Social Learning with Endogenous Information

S. Nageeb Ali*

March 21, 2016

Abstract

I study observational learning with costly information acquisition. Once information is costly, individuals would like to economize on the cost of buying information and free-ride on the public history. Necessary and sufficient conditions for complete learning follow from an elementary principle: a player purchases only that information that can change her mind. With a "coarse" action space, learning is complete if and only if for every cost $c > 0$, a positive measure of types can acquire, at cost less than c, an experiment that can overturn the public history. With a "rich" action space, learning is complete if and only if for every cost $c > 0$, a positive measure of types can acquire an informative signal at cost weakly less than c. Unlike the environment with *costless information*, these results apply even if each individual directly observes all information acquired by her predecessors. I apply this framework to financial trading and show that prices converge to the truth if and only if information is affordable. Otherwise, the bid-ask spread may vanish while leaving a wedge between long-run prices and the value of an asset.

*Pennsylvania State University (email: nageeb@psu.edu). I thank Chris Chambers, Ben Golub, Johannes Horner, Vijay Krishna, Manuel Mueller-Frank, Omar Nayeem, Paul Niehaus, Henrique de Oliveira, Mallesh Pai, Andres Santos, Ron Siegel, Joel Sobel, Tom Wiseman, and especially Aislinn Bohren and Navin Kartik. Tetsuya Hoshino and Erik Lillethun provided excellent proofreading. This work is financially supported by NSF grant SES-1127643.

Contents

1 Introduction **1**

2 Model **4**

3 When is Learning Complete? **5**

3.1 The Affordability of Information 5

3.2 The Responsiveness of a Decision Problem 6

3.3 Main Result 7

3.4 The Benchmark of Costless Information 10

3.5 Observability of Predecessors' Information 10

4 Long-run prices in a financial market with costly information **11**

5 Extensions **12**

5.1 Comparing Decentralized Learning to Centralized Learning 12

5.2 Sequential Information Acquisition 13

5.3 Heterogeneous Priors and Preferences 13

5.4 Multiple States 14

6 Conclusion **14**

A Omitted Proofs **16**

References **19**

1 Introduction

The rational herding literature explores why individuals imitate others, and how a large group may make the wrong choice even if their collective information resolves all uncertainty. It is now well-understood that two frictions generate inefficiencies:

a) *Beliefs are unobservable to successors*: When choosing from a coarse set of actions, an individual's action may only partly reveal her posterior belief to successors. Were the information or beliefs of each player publicly observable or inferable by her successors, learning is necessarily complete.

b) *Private beliefs are bounded*: In a coarse action space, individuals herd with strictly positive probability on the wrong action if all signals induce bounded private beliefs, but such possibilities are precluded if signals induce unbounded likelihood ratios (Smith and Sørensen 2000).

These two principles emerge from frameworks in which information is acquired at no cost. Yet, social learning is sometimes relevant because information is costly: each individual has a motive to acquire less information herself and to free ride instead on the wisdom of others. In this vein, Banerjee (1992) suggests the importance of studying costly information acquisition:

> "The most serious departure of our model from reality is probably our assumption that signals to the agents are essentially free; a more realistic analysis would combine the question of incentives for obtaining these signals with the kinds of considerations we discuss...."

This paper investigates that question building directly on the conventional observational learning framework. I offer a general characterization of long-run learning across a range of environments, without imposing a parametric structure on information or its cost, and while permitting players to be heterogeneous in their information acquisition costs. My results use the most basic property of any consequentialist theory of information demand: information is valuable only if it can change one's action.[1]

Summary of Model and Results: The unobservable state of the world, ω, is low (0) or high (1). Players choose an action from a compact one-dimensional menu A and each prefers to take a higher (resp. lower) action in the high (resp. low) state. Each player first observes actions taken by predecessors, then chooses how much information to acquire about the state, and then chooses her own action. Information is modeled non-parametrically, and corresponds to choosing an an experiment to conduct that comes at a cost. Players may be heterogeneous in the costs they face in conducting experiments, and are not necessarily ordered in their information acquisition costs.

Once a player acquires information, her decision problem corresponds to choosing the action that maximizes her expected utility, conditioning on the history of actions that she observes and her information. In other words, if her belief that the state is high—based on the public history and her

[1]This fundamental property is implicit in Blackwell's comparison of information structures (Blackwell 1951, 1953) and is described by Arrow (1971) who attributes it to Marschak (1959). This property is also used by Schlee (1990) and Grant, Kajii, and Polak (1998) as a point of contrast for theories in which information has intrinsic value.

information—is μ, then she chooses $\max_{a \in A}[\mu u(a, 1) + (1 - \mu)u(a, 0)]$. When that menu of actions A is finite, my study corresponds exactly to incorporating information costs within Bikhchandani, Hirshleifer, and Welch (1992), which is the standard framework for observational learning. That is an important special case of my results. But I wish to consider a class of games with a general menu of actions so that the results may speak to a broader set of applications.

For example, suppose that players are sequentially choosing *how much* to invest in a risky vs. safe technology (or asset), observing the prior history of investment choices, and have the choice to acquire information about the quality of the risky technology. With costless information, we already know that such a continuous-action game can sometimes lead to results that differ sharply from finite-action games because one's information is revealed in one's action (see Lee 1993 and papers that build on this study[2]), but under very natural conditions, continuous action games may not engender complete learning. Accordingly, I take the decision problem as my primitive and assume properties directly on that decision problem to understand when learning is complete. This approach is unconventional, but it simultaneously yields ingredients for complete learning across finite and continuous action games without complicating the analysis.

These ingredients pivot on the *responsiveness* of the decision problem to information, a criterion that I have not seen explored in the prior literature. A decision problem is *responsive* if a player's optimal action is different at different beliefs, and is *unresponsive* otherwise. An unresponsive decision problem is *unresponsive at certainty* (henceforth UAC) if for at least one state ω, her optimal action when she is certain of ω coincides with her optimal action whenever she is sufficiently confident that the state is ω. The finite-action model necessarily satisfies UAC because each undominated action is chosen over an interval of beliefs. But so might a continuous action choice model, including the investment model that I describe above, under natural conditions (see Example 1 below and Chari and Kehoe 2004). The paragon of a responsive decision problem is the conventional continuous-action quadratic loss environment proposed by Lee (1993) in which each player's optimal action matches her belief.

Responsiveness has direct implications for the demand for information. If the decision problem is responsive, a player always "tweaks" her actions based on what she learns. Hence, she values any experiment that offers information with positive chance. By contrast, in a UAC decision problem, once a player is sufficiently confident about the state, she modifies her action only if information overturns her belief. Accordingly, she values only those experiments that overturn her belief with positive probability.

I couple these properties for information demand with an assessment of the cost of different kinds of experiments to determine when long-run learning is complete. I say that an experiment X is *affordable* if for every cost $c > 0$, a positive measure of types can conduct that experiment at a cost less than c. Affordability is a condition on the support of the distribution of costs for a particular experiment, and an experiment is unaffordable if there is a strictly positive minimum cost incurred in conducting it.[3]

[2] A partial list is Vives (1997), Burguet and Vives (2000), Eyster and Rabin (2010, 2014), and Guarino and Jehiel (2013).

[3] Affordability is analogous to conditions imposed on the support of distribution of voting costs in models of endogenous turnout, e.g., Palfrey and Rosenthal (1985), Levine and Palfrey (2007), and Krishna and Morgan (2011).

Theorem 1 states:

> *If players face a responsive decision problem, learning is complete if and only if information is affordable. If players face a UAC decision problem, learning is complete if and only if for every non-degenerate public belief, there is affordable information that can overturn it.*

For responsive decision problems, learning obtains whenever some signal, even binary, is affordable. By contrast, if the decision problem is UAC, whenever the public belief about the state is interior, there must be an affordable signal that can "overturn that belief" and influence a player's action. A sufficient condition is that the affordable signal induces unbounded likelihood ratios. If all affordable information is uniformly bounded in its ability to sway actions, learning is incomplete in a UAC decision problem.

The result dovetails with the insight of Smith and Sørensen (2000) that a finite action decision problem may require unbounded signals to overturn a herd, whereas we know that with the decision problem modeled by Lee (1993), learning is complete whenever informative signals are available. So as to appropriately assess the impact of information being costly, I prove a benchmark result—with costless information—that also uses the language of responsiveness. Theorem 2 establishes that for a responsive decision problem, learning is complete if and only if information is available, and in a UAC decision problem, learning is complete if and only if information induces unbounded likelihood ratios. Formally, this result amounts to an extension of the seminal result of Smith and Sørensen (2000) to all UAC decision problems, including those in which the action space is continuous.

