Overabundant Information and Learning Traps*

Annie Liang† Xiaosheng Mu‡

November 5, 2018

Abstract

We develop a model of social learning from complementary information: Short-lived agents sequentially choose from a large set of (flexibly correlated) information sources for prediction of an unknown state, and information is passed down across periods. Will the community collectively acquire the best kinds of information? Long-run outcomes fall into one of two cases: (1) efficient information aggregation, where the community eventually learns as fast as possible; (2) "learning traps," where the community gets stuck observing suboptimal sources and information-aggregation is inefficient. Our main results identify a simple property of the signal correlation structure that determines which occurs. In both regimes, we characterize which sources are observed in the long run and how often. These results hold both for persistent and for slowly changing states.

1 Introduction

We consider the role of informational complementarities for social learning. Informational complementarities are a topic of classic (Milgrom and Weber, 1982*a*,*b*; McLean and Postlewaite, 2002) and recent (Borgers, Hernando-Veciana and Krahmer, 2013; Chen and Waggoner, 2016; Chade and Eeckhout, 2018) interest, but have primarily been studied in static environments, whereas social learning is inherently dynamic. When information is acquired over time, complementarities cause past information acquisitions to reshape the relative value of information sources for later agents. Our main contribution is to characterize the implications of this externality for long-run aggregation of information.

*We thank Vasilis Syrgkanis for insightful comments in early conversations about this project. We are also grateful to Nageeb Ali, Aislinn Bohren, Ben Golub, Paul Milgrom, Andrew Postlewaite, Ilya Segal, Carlos Segura, Rajiv Sethi, Andrzej Skrzypacz, and Yuichi Yamamoto for comments that improved this paper.

†University of Pennsylvania

‡Cowles Foundation and Columbia University

To fix ideas, consider researchers studying the impact of stress on depression. There are many kinds of information relevant to this question, and researchers choose what kind of information to acquire—e.g. what kind of neurochemical data to obtain, or what kind of models to explore. These studies are often complements: For example, measurement of cortisol production under stress is more valuable if we understand the role that cortisol plays in activating depression, and vice versa.

Characterizing how these complementarities shape information acquisitions is important for understanding outcomes in environments where the information acquisition decisions of strategic players interact, and also for contemplating interventions for shaping these acquisitions. Whether these complementarities are helpful for social aggregation of information is not clear from intuition alone. In some cases we might expect a positive spillover effect: For example, theoretical research about cortisol increases the marginal value to measurement of cortisol, and measurement of cortisol increases the value to sharpening theoretical understanding. Thus, informational complementarities could encourage researchers to collectively acquire the best body of knowledge. But there is also an argument to be made in the opposite direction: For example, *lack* of prior research on cortisol renders current measurements uninformative, and without researchers actively measuring cortisol, the demand for theoretical research about cortisol is also low.

Our main results demonstrate that a simple property of the informational environment determines whether there is guaranteed *efficient information aggregation*—past information pushes agents toward the best kinds of information—or whether there are *learning traps*—early suboptimal information acquisitions propagate across time, and there are persistent inefficiencies in information gathering. Our primary welfare criterion is whether the community obtains the fastest possible speed of learning, a focus that we share with Vives (1992), Golub and Jackson (2012), Sethi and Yildiz (2016), Hann-Caruthers, Martynov and Tamuz (2017), and Harel et al. (2018) among others. Additionally, we characterize payoff losses as evaluated by a patient social planner (see especially Section 5).

Our framework is a social learning model where agents, indexed by discrete time, acquire information and take actions (prediction of a payoff-relevant state). We depart from the classic model (Banerjee, 1992; Bikhchandani, Hirshleifer and Welch, 1992; Smith and Sorenson, 2000) in two key ways: First, we suppose that all information is public, so that predictions are based on the history of signal realizations so far. This departure turns off the inference problem essential to the existence of cascades in standard herding models (and indeed trivializes the standard model). Second, we assume endogenous information acquisition—specifically, agents choose from a large number of information sources, each associated with a signal about a payoff-relevant state.[1] We allow for flexible correlation across the

[1]Here we build on Burguet and Vives (2000), Mueller-Frank and Pai (2016), and Ali (2018), who introduce

sources by modeling the available information as different (noisy) linear combinations of the payoff-relevant state and a set of "confounding" variables. These assumptions jointly allow us to focus on inefficiencies that result from informational complementarities, as opposed to the classic friction of inference of information from actions.

