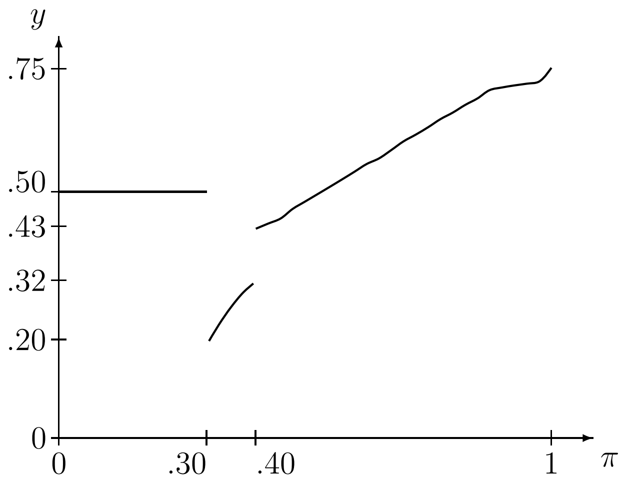

Figure 1: **Action ordering.** The unconditional private belief density is $f(\sigma) = K\sigma^6$ over the range $(0, 4/7)$, where $K = 7^8/4^7$, with mean belief 1/2. Action $a = 1, 2$ has payoff a in state H and $3 - a$ in state L. The myopic ($\delta = 0$) basin for action 1 is $(0, 3/7)$. When $\delta = .95$, the graph shows the individual's posterior belief y against his prior belief π at the threshold $x(\pi)$. 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.

in state H, but much worse in state L. Consider the policy that takes action 1 for all but very low beliefs. Then swapping the intervals raises the chance that the mostly worse action 2 is taken, and so might be suboptimal. Figure 1 illustrates this possibility.

Dominated Actions. Lemma 2 also does not rule out using dominated actions for more information is transmitted with more actions. To see this, assume A actions, discount factor $\delta > 0$, and bounded beliefs. Since action A is uniquely optimal in state H, this remains true for any public belief π in the interior of the cascade set $J_A = [\bar{\pi}, 1]$. The associated value function v is then affine on J_A, but not on any interval extending J_A to the left — i.e. it is strictly convex at $\bar{\pi}$. Now add an extra action $A + 1$ with payoff $u(A + 1, \omega) = u(A, \omega) - \varepsilon$, for all states ω, where $\varepsilon > 0$. Even though $A + 1$ is strictly dominated by A, we argue that $A + 1$ is used with positive chance for small enough $\varepsilon > 0$.

If action $A + 1$ is never used, then v remains the value function of the new problem. One optimal policy at $\bar{\pi}$ is then to take A almost surely. Instead map private signals below 1/2 into A, and above 1/2 into $A + 1$. This strictly spreads the posterior beliefs when $q_A < \bar{\pi} < q_{A+1}$. Since the value function is strictly convex at $\bar{\pi}$, the expected continuation value exceeds $v(\bar{\pi})$ by some $\eta > 0$. On the other hand, the policy change implies a myopic loss less than ε. When $\delta\eta - (1-\delta)\varepsilon > 0$, the modified policy beats the optimal policy — a contradiction. But then v cannot be the value function of the modified problem, and so the new action $A+1$ must be used with positive chance.

Optimal Taxes? Lemma 2 ignores implementation. The team equilibrium assumes everyone cares about posterity. Can the planner's outcome be decentralized with selfish individuals? Since the only observables are actions, action-dependent transfer payments are the only available policy instrument — assuming that only contemporaneous actions are rewarded/punished. Suppose that everyone maximizes his expected one-shot myopic payoff $u(a, \pi)$ plus incurred transfers. Faced with any such incentives, our proof of Lemma 2 that interval rules are optimal is still valid, since our indices affinely adjust. But since such transfers can never reverse the myopic ordering of actions, and so are not sufficient if the optimal action ordering differs from the optimal action ordering. We leave as an open whether sufficient conditions exist for which transfers decentralize the planner's solution.

Below, we explore how patience affects the informational herding problem. Section 4 derives a novel comparative static valid at any public belief, showing how individuals lean against the conventional wisdom when $\delta > 0$. Section 5 then shows how cascade sets shrink as individuals grow more patient.

4 THE SHORT RUN COMPARATIVE STATIC

4.1 Two Actions

We now derive a local comparative statics result. For clarity, we first simply assume two actions: a *low action* ℓ taken at low beliefs and a *high action* h, taken at high beliefs. While it seems intuitive that action ℓ is myopically best for low beliefs and action h for high beliefs, this intuition is misleading (see Figure 1). This is a local notion, and in principle the low and high actions could depend on the public belief π.

By Lemma 4, both actions are taken with positive probability when π is in the open non-empty interval $M = [0,1] \setminus (J_\ell(\delta) \cup J_h(\delta))$. We first rule out an annoying possibility.

A-1 *A unique rule $x(\pi)$ is optimal in some open neighbourhood $N(\bar{\pi}) \subseteq M$ of $\bar{\pi}$.*

Our decision rule x implies a *private belief threshold* θ, separating the belief intervals for the two actions. The action ordering cannot switch in the open neighbourhood $N(\bar{\pi})$.[7]

The public, private, and posterior beliefs π, σ, and ρ are related by the monotone maps:

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

The rule x is thus locally summarized by the *posterior belief threshold* $y \equiv r(\pi,\theta)$.

Clearly, in the myopic case when $\delta = 0$, the optimal posterior threshold $y(\pi)$ between the intervals of private posterior beliefs for actions ℓ and h 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 superscript.) Specifically we prove that $y(\pi)$ *increases* in π. So for higher public beliefs π, the $\mathcal{DM}$ is discouraged from choosing the increasingly popular action h versus ℓ. This experimentation benefits successors, since the less likely action moves public beliefs more.

A critical hurdle that must be overcome is the following: It is natural to suppose, and crucial for the argument, that the posterior public belief following each action is monotone in the prior. Recall that actions constitute a discrete filter on the underlying signals. Thus, such a public posterior is an average of unobserved private posteriors that could lead to the action. Action ℓ (or h) is taken whenever such a private posterior falls below (or above) a threshold. Thus, we require monotonicity of a truncated random variable. In the spirit of Burdett (1996), we find that a log-concavity assumption is the missing ingredient.

Let f be the *unconditional private belief density* for μ. The private belief log-likelihood ratio $\Lambda(\sigma) = \log(\sigma/(1-\sigma))$ then has density $\phi(\Lambda) \equiv f(\sigma(\Lambda))\sigma'(\Lambda) = f(\sigma(\Lambda))e^{\Lambda}/(1+e^{\Lambda})^2$.

A-2 *The log-likelihood ratio density $\phi(\Lambda)$ exists, and is strictly log-concave.*

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

Proof: The proof exploits the first order condition and the value function convexity. It works even when v is not globally differentiable,[8] because differentiability obtains at precisely those points where it is required. In light of (1), and our assumed action order,

[7] Indeed, local uniqueness and a uhc rule correspondence precludes policy jumps, and thus in particular, any switch between $h = 2$ and $h = 1$. This can be seen, eg., using the metric introduced proof of Lemma 3.

[8] Differentiability of the value function in problems of learning is an open problem. 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.

the maximand of the Bellman equation in this two action world is

$$B(\pi,y)=\psi(\pi,y)[(1-\delta)\bar{u}_\ell(q_\ell)+\delta v(q_\ell)]+(1-\psi(\pi,y))[(1-\delta)\bar{u}_h(q_h)+\delta v(q_h)] \tag{3}$$

where[9] $\bar{u}_a(y)=\bar{u}(a,y)$, $\psi(\pi,y)=\psi(\ell,\pi,y)$, and $q_\ell=q(\ell,\pi,y)<y<q_h=q(h,\pi,y)$.

Step 1 (First Order Condition) *If y solves $\max_y B(\pi,y)$, then the value function v is differentiable at the continuation beliefs q_a after actions $a=1,2$. Moreover,*

$$B_y(\pi,y)=\frac{\partial\psi}{\partial y}\{(1-\delta)\bar{u}_\ell(y)+\delta[v(q_\ell)+v'(q_\ell)(y-q_\ell)]-(1-\delta)\bar{u}_h(y)-\delta[v(q_h)+v'(q_h)(y-q_h)]\} \tag{4}$$

Proof: If we assume first that v is differentiable at q_ℓ and q_h, then $B_y(\pi,y)$ equals

$$\frac{\partial\psi}{\partial y}((1-\delta)\bar{u}_\ell(q_\ell)+\delta v(q_\ell))+\psi\left((1-\delta)\frac{\partial\bar{u}_\ell(q_\ell)}{\partial q_\ell}\frac{\partial q_\ell}{\partial y}+\delta v'(q_\ell)\frac{\partial q_\ell}{\partial y}\right)$$
$$-\frac{\partial\psi}{\partial y}((1-\delta)\bar{u}_h(q_h)+\delta v(q_h))+(1-\psi)\left((1-\delta)\frac{\partial\bar{u}_h(q_h)}{\partial q_h}\frac{\partial q_h}{\partial y}+\delta v'(q_h)\frac{\partial q_h}{\partial y}\right)$$

Recalling that $\bar{u}_\ell$ and $\bar{u}_h$ are affine functions, and using Lemma 5 below, this produces (4).