So, at first glance, Theorems 1-2 should appear similar. But the channel through which responsiveness plays a role is different in these two settings. When information is costless, a responsive decision problem ensures that the best-response that maps beliefs to actions is invertible, and therefore, actions are not filtering information; when the decision problem is UAC, that best-response mapping is non-invertible, and once the public history is sufficiently confident, actions begin to filter information. Therefore, when information is costless, the key issue is *inference*. When information is costly, by contrast, the channel through which responsiveness matters is the value of information, and as describe above, a responsive decision problem generates a much stronger demand for information than a decision problem that is UAC.

The best way to distinguish these two mechanisms is to suppose that each player directly observes not only the actions of predecessors but also their information. With costless information, learning trivially reduces to a statistical learning problem, regardless of responsiveness, and therefore, learning is complete in all decision problems. By contrast, with costly information, responsiveness matters *exactly* as described by Theorem 1. Observability of the information of predecessors does not change the implications of responsiveness for information demand.[4]

[4] Because the same results follow even if the information of predecessors is observable, Theorem 1 speaks to contexts where individuals communicate their private information to successors (e.g. Shiller 2000; Çelen, Kariv, and Schotter 2010). It also suggests that the focus on actions filtering information, which is the obviously right focus when information is costless, may be less relevant for long-run learning when information is costly.

2 Model

Actions and Payoffs: Each of an infinite sequence of players $t = 1, 2, 3, \ldots$ makes a single choice from $\mathcal{A}$, a compact subset of $\Re$, in which $|\mathcal{A}| \geq 2$. The payoff from action a in state of the world $\omega \in \Omega \equiv \{0, 1\}$ is $u(a, \omega)$, which is continuous in a for each ω. No action is weakly dominated: if $u(a, 0) > u(a', 0)$, then $u(a, 1) < u(a', 1)$. Therefore, there is no loss of generality in assuming that $u(a, 0)$ is strictly decreasing in a and $u(a, 1)$ is strictly increasing in a. The lowest and highest actions in $\mathcal{A}$ are denoted by $\underline{a}$ and $\overline{a}$ respectively, and these are the uniquely optimal actions in $\omega = 0$ and $\omega = 1$ respectively. The gap in payoffs between the best and worst actions in any state is bounded from above by $\gamma \equiv \max_{\omega \in \Omega} |u(\underline{a}, \omega) - u(\overline{a}, \omega)|$, which corresponds to an upper-bound for the value of information.

Beliefs and Information: Players are uncertain about ω and share a common prior that attributes probability $\pi \in (0, 1)$ to $\omega = 1$. Each player can acquire information about the state. If she chooses to not acquire information, then she faces no information acquisition cost. If she acquires information, she chooses an experiment X from a compact set $\mathcal{X}$, endowed with the weak topology. An experiment X generates a realization in $[0, 1]$ governed by the cumulative distribution function $F_X(\cdot, \omega)$ in state ω, independently of the realizations observed by other players.[5] No realization perfectly reveals the state: $F_X(\cdot, 0)$ and $F_X(\cdot, 1)$ have common support $\Lambda(X)$. The realizations of experiments are normalized by the posterior beliefs that they induce with a neutral prior, so that $\Lambda(X)$ is the range of private posterior beliefs that X may induce with a neutral prior.[6] Every experiment X is at least partially informative, i.e., there exists $p \in [0, 1]$ such that $F_X(p, 0) \neq F_X(p, 1)$. The convex hull of $\Lambda(X)$ is $\left[\underline{p}(X), \overline{p}(X)\right]$. Experiment X is *bounded* if $0 < \underline{p}(X) \leq \overline{p}(X) < 1$ and *unbounded* if $\underline{p}(X) = 0$ and $\overline{p}(X) = 1$. For simplicity, I assume that an experiment is either bounded or unbounded.[7]

A player's cost of running experiments depends on her *information cost type*, θ, which is drawn i.i.d. by a Borel measure ρ that has support Θ, a compact subset of a metric space endowed with the Borel σ-algebra. The cost of running experiment X for a player of type θ is $c(X, \theta)$, which is non-negative and continuous in both arguments.

Timing and Observation: Each player observes actions of all predecessors but not their information. The public history before player $t > 1$ chooses to acquire information is $h^t \equiv (a_i)_{i=1,\ldots,t-1}$. After observing the public history and her type, a player chooses an experiment X to conduct and after observing its realization, chooses an action a.

Payoffs: She earns a payoff of $u(a, \omega)$ in state ω from her action, and therefore, this corresponds to her overall payoff in the game if she acquired no information; if she her information cost type is θ, she

[5] In other words, even if two players conduct the same experiment—or observe the same signal process—the realizations that they observe are i.i.d. conditional on the state.

[6] In other words, if F_X is differentiable at p, and $p \in \Lambda(X)$, $p = \frac{f_X(p,1)}{f_X(p,1)+f_X(p,0)}$. If an individual has prior π, the signal realization p generates a posterior likelihood ratio of $\frac{p\pi}{(1-p)(1-\pi)}$.

[7] Allowing for an experiment X where $\underline{p}(X) = 0$ and $\overline{p}(X) < 1$, or the opposite, does not change the results but would require separate analysis for those cases.

conducted experiment X and then chose a, her payoff is $u(a, \omega) - c(X, \theta)$.

Solution-Concept and Learning: I study Perfect Bayesian equilibria in which indifference is broken by the player choosing the lowest among all her optimal actions.[8] In a PBE, σ, and history h^t, let $\mu_t(h^t) = Pr(\omega = 1|h^t)$ summarize the *public belief* after history h^t. Consider a set $\mathcal{H}$ of infinite length histories, and for such a history h_∞, let h^t_∞ be its truncation to actions in periods $1, \ldots, t-1$. For $\omega \in \{0, 1\}$, let $\mathcal{H}_\omega$ denote the set of histories in $\mathcal{H}$ such that $\lim_{t\to\infty} \mu^t(h^t_\infty) = \omega$. Learning is *complete* if for each $\omega \in \{0, 1\}$, $Pr\left(h_\infty \in \mathcal{H}_\omega \mid \omega\right) = 1$, and otherwise, learning is *incomplete*.

3 When is Learning Complete?

3.1 The Affordability of Information

For each experiment X, consider the distribution of costs of conducting that experiment: $G(X, k) \equiv \rho\left(\{\theta \in \Theta : c(X, \theta) < k\}\right)$.

Definition 1. Experiment X is *affordable* if $G(X, k) > 0$ for every $k > 0$. *Information is affordable* if $\mathcal{X}$ contains an affordable experiment, and otherwise, information is unaffordable.

By continuity, experiment X is affordable if and only if the support of the distribution of costs of conducting X includes 0. The following result establishes that if information is unaffordable, there must be fixed or lumpy costs in information acquisition.

Lemma 1. *If information is unaffordable, there exists $\varepsilon > 0$ such that $c(X, \theta) > \varepsilon$ for every X, θ.*

A stronger condition than affordable information is that the affordable experiments include those that can overturn confident public beliefs. Recall that the convex hull of the support of X is $[\underline{p}(X), \overline{p}(X)]$.

Definition 2. *Overturning information is affordable* if for every $p_*, p^* \in (0, 1)$, there exists an affordable experiment X such that $\underline{p}(X) \leq p_*$ and an affordable experiment Y such that $\overline{p}(Y) \geq p^*$.

The affordability of overturning information implies that for any required level of confidence, there are affordable experiments that can (with positive chance) induce at least that level of confidence; the condition owes its name from there being affordable experiments that can "overturn" the public history. A sufficient, but not necessary, condition is that an unbounded experiment is affordable. Rather than negating Definition 2, I define the unaffordability of overturning information to be *slightly* stronger so as to simplify exposition.

Definition 3. *Overturning information is unaffordable* if there exists $\varepsilon > 0$, $p_* > 0$, and $p^* < 1$ such that $c(X, \theta) < \varepsilon$ implies that $\Lambda(X) \subset [p_*, p^*]$.

I assume that overturning information is either affordable or unaffordable.[9]

[8] The results do not depend on this tie-breaking rule.