As a benchmark, we first derive the *optimal* long-run frequency of signal acquisitions. These correspond to the choices that maximize information revelation about the payoff-relevant state, and also to the choices that maximize a discounted sum of agent payoffs (in a patient limit).

Our main results demonstrate that whether society's acquisitions converge to this optimal long-run frequency depends critically on how many signals are needed to identify the payoff-relevant state. The key intuition refers back to an observation made in Sethi and Yildiz (2016): An agent who repeatedly observes a source confounded by an unknown parameter learns *both* about the payoff-relevant state and also about the confounding term, and hence improves his interpretation of *this* source over time. In our setting, where a single confounding term can affect multiple sources, there is a further spillover effect: Learning from one source helps agents to interpret information from *all* sources confounded by the same parameters.

Suppose that in order to learn the payoff-relevant state, agents must observe a set of sources that reveals all of the confounding terms as well. Then endogenously, agents will acquire information that (collectively) reveals all of the unknowns. This will lead agents to evaluate *all* sources by a prior-independent asymptotic criterion, which identifies the best set of sources. Formally, we obtain the following result: If K sources are required to recover the payoff-relevant state (where K is also the number of unknown states), then long-run acquisitions are optimal starting from any prior belief.

In contrast, if it is possible to learn the payoff-relevant state without recovering all of the confounding terms, then agents can persistently undervalue sources that provide information confounded by these remaining variables, and long-run learning may be inefficient. Our second main result says that any set of fewer than K sources that recovers the payoff-relevant state creates a "learning trap" under some set of prior beliefs.

We then consider the welfare losses associated with learning traps. We show that the speed of learning can be arbitrarily slow (relative to the efficient benchmark), and payoffs can be arbitrarily inefficient when we consider the *ratio* of achieved and feasible payoffs. However, because we assume that ω is persistent across time, agents learn ω even while in a learning trap, and thus payoff losses are negligible when measured as *difference* of achieved

endogenous information acquisition to a classic social learning setting. Relative to this work, our paper considers choice from a fixed set of information sources (with a capacity constraint), in contrast to choice from a flexible set of information sources (with a cost on precision).

and feasible payoffs. We demonstrate next that in nearby models in which the state is not fully persistent, this conclusion fails and average payoff losses can be arbitrarily large.

To show this, we consider a generalization of the model where the state vector is changing over time. This is a substantially harder setting to analyze, and correspondingly prior work is very limited. (The recent works of Frongillo, Schoenebeck and Tamuz (2011), Vivi Alatas and Olken (2016), and Dasaratha, Golub and Hak (2018) are the only social learning settings with a dynamic state that we know of.) We first demonstrate the robustness of our main insights from the perfectly persistent case by considering a sequence of autocorrelated models that converge to our main model. Under some restrictions, we show that signal sets constituting potential learning traps also constitute learning traps for autocorrelation sufficiently close to 1. In contrast, those signal structures that guarantee efficient long-run learning in the persistent model are also shown to guarantee efficient long-run learning in nearby autocorrelated models. Welfare losses in learning traps (measured either as payoff ratio or payoff difference) can be arbitrarily large when the state is nearly (but not perfectly) persistent.

Technically, the main challenge in analyzing this extension is the failure of signal exchangeability. Unlike in the main model, posterior variance can no longer be expressed as a function of counts for how often each source has been observed. Our approach is to study a controlled dynamic system in the covariance matrix. We construct a Lyapunov function in the covariance matrix that is monotonic over time, and use this to bound the speed of learning. We do not pursue a full characterization of the autocorrelated model, although this is an interesting question for future work.

In the final part of our paper, we study interventions for breaking learning traps. We show that policymakers can restore efficient information aggregation by providing sufficiently many kinds of free information, or by reshaping the reward structure so that agents' predictions are based on information that they acquire over many periods. The success of these interventions depends on specific features of the informational environment.