If v is not differentiable at q_ℓ or q_h or both, then its right derivative strictly exceeds its left — i.e. a 'convex kink'. Since q_ℓ and q_h are increasing functions of y, and since $\partial\psi/\partial y>0$ and $q_\ell<y<q_h$, (4) applies to one-sided derivatives — in other words, $B_{y-}(\pi,y)<B_{y+}(\pi,y)$. But optimality of y implies that the left derivative is nonnegative, and the right derivative nonpositive. The convex kink cannot then obtain. □

Lemma 5 *$(y-q_1)\partial\psi/\partial y=\psi\partial q_1/\partial y$ and $(q_2-y)\partial\psi/\partial y=(1-\psi)\partial q_h/\partial y$.*

The lemma admits a simple intuition. Observe that y and q_a are the *marginal* and *average* private beliefs leading to action a, while ψ_a is the total chance of taking a. Think of these as a firm's marginal cost MC, average cost AC, and quantity Q. Then $(\rho-q)\partial\psi/\partial\rho=\psi\partial q/\partial\rho$ reduces to 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 $\partial\psi/\partial y>0$, the first order condition says that the indices coincide at knife-edge beliefs: $w_\delta(\ell,\pi,y)=w_\delta(h,\pi,y)$. This could also be rewritten as the equality at belief y of a myopic gain and a dynamic loss from a marginal shift in the threshold belief (see Figure 2):

$$(1-\delta)\left[\bar{u}_\ell(y)-\bar{u}_h(y)\right]=\delta\left[(v(q_h)+v'(q_h)(y-q_h))-(v(q_\ell)+v'(q_\ell)(y-q_\ell))\right] \tag{5}$$

[9]As the policy rule x is equivalently represented by posterior threshold y, this proof replaces the function $q_a(\pi,x)$ with $q_a(\pi,y)$.

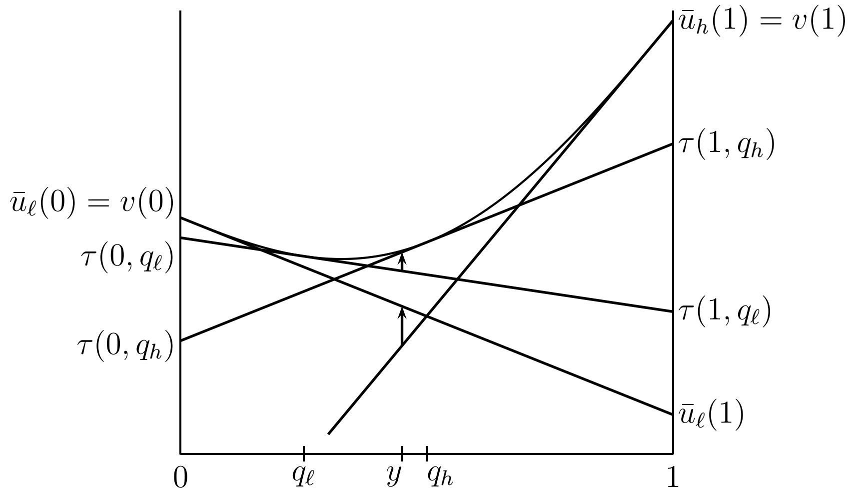

Figure 2: **First order condition.** The convex curve is the value function v as a function of beliefs. It intersects the myopic value functions at 0 and 1. Given the threshold posterior belief y, we show the tangents to v at posteriors q_ℓ, q_h. The vertical arrows over y indicate the myopic loss and dynamic gain from a marginal change in y. The first order condition states that the lengths of these arrows stand in proportion $\delta/(1-\delta)$.

When $n = 2$, the first agent's decision merely conveys to the second agent whether the belief ρ is above or below the threshold y. From (5) it follows that the optimal threshold lies at the intersection of the two tangents.[10]

Let us now turn to the comparative static exercise. By Lemma 6 below, the public posteriors q_ℓ, q_h are increasing in the public prior belief π. Next, by Lemmas 7 and 8 (as in Figure 2), the RHS of (5) is falling in each of q_ℓ, q_h, and thus π, by the convexity of the value function. This yields a sufficient condition for an increasing optimizer $y(\pi)$.

Step 2 (Single-Crossing Property) *For all $\pi, \pi' \in N(\bar{\pi})$ with $\pi < \pi'$: (a) the function $B(\pi, y)$ obeys the single crossing property $B_{y-}(\pi', y(\pi)) > 0$, and so (b) $y(\pi') > y(\pi)$.*

Proof: For (*a*), rewrite $B_{y-}(\pi, y) = \Upsilon(y, q_\ell(\pi, y), q_h(\pi, y))\partial\psi(\pi, y)/\partial y$, where

$$\Upsilon(y, q_\ell, q_h) \equiv (1-\delta)[\bar{u}_\ell(y) - \bar{u}_h(y)] + \delta[\tau(y, q_\ell) - \tau(y, q_h)]$$

and where $\tau(y, q) \equiv v(q) + v'(q-)(y-q)$ is the affine function left-tangent to v at q.

[10]Our result generalizes Dow's 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.

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

$$\Upsilon(y(\pi), q_\ell(\pi', y(\pi)), q_h(\pi', y(\pi))) > \Upsilon(y(\pi), q_\ell(\pi, y(\pi)), q_h(\pi, y(\pi)))$$

Since only the continuation beliefs change, this is equivalent to

$$\tau(y(\pi), q_\ell(\pi', y(\pi))) - \tau(y(\pi), q_h(\pi', y(\pi))) > \tau(y(\pi), q_\ell(\pi, y(\pi))) - \tau(y(\pi), q_h(\pi, y(\pi)))$$

When $\pi < \pi'$, we have $q_\ell(\pi, y(\pi)) < q_\ell(\pi', y(\pi)) < y(\pi) < q_h(\pi, y(\pi)) < q_h(\pi', y(\pi))$ by Lemma 6. From Lemmas 7 and 8, we find that $\tau(y(\pi), q_h(\pi, y(\pi))) \geq \tau(y(\pi), q_h(\pi', y(\pi)))$ and $\tau(y(\pi), q_\ell(\pi', y(\pi))) \geq \tau(y(\pi), q_\ell(\pi, y(\pi)))$. Also, at least one inequality is strict since some continuation belief q_a lies inside M, by Claim 10. This proves the needed inequality.

For (b), suppose first that $y(\pi') = y(\pi)$. Then $B_{y-}(\pi', y(\pi')) = B_{y-}(\pi', y(\pi)) > 0$ by (a), contradicting the optimality of $y(\pi')$. Or suppose that $y(\pi') < y(\pi)$. In this case, by a variant of the Fundamental Theorem of Calculus for convex functions (see Claim 1),

$$B(\pi', y(\pi)) - B(\pi', y(\pi')) = \int_{y(\pi')}^{y(\pi)} B_{y-}(\pi', z)dz$$

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

The next step provides conditions for the posterior belief after any action a to increase in the public belief π. While at first glance intuitive, it is generally violated by the standard atomic signal distributions. The log-concavity assumption comes crucially into play here.

Lemma 6 *For fixed y, the continuation public belief $q_a(\pi, y)$ is strictly increasing in π.*

Proof: Let $\theta(\pi)$ solve $y = r(\pi, \theta)$. Since an increase in $q_\ell(\pi, y)$ is equivalent to an increase in $q_\ell(\pi, y)/[1 - q_\ell(\pi, y)]$, we need only show the latter rises in y. Now,

$$\frac{q_\ell(\pi, y)}{1 - q_\ell(\pi, y)} = \left(\frac{\pi}{1-\pi}\right) \frac{\int_0^{\theta(\pi)} f^H(\sigma)d\sigma}{\int_0^{\theta(\pi)} f^L(\sigma)d\sigma} = \left(\frac{\pi}{1-\pi}\right) \frac{\int_0^{\theta(\pi)} \sigma\, f(\sigma)d\sigma}{\int_0^{\theta(\pi)} (1-\sigma)\, f(\sigma)d\sigma} \tag{6}$$

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. Changing densities via the posterior likelihood ratio $\lambda = \pi\sigma/[(1-\pi)(1-\sigma)]$, we find (see

§A.6) that $(1-\pi)(1-\sigma)f(\sigma)d\sigma = g(\pi,\lambda)d\lambda$ and $\pi\sigma f(\sigma)d\sigma = \lambda g(\pi,\lambda)d\lambda$ for the function

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

given our density $\phi(\Lambda) = f(\sigma(\Lambda))e^{\Lambda}/(1+e^{\Lambda})^2$. This allows us to rewrite (6) as:

$$\frac{q_\ell(\pi,y)}{1-q_\ell(\pi,y)} = \frac{\int_0^{y/(1-y)} \lambda g(\pi,\lambda)\,d\lambda}{\int_0^{y/(1-y)} g(\pi,\lambda)\,d\lambda} = \frac{\int_0^1 \lambda g(\pi,\lambda)\mathbb{I}(\lambda)\,d\lambda}{\int_0^1 g(\pi,\lambda)\mathbb{I}(\lambda)\,d\lambda} = E\left[\lambda \middle| 0 \le \lambda \le \frac{y}{1-y}\right]$$

where $\mathbb{I}$ is the indicator function of $\lambda \in [0, y/(1-y)]$. The above increases in π provided $g(\pi,\lambda)$ obeys the monotone likelihood ratio property, which it does (see §A.6).

A similar proof works for action h, given $q_h(\pi,y)/(1-q_h(\pi,y)) = E[\lambda|\lambda \ge y/(1-y)]$. □

Convex functions obey a useful tangent property, summarized here (proved in §A.4).

Lemma 7 (Tangents to a Convex Function) *Fix* $0 \le z_1 < z_2 < z_3 < z_4 < z_5 \le 1$. *Let* $\tau_i : \mathbb{R} \to \mathbb{R}$ *be tangent functions to* v *at* z_i *for* $i = 1, 2, 4, 5$. *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.