[9] The discrepancy between Definition 3 and the negation of Definition 2 is that the former bounds the likelihood ratio of beliefs for an experiment X from both below and above, whereas the latter is compatible with bounds being only from below or above. Below, I describe why I impose this stronger condition rather than negating Definition 1.

3.2 The Responsiveness of a Decision Problem

An innovation of this paper is the notion of responsiveness, which I develop here. For each belief μ, let $a^*(\mu)$ denote the action that maximizes $\mu u(a, 1) + (1-\mu)u(a, 0)$. When there are multiple optimal actions at a belief, I select the lowest one for convenience, but this tie-breaking rule does not affect the analysis. Necessarily, $a^*(\mu)$ is non-decreasing in μ. Because $\mathcal{A}$ contains no weakly dominated actions, $\underline{a}$ is uniquely optimal when $\mu = 0$ and $\overline{a}$ is uniquely optimal when $\mu = 1$.

Definition 4. The decision problem $(\mathcal{A}, u)$ is *responsive* if $a^*(\mu) \neq a^*(\nu)$ whenever $\mu \neq \nu$, and otherwise, it is *unresponsive*. $(\mathcal{A}, u)$ is *unresponsive at certainty* (UAC) if there exists $\underline{\mu} > 0$ and $\overline{\mu} < 1$ such that at least one of the following holds: (i) $a^*(\underline{\mu}) = \underline{a}$, (ii) $a^*(\overline{\mu}) = \overline{a}$. This classification corresponds to the *responsiveness* of a decision problem.

The learning model of Lee (1993) typifies a responsive decision problem being that a player's belief can be perfectly inferred from her action. The standard finite action environment satisfies UAC, but UAC is equally compatible with a continuous action problem. All that UAC captures is a failure of invertibility of a^* in a neighborhood of certainty, which can occur even with continuous actions. I offer first a minimal modification to Lee (1993) that translates the paradigmatic example of a responsive decision problem into one that is UAC.

Example 1. *Suppose, as in Lee (1993), that $u(a, \omega) = -(a-\omega)^2$ and $\mathcal{A} = [0, 1]$. This decision problem is responsive because $a^*(\mu) = \mu$. But now modify that action space to $\mathcal{A}_{mod} = [\frac{1}{4}, 1]$. In that case, the optimal action $a^*_{mod}(\mu) = \frac{1}{4}$ whenever $\mu < \frac{1}{4}$.*

We see in this case that a decision problem may be UAC *even though the space of actions in each case is isomorphic to the space of beliefs*, and payoffs are continuous in actions. Continuity, while necessary for a decision problem to be responsive, is insufficient.

Portfolio choice problems naturally involve both continuous actions and generate UAC decision problems. I offer two examples of such problems below.

Example 2. *Suppose, as in Chari and Kehoe (2004), that players choose how much to invest in a risky project. The menu of actions is $\mathcal{A} = [0, 1]$ and investing a in the risky project generates a payoff $h(a)$ if the technology is good ($\omega = 1$) and 0 otherwise; however much is left is invested in a safe asset, which offers $1:1$ returns corresponding to the investment. The payoff function then is $u(a, \omega) = h(a)1_{\omega=1} + 1 - a$, in which h is strictly concave, $1 < h'(0) < \infty$, and $h(0) = 0$. Observe that $a^*(\mu) = 0$ if $\mu < \frac{1}{h'(0)}$, and therefore, the decision problem is UAC.*[10]

Example 3. *Suppose that players choose, sequentially, how much to invest in a risky asset where each individual has wealth 1 and a strictly concave utility function, which for simplicity is $\log(\cdot)$. Investing a in the risky asset yields a payoff of ka where $k > 1$ if the asset is good ($\omega = 1$), and 0 otherwise. The*

[10] A tractable example of such an investment function is $h(a) = \log(2a+1)$, in which case $a^*(\mu) = \max\{\mu - \frac{1}{2}, 0\}$.

optimal allocation is $a^*(\mu) = \max\left\{\frac{k\mu-1}{k-1}, 0\right\}$*. Therefore, a player chooses to invest nothing in the risky asset if* $\mu < \frac{1}{k}$*, and therefore, the decision problem is UAC.*

Responsiveness has implications for the demand for information, which disciplines learning when information is costly, and for the invertibility of the best-response mapping (from beliefs to actions), which influences learning when information is costless. Theorems 1-2 derive these implications below.

3.3 Main Result

In every decision problem, individuals would like to economize on the cost of buying information and free-ride on the public history. Only information that can change one's mind can possibly have value. Using that obvious fact, responsiveness determines the demand for information, which when coupled with affordability conditions outlined in Section 3.1 characterizes when learning is complete.

Theorem 1. *Fix a prior* $\pi \in (0,1)$.

a) *If* $(\mathcal{A}, u)$ *is responsive, learning is complete if and only if information is affordable.*
b) *If* $(\mathcal{A}, u)$ *is UAC, learning is complete if and only if overturning information is affordable.*
c) *If* $(\mathcal{A}, u)$ *is unresponsive, there exists an open set of priors,* Π*, and a set of experiments* $\mathcal{X}$ *such that if* $\pi \in \Pi$*, learning is incomplete even if information is affordable.*

Let me illustrate this logic through UAC decision problems. When the public history places $1-\varepsilon$ probability on a state, a player's maximum willingness to pay for information is $O(\varepsilon)$. If an experiment that can induce a player to take different actions is affordable, then there exists a positive measure of types that could purchase this information for less than ε and still find it valuable. Because their actions reflect the realization of these experiments, the public history of actions continues to be informative (just as in Smith and Sørensen 2000). Therefore, incorrect herds are overturned permitting behavior to settle on only the correct action. Now suppose that overturning information is unaffordable. When the public history is sufficiently informative, no one is willing to purchase information that could overturn the herd because these overturning experiments are simply too costly. Any other experiment is worthless because it does not change one's action. Learning is therefore incomplete.[11]

Proof of Theorem 1. Let $B(\mu, p)$ be the Bayesian posterior when a player has a prior μ and observes a signal realization p. For each state ω, consider the likelihood ratio with respect to the other state: $l_1^t(h^t) = \frac{1-\mu^t(h^t)}{\mu^t(h^t)}$ and $l_0^t(h^t) = 1/l_1^t(h^t)$. I treat $\langle l_i^t(\cdot)\rangle_{t=1}^{\infty}$ as a stochastic process, and it is straightforward to see that conditioning on $\omega = i$, it is a non-negative martingale. The Martingale Convergence Theorem ensures that it converges almost-surely to a random variable l_i^∞ whose support is in $[0, \infty)$.

[11] I assume Definition 3 rather than negate Definition 2 because it might otherwise be possible that the decision problem is unresponsive for certainty of $\omega = 0$, but not $\omega = 1$, but the affordability criterion in Definition 2 fails only for experiments that generate beliefs below some p_* but not for experiments that generate beliefs above p^*. Thus, learning could be complete in that case because the unresponsiveness lines up exactly with where an affordable experiment can overturn public beliefs. The gap between Definition 3 and negating Definition 2 disappears when $\mathcal{A}$ or $\mathcal{X}$ is finite.

Case 1: $(\mathcal{A}, u)$ **is Responsive:** Suppose that information is affordable, and let X be an affordable experiment. Observe that when the public belief is $\mu \in (0,1)$, the value of experiment X for a responsive individual is

$$V(X,\mu) \equiv \mu \int_0^1 u(a^*(B(\mu,p)),1)dF_X(p,1) + (1-\mu)\int_0^1 u(a^*(B(\mu,p)),0)dF_X(p,0)$$

whereas the value of not acquiring information is

$$\bar{V}(\mu) \equiv \mu \int_0^1 u(a^*(\mu),1)dF_X(p,1) + (1-\mu)\int_0^1 u(a^*(\mu),0)dF_X(p,0). \tag{1}$$

Because X is informative, $B(\mu,p)$ differs from μ with strictly positive probability, and because u is responsive, $a^*(B(\mu,p)) \neq a^*(\mu)$. I argue by revealed preference that $V(X,\mu)$ is strictly higher than $\bar{V}(\mu)$: the decision maker can upon choosing experiment X guarantee herself the payoff of $\bar{V}(\mu)$ by choosing $a^*(\mu)$ for every realization of p. He deviates from this plan with strictly positive probability, which implies that $V(X,\mu) - \bar{V}(\mu) > 0$. Therefore, there exists $\varepsilon > 0$ such that $V(X,\mu) - \bar{V}(\mu) > \varepsilon$. Since X is affordable, there exists a set of types Θ_ε^X of strictly positive ρ-measure, each of which strictly prefers experiment X to not acquiring any information.