Besides the papers mentioned above, this paper builds on a recent literature that studies choice from a discrete and fixed set of information sources—see for example Che and Mierendorff (2017) and Mayskaya (2017), who study choice between two Poisson sources, and Sethi and Yildiz (2016) and Fudenberg, Strack and Strzalecki (2017), who study choice between multiple Gaussian sources. For the most part, these models have not allowed for flexible correlations across the available kinds of information, and thus preclude complementarities across sources. Our complementary paper Liang, Mu and Syrgkanis (2017) also studies information acquisition from flexibly correlated Gaussian signals, but assumes that all of the available sources must be observed infinitely often in order to learn the payoff-relevant state. This restriction returns a special case of those environments that we show here lead to efficient long-run learning, and the main interest in Liang, Mu and Syrgkanis (2017) is exact characterization of the dynamically optimal strategy.

Finally, our paper is related to Sethi and Yildiz (2017), which also considers information acquisition from a large number of correlated Gaussian signals. In Sethi and Yildiz (2017), the sources of information are people, who themselves listen to other sources (people) over time, and the focus is on correlation structures emerging from disjoint communities with community-specific bias terms. From a modeling standpoint, their setting differs from ours in having time-varying signal precisions and free signals. They show that individuals can exhibit homophily in the long run, listening only to other individuals from the same community; this phenomenon is related to the observation of learning traps in the present paper.

2 Setup

2.1 Informational Environment

There are K persistent unknown states: a (real-valued) payoff-relevant state ω and $K-1$ (real-valued) confounding states $b_1, \dots, b_{K-1}$. We assume that the state vector $\theta := (\omega, b_1, \dots, b_{K-1})'$ follows a multivariate normal distribution $\mathcal{N}(\mu^0, \Sigma^0)$ where $\mu^0 \in \mathbb{R}^K$, and the prior covariance matrix Σ^0 has full rank.[2,3]

There are N (fixed) *kinds* or *sources* of information available at each discrete period $t \in \mathbb{Z}_+$. Observation of source i in period t produces an independent realization of the random variable

$$X_i^t = \langle c_i, \theta \rangle + \epsilon_i^t, \quad \epsilon_i^t \sim \mathcal{N}(0,1)$$

where $c_i = (c_{i1}, \dots, c_{iK})'$ is a (persistent) vector of constants, and the error terms ϵ_i^t are independent from each other and across periods. Normalization of the error distributions is without loss of generality, since the coefficients c_i are unrestricted (and rescaling the coefficients is equivalent to rescaling precision). Throughout, we take C to be the $N \times K$ coefficient matrix whose i-th row is c_i'. We will often drop the time indices on the random variables, equating X_i with source i and understanding that the error term is realized anew with each observation.

The payoff-irrelevant states $b_1, \dots, b_{K-1}$ produce correlations across the sources, and can be interpreted for example as:

[2] The full rank assumption is without loss of generality: If there is linear dependence across the states, the model can be mapped into an equivalent setting that has a lower dimensional state space and that satisfies the full rank condition.

[3] All vectors in this paper are column vectors.

- *Confounding explanatory variables*: Observation of signal i produces the (random) outcome $y = \omega c_i^1 + b_1 c_i^2 + \ldots b_{K-1} c_i^K + \epsilon_i$, which depends linearly on an observable characteristic vector c_i. For example, y might be the average incidence of depression in a group of individuals with characteristics c_i. The state of interest ω is the coefficient on a given characteristic c_i^1, and the payoff-irrelevant states are the unknown coefficients on the auxiliary characteristics $c_i^2, \ldots, c_i^K$. Different sources represent subpopulations with different characteristics.

- *Knowledge and technologies that aid interpretation of information*: Interpret the confounding states as "disturbance" terms. For example, measurement of a neurochemical in blood samples may correspond to observations of the signal $X = \omega + b + \epsilon$, where the variance of the confounding state b corresponds to the quality of the technology (and can be improved for example by acquiring some other signal $b + \epsilon$.)

Throughout, the set of sources is indexed by $[N] = \{1, \ldots, N\}$. We call a subset of sources $\mathcal{S} \subset [N]$ *spanning* if the vectors $\{c_i : i \in \mathcal{S}\}$ span the coordinate vector $e_1 = (1, 0, \ldots, 0)' \in \mathbb{R}^K$, so that it is possible to completely learn the payoff-relevant state ω by repeatedly observing signals from only $\mathcal{S}$. We call $\mathcal{S}$ *minimally spanning* if it is spanning, and no proper subset of $\mathcal{S}$ is spanning.