We prove in §A.4 that if small information is worthless today, then it is also tomorrow. But as seen in the next example, the optimization today is simple, and yields a constant posterior belief threshold. Hence, the information has value today — i.e., strict convexity.

Lemma 8 (Strict Convexity) *The value function* v *is strictly convex on* M.

A Two Period Example. While our theory was proven for infinite horizon models, we illustrate it in a simple closed-form two period example.[11] A professor and a student share a common prior π, and observe conditionally iid signals σ_p and σ_s with common state-dependent densities $f^H(\sigma) = 2\sigma$ and $f^L(\sigma) = 2(1-\sigma)$. The unconditional density $f(\sigma) = 1$ over private beliefs is uniform $(0,1)$, and so the induced unconditional density over log-likelihood ratios is log-concave, and assumption $(*)$ holds. The professor sees a signal, and takes an action; his student observes his action, but not his signal. The professor is totally self-sacrificial: He takes action $a \in \{1,2\}$ to maximize his student's expected payoff alone, where $u(2,H) = u(1,L) = 1, u(1,H) = u(2,L) = 0$.

[11] We could have brought finite horizon models under the umbrella of this theory, but it would have greatly complicated both the notation and the dynamic programming proof, having to deal with both finite and infinite horizon cases.

If the student starts with a continuation public belief q, then she will take action 2 exactly when her signal $\sigma_s \geq 1-q$. Given the conditional signal densities, her value function is $V_s(q) = (1-q)\int_0^{1-q} 2(1-\sigma_s)d\sigma_s + q\int_{1-q}^1 \sigma_s d\sigma_s = 1-q+q^2$. The professor obviously employs a threshold rule $\theta = \theta(\pi)$: He chooses 1 if his signal $\sigma_p < \theta$, and 2 if $\sigma_p \geq \theta$. His problem is to maximize $V(\pi) = E_\pi V_s(q)$, where $q = q(a, \pi, \theta)$ is his student's public belief (denoted q_a later on). Beliefs are a martingale, or $\pi = E[q|\pi]$, so that

$$V(\pi) = E[V_s(q)|\pi] = E(1-q+q^2|\pi) = 1 - E[q|\pi] + E[q|\pi]^2 = 1 - \pi + \pi^2 + E[(q-\pi)^2|\pi]$$

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

$$E[(q-\pi)^2|\pi] = \frac{\pi - q_1}{q_2 - q_1}(q_2 - \pi)^2 + \frac{q_2 - \pi}{q_2 - q_1}(\pi - q_1)^2 = (q_2 - \pi)(\pi - q_1) \tag{7}$$

by the martingale property of beliefs, where the two continuation public beliefs are

$$q_1 = \frac{\pi\theta^2}{\pi\theta^2 + (1-\pi)(2\theta - \theta^2)} \qquad \text{and} \qquad q_2 = \frac{\pi(1-\theta^2)}{\pi(1-\theta^2) + (1-\pi)(1 - 2\theta + \theta^2)}$$

Maximizing (7), we find the optimal threshold $\theta(\pi) = (\pi - 1 + \sqrt{\pi - \pi^2})/(2\pi - 1)$ if $\pi \neq 1/2$, with limit $\theta(1/2) = 1/2$. The posterior belief

$$r(\pi, \theta(\pi)) = \frac{\pi\theta(\pi)}{\pi\theta(\pi) + (1-\pi)(1-\theta(\pi))} = \frac{1}{2} + \frac{1 - 2\sqrt{\pi - \pi^2}}{2(2\pi - 1)} \gtrless 1/2 \quad \text{as } \pi \gtrless 1/2$$

This expression is increasing in π, thus illustrating our contrarianism conclusion. 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 $\theta = 1/2$ away from the myopic threshold $\theta = 1-\pi$. Indeed, one can easily check that either $\pi > 1/2$, whereupon $1/2 < \theta(\pi) < 1 - \pi$, or $\pi < 1/2$, and so $1/2 > \theta(\pi) > 1 - \pi$. Finally, the professor's optimal value is easily computed from (7):

$$V(\pi) = 1 - \pi + \pi^2 + E[(q-\pi)^2|\pi] = 1 - \pi + \pi^2 + \left(\frac{2\pi(1-\pi)}{2\sqrt{\pi - \pi^2} + 1}\right)^2$$

The latter term captures the value of the information conveyed by the professor's action.

It measures the excess of the team equilibrium payoff over the herding equilibrium payoff. It is maximized at $\pi = 1/2$.

4.2 Multiple Actions

The proof of the last subsection avoided considering the cross partial derivative of the Bellman function in beliefs and thresholds. Instead, we firstly proceeded by considering just the single partial derivative in the threshold in Step 2, and crucially, we operated at the optimal threshold. Our attempts to extend this approach have fallen short. With more than two actions, we proceed with a more traditional approach to a comparative statics exercise. Our approach will therefore be intuitive, but will necessitate a slightly bolder assumption about the value function.

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

With multiple actions, some may be taken with probability zero (*inactive*). Posit that:

A-1′ *A unique rule $x(\pi)$ employing the same set of $\alpha > 2$ active actions is optimal in an open neighbourhood $N(\bar{\pi}) \subseteq M$ of $\bar{\pi}$.*

Given the revised assumption A-1′, we excise (and ignore) all inactive actions. Denote the posterior belief thresholds by $y = (y_1, \ldots, y_{\alpha-1})$, with $y_1 < \cdots < y_{\alpha-1}$.

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

$$B(\pi, y) \equiv \sum_{a=1}^{\alpha} \psi(a, \pi, y) \left[(1-\delta)\bar{u}(a, q(a, \pi, y)) + \delta v(q(a, \pi, y))\right]$$

We need a traditional comparative statics assumption:

A-4 *The matrix of second derivatives $B_{yy}(\pi, y(\pi))$ is negative definite for $\pi \in N(\bar{\pi})$.*

Proposition 3 *Assume $\bar{\pi} \in M$ and $\delta > 0$. Under A-1′, A-2, A-3, and A-4, contrarian behaviour is encouraged: the threshold vector $y(\pi)$ strictly increases in $\pi \in N(\bar{\pi})$.*

The first order condition for y is again rewritten

$$B_{y_a}(\pi, y) \equiv \frac{\partial \psi(a, \pi, y)}{\partial y_a} \left[(1-\delta)\bar{u}_a(y_a) + \delta \tau_a(y_a) - (1-\delta)\bar{u}_{a+1}(y_a) - \delta \tau_{a+1}(y_a)\right] \qquad (8)$$

where $\tau_a(y)$ is the tangent function to v at q_a. A marginal change in y affects B_{y_a} only through y_a, y_{a-1} (affecting $q(a,\pi,y)$), and y_{a+1} (affecting $q(a+1,\pi,y)$). An increase in y_{a-1} will increase $q(a,\pi,y)$, and since $q(a,\pi,y) < y_a$ and v is convex, the tangent value $\tau(y_a, q(a,\pi,y))$ will increase. In turn, $B_{y_a y_{a-1}} \geq 0$ from (8). This supermodularity property, together with negative definiteness implies that the inverse matrix is nonpositive:

Lemma 9 *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 $z \geq 0$, consider the function $F(x,t) = x'Bx + tz'x$ where $t \in \mathbb{R}$. Note that $F_x(x,t) = 2x'B + tz'$. Then F is supermodular in x since B has all non-negative off-diagonal elements. Also, F has increasing differences in (x,t) since $F_{xt} = z' \geq 0$. Since B has full rank, the first order condition $2Bx = -tz$ has a unique solution $x^*(t)$. Since $F_{xx} = B$ is negative definite, $x^*(t)$ is the unique maximizer of F. The monotone comparative statics result of Topkis implies that $x^*(t)$ is weakly increasing. Application of the implicit function theorem to $2Bx + tz = 0$ gives $x_t^*(t) = -(1/2)B^{-1}z$. Since $x^*(t)$ is increasing, $B^{-1}z \leq 0$. Since $z \geq 0$ is arbitrary, we conclude that $B^{-1} \leq 0$. □

To complete the proof of the Proposition, apply the implicit function theorem to the first order equation $B_y(\pi, y(\pi)) = 0$ to get $y_\pi(\pi) = -B_{yy}^{-1}(\pi, y(\pi))B_{\pi y}(\pi, y(\pi))$. By Lemma 9, $-B_{yy}^{-1}(\pi, y(\pi))$ has all non-negative elements, strictly positive on the diagonal. Our proof of the single-crossing property extends to show that all elements of the vector $B_{\pi y}(\pi, y(\pi))$ are strictly positive. Thus $y_\pi(\pi)$ has all strictly positive entries.

5 LONG RUN COMPARATIVE STATICS

5.1 Long Run Behaviour in the Optimal Solution

We now ask what happens after an infinite time, which has been the focus of the informational herding literature. Before developing our more novel result here, the first subsection below adds patience to the main limit characterization results of SS.

As usual, the belief process converges by the martingale convergence theorem. Not only must beliefs settle down, but also the planner is never dead wrong about the state.