I use this to now prove that learning is complete, i.e., that conditional on $\omega = i$, the support of l_i^∞ is $\{0\}$. Suppose towards a contradiction that $l > 0$ is in the support of l_i^∞. Consider any $\tilde{l} \in (l-\varepsilon, l+\varepsilon)$, in which $\varepsilon < l$. By above, when the public belief is $\tilde{l}$, a strictly positive measure of types obtains information and takes actions that perfectly reveal their signal realizations. The Strong Law of Large Numbers implies then that the probability that the public likelihood ratio remains perpetually in $(l-\varepsilon, l+\varepsilon)$ is 0, and therefore, $l > 0$ cannot be in the support of l_i^∞. Therefore, the support of l_i^∞ is $\{0\}$.

Suppose instead that information is unaffordable. Observe that by definition of $a^*(\mu)$,

$$\max\{\mu u(\bar{a},1) + (1-\mu)u(\bar{a},0), \mu u(\underline{a},1) + (1-\mu)u(\underline{a},0)\} \leq \mu u(a^*(\mu),1) + (1-\mu)u(a^*(\mu),0),$$

and therefore, for every μ,

$$[\mu u(\bar{a},1) + (1-\mu)u(\underline{a},0)] - [\mu u(a^*(\mu),1) + (1-\mu)u(a^*(\mu),0)] \leq \min\{(1-\mu)\gamma, \mu\gamma\},$$

where recall that γ is an upper bound for the greatest payoff difference between the best and worst action in any state. The LHS describes the value from obtaining a perfect signal, and the RHS offers a bound on that value. Select $\varepsilon < \gamma$ such that accordingly to Lemma 1, $c(X,\theta) > \varepsilon$ for every $X \in \mathcal{X}$. Observe that if $(1-\mu)\gamma < \varepsilon$ or $\mu\gamma < \varepsilon$, then no player is willing to incur a cost of ε for even a perfectly revealing experiment. Define $\mu^* = 1 - \frac{\varepsilon}{\gamma}$ and $\mu_* = \frac{\varepsilon}{\gamma}$. Therefore, if the public belief ever crosses outside of $[\mu_*, \mu^*]$, no player acquires any information. These are cascade regions and if learning does not stop earlier, it stops when μ^t enters $[0,\mu_*] \cup [\mu^*, 1]$. I argue that these cascade regions generate incomplete learning.

Suppose that $\omega = i$, and towards a contradiction that learning is complete. Then, l_i^t cannot exceed $\frac{\gamma-\varepsilon}{\varepsilon}$ since players stop acquiring information once it does so. The Bounded Convergence Theorem then ensures that $E[l_i^\infty] = \lim_{t\to\infty} E[l_i^t]$, which because $\langle l_i^t(\cdot)\rangle_{t=1}^\infty$ is a martingale, coincides with $l_i^0 > 0$. This yields a contradiction to the claim that $Pr(l_i^\infty = 0 \mid \omega = i) = 1$.

Case 2: $(\mathcal{A}, u)$ **is UAC:** Suppose that overturning information is affordable, and towards a contradiction that for $\mu \in (0,1)$, $\frac{\mu}{1-\mu}$ or its reciprocal is in the the support of l_i^∞. Let X be an affordable experiment such that there exists a set of signal realizations $\tilde{P}$ such that for every $p \in \tilde{P}$, $a^*(B(\mu,p)) \neq a^*(\mu)$ and $\int_{\tilde{P}} dF_X(p,\omega) > 0$ for every $\omega \in \Omega$. The value of experiment X for a responsive individual is $V(X,\mu)$ as characterized in (1). Mirroring the argument of Case 1, it follows that by revealed preference, $V(X,\mu) - \bar{V}(\mu) > 0$, and therefore, there exists $\varepsilon > 0$ such that $V(X,\mu) - \bar{V}(\mu) > \varepsilon$. Because X is affordable, $\rho(\Theta_\varepsilon^X) > 0$, and each type in Θ_ε^X strictly prefers to conduct experiment X than to not acquire information. Therefore, a strictly positive measure of types are conducting experiments that influences their actions. Since the posterior beliefs after these actions depart from μ with strictly positive probability, it follows that when $\omega = i$, $\langle l_i^t\rangle$ converges almost-surely to the random variable l_i^∞ with support $\{0\}$.

Suppose that overturning information is unaffordable. Consider ε, p_*, p^* such that $c(X,\theta) < \varepsilon$ implies $\Lambda(X)$ is a subset of $[p_*, p^*]$. Let $\underline{\mu}$ and $\overline{\mu}$ be the highest and lowest beliefs such that $a^*(\underline{\mu}) = \underline{a}$ and $a^*(\overline{\mu}) = \overline{a}$. Since the decision problem is UAC, either $\underline{\mu} > 0$ or $\overline{\mu} < 1$ or both. Define

$$\mu_+ = \min\left\{\frac{\varepsilon}{\gamma}, \frac{\underline{\mu}(1-p^*)}{\underline{\mu} + p^* - 2\underline{\mu}p^*}\right\},$$

$$\mu^+ = \max\left\{1 - \frac{\varepsilon}{\gamma}, \frac{\overline{\mu}(1-p_*)}{\overline{\mu} + p_* - 2\overline{\mu}p_*}\right\}$$

Observe that for a public belief $\mu \in [0,\mu_+] \cup [\mu^+, 1]$, the most that a player would pay for a perfect signal is ε. So if a player acquires information at all, she acquires a bounded experiment X in which $\Lambda(X)$ is a subset of $[p_*, p^*]$. However, beliefs are sufficiently concentrated around a state i that for all $p \in [p_*, p^*]$, $a^*(B(\mu,p)) = a^*(\delta_i)$, where δ_i is the belief that puts probability 1 on state i. Since her actions are unaffected by this information, she has no incentive to acquire such signals either. Therefore, once the public belief enters $[0,\mu_+] \cup [\mu^+, 1]$, no player has any incentive to acquire information.

Towards clarifying the impact on learning, suppose that $\omega = 1$ and $\mu_+ > 0$. If $\pi \leq \mu_+$, then we are done since no player ever acquires information. Otherwise, if learning has not stopped before, it does so once $\langle l_1^t\rangle$ exceeds $\frac{1-\mu_+}{\mu_+}$. Suppose towards a contradiction that learning is complete at $\omega = 1$. It must be that $l_1^t < \frac{1-\mu_+}{\mu_+}$ for every t since otherwise players would stop acquiring information. The Bounded Convergence Theorem implies that $E[l_1^\infty] = \lim_{t\to\infty} E[l_1^t] = l_1^0 > 0$, which contradicts $Pr(l_1^\infty = 0 \mid \omega = 1) = 1$. An analogous argument applies if $\omega = 0$ and $\mu^+ < 1$.

Case 3: $(\mathcal{A}, u)$ **is Unresponsive:** Suppose that $(\mathcal{A}, u)$ is unresponsive: trivially, consider the setting described in Theorem 2, letting the signal described therein be the only informative signal, and suppose that it is affordable. □

3.4 The Benchmark of Costless Information

Suppose that information, instead of being costly and endogenous, were costless and exogenous (as in the standard herding framework). Each player observes independent realizations from an informative experiment X. For simplicity, I suppose that $F_X(\cdot, \omega)$ is continuously differentiable on its support.

Theorem 2. *Fixing a prior $\pi \in (0, 1)$, the following describes conditions for complete learning when information is costless:*

a) *If $(\mathcal{A}, u)$ is responsive, then learning is complete.*
b) *If $(\mathcal{A}, u)$ is UAC, then learning is complete if and only if X is unbounded.*
c) *If $(\mathcal{A}, u)$ is unresponsive, then there exists an open set of prior-signal combinations (Π, X) such that learning is incomplete if $\pi \in \Pi$ and the available experiment is X.*

When information is costless, responsiveness plays an intuitive role: do actions fully reveal information by making $a^*(\mu)$ invertible? When $(\mathcal{A}, u)$ is responsive, actions are fully revealing and this perpetual accumulation of information inexorably concentrates public belief to the truth (a.s.), as in Lee (1993). Otherwise, if the decision problem is unresponsive, information is lost because an individual's action is a coarse signal of his beliefs (even if his action space is a continuum). When this coarseness manifests at extreme beliefs, extreme signal realizations are needed to sway actions from an inefficient herd, exactly as in the finite-action case studied by Smith and Sørensen (2000).