We assume throughout that the complete set of signals $[N]$ is spanning, so that the payoff-relevant state can be recovered by observing all signals infinitely often.[4] This assumption nests two interesting cases. Say that the informational environment has *exactly sufficient information* if $[N]$ is minimally spanning. Then, it is possible to recover ω by observing each information source infinitely often, but not by exclusively observing any proper subset of sources.

We are primarily interested in settings of *informational overabundance*, where $[N]$ is spanning but not minimally spanning. Multiple different subsets of signals allow for recovery of ω, and a key point of our analysis is to compare the set of sources that "should" be observed in the long run with the set of sources that is in fact observed in the long run. Except for trivial cases, informational overabundance corresponds to $N > K$ (more signals

[4]This assumption is without loss, and our results do extend to situations where ω is *not* identified from the available signals. To see this, we first take a linear transformation and work with the following equivalent model: The state vector $\tilde{\theta}$ is K-dimensional *standard Gaussian*, each signal $X_i = \langle \tilde{c}_i, \tilde{\theta} \rangle + \epsilon_i$, and the payoff-relevant parameter is $\langle u, \tilde{\theta} \rangle$ for some fixed vector u. Let R be the subspace of $\mathbb{R}^K$ spanned by $\tilde{c}_1, \ldots, \tilde{c}_N$. Then project u onto R: $u = r + w$ with $r \in R$ and w orthogonal to R. Thus $\langle u, \tilde{\theta} \rangle = \langle r, \tilde{\theta} \rangle + \langle w, \tilde{\theta} \rangle$. By assumption, the random variable $\langle w, \tilde{\theta} \rangle$ is independent from any random variable $\langle c, \tilde{\theta} \rangle$ with $c \in R$ (because they have zero covariance). Thus the uncertainty about $\langle w, \tilde{\theta} \rangle$ cannot be reduced upon any signal observation. Consequently, agents only seek to learn about $\langle r, \tilde{\theta} \rangle$, returning to the case where the payoff-relevant parameter *is* identified.

than states).[5]

2.2 Information Acquisitions

A sequence of agents indexed by time t move sequentially. Each agent chooses one of the N sources and observes a realization of the corresponding signal. He then predicts ω, selecting an action $a \in \mathbb{R}$ and receiving the payoff $-(a-\omega)^2$. We assume throughout that all signal realizations are public. Thus, each agent t faces a history $h^{t-1} \in ([N] \times \mathbb{R})^{t-1} = H^{t-1}$ consisting of all past signal choices and their realizations, and his signal acquisition strategy is a function from histories to sources. The agent's optimal prediction of ω is his posterior mean, and his expected payoff is the negative of his posterior variance about ω. At every history h^{t-1}, the agent's expected payoffs are maximized by choosing the signal that minimizes his posterior variance about ω.

In our Gaussian environment, posterior variance about ω is a deterministic function $V(q_1, \dots, q_N)$ of the number of times q_i that each signal i has been observed so far.[6] So each agent's signal acquisition is a function of past signal acquisitions only (and not of the signal realizations). This allows us to track society's acquisitions as deterministic *count vectors*

$$m(t) = (m_1(t), \dots, m_N(t)) \in \mathbb{Z}_+^N$$

where $m_i(t)$ is the number of times that source i has been observed up to and including period t. Specifically, $m(t)$ evolves according to the following rule: $m(0)$ is the zero vector, and for each time $t \geq 0$ and source i,

$$m_i(t+1) = \begin{cases} m_i(t)+1 & \text{if } V(m_i(t)+1, m_{-i}(t)) \leq V(m_j(t)+1, m_{-j}(t)) \ \forall j \\ m_i(t) & \text{otherwise} \end{cases}$$

so that the count vector increases by 1 in the coordinate corresponding to the signal that allows for the greatest immediate reduction in posterior variance. We allow ties to be broken arbitrarily, and there may be multiple possible paths $m(t)$. We are interested in the *long-run frequencies* of observation $\lim_{t\to\infty} m_i(t)/t$ for each source i—that is, the fraction of periods eventually devoted to each source. As we show later in Section 4, these limits exist under a mild technical assumption. Note the possibility for some signals to have zero long-run frequency.