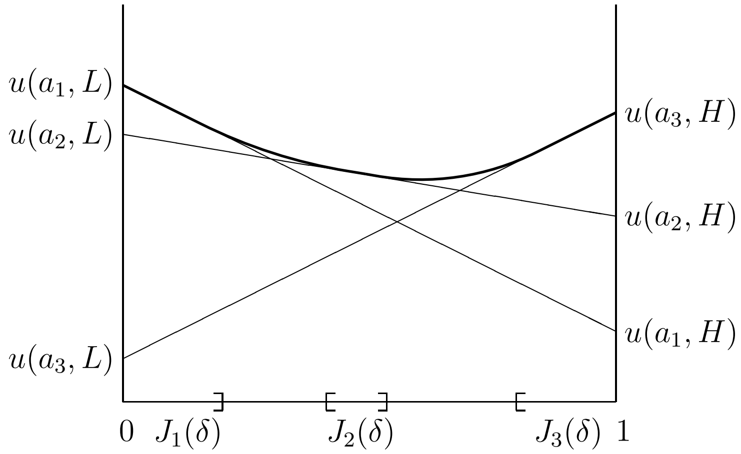

Figure 3: **Typical value function.** Stylized graph of $v(\pi, \delta)$, $\delta \geq 0$, with three actions.

Lemma 10 *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.*

A proof of this folk result is found in SS. The next result is an expression of EK's Theorem 5 that the limit belief π_∞ precludes further learning. In the informational herding model, this is only possible during a cascade, when one action is chosen almost surely, and thus is uninformative. The next characterization of the stationary points of the stochastic process of beliefs $\langle \pi_n \rangle$ directly generalizes the analysis for $\delta = 0$ in SS. See Figure 3 for an illustration of how the shape of the optimal value function reflects the cascade sets.

Lemma 11 (Cascade Sets)
(*a*) *For all $\delta \in [0, 1)$, the support of the limit belief π_∞ is in the sets $J_1(\delta) \cup \cdots \cup J_A(\delta)$.*
(*b*) *With unbounded private beliefs, only the extreme cascade sets $J_1(\delta), J_1(\delta)$ are nonempty.*
(*c*) *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: All but the initial limit belief result are established in the appendix. To see why that one is true — that a *limit cascade* must occur, as SS call it — observe that for any belief $\hat{\pi}$ not in any cascade set, at least two signal outcomes (i.e. actions in A) are realized with positive probability. By the interval structure, the highest such signal is more likely in state H, and the lowest more likely in state L. So the next period's belief differs from $\hat{\pi}$ with positive probability. Intuitively, or by the characterization result for Markov-martingale processes in appendix B of SS, $\hat{\pi}$ cannot lie in the support of π_∞. □

The proof of this result also shows that the larger is δ, the weakly smaller are all cascade sets: Indeed, this drops out rather easily from the monotonicity of the value function in δ. We defer asserting this result for now, as Proposition 5 (later) leverages weak monotonicity.

Lemma 12 (Convergence of Beliefs) *Consider a solution of the planner's problem.*
(a) *For unbounded private beliefs, π_∞ is concentrated on the truth for any $\delta \in [0, 1)$.*
(b) *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 $\pi_\infty \notin J_A(\delta)$.*
(c) *With bounded private beliefs, the chance of complete learning ($\pi_\infty \in J_A(\delta)$ in state H) tends to 1 as $\delta \uparrow 1$.*
(d) *In the two state informational herding model, if the unconditional log-likelihood ratio density $\phi(\Lambda)$ is strictly log-concave, then beliefs can never enter a cascade set from outside.*

Proof: Part (a) follows from Lemma 10 and Lemma 11-a,b.

Part (b) follows just as in Theorem 1 of SS. We now extend that proof to establish the limiting result for $\delta \uparrow 1$ in part (c). First, Lemma 11 assures us that for δ close enough to 1, π_∞ places all weight in $J_1(\delta)$ and $J_A(\delta)$. The likelihood ratio $\Lambda_n \equiv (1 - \pi_n)/\pi_n$ is a martingale conditional on state H. Because the likelihood ratio $(1 - \sigma)/\sigma$ is bounded above by some $\bar{\Lambda} < \infty$ for all private beliefs σ, the sequence $\langle \Lambda_n \rangle$ is bounded above by $\bar{\Lambda}(1 - \underline{\pi}(\delta))/\underline{\pi}(\delta)$, and the mean of Λ_∞ must equal its prior mean $(1 - \pi_0)/\pi_0$. Since $\lim_{\delta \to 1} \underline{\pi}(\delta) = 0$, the weight that π_∞ places on $J_1(\delta)$ in state H must vanish as $\delta \to 1$.

For part (d), note that assumption A-2 precludes non-monotonic continuation beliefs, by Lemma 6. By SS, these are *necessary* for the existence of nontrivial cascades. □

Part (d) is a contribution about cascade sets quite apart from our contrarianism thrust. It shows how Lemma 6 in §4 is really an independent result in its own right. The discrete signal examples of cascades in BHW trivially all violate assumption A-2.

Observe how incomplete learning besets even an extremely patient planner. So this problem does not fall under the purview of EK's Theorem 9, where it is shown that if the optimal value function v is *strictly convex* in beliefs π, learning is complete for δ near 1. For here, the planner optimally behaves myopically for very extreme beliefs: That is, $v(\pi) = \bar{u}_1(\pi)$ for π near 0, and $v(\pi) = \bar{u}_A(\pi)$ for π near 1, both *affine* functions. This points to the source of the incomplete learning: lumpy signals (actions, in the herding model) rather than impatience. It is simply individuals' inability to properly signal their private information via a finite action mesh that eventually frustrates the learning process.

We are now positioned to reformulate the learning results of the last section at the level of actions. For this, we recall that a *herd* obtains on action a at stage N if all individuals $n = N, N+1, N+2, \ldots$ choose action a. As SS shows, while a cascade implies a herd, the converse is false. To show that herds arise, we can generalize the *Overturning Principle* of SS to this case: The appendix establishes (in Claim 13) that for beliefs π near $J_a(\delta)$, actions other than a will push the updated public belief far from its current value. Thus, convergence of beliefs implies convergence of actions — or, a limit cascade implies a herd. Since cascade sets are constrained efficient (Lemma 12) and beliefs must converge (Lemma 12), herding is likewise constrained best (proof appendicized):

Proposition 4 (Herding is Constrained Efficient) *Consider any planner's solution:*
(*a*) *An ex post optimal herd eventually starts for* $\delta \in [0,1)$ *and unbounded private beliefs.*
(*b*) *With bounded private beliefs, a herd on an action eventually starts. Unless* $\pi_0 \in J_A(\delta)$*, a herd arises on an action other than* A *with positive chance in state* H *for any* $\delta \in [0,1)$.
(*c*) *The chance of an incorrect herd with bounded private beliefs vanishes as* $\delta \uparrow 1$.

It is no surprise that the planner ends up with full learning 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. Herding is truly a robust property of the observational learning paradigm.

5.2 Monotonic Cascade Sets

We finally establish a strict limit comparative static: If the planner is indifferent about learning at some belief (i.e. at the edge of a cascade set), then he strictly prefers to learn if he is slightly more patient.

Proposition 5 *Assume bounded beliefs. All non-empty cascade sets shrink strictly when* δ *rises:* $\forall a \in A$*, if* $\delta_2 > \delta_1$ *and* $J_a(\delta_1) \neq \emptyset$*, then* $J_a(\delta_2) \subset J_a(\delta_1)$.

In other words, as individuals grow more patient, the set of cascade beliefs which foreclose on learning strictly shrinks. In particular, the cascade set is strictly smaller than the myopically efficient level for even slightly patient individuals. This is an analogue of our short-run comparative static Proposition 2, since it means that for any positive discount factor, some individuals should behave in a contrarian fashion, violating the cascade. However, here no log-concavity assumption is needed.

The proof makes use of two key results, with sufficient independent interest that we hereby summarize them here (but appendicize their proofs).

Lemma 13 (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, provided the planner's strategy in some future eventuality strictly prefers a non-myopic action, his continuation value must strictly exceed his myopic value. We show that this holds for any continuation public belief outside both cascade sets $J_a(\delta_1) \supseteq J_a(\delta_2)$. So a more patient player who weights the continuation value more, will enjoy a higher value.

Call two rules *equivalent* if they are represented by the same thresholds, with the same associated mixing, if any, at the endpoints. Comparing with Step 1 of §4, we now partially characterize the differentiability of the value function at the edge of cascade sets.

Lemma 14 (Differentiability) *Let $\hat{\pi} \in (0,1)$ be an endpoint of cascade set $J_a(\delta)$. If all rules optimal at $\hat{\pi}$ are equivalent, then $v_\delta(\cdot)$ is differentiable at the belief $\hat{\pi}$.*

Since cascade sets *do* change with the discount factor, we now explain why we proved a Second Welfare Theorem (Lemma 1) but not a First. For there exists a team equilibrium which is not optimal. To see why, assume an initial public belief $\pi \in J_a \setminus J_a(\delta)$. Then herding on action a is a team equilibrium: Given that everyone subsequent believes that individual n's action is uninformative, the best he can do is to maximize his own current payoff, and when $\pi \in J_a$ it is myopically optimal to take action a with probability one. But since $\pi \notin J_a(\delta)$, this is not optimal.

6 CONCLUSION

This paper has discovered and explored the fact that informational herding is simply incomplete learning by a single experimenter, suitably concealed. Our mapping, recasting everything in rule space, has led us to an equivalent social planner's problem. In so doing, we have contributed to the small literature on optimal interval rules. While the revelation principle in mechanism design also uses such a 'rule machine', the exercise is harder for multi-period, multi-person models with uncertainty, since the planner must respect the agents' belief filtrations. While this is trivially achieved in rational expectation price settings, one must here exploit the public beliefs, and largely invert the model.

Once informational herding is understood as single-person Bayesian experimentation, it no longer seems so implausible that incorrect herds may be constrained efficient. For incomplete learning is if anything the hallmark of optimal experimentation models, even with forward-looking behaviour. This is the setting we then explored.