3.5 Observability of Predecessors' Information

When information is costless, the essence of observational learning is that the information held by predecessors is not directly observable, and can at best, be inferred from actions. Responsiveness matters in Theorem 2 because of its implications for that inference problem, and were past information directly observable, learning would be complete regardless of responsiveness. By contrast, when information is costly, responsiveness continues to matter even if past signals are observable, so long as one makes an additional assumption.

Corollary 1. *Suppose that each player observes the actions and information of predecessors. If information were costless, then learning is complete regardless of the responsiveness of the decision problem. However, if for ρ-almost every θ, $c(X, \theta) > 0$ for every informative experiment X, long-run learning follows the characterization of Theorem 1: if the decision problem is responsive (resp. UAC), then learning is complete if and only if information (resp. overturning information) is affordable.*

Because for almost every θ, $c(X, \theta) > 0$ for every X, almost-surely, no type is obtaining information for free. Therefore, individuals' incentive to acquire information are exactly as in Theorem 1. The intuition is that with costly information acquisition, the responsiveness of the decision problem influences the demand for information, which is separate from the issue of inference. If the decision problem is UAC and overturning information is unaffordable, public beliefs—generated from both past actions and information—that are sufficiently confident will induce players to acquire neither overturning experiments (because those are too costly) nor non-overturning experiments (because they are worthless). Learning is therefore incomplete.

A byproduct of this analysis is that once information is costly, the insights and approach of observational learning apply to social learning contexts where players may communicate their information to successors (e.g. Shiller 2000; Çelen et al. 2010). While such communications foster complete learning when information is costless, learning need not be complete when information is costly.

4 Long-run prices in a financial market with costly information

I re-visit the question of Grossman and Stiglitz (1980): when information is costly, can markets be informationally efficient? I re-visit this question in a Glosten and Milgrom (1985) framework with costly information, and I show that the price of a security converges to its fundamental value if and only if information is affordable.

Consider a single security whose true value, $V(\omega)$, depends on the realization of $\omega \in \{0, 1\}$, and satisfies $0 \leq V(0) < V(1) < \infty$. Players share a common prior that $\omega = 1$ with probability $\pi \in (0, 1)$. There is a pool of traders, some of whom are *noise traders*, and the rest are *speculative traders*. In a period, the trader is a speculative trader with probability $q \in (0, 1)$, independently of the type of other traders. Traders arrive in a fixed sequential order, and each faces a one-time trading decision: she can buy or sell one unit of security, or remain inactive. Conditional on a trader being a noisy trader, she chooses each of these actions with probability $\frac{1}{3}$. A speculative trader has a payoff of $V(\omega) - \overline{P}$ from buying a security at price $\overline{P}$, and a payoff of $\underline{P} - V(\omega)$ from selling a security at price $\underline{P}$, and a payoff of 0 from being inactive. These prices are set by a risk-neutral market-maker who has no private information and is subject to competition. After each history h^t, the market-maker posts a bid price $\underline{P}(h^t)$ at which he is willing to buy a security and an ask price $\overline{P}(h^t)$ at which he is willing to sell a security. The zero profit condition implies that $\overline{P}(h^t) \geq \mu(h^t)V(1) + (1 - \mu(h^t))V(0) \geq \underline{P}(h^t)$, where $\mu(h^t)$ is the public belief that $\omega = 1$ after history h^t.

Each speculative trader can conduct an experiment X that offers information about ω, and incurs cost $c(X, \theta)$, where θ is the information cost type of that speculative trader. Individuals observe the full history of past trading behavior and prices. As in Section 2, I consider a set of $\mathcal{H}$ of infinite length histories, and for such a history, let h^t_∞ be its truncation to periods $1, \ldots, t - 1$.

Definition 5. *The bid-ask spread vanishes* if for every $\omega \in \{0,1\}$,

$$Pr\left(lim_{t\to\infty}(\overline{P}(h^t_\infty) - \underline{P}(h^t_\infty)) = 0 \mid \omega\right) = 1.$$

Prices converge to value if for every $\omega \in \{0,1\}$,

$$Pr\left(lim_{t\to\infty}\max\{\overline{P}(h^t_\infty) - V(\omega), \underline{P}(h^t_\infty) - V(\omega))\} = 0 \mid \omega\right) = 1.$$

Theorem 3. *Prices converge to value if and only if information is affordable, but the bid-ask spread vanishes independently of the affordability of information.*

Here is the intuition. With prices being set by a competitive market-maker, the appropriate version of responsiveness is that of "responsiveness in equilibrium": when the public belief is μ (which feeds into the prices set by the market-maker), and a trader's private belief is ν, does a trader's optimal action $a^*(\nu,\mu)$ change as her belief ν varies in the interval $[\mu-\varepsilon,\mu+\varepsilon]$? That answer is yes because she wishes to sell the security whenever $\nu < \mu$ and buy it whenever $\nu > \mu$. Accordingly, a speculative trader values any informative experiment, which leads to complete learning if information is affordable.

Two insights emerge from Theorem 3. First, the role of prices in flexibly adjusting the tradeoffs between buying and selling securities enhances informational efficiency not only with costless information (as highlighted by Glosten and Milgrom 1985 and Avery and Zemsky 1998) but also when information is costly. Were prices fixed, complete learning would require overturning information to be affordable. Second, these results offer a herding analogue to the failures of information aggregation conjectured by Grossman and Stiglitz (1980): once the price becomes too informative, individuals may stop acquiring information if there are "fixed" or lumpy costs in information acquisition, and that force generates a long-run wedge between the price of an asset and its fundamental value.

5 Extensions

5.1 Comparing Decentralized Learning to Centralized Learning

Since the central force is that players free-ride in their efforts to acquire information, I should compare the decentralized behavior to that of a centralized social planner. Therefore, consider the same decision maker choosing actions at $t = 1, 2, 3, \ldots$ and with costs of acquiring information in period t drawn i.i.d. according to the measure ρ. She discounts payoffs from period t by δ^{t-1} where $\delta < 1$, but does not observe any payoff until the end of the game.[12] When is her long-run learning complete?

Theorem 4. *Consider a centralized Social Planner who is not perfectly patient ($\delta < 1$). Learning is complete either if $(\mathcal{A}, u)$ is responsive and information is affordable or if $(\mathcal{A}, u)$ is UAC and overturning*

[12] One would wish to to fix δ rather than studying the limits as $\delta \to 1$ because in that latter case, she would learn under a wide range of cost conditions, analogous to the complete learning results that emerge in bandit environments with perfectly patient players (e.g. Aghion, Bolton, Harris, and Jullien 1991; Ali 2011).

information is affordable. If information is unaffordable, learning is incomplete.

The sufficient conditions for decentralized learning to be complete are also sufficient for centralized learning to be complete, and information being unaffordable is sufficient to guarantee that centralized learning is incomplete. The main difference between the decentralized and centralized solutions is when the decision problem is UAC, overturning information is unaffordable, but information is affordable. In such cases—for example, if bounded information were free—the centralized Social Planner would learn completely while the decentralized solution still features incomplete learning.[13]

5.2 Sequential Information Acquisition

I model information acquisition as a single-stage process but results extend to a sequential information acquisition environment (e.g. Wald 1947; Moscarini and Smith 2001) in which each individual can choose how long to acquire information.[14] Suppose that each player can acquire multiple signals sequentially, conditioning the acquisition of a signal on the realizations of signals she has already acquired. The model of Section 2 then offers a reduced-form problem of this richer setting where every feasible strategy is represented as an experiment with its (expected) cost. The implications for learning then emerge from the affordability of information or overturning information when the decision problem is responsive or UAC. If, for example, a bounded experiment X is free but overturning information is unaffordable, and an individual can conduct it countless many times, then overturning information is affordable in the reduced-form. By contrast, if individuals can conduct at most $\mathcal{T}$ experiments, then overturning information is unaffordable in the reduced-form.