The thrust of our paper consists of two strict comparative static results, which formalize and illustrate the social efficiency of contrarian behaviour in an informational herding model. First, in the short-run — that is, at any particular finite stage — given a log-concave log-likelihood ratio density, as an action becomes more popular, individuals should more strongly lean against taking it on the basis of their private information. This log-concavity condition is new in information economics. Very loosely, its motivation owes to the fact that Bayes rule entails a linear aggregation of log-likelihood ratios.

Second, we showed that with a finite action space, informational herding is not in itself bad, but is in fact a constrained efficient long-run outcome of the social planner's problem, and is robust to changing the planner's discount factor. For any bounded signal distribution, cascade sets for actions strictly shrink in the discount factor $\delta < 1$. So (i) for some public beliefs, the team solution calls for violating a myopic cascade, but (ii) the cascade set never vanishes for any discount factor δ.

Our exploration has hinged crucially on an extension of the Gittins index of the social present value of each action. This formulation differs from the standard bandit index because of the agency problem: Since private signals are privately observed, aligning private and social incentives entails a translation using the marginal social value.

En route to our main results, we have found two useful cases where we can prove differentiability of the value function in problems of learning. We are aware of no other examples in the literature of such conditions. We have also found a simple robust condition under which cascades cannot be entered in two state herding models.

We close with remarks on related literature that has appeared after we first wrote this paper. Sgroi (2002) proposes that late individuals in the herding model can benefit from a change in the observational regime, so that the first N individuals cannot see their predecessors' actions. This policy is available to the social planner in our model: Individuals need simply use the myopically optimal threshold $x(\pi_0)$, where π_0 is the initial public belief. In general, this special policy is sub-optimal. Doyle (2002) considers the social planner's problem in the investment delay problem of Chamley and Gale (1994).

A APPENDIX: OMITTED PROOFS

A.1 Upper-hemicontinuity: Proof of Lemma 3-(b)

Given the interval structure of Lemma 2, the planner simply determines the chances $\psi(\pi, \xi^\delta)$ of choosing each action, and their order. Represent this by the product of an

$A \times A$ permutation matrix and the vector $(\psi(\pi, \xi^\delta), a = 1, \ldots, A)$. 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.

A.2 On the Log-Concavity Assumption

We note that for a twice differentiable signal density f, Assumption A-2 reduces to

$$\frac{f(\sigma)f''(\sigma) - [f'(\sigma)]^2}{f(\sigma)^2} + \frac{1}{\sigma(1-\sigma)}\left[\frac{f'(\sigma)}{f(\sigma)}(1-2\sigma) - 2\right] \leq 0$$

This neither implies nor is implied by the common assumption of a log-concave signal density. It is satisfied if f is log-concave and $f'(\sigma)/f(\sigma) \lesseqgtr 2/(1-2\sigma)$ for $\sigma \lesseqgtr 1/2$. Since $f'(\sigma)/f(\sigma)$ falls in σ for f log-concave, it suffices that f be log-concave and $f'(0)/f(0) \leq 2$ and $f'(1)/f(1) \geq -2$. The canonical family of Beta distributions on $[0,1]$ with non-explosive tails satisfies it. The distribution of the log-likelihood ratio is an equally natural primitive of our model, and admits all well-known log-concave distributions. (See Marshall and Olkin (1979), §18.*B.2.d*, for a partial list of the families with log-concave densities.) Any concavity assumption is defined relative to the additive structure of the variable's space, and the log-likelihood ratio space is natural for then Bayes' rule is additive.

A.3 Marginal and Average Private Signals: Proof of Lemma 5

We prove the first relation — the other is symmetric. Put $\psi = \pi\psi^H + (1-\pi)\psi^L$, where $\psi^\omega = \mu^\omega([0,\theta])$ for the private belief threshold $\theta = r(1-\pi, y)$. Recalling that $d\mu^H/d\mu = 2\sigma$ and $d\mu^L/d\mu = 2(1-\sigma)$, we have $\partial\psi/\partial y = 2f(\theta)[\pi\theta + (1-\pi)(1-\theta)](\partial\theta/\partial y)$. Differentiating the quotient $q_1 = \pi\psi^H/\psi$, we find that

$$\psi\frac{\partial q_1}{\partial y} = \psi\frac{\partial}{\partial\theta}\left(\frac{\pi\psi^H}{\pi\psi^H + (1-\pi)\psi^L}\right)\frac{\partial\theta}{\partial y} = \frac{\pi(1-\pi)[\theta\psi^L - (1-\theta)\psi^H]f(\theta)}{\psi}\frac{\partial\theta}{\partial y}$$

Finally, the relation obtains once we see that

$$(y - q_1)\frac{\partial\psi}{\partial y} = 2\left(\frac{\pi\theta}{\pi\theta + (1-\pi)(1-\theta)} - \frac{\pi\psi^H}{\psi}\right)f(\theta)[\pi\theta + (1-\pi)(1-\theta)]\frac{\partial\theta}{\partial y} \qquad \square$$

A.4 Properties of Convex Functions

Proof of Lemma 7: When v is affine on $[z_1, z_2]$, tangents τ_ℓ and τ_h coincide, and so $\tau_\ell(z_3) = \tau_h(z_3)$. Otherwise, the tangents differ, with τ_h steeper than τ_ℓ. Since v is convex, the tangent at z_ℓ lies below v at z_h, and so $\tau_\ell(z_h) \le v(z_h) = \tau_h(z_h)$. Also, $\tau_h(z_3) - \tau_h(z_h) > \tau_\ell(z_3) - \tau_\ell(z_h)$. Altogether, $\tau_\ell(z_3) < \tau_h(z_3)$. The other case follows similarly. □

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

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

A.5 Strict Convexity: Proof of Lemma 8

Claim 2 *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 $\bar{v}(L)$. □

Claim 3 (An Auxiliary Affine Optimization) *Let $y^* \in [0, 1]$ be an optimal posterior threshold at belief $\pi \in [0, 1]$, inducing continuation beliefs q_a after seeing action $a = \ell, h$. Define the affine function $\tilde{u}_a(y) = (1-\delta)\bar{u}_a(y) + \delta\tau_a(y)$ for action $a = \ell, h$, where τ_a is the tangent function to v at q_a. Then $y \in [0, 1]$ is an optimal posterior threshold belief at π if y maximizes*

$$\tilde{B}(\pi, y) = \psi(\pi, y)\tilde{u}_\ell(y) + (1 - \psi(\pi, y))\tilde{u}_h(y).$$

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

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

Proof of Lemma 8: Contrary to strict convexity, assume that v is affine on some $[z_1, z_2] \subseteq M$. By Claim 2, 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 $[q(a, z_1, x), q(a, z_2, x)]$. At this point, the optimal contingent strategy for all the future is still constant, so that — by Claim 2 — the value function is affine on this image interval.

By Claim 3, at any belief $\pi \in [z_1, z_2]$, every solution to $\max_{y\in[0,1]} \tilde{B}(\pi, y)$ is an optimal posterior threshold. The problem $\max_{y\in[0,1]} \tilde{B}(\pi, y)$ can be solved easily ex post: At any posterior belief ρ, take the action a with the greatest $\tilde{u}_a(\rho)$, i.e. the same action on one side of the intersection point y of the affine functions $\tilde{u}_\ell$ and $\tilde{u}_h$. (Since $\pi \in M$, both actions must be taken, precluding $y = 0, 1$.) Since the tangent functions at the continuation beliefs, and thus $\tilde{u}_a$, 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 θ. But the relation $y = r(\pi, \theta)$ uniquely determines π, and so cannot be satisfied on all of $[z_1, z_2]$. Contradiction.

A.6 The Function g

We show that g obeys $(1-\pi)(1-\sigma)f(\sigma)d\sigma = g(\pi, \lambda)d\lambda$. Since $\phi(\Lambda) \equiv f(\sigma(\Lambda))\sigma'(\Lambda)$:

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

To see the MLRP Property of g, we show that $\partial^2 \log(g(\pi, \lambda))/(\partial\lambda\partial\pi) > 0$. 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.

A.7 Cascade Sets: Proof of Lemma 11

Let the Bellman operator T_δ be 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 $v_\delta(\cdot)$ is its unique fixed point in the space of bounded, continuous, weakly convex functions.