5.3 Heterogeneous Priors and Preferences

I have focused on a setting in which all players share the same payoffs. The results extend seamlessly to including dominant strategy or "crazy" types whose preferred action is independent of ω. A more interesting extension is that in which players have "monotone" private preferences (every player prefers weakly higher actions in the higher state and weakly lower actions in the lower state), as in Goeree, Palfrey, and Rogers (2006) and Wiseman (2008). A diversity of preferences, or beliefs, can lead to complete learning when information is costless, if there is "sufficient diversity." I show that the same conclusion applies with costly information acquisition.

I model this issue via the lens of heterogeneous beliefs.[15] Suppose that in addition to player i's cost

[13] I have compared behavior to that of a centralized Social Planner who observes the entire sequence of signals. As highlighted by Smith, Sorensen, and Tian (2014), a more subtle comparison is to a "teams-problem" whereby the Social Planner is designing the optimal rule for a team to follow that cannot observe each other's information. With costless information, they show that the optimal solution rewards contrarianism. With costly information acquisition, the Social Planner has to reward the right information cost types to engage in information acquisition. I leave it as a question for future study to characterize when such a teams-problem features complete learning, and whether its solution can be implemented via a simple scheme.

[14] I thank Thomas Wiseman for suggesting this extension.

[15] There is an equivalent model in which players share common priors and their payoffs for each action are augmented by individual private shocks.

type, she has a prior belief $\pi_i \in [0,1]$ that the state of the world is $\omega = 1$, and a player's belief is independent of her cost type. Suppose π_i is distributed independently according to an atomless cdf H that has a strictly positive density on its support $[\underline{\pi}, \overline{\pi}]$, and this is common knowledge. Consider a case in which social learning fails when players share common priors and preferences: each player has the same payoff function, the decision problem is UAC with critical beliefs $\underline{\mu} > 0$ and $\overline{\mu} < 1$ at which a decisionmaker adopts actions $\underline{a}$ and $\overline{a}$ respectively. Finally, let overturning information be unaffordable but information overall be affordable.

Theorem 5. *Learning is complete if* $[\underline{\pi}, \overline{\pi}] = [0,1]$ *and incomplete if* $\underline{\pi} > 0$ *and* $\overline{\pi} < 1$. .

Long-run learning is facilitated by the population having a diversity of opinions because at every interior belief, there is someone who values information even if the decision problem is UAC. As these individuals acquire and inject information, the public belief converges almost-surely to almost all types being close to certainty of the state.[16]

5.4 Multiple States

Theorem 1 generalizes to a richer finite state space with more than two states and a general (multidimensional) action space for the case in which $(\mathcal{A}, u)$ is responsive. When $(\mathcal{A}, u)$ is UAC, the challenge with multiple states is that without greater structure, players cannot order actions or the set of signal realizations that induce those actions. This issue does not arise if $\mathcal{A}$ is a compact subset of $\Re$ and standard monotonicity assumptions are imposed on payoffs and information. Suppose that $u(a, \omega)$ satisfies the single-crossing property so that higher actions are preferred in higher states, and that the realizations of informative signals can be ordered according to the monotone likelihood ratio property so that higher realizations are relatively more likely in higher states. In such cases, higher signal realizations induce higher actions (Athey 2002). With this structure, Theorem 1's characterization of learning when $(\mathcal{A}, u)$ is UAC extends to finite state spaces with more than two states and a one dimensional action space.

6 Conclusion

I conclude by discussing the relationship of the above results to prior work at the intersection of observational learning and costly information acquisition.

Burguet and Vives (2000) and Chamley (2004) study models of social learning in which players face the decision problem studied by Lee (1993), each chooses the precision of a normally distributed signal at a cost that is increasing in its precision, and players are identical in their information acquisition costs. Using the parametric structure of this model, they establish that learning is complete if and only if the marginal cost of precision for a completely noisy signal is zero. Because they restrict attention

[16] I study heterogeneous priors and preferences when players share the same order on actions, since such heterogeneity changes the results of the prior section. Dispensing with this form of monotonicity generates the form of confounded learning studied by Smith and Sørensen (2000) even with costless information, and adding costs to information only cements this possibility.

to responsive decision problems, their insights do not apply to the standard finite-action observational learning environment. Moreover, information in their environment has unbounded likelihood ratios, and so their characterization of complete learning in their environment departs from mine (in which learning is complete if even bounded information is affordable). Our results are complementary insofar as theirs' emphasize the marginal cost of information whereas mine emphasize the distribution of costs of conducting experiments.

Hendricks, Sorensen, and Wiseman (2012) study a sequential search problem in which each player faces a choice to purchase a product, freely obtains some private information about her value for the product, can choose to acquire at some cost a perfect signal about their value for the product, and observe only the fraction of people who have purchased the product. They prove that when the product's quality is low, learning is complete, but when it is high, learning may be incomplete. They derive a closed-form for the probability of an incorrect herd that they use to generate comparative statics and compare to data from an online music market. Because players do not observe the completely ordered history, their analysis does not afford an elementary martingale treatment, and hence, their analysis and setting aren't comparable to that here.

Also restricting attention to finitely many actions, Mueller-Frank and Pai (2016) study a costly search process in which the payoff from each action is distributed i.i.d. and when a player "searches an action," he observes the value of that action perfectly without learning that of any other. The state space is uncountable and multidimensional, and information takes a special form, identifying the value *perfectly* along a dimension of the state. Search is costly, and types are ordered in the cost of search. Complete learning obtains if and only if some type in the support can obtain perfect noise-free information for free, and otherwise, learning is incomplete. Our frameworks, approach, and focus are considerably different, and to the extent that there is an overlap, the content of our results differs, in that mine requires that only overturning information be affordable rather than perfect information.

To summarize: the most closely related literature has proposed environments that have substantially differed from each other, none of which nests the standard observational learning framework.[17] Being that each of these prior papers uses its particular parametric structure to derive conclusions, it is difficult to see what might be needed for complete learning when information costs are introduced into the standard herding environment. This note fills that lacuna by using an elementary approach to explore the implications of costly information across discrete and continuous action spaces, permitting heterogeneity in the cost of information, and generating several extensions that I have explored here.

I have found it useful to frame these results in the language of responsiveness. With costless information, responsiveness captures whether the best-response mapping $a^*(\mu)$ is invertible. With costly information, responsiveness captures the essence of information demand: if the decision problem is responsive, no information is ever worthless, and otherwise, only overturning information is valuable.

[17] The focus of Mueller-Frank and Pai (2016) on some type obtaining *perfect* signals at no cost in a *discrete* action space contrasts with the focus of Burguet and Vives (2000) and Chamley (2004) on *noisy* signals with a *continuous* action space.

A Omitted Proofs

Proof of Lemma 1. Suppose that the conclusion is false. Then there exists a sequence $(X_n, \theta_n)_{n=1,2,...}$ such that $c(X_n, \theta_n) < \frac{1}{n}$. Because $\mathcal{X} \times \Theta$ is compact, the sequence has a convergent subsequence (Aliprantis and Border 2006, Theorem 2.3.1). Let (X^*, θ^*) be a limit point of that subsequence; being a limit point implies that $c(X^*, \theta^*) < \frac{1}{n}$ for each n, and therefore, equals 0. Because c is continuous, it follows that for every $\varepsilon > 0$, there exists a sufficiently small neighborhood $N(\theta^*)$ such that for each $\theta' \in N(\theta^*)$, $c(X^*, \theta') < \varepsilon$. Since θ^* is in the support of ρ, it follows that $\rho(N(\theta^*)) > 0$, and therefore, X^* is affordable. Taking the contrapositive establishes the result. □

Proof of Theorem 2. For a measurable set of actions A, let

$$P(A, \mu) \equiv \{p \in \Lambda(X) : a^*(B(\mu, p)) \in A\},$$
$$\alpha(A, \mu, \omega) \equiv \int_{P(A,\mu)} dF_X(p, \omega).$$

An action a is in the support, $\overline{\mathcal{A}}(\mu)$, if for every $\varepsilon > 0$, $\alpha((a - \varepsilon, a + \varepsilon), \mu, \omega) > 0$ for every ω. Let $p^*(a, \mu)$ and $p_*(a, \mu)$ be the *sup* and *inf* of $P(\{a\}, \mu)$ respectively. Since $F_X(\cdot, \omega)$ is continuously differentiable, it follows that for every measurable subset A, $\alpha(A, \mu, \omega)$ is continuous in μ. Let $\beta(a, \mu)$ be the updated public belief when action $a \in \overline{\mathcal{A}}(\mu)$ is chosen at public belief μ; for every action a such that $p^*(a, \mu) \neq \underline{p}(X)$, $\beta(a, \mu)$ is continuous in μ. Define the cascade set of beliefs to be $\mathcal{C}(X) \equiv \bigcup_{a \in \mathcal{A}} \{\mu \in [0, 1] : P(\{a\}, \mu) = \Lambda(X)\}$.