Let $\bar{u}(\pi) = \max_a \bar{u}_a(\pi)$ be the myopic expected utility frontier function.[12]

Claim 4 (Iterates and Limit) *The sequence $\{T_\delta^n \bar{u}\}$ consists of pointwise increasing weakly convex functions that converge to $v_\delta(\cdot)$. The value $v_\delta(\cdot)$ weakly exceeds $\bar{u}$, and strictly so outside the cascade sets: $v_\delta(\pi) > \bar{u}(\pi)$ $\forall \delta \in [0, 1)$ and $\forall \pi \notin \cup_{a=1}^{A} J_a(\delta)$.*

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

We have not found a published or cited proof of the following folk result.[13]

Claim 5 (Weak Value Monotonicity) *The value function is weakly increasing in δ: Namely, for $\delta_1 > \delta_2$, $v_{\delta_1}(\pi) \geq v_{\delta_2}(\pi)$ for all π.*

Proof: Clearly, $\sum_{a \in A} \psi(\pi, x) \bar{u}_a(q(a, \pi, x)) \leq \sum_{a \in A} \psi(\pi, x) v(q(a, \pi, x))$ for any x and any function $v \geq \bar{u}$. If $\delta_1 > \delta_2$, then $T_{\delta_1} \bar{u} \geq T_{\delta_2} \bar{u}$, 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 $T_{\delta_1}^n \bar{u} \geq T_{\delta_2}^n \bar{u}$. Let $n \to \infty$ and apply Claim 4. □

Claim 6 (Weak Inclusion) *All cascade sets weakly shrink when δ increases: In other words, $\forall a \in A$, if $1 > \delta_1 > \delta_2 \geq 0$, then $J_a(\delta_1) \subseteq J_a(\delta_2)$.*

Proof: As seen in Claims 4 and 5, $v_{\delta_1}(\pi) \geq v_{\delta_2}(\pi) \geq \bar{u}(\pi) \geq \bar{u}_a(\pi)$ for all π, when $\delta_1 > \delta_2$. For $\pi \in J_a(\delta_1)$, we know $v_{\delta_1}(\pi) = \bar{u}_a(\pi)$ and thus $v_{\delta_2}(\pi) = \bar{u}_a(\pi)$. The optimal value can thus be obtained by inducing a a.s., so that $\pi \in J_a(\delta_2)$. □

Claim 7 (Unbounded Beliefs) *With unbounded private beliefs, the cascade sets for the extreme actions are $J_1(\delta) = \{0\}$ and $J_A(\delta) = \{1\}$. All other cascade sets $J_a(\delta)$ are empty.*

Proof: SS establish for the myopic model that all $J_a(0)$ are empty, except for $J_1(0) = \{0\}$ and $J_A(0) = \{1\}$. Now apply Lemma 4 and Claim 6. □

[12]Observe how this differs from $v_0(\pi) \equiv \sup_x \sum_a \psi(\pi, x) \bar{u}_a(q(a, \pi, x))$. In other words, $v_0(\pi)$ allows the myopic individual to observe one signal σ before obtaining the ex post value $\bar{u}(r(\pi, \sigma))$. In our example of §4.1, we have $\bar{u}(\pi) = \max\langle \pi, 1-\pi \rangle$ and $v_0(\pi) = 1 - \pi + \pi^2$.

[13]But Aghion, Bolton, Harris, and Jullien (1991) do assert without proof (p. 625) that the patient value function exceeds the myopic one.

Claim 8 (Bounded Beliefs) *If the private beliefs are bounded, then* $J_1(\delta) = [0, \underline{\pi}(\delta)]$ *and* $J_A(\delta) = [\bar{\pi}(\delta), 1]$, *where* $0 < \underline{\pi}(\delta) < \bar{\pi}(\delta) < 1$.

Proof: We prove that for sufficiently low beliefs it is optimal to choose a rule x that almost surely induces 1; the argument for high beliefs is very similar. Since action 1 is optimal at belief $\pi = 0$, and is not weakly dominated, it must be the optimal choice for beliefs $\pi \leq \tilde{\pi}$, for some $\tilde{\pi} > 0$. Thus, $\bar{u}_1(\pi) = \bar{u}(\pi)$ on $[0, \tilde{\pi}]$. Since each $\bar{u}_a$ is affine, $\bar{u}_1(\pi) > \bar{u}_a(\pi) + \eta$ for all $a \neq 1$ for some $\eta > 0$, and for all beliefs π in the interval $[0, \tilde{\pi}/2]$.

No observation $a \in A$ can produce a stronger signal than any $\sigma \in \text{supp}(\mu) \subseteq [\underline{\sigma}, \bar{\sigma}] \subset (0,1)$. So any initial belief π is updated to at most $\bar{q}(\pi) = \pi\bar{\sigma}/[\pi\bar{\sigma} + (1-\pi)(1-\bar{\sigma})]$. For π small enough, $\bar{q}(\pi) \in [0, \tilde{\pi}/2]$ and $\bar{q}(\pi) - \pi$ is arbitrarily small, and so is $v_\delta(\bar{q}(\pi)) - v_\delta(\pi)$ small by continuity of v — in particular, 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. □

Claim 9 (Limiting Patience) *For large enough* δ, *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: Fix $\delta \in [0, 1)$, and an action index a ($1 < a < A$) for which $J_a(\delta) = [\pi_1, \pi_2]$, for some $0 < \pi_1 \leq \pi_2 < 1$. Since there are informative private beliefs, $\exists \bar{\sigma}^* \in (1/2, 1)$ with $1 > \mu^H([\bar{\sigma}^*, 1]) > \mu^L([\bar{\sigma}^*, 1]) > 0$. We now consider an alternative rule x taking $a-1$ and a when σ respectively lands in intervals $I_{a-1} = [0, \bar{\sigma}^*]$ and $I_a = [\bar{\sigma}^*, 1]$ (see Lemma 2).

Updating the prior π with the event $\{\sigma \in [\bar{\sigma}^*, 1]\}$ results in the posterior belief $q(\pi) = \pi\mu^H((\bar{\sigma}^*, 1])/[\pi\mu^H((\bar{\sigma}^*, 1]) + (1-\pi)\mu^L((\bar{\sigma}^*, 1])]$ in state H. For any compact subinterval $I \subset (0, 1)$, in particular one with $I \supseteq J_a(\delta)$, there exists $\varepsilon \equiv \varepsilon(I) > 0$ with $q(\pi) - \pi \geq \varepsilon$ for all $\pi \in I$. Thus, q maps the interval $[\pi_2 - \varepsilon/2, \pi_2]$ into (but not necessarily onto) $[\pi_2 + \varepsilon/2, 1]$. Choose $\gamma > 0$ so large that $\bar{u}_a(\pi) < \bar{u}_{a+1}(\pi) + \gamma$ for all $\pi \in [0, 1]$. Since $v_\delta(\pi) > \bar{u}_a(\pi)$ outside $J_a(\delta) = [\pi_1, \pi_2]$, and both are continuous in π, we may also choose $\eta > 0$ so small that $v_\delta(\pi) > \bar{u}_a(\pi) + \eta$ for all $\pi \in [\pi_2 + \varepsilon/2, 1]$. By Claim 5, we thus have $v_{\delta'}(\pi) > \bar{u}_a(\pi) + \eta$ for all $\delta' > \delta$. If $\delta' > \delta$ is so large that $(1-\delta')\bar{u} < \delta'\eta$, then the Bellman equation (1) reveals that our suggested rule x beats inducing a a.s. when $\pi \in [\pi_2 - \varepsilon/2, \pi_2]$. 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. □

A.8 Strict Value Monotonicity: Proof of Lemma 13

We first consider a stronger version of Claim 5. Call the private signal distribution *2S (Two Signals)* if its support contains only two isolated points.

Claim 10 (Unreachable Cascade Sets) *Fix $\delta \geq 0$. If 2S fails, then for any π not in any δ-cascade set (★): an action a is taken with positive chance inducing a posterior belief $q(a, \pi, x)$ not in any δ-cascade set. If 2S holds, then (★) obtains for all non-cascade beliefs π except possibly at most $A - 1$ points, each the unique belief between any pair of nonempty cascade sets $J_{a'}(0)$ and $J_a(0)$ from which both cascade sets can be reached.*

Proof: At a non-cascade belief π, at least two actions are taken with positive chance, and by the interval structure, some action shifts the public belief upwards while another shifts it downwards. With unbounded beliefs, $q(a, \pi, x)$ never lies in a cascade set; so assume bounded beliefs. Let $\text{co}(\text{supp}(F)) = [\underline{\sigma}, \bar{\sigma}]$. Assume that π lies between the nonempty cascade sets $J_{a'}(0) < J_a(0)$, and let $\underline{\pi} = \sup J_{a'}(0)$ and $\bar{\pi} = \inf J_a(0)$. By definition of these cascade sets, $r(\underline{\pi}, \bar{\sigma}) \leq r(\bar{\pi}, \underline{\sigma})$. If all possible actions at π led into a cascade set, then $r(\pi, \underline{\sigma}) \leq \underline{\pi}$ and $r(\pi, \bar{\sigma}) \geq \bar{\pi}$. But these inequalities can only hold with equality since

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

and because the outer terms coincide, as Bayes-updating commutes. So, between $J_{a'}(0)$ and $J_a(0)$ there exists at most one point $\hat{\pi}$ which can satisfy both equations; moreover, such a point exists iff *2S* holds. Indeed, given *2S,* we may simply choose $\hat{\pi}$ to solve $r(\hat{\pi}, \bar{\sigma}) = \bar{\pi}$, while if *2S* fails, then with positive chance, a nonextreme signal is realized, and the posterior q is not in a cascade set. With $\delta > 0$ we have weakly smaller cascade sets by Claim 6 of the Lemma 11 proof, so a $\hat{\pi}$ failing (★) is even less likely to exist — in fact it would further require $\sup J_{a'}(\delta) = \sup J_{a'}(0)$ and $\inf J_a(\delta) = \inf J_a(0)$.

Finally, assume *2S*. Consider any $\hat{\pi}$ with reachable cascade sets $J_{a'}(\delta)$ and $J_a(\delta)$. Then the rule $\hat{x}$ mapping $\underline{\sigma}$ into a' (low signal to $\underline{\pi}$) and $\bar{\sigma}$ into a (high signal to $\bar{\pi}$) is indeed optimal. By convexity, $v_\delta(\pi)$ is at most the average of $v_\delta(\bar{\pi})$ and $v_\delta(\underline{\pi})$ (weights given by transition chances), and $\hat{x}$ achieves this average. So $v_\delta(\cdot)$ is affine on $(\bar{\pi}, \underline{\pi})$. □

We now finish proving Lemma 13. By Claim 4 of Lemma 11's proof, $v_{\delta_1}(\pi) > \bar{u}(\pi)$ 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 $v_{\delta_1}(\pi) = \bar{u}(\pi) < v_{\delta_2}(\pi)$. Suppose π lies outside the δ_1-cascade sets.

Assume first that π satisfies (★) of Claim 10 for δ_1 (and thus also for δ_2). Then at least one action a is taken with positive chance inducing a belief $q(\pi, \xi^{\delta_1}(\pi), a)$ not in a

δ_1-cascade set. Thus, $v_{\delta_1}(q(\pi, \xi^{\delta_1}(\pi), a)) > \bar{u}(q(\pi, \xi^{\delta_1}(\pi), a))$. Since $\delta_2 > \delta_1$,

$$v_{\delta_1}(\pi) = (T_{\delta_1} v_{\delta_1}(\cdot))(\pi) < (T_{\delta_2} v_{\delta_1}(\cdot))(\pi) \leq (T_{\delta_2} v_{\delta_2}(\cdot))(\pi) = v_{\delta_2}(\pi) \tag{9}$$

Next assume that some $\hat{\pi}$ between $J_{a'}(\delta_1)$ and $J_a(\delta_1)$ fails (★) for δ_1. If (9) holds at $\hat{\pi}$, we are done. Assume not. Claim 10 noted that between consecutive cascade sets such $\hat{\pi}$ must be unique, and that it implied *2S*. In that case, (9) holds in a punctured neighbourhood $(\underline{\pi}, \pi) \cup (\pi, \bar{\pi})$ of $\hat{\pi}$, where $\underline{\pi} = \sup J_{a'}(\delta_1)$ and $\bar{\pi} = \inf J_a(\delta_1)$. Also, from the last paragraph of Claim 10's proof, $v_{\delta_1}(\cdot)$ was everywhere an affine function on $[\underline{\pi}, \bar{\pi}]$, which in turn, is a supporting tangent line to the convex function $v_{\delta_2}(\cdot)$ at $\hat{\pi}$ (see Claim 5). As it touches $v_{\delta_2}(\cdot)$ at $\hat{\pi}$ only, $v_{\delta_2}(\underline{\pi}) > v_{\delta_1}(\underline{\pi})$ and $v_{\delta_2}(\bar{\pi}) > v_{\delta_1}(\bar{\pi})$.

To find a lower bound to $v_{\delta_2}(\hat{\pi})$, apply rule $\hat{x}$ from Claim 10's proof at the belief $\hat{\pi}$. Since $\hat{x}$ maps $\underline{\sigma}$ into $\underline{\pi} \in J_{a'}(\delta_1)$ and $\bar{\sigma}$ into $\bar{\pi} \in J_a(\delta_1)$, it yields myopic first-period values $\bar{u}_{a'}(\underline{\pi}) = v_{\delta_1}(\underline{\pi})$ and $\bar{u}_a(\bar{\pi}) = v_{\delta_1}(\bar{\pi})$, and continuation values $v_{\delta_2}(\underline{\pi})$ and $v_{\delta_2}(\bar{\pi})$. By the right hand side of (1), this mixture is worth strictly more than $v_{\delta_1}(\hat{\pi})$:

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

which is clearly at most $v_{\delta_2}(\hat{\pi})$. Given this contradiction, (9) must hold at $\hat{\pi}$. □

A.9 Differentiability: Proof of Lemma 14

We proceed in part by establishing claims. Assume that the value v is not differentiable at $\hat{\pi}$, and that all optimal rules at $\hat{\pi}$ are equivalent. We show this leads to a contradiction. Namely, if all optimal rules at $\hat{\pi}$ are equivalent, then v is differentiable at $\hat{\pi}$, as asserted.

Claim 11 $\forall \varepsilon > 0\ \exists \eta > 0$ *such that when* $|\pi - \hat{\pi}| < \eta$*, any optimal rule at* π *induces action* a *with chance at least* $1 - \varepsilon$ *in states* H, L.

Proof: Since $\hat{\pi} \in J_a(\delta)$, one optimal rule at $\hat{\pi}$ induces a with chance one. This property is shared by all rules optimal at $\hat{\pi}$. Next, if $\psi(\pi, \xi^\delta) = \pi \psi(H, \xi^\delta) + (1 - \pi)\psi(L, \xi^\delta)$ is near 1, so are both $\psi(H, \xi^\delta)$ and $\psi(L, \xi^\delta)$. The claim follows from Lemma 3-(b). □

Claim 12 $\forall N \in \mathbb{N}\ \forall \varepsilon > 0\ \exists \eta > 0$ *so that if* $|\pi - \hat{\pi}| < \eta$ *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 $\eta < 1/2$. By Claim 11, for π_n near $\hat{\pi}$, action a occurs with chance at least $1-\eta$ in each state starting from π_n. If a occurs, then π_{n+1} obeys $|\pi_{n+1} - \pi_n| \leq 4\hat{\pi}(1-\hat{\pi})\eta$, by Bayes rule. So $|\pi_{n+1} - \pi_n|$ can be chosen arbitrarily small when a occurs, for π_n near $\hat{\pi}$.

Choose the initial π so close to $\hat{\pi}$ that if a occurs for the next N periods, the posterior stays so close to $\hat{\pi}$ that a occurs with conditional chance at least $(1-\varepsilon)^{1/N}$ each period. Namely, let $\rho_1 \neq \hat{\pi}$ be so close to $\hat{\pi}$ that $\rho_1(1-\rho_1) \leq 3\hat{\pi}(1-\hat{\pi})/2$ and at any π within $|\rho_1 - \hat{\pi}|$ of $\hat{\pi}$, all optimal rules take a with chance at least $(1-\varepsilon)^{1/N}$ in each state. Let $\eta_1 = |\rho_1 - \hat{\pi}|$ and choose $\rho_2 \neq \hat{\pi}$ within $\eta_1/[8\hat{\pi}(1-\hat{\pi})]$ of $\hat{\pi}$ and so close to $\hat{\pi}$ that a occurs with chance at least $1-\eta_1$ in each state from any π within $|\rho_2 - \hat{\pi}|$ of $\hat{\pi}$. Iterate this to choose $\eta_2 = |\rho_2 - \hat{\pi}|$ and then ρ_3, and then $\rho_4, \ldots, \rho_N$. If the initial π lies within $|\rho_N - \hat{\pi}|$ of $\hat{\pi}$, then it stays within $|\rho_1 - \hat{\pi}|$ of $\hat{\pi}$ the next N periods when a occurs in each period. $\square$

Finally, we finish the proof of Lemma 14. Any optimal strategy starting at belief π yields some state-contingent values $\bar{v}^L$ and $\bar{v}^H$. The affine function $\phi(\rho)$ through $\phi(0) = \bar{v}^L$ and $\phi(1) = \bar{v}^H$ is then tangent to the value function at π.

Since it is optimal to take a forever at $\hat{\pi}$, one tangent to v at $\hat{\pi}$ is the affine function h which intersects $u(a, L)$ at $\pi = 0$ and $u(a, H)$ at $\pi = 1$. Consider the left and right derivatives of v at $\hat{\pi}$, with corresponding tangent lines $h_1(\rho)$ and $h_2(\rho)$ at belief ρ. One of those tangents — say, h_1 — must differ from h (when h_1 differs, necessarily $a > 1$). Define $v_1^L = h_1(0) > \phi(0) = u(a, L)$ and $v_1^H = h_1(1) < \phi(1) = u(a, H)$. Since $u(1, L) \geq v_1^L > u(a, L)$, a unique $m > 0$ exists satisfying $v_1^L = mu(1, L) + (1-m)u(a, L)$.

As v is convex, it is differentiable almost everywhere. So let $\pi_k \uparrow \hat{\pi}$ be a sequence of beliefs converging up to $\hat{\pi}$, with the value function differentiable at each π_k. The tangent function is then uniquely determined for each π_k, and its intercepts at $\rho = 0, 1$ are the state-dependent payoffs of any optimal strategy started at π_k, namely $v^L(\pi_k) \geq v_1^L$ and $v^H(\pi_k) \leq v_1^H$. The inequalities of course follow by convexity of v and $\pi_k < \hat{\pi}$.

Now choose N so large and ε so small that $m/2 \geq 1 - (1-\delta^N)(1-\varepsilon)$. Note that action 1 is strictly the best action in state L. Then by Claim 12, for all large enough k, the expected value $v^L(\pi_k)$ in state L of the optimal strategy starting at π_k is at most

$$
\begin{aligned}
v^L(\pi_k) &\leq (1-\delta^N)(1-\varepsilon)u(a, L) + [1 - (1-\delta^N)(1-\varepsilon)]u(1, L) \\
&\leq (1 - m/2)u(a, L) + (m/2)u(1, L) \\
&< (1-m)u(a, L) + mu(1, L) = v_1^L \leq v^L(\pi_k)
\end{aligned}
$$

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

A.10 Constrained Efficient Herding: Proof of Proposition 4

Near $J_a(\delta)$ we should expect to observe action a. The next lemma states that when other actions are observed they lead to a drastic revision of beliefs, *or* there was a non-negligible probability of observing some other action which would overturn the beliefs.

Claim 13 (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:*
(a) $\psi(\pi, \xi^\delta(\pi)) \geq 1-\varepsilon$, and $|q(b,\pi,\xi^\delta(\pi)) - \pi| > \varepsilon$ for all $b \neq a$ that occur; or
(b) $\psi(\pi, \xi^\delta(\pi)) < 1-\varepsilon$, and $\psi(\pi, \xi^\delta(\pi)) \geq \varepsilon/A$, $|q(a,\pi,\xi^\delta(\pi)) - \pi| > \varepsilon$ for some $a \in A$.