Lemma 2. *A public belief $\mu \in \mathcal{C}(X)$ if and only if $\alpha(A, \mu, 0) = \alpha(A, \mu, 1)$ for every measurable A.*

Proof. If $\mu \in \mathcal{C}(X)$, it trivially follows that there exists an action a such that $F_X(p^*(a, \mu), \omega) - F_X(p_*(a, \mu), \omega) = 1$ for each ω. Suppose that $\mu \notin \mathcal{C}(X)$. Then there exists an action $\tilde{a}$ such that $F_X(p^*(\tilde{a}, \mu), \omega) \in (0, 1)$, and consider the set of actions $[\underline{a}, \tilde{a}]$: by Lemma A.1 of Smith and Sørensen (2000), it follows that $F_X(p^*(\tilde{a}, \mu), 1) < F_X(p^*(\tilde{a}, \mu), 0)$. □

For each state ω, consider the likelihood ratio with respect to the other state: $l_1^t(h^t) = \frac{1 - \mu^t(h^t)}{\mu^t(h^t)}$ and $l_0^t(h^t) = 1/l_1^t(h^t)$. I treat $\langle l_i^t(\cdot) \rangle_{t=1}^{\infty}$ as a stochastic process, and it is straightforward to see that it is a non-negative martingale conditioning on $\omega = i$. The Martingale Convergence Theorem ensures that it converges almost-surely to a random variable l_i^{∞} whose support is in $[0, \infty)$.

Lemma 3. *Conditional on $\omega = i$, the likelihood ratio l is in the support of l_i^{∞} implies that $\frac{1}{1+l}$ is a subset of $\mathcal{C}(X)$ if $i = 1$, and $\frac{l}{1+l}$ is a subset of $\mathcal{C}(X)$ if $i = 0$.*

Proof. Suppose towards a contradiction that the support of l_1^{∞} includes l such that $\mu = \frac{1}{1+l}$ is not in $\mathcal{C}(X)$. Consider action $\tilde{a}$ such that $F_X(p^*(\tilde{a}, \mu), \omega) \in (0, 1)$, and $\beta(\tilde{a}, \mu) < \mu$; such an action must exist by Lemma 2 and the law of iterated expectations. By monotonicity, for each $a \in [\underline{a}, \tilde{a}] \cap \overline{\mathcal{A}}(\mu)$, $|\beta(a, \mu) - \mu| \geq |\beta(\tilde{a}, \mu) - \mu|$. Let $\tilde{\alpha} = \frac{\alpha([\underline{a}, \tilde{a}], \mu, 1)}{2}$. Since $\alpha(\cdot, \mu, \omega)$ and $\beta(\cdot, \mu)$ are continuous in μ, it follows that there exists $\varepsilon > 0$ such that for every $\mu' \in (\mu-\varepsilon, \mu+\varepsilon)$, the updated belief is in $(\mu-\varepsilon, \mu+\varepsilon)$ with probability at most $1-\tilde{\alpha}$, yielding a contradiction. An analogous argument applies for l_0^∞. □

Now suppose $(\mathcal{A}, u)$ is responsive. Then $\mathcal{C}(X) = \{0,1\}$, and since the Martingale Convergence Theorem ensures that l_i^∞ has support in $[0,\infty)$, Lemma 3 implies that $Pr(l_i^\infty = 0 \mid \omega = i) = 1$.

Suppose that $(\mathcal{A}, u)$ is UAC. For an unbounded experiment X, $\mathcal{C}(X) = \{0,1\}$, and so as above, $Pr(l_i^\infty = 0 \mid \omega = i) = 1$. Now suppose that X is bounded. Let $\underline{\mu}$ and $\overline{\mu}$ be the highest and lowest beliefs respectively such that $a^*(\underline{\mu}) = \underline{a}$ and $a^*(\overline{\mu}) = \overline{a}$. We consider the following cases below.

1. Suppose that $0 < \underline{\mu} < \overline{\mu} < 1$. Define

$$l_* \equiv \left(\frac{1-\underline{\mu}}{\underline{\mu}}\right)\left(\frac{1-\underline{p}(X)}{\underline{p}(X)}\right), l^* \equiv \left(\frac{1-\overline{\mu}}{\overline{\mu}}\right)\left(\frac{1-\overline{p}(X)}{\overline{p}(X)}\right).$$

   It follows that once l_1^t enters $[0, l^*] \cup [l_*, \infty]$, all subsequent players choose the same action regardless of their signal realization. Learning is incomplete in both $\omega = 0, 1$.
2. Suppose that $\underline{\mu} > 0$ but $\overline{\mu} = 1$. To show that learning is incomplete with strictly positive probability, it suffices to establish that there exists l such that $Pr(l_1^\infty > l \mid \omega = 1) > 0$. Suppose otherwise. Then, $E[l_1^\infty \mid \omega = 1] = 0$. However, it must also be that for every t, $Pr(l_1^t < l_*) = 1$ since otherwise, there is positive probability that the public likelihood ratio converges to a positive number. Since l_1^t is dominated by l_*, we can apply the Bounded Convergence Theorem to establish that $E[l_1^\infty \mid \omega = 1] = \lim_{t\to\infty} E[l_1^t \mid \omega = 1]$, which equals $l_1^0 > 0$ since $\langle l_1^t \rangle$ is a martingale, yielding a contradiction.
3. The argument is analogous for $\underline{\mu} = 0$ but $\overline{\mu} < 1$ by considering the stochastic process $\langle l_0^t \rangle$.

Finally, suppose that $(\mathcal{A}, u)$ is unresponsive. Consider an action a and a range of beliefs $[\underline{\mu}, \overline{\mu})$ such that for every $\mu \in [\underline{\mu}, \overline{\mu})$, $a^*(\mu) = a$. Consider any combination of prior-signal combination (π, s) such that $\pi \in [\underline{\mu}, \overline{\mu})$, and $B(\pi, \overline{p}(X)) < \overline{\mu}$ and $B(\pi, \underline{p}(X)) > \underline{\mu}$. For such combinations, every player chooses action a regardless of her signal realization. □

Proof of Theorem 3. Fix an equilibrium of the game. Each trader's best response at a history h^t depends only on the public belief $\mu(h^t)$. Let $\overline{P}(\mu)$ and $\underline{P}(\mu)$ denote the ask and bid prices for a history that generates public belief μ.

As before, $\langle l_i^t \rangle$ is a martingale conditioning on $\omega = i$, and converges almost-surely to a random variable l_i^∞. Suppose that $\mu \in (0,1)$, $\frac{\mu}{1-\mu}$ or its reciprocal is in the support of l_i^∞. For μ to be a limit belief, it must be that there is no more trading done on the basis of information, and therefore, $\overline{P}(\mu) = \underline{P}(\mu) = \mu V(1) + (1-\mu)V(0)$. Therefore, the bid-ask spread vanishes.

If a speculative trader does not acquire information, then it is optimal at those prices for her to remain inactive (or at least, she is indifferent between trading and not). Suppose that information is affordable. Let X be an informative experiment, which implies that with strictly positive probability,

X generates realizations p such that $B(\mu, p) \neq \mu$. But if $B(\mu, p) < \mu$, then it is strictly optimal for her to sell the security, and if $B(\mu, p) > \mu$, then it is strictly optimal for her to buy the security. Therefore, it follows, as in the proof of Theorem 1 that the value of acquiring signal X strictly exceeds that of not acquiring any information. Since X is affordable, a strictly positive measure of types acquire information and trade speculatively. But then setting $\overline{P}(\mu) = \underline{P}(\mu) = \mu V(1) + (1-\mu)V(0)$ is not an equilibrium price for the market maker.