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

First, assume bounded private beliefs. By (c) of Lemma 11, 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 $(0,1)$. Let $\operatorname{co}(\operatorname{supp}(\mu)) = [\underline{\sigma}, \bar{\sigma}]$. 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 $\psi(\pi, \xi^\delta(\pi)) \geq 1-\varepsilon$ for some $\pi \in I$. Then any action $b \neq a$ 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 b implies case (a) (selecting, if necessary, ε even smaller).

If instead $\psi(\pi, \xi^\delta(\pi)) < 1-\varepsilon$, 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 (b).

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 $|\bar{u}(a,\rho) - \bar{u}(a',\rho)|$ exceeds $\kappa\zeta$ for all $a \neq a'$, for all $\rho \in [0,\zeta]$.

Assume that π is near the cascade sets $\{0\}$ or $\{1\}$ — say π near 0. Then only one $\hat{a} \in A$ can have low continuation belief $q(\hat{a},\pi,\xi^\delta(\pi)) \in [0,\zeta]$. 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 q remains in $[0,\zeta]$). So the altered policy is a strict improvement.

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

Finally, assume $\psi(1, \pi, \xi^\delta(\pi)) < 1 - \varepsilon$. Then $\psi(\hat{a}, \pi, \xi^\delta(\pi)) < 1 - \varepsilon$. Otherwise, $\hat{a} \neq 1$ and a myopic gain of at least $(1-\varepsilon)\zeta - \varepsilon U$ obtains from swapping the private beliefs for 1 and $\hat{a}$, 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 $\psi(\hat{a}, \pi, \xi^\delta(\pi)) < 1 - \varepsilon$ there must exist some other action taken with chance at least ε/A yielding continuation belief outside $[0, \zeta]$. Thus, case (b) holds. □

For the proof of Proposition 4, 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

$$\{\omega | \omega \in D_n \text{ infinitely often (i.o.)}\} = \{\omega | \sum_1^\infty P(D_{n+1} | Y_n, \ldots, Y_1) = \infty\}$$

Fix an optimal policy ξ^δ. Choose $\varepsilon > 0$ to satisfy Claim 13 for all actions $1, 2, \ldots, A$. For fixed a, define events $E_n = \{\pi_n \text{ is } \varepsilon\text{-close to } J_a(\delta)\}$, $F_n = \{\psi(\pi_n, \xi^\delta(\pi_n)) < 1 - \varepsilon\}$, and $G_{n+1} = \{|\pi_{n+1} - \pi_n| > \varepsilon\}$. If $E_n \cap F_n$ is true, then Claim 13 scenario (b) must obtain, 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 10, G_n i.o. is a probability zero event. By implication, $E_n \cap F_n$ i.o. has probability zero.

Consider the event H that $\langle \pi_n \rangle$ has a limit in $J_a(\delta)$ and $E_n \cap F_n$ occurs only finitely often. By definition, H implies that eventually $E_n \cap F_n^c \cap G_{n+1}^c$. But $E_n \cap F_n^c$ implies that every action $b \neq a$ leads to G_{n+1}, by Claim 13 (a). Action a is then eventually taken on H. Sum over all a to get a chance one event, by Lemmas 10 and 11. □

A.11 Monotonic Cascade Sets: Proof of Proposition 5

Since $J_a(\delta_2) \subseteq J_a(\delta_1)$ by Claim 6 of §A.7, and $J_a(\delta_1) = \{\pi | v_{\delta_1}(\pi) - \bar{u}_a(\pi) = 0\}$ is closed by continuity of $v_{\delta_1}(\pi) - \bar{u}_a(\pi)$ in π, we need only prove $p = \inf J_a(\delta_1) \notin J_a(\delta_2)$.

Case 1. Assume that at public belief p and with discount factor δ_1, some optimal rule $\hat{x}$ does not a.s. take action a. Instead, with positive chance, $\hat{x}$ takes some action b producing a posterior $q(a_k, \pi, x)$ not in any δ_1-cascade set. [Since a is myopically optimal at $p \in J_a(\delta_1) \subseteq J_a(0)$, the optimal rule $\hat{x}$ cannot induce any other myopically suboptimal action $a' \neq a$ at a stationary belief.] So from Lemma 13, $v_{\delta_2}(q(b, \pi, \hat{x})) > v_{\delta_1}(q(a_k, \pi, \hat{x})) \geq$

$\bar{u}_k(q(a_k, \pi, \hat{x}))$, and as we can always employ the rule $\hat{x}$ with the discount factor δ_2,

$$\begin{aligned} v_{\delta_2}(p) &\geq \textstyle\sum_{a_j \in A} \psi(a_j, p, \hat{x}) \left[(1-\delta_2)\bar{u}_j(q(a_j, p, \hat{x})) + \delta_2 v_{\delta_2}(q(a_j, p, \hat{x}))\right] \\ &> \textstyle\sum_{a_j \in A} \psi(a_j, p, \hat{x}) \left[(1-\delta_1)\bar{u}_j(q(a_j, p, \hat{x})) + \delta_1 v_{\delta_1}(q(a_j, p, \hat{x}))\right] = v_{\delta_1}(p) \end{aligned}$$

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

Case 2. Next suppose that the optimal rule at public belief p with discount factor δ_1 is unique. Then the partial derivative $v_{\delta_1}(p)$ exists by Lemma 14. By the convexity of the value function, any selection from the subdifferential $\partial v(\pi)$ converges to $v_{\delta_1}(p)$ as π increases to p. Since the optimal rule correspondence is upper hemicontinuous by the Maximum Theorem, and uniquely valued at p, the posterior belief $q(b, \pi, \xi^{\delta_1}(\pi))$ is continuous in π at p for any rule optimal selection ξ^{δ_1} and any action b.

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

$$\bar{u}_a(r) - \bar{u}_{a'}(r) = \delta_1 \left[\bar{u}_a(r) - \bar{u}_{a'}(r) + v_{\delta_1}(r) - v_{\delta_1}(p) - m_a^{\delta_1}(r - p)\right] \tag{10}$$

Moreover, from the previous proof of Proposition 1, $m_a^{\delta_1}$ is the slope of $\bar{u}_a$, because the function $h_a(\rho) = v_{\delta_1}(p) + m_a^{\delta_1}(\rho - p)$ evaluates the prospect of taking action a forever.

We prove $w_a^{\delta_2}(p, \underline{\sigma}) < w_{a'}^{\delta_2}(p, \underline{\sigma})$, and so conclude $p \notin J_a(\delta_2)$. If not, assume $w_a^{\delta_2}(p, \underline{\sigma}) \geq w_{a'}^{\delta_2}(p, \underline{\sigma})$, i.e. $p = \inf J_a(\delta_2)$. Subtracting $w_a^{\delta_1}(p, \underline{\sigma}) \geq w_{a'}^{\delta_1}(p, \underline{\sigma})$, we have the contradiction:

$$\begin{aligned} 0 &\geq [w_a^{\delta_2}(p, \underline{\sigma}) - w_{a'}^{\delta_2}(p, \underline{\sigma})] - [w_a^{\delta_1}(p, \underline{\sigma}) - w_{a'}^{\delta_1}(p, \underline{\sigma})] \\ &= (\delta_2 - \delta_1) \left[\bar{u}_a(r) - \bar{u}_{a'}(r) - v_{\delta_1}(p)\right] + \delta_2 v_{\delta_2}(r) - \delta_1 v_{\delta_1}(r) - \delta_2 m_a^{\delta_2}(r - p) + \delta_1 m_a^{\delta_1}(r - p) \\ &> (\delta_2 - \delta_1) \left[\bar{u}_a(r) - \bar{u}_{a'}(r) - v_{\delta_1}(p)\right] + \delta_2 v_{\delta_1}(r) - \delta_1 v_{\delta_1}(r) - \delta_2 m_a^{\delta_1}(r - p) + \delta_1 m_a^{\delta_1}(r - p) \\ &= (\delta_2 - \delta_1) \left[\bar{u}_a(r) - \bar{u}_{a'}(r) - v_{\delta_1}(p) + v_{\delta_1}(r) - m_a^{\delta_1}(r - p)\right] \\ &= (\delta_2 - \delta_1) \left[\bar{u}_a(r) - \bar{u}_{a'}(r)\right] / \delta_1 \geq 0 \end{aligned}$$

Indeed, when $p \in J_a(\delta_1)$, one optimal policy ξ_1^{δ} induces a almost surely at belief p, so that

$q(a', p, \xi^{\delta_1}(p)) = q(p, \xi^{\delta_2}(p), a') = r$. The first equality then follows from (2) for each index, and $v_{\delta_2}(p) = v_{\delta_1}(p)$ when $p \in J_a(\delta_1) \cap J_a(\delta_2)$. The second exploits $m_a^{\delta_1} = m_a^{\delta_2}$ (true as both are the slope of $\bar{u}_a$), and $v_{\delta_2}(r) > v_{\delta_1}(r)$, as given by Lemma 13. The final inequality follows since $p \in J_a(\delta_1) \subseteq J_a(0)$, so that a is myopically optimal at r.

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.

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.

Chernoff, H. (1952): "A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations," *Annals of Mathematical Statistics*, 23, 493–507.

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*, 14, 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.

McLennan, A. (1984): "Price Dispersion and Incomplete Learning in the Long Run," *Journal of Economic Dynamics and Control*, 7, 331–347.

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.

Smith, L., and P. Sørensen (2000): "Pathological Outcomes of Observational Learning," *Econometrica*, 68, 371–398.

Sobel, M. (1953): "An Essentially Complete Class of Decision Functions for Certain Standard Sequential Problems," *Annals of Mathematical Statistics*, 24, 319–337.

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.