Suppose instead that information is unaffordable, and let $\varepsilon > 0$ be a cost such that any informative experiment costs more than ε. If the public belief is μ, an upper-bound on how much a speculative trader values perfect information is

$$\mu\left[V(1) - (\mu V(1) + (1-\mu)V(0))\right] + (1-\mu)\left[\mu V(1) + (1-\mu)V(0) - V(0)\right] = 2\mu(1-\mu)(V(1) - V(0)).$$

Therefore, there exists μ_* and μ^* such that if $\mu < \mu_*$ or $\mu > \mu^*$, the value of perfect information is smaller than ε. Consider towards a contradiction an infinite history h_∞ such that in state ω, $lim_{t\to\infty} \max\{\overline{P}(h^t_\infty) - V(\omega), \underline{P}(h^t_\infty) - V(\omega))\} = 0$. Such a history would require players to acquire information and trade speculatively for μ arbitrarily close to 0 or 1. But once μ escapes $[\mu_*, \mu^*]$, no speculative trader has any incentive to conduct an experiment, leading to a contradiction. □

Proof of Theorem 4. First consider the case where (resp. overturning) information is affordable, and $(menu, u)$ is responsive (resp. UAC). Suppose that $\omega = 0$ (the argument is identical for $\omega = 1$) and suppose towards a contradiction that learning were incomplete. So there exists $\mu \in (0,1)$ such that $\frac{\mu}{1-\mu}$ is in the support of l_0^∞. I use the one-shot deviation principle to establish a contradiction. Consider an affordable informative (resp. overturning) experiment X. We have already established in the proof of Theorem 1 that there exists $\varepsilon > 0$ such that $V(X, \mu) - \bar{V}(\mu) > \varepsilon$. Because $\rho(\Theta^X_\varepsilon) > 0$, it follows that whenever the Social Planner's cost type is in Θ^X_ε, she would strictly benefit by doing a one-shot deviation where she conducts experiment X rather than not acquiring information, which is a contradiction.

Now suppose that information is unaffordable. Let $\varepsilon < \gamma$ be such that $c(X, \theta) > \varepsilon$. Consider beliefs $\mu^* = 1 - \frac{(1-\delta)\varepsilon}{\gamma}$ and $\mu_* = \frac{(1-\delta)\varepsilon}{\gamma}$. Notice that for beliefs $\mu \in [\mu^*, 1] \cup [0, \mu_*]$, the Social Planner would not be willing to pay ε once for a fully revealing signal. Therefore, if learning does not stop beforehand, it stops once beliefs reach this region. □

Proof of Theorem 5. Normalize the public belief $\mu^t(h^t)$ as if the prior were neutral, and let $B(\mu, p, \pi_i)$ denote the posterior belief of player i when the public belief is μ, she observes signal realization p, and her prior belief is π_i. Suppose that $\underline{\pi} = 0$ and $\overline{\pi} = 1$, and a bounded experiment X such that $c(X, \theta) = 0$, and $\Lambda(X) \subset [p_*, p^*]$ including the end-points. At a public belief μ, consider a prior π_i that satisfies

$$\frac{\overline{\mu}}{1-\overline{\mu}}\left(\frac{\mu}{1-\mu}\right)^{-1}\left(\frac{\overline{p}}{1-\overline{p}}\right)^{-1} < \frac{\pi_i}{1-\pi_i} < \frac{\overline{\mu}}{1-\overline{\mu}}\left(\frac{\mu}{1-\mu}\right)^{-1}\left(\frac{\underline{p}}{1-\underline{p}}\right)^{-1}.$$

Observe that a player with prior π_i has a strictly positive value for the experiment X since she chooses different actions in every small neighborhood of $\overline{p}$ than she does in every small neighborhood of $\underline{p}$. Since there is a strictly positive measure of players with such priors, there exists for every $\varepsilon > 0$, a strictly positive measure of types (θ_i, π_i) that strictly prefer to obtain the bounded experiment X at cost no less than ε to no information at all. Since the public belief departs from μ with strictly positive probability, it follows that l_i^∞ has support $\{0\}$. By contrast if $\underline{\pi} > 0$ and $\overline{\pi} < 1$, this case of Theorem 1 is easily extended by re-defining μ_+ and μ^+ to account for the range of possible priors. □

References

Aghion, P., P. Bolton, C. Harris, and B. Jullien (1991): "Optimal learning by experimentation," *The review of economic studies*, 58, 621–654.

Ali, S. N. (2011): "Learning self-control," *The Quarterly Journal of Economics*, 126, 857–893.

Aliprantis, C. D. and K. Border (2006): *Infinite dimensional analysis: a hitchhiker's guide*, vol. 3rd Edition, New York, N.Y.: Springer.

Arrow, K. J. (1971): "The value of and demand for information," *Decision and organization*, 2, 131–139.

Athey, S. (2002): "Monotone comparative statics under uncertainty," *The Quarterly Journal of Economics*, 117, 187–223.

Avery, C. and P. Zemsky (1998): "Multi-Dimensional Uncertainty and Herd Behavior in Financial Markets," *American Economic Review*, 88, 724–48.

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

Bikhchandani, S., D. Hirshleifer, and I. Welch (1992): "A theory of fads, fashion, custom, and cultural changes as informational cascades," *Journal of Political Economy*, 100, 992–1026.

Blackwell, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, ed. by J. Neyman, Berkeley, CA: University of California Press, 93–102.

——— (1953): "Equivalent comparisons of experiments," *The Annals of Mathematical Statistics*, 24, 265–272.

Burguet, R. and X. Vives (2000): "Social learning and costly information acquisition," *Economic theory*, 15, 185–205.

Çelen, B., S. Kariv, and A. Schotter (2010): "An experimental test of advice and social learning," *Management Science*, 56, 1687–1701.

Chamley, C. P. (2004): *Rational Herds*, Cambridge, UK: Cambridge University Press.

Chari, V. V. and P. J. Kehoe (2004): "Financial crises as herds: overturning the critiques," *Journal of Economic Theory*, 119, 128–150.

Eyster, E. and M. Rabin (2010): "Naive herding in rich-information settings," *American Economic Journal: Microeconomics*, 2, 221–243.

——— (2014): "Extensive imitation is irrational and harmful," *The Quarterly Journal of Economics*, 129, 1861–1898.

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

Goeree, J. K., T. R. Palfrey, and B. W. Rogers (2006): “Social learning with private and common values,” *Economic Theory*, 28, 245–264.

Grant, S., A. Kajii, and B. Polak (1998): “Intrinsic preference for information,” *Journal of Economic Theory*, 83, 233–259.

Grossman, S. J. and J. E. Stiglitz (1980): “On the impossibility of informationally efficient markets,” *American Economic Review*, 70, 393–408.

Guarino, A. and P. Jehiel (2013): “Social learning with coarse inference,” *American Economic Journal: Microeconomics*, 5, 147–174.

Hendricks, K., A. Sorensen, and T. Wiseman (2012): “Observational learning and demand for search goods,” *American Economic Journal: Microeconomics*, 4, 1–31.

Krishna, V. and J. Morgan (2011): “Overcoming ideological bias in elections,” *Journal of Political Economy*, 119, 183–211.

Lee, I. H. (1993): “On the convergence of informational cascades,” *Journal of Economic Theory*, 61, 395–411.

Levine, D. K. and T. R. Palfrey (2007): “The paradox of voter participation? A laboratory study,” *American political science Review*, 101, 143–158.

Marschak, J. (1959): “Remarks on the Economics of Information,” in *Contributions to Scientific Research in Management*, University of California, 79–98.

Moscarini, G. and L. Smith (2001): “The optimal level of experimentation,” *Econometrica*, 69, 1629–1644.

Mueller-Frank, M. and M. M. Pai (2016): “Social Learning with Costly Search,” *American Economic Journal: Microeconomics*, 8, 83–109.

Palfrey, T. R. and H. Rosenthal (1985): “Voter participation and strategic uncertainty,” *American Political Science Review*, 79, 62–78.

Schlee, E. (1990): “The value of information in anticipated utility theory,” *Journal of Risk and Uncertainty*, 3, 83–92.

Shiller, R. J. (2000): “Conversation, Information, and Herd Behavior,” *American Economic Review*, 85, 181–185.

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

Smith, L., P. Sorensen, and J. Tian (2014): “Informational Herding, Optimal Experimentation, and Contrarianism,” University of Wisconsin.

Vives, X. (1997): “Learning from others: a welfare analysis,” *Games and Economic Behavior*, 20, 177–200.

Wald, A. (1947): “Foundations of a general theory of sequential decision functions,” *Econometrica*, 279–313.

Wiseman, T. (2008): “Disagreement leads to complete learning: Sequential choice with continuous types,” *Economics Letters*, 100, 53–55.