[5]It is possible for ω to be "overidentified" from a set of $N \leq K$ signals, e.g. $X_1 = \omega + \epsilon_1$, $X_2 = \omega + b_1 + b_2 + \epsilon_2$, and $X_3 = b_1 + b_2 + \epsilon_3$. In this case, the set $\{X_1, X_2, X_3\}$ is spanning, but not minimally spanning since both of its subsets $\{X_1\}$ and $\{X_2, X_3\}$ are also spanning. Although $N = K = 3$ in this example, it is equivalent to a model in which there is a single confounding term $\tilde{b}_1 = b_1 + b_2$, and the three signals are rewritten $X_1 = \omega + \epsilon_1$, $X_2 = \omega + \tilde{b}_1 + \epsilon_2$ and $X_3 = \tilde{b}_1 + \epsilon_3$. Then we do have $N > K$ in this equivalent model.

[6]For a normal prior and normal signals, the posterior covariance matrix does not depend on signal realizations. See Appendix A.1 for the complete (closed-form) expression for V.

3 Optimal Benchmark

3.1 Planner's Problem

We begin by characterizing the long-run information acquisitions that maximize social welfare; this characterization will be used subsequently as a benchmark against which to determine the efficiency of society's actual information acquisitions. We consider two formulations of a social planner's objective for this setting:

Speed of Learning: The planner desires for society to aggregate information about the payoff-relevant state ω at the fastest asymptotic rate. Define

$$q^{OPT}(t) \in \underset{(q_1,\ldots,q_K):q_i\in\mathbb{Z}^+,\sum_i q_i=t}{\operatorname{argmin}} V(q_1,\ldots,q_K).$$

to be any allocation of t observations across the N signals that minimize posterior variance about ω.[7]

Define $\lambda^{OPT} = \lim_{t\to\infty} q^{OPT}(t)/t$ to be the limiting frequency over signals corresponding to these allocations, and say that a signal acquisition strategy *maximizes speed of learning* if it almost surely leads to long-run frequency vector λ^{OPT}.

δ-Aggregated Payoffs (in Patient Limit): The social planner chooses a signal acquisition strategy to maximize a discounted average of payoffs across individuals

$$U_\delta := \mathbb{E}\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}\cdot -(a_t-\omega)^2\right] \tag{1}$$

where actions are taken optimally as in Section 2.2.[8] We will be interested in the patient limit where $\delta \to 1$.

3.2 Key Object: ϕ

We now define an object that will play a central role in the remaining analyses. Consider any minimal spanning set $\mathcal{S}$. Since the set is spanning, some linear combination of the signals in

[7]This is equivalent to the notion of **c**-optimality in the experimental design literature: Let $\theta \in \mathbb{R}^K$ be an unknown state vector and suppose an analyst observes the vector of outcomes $y \sim \mathcal{N}(X\theta, \sigma^2 I)$, where X is a $K \times K$ design matrix. The analyst chooses X to minimize the variance of a linear combination of the parameters in θ. Our problem corresponds to the case in which the analyst seeks to minimize the variance of the first coordinate of θ. Note that relative to the classic experimental design problem, our setting differs in that we additionally impose a restriction that each row of X comes from a permissible set $\{c_i\}$.

[8]Due to the specification of quadratic loss payoffs, the signal acquisition strategy neen not condition on signal realizations.

$\mathcal{S}$ is equivalent to an unbiased signal about ω; since the set is *minimally* spanning, the linear combination assigns nonzero weight to all signals in the set. Formally, the first coordinate vector e_1 (corresponding to the payoff-relevant state ω) can be decomposed as a unique linear combination of signals in $\mathcal{S}$:

$$e_1 = \sum_{i \in \mathcal{S}} \beta_i^{\mathcal{S}} \cdot c_i \qquad \text{where each } \beta_i^{\mathcal{S}} \neq 0.$$

The sum of (the absolute values of) these coefficients is important, and we name it ϕ below.

Definition 1. *Define* $\phi(\mathcal{S}) := \sum_{i \in \mathcal{S}} |\beta_i^{\mathcal{S}}|$.

This is loosely the "asymptotic standard deviation" of beliefs about ω given optimal sampling from $\mathcal{S}$, as shown below.

Fact 1. *For each* t, *let* $q^{\mathcal{S}}(t) = (q_1^{\mathcal{S}}(t), \ldots, q_N^{\mathcal{S}}(t)) \in \mathbb{Z}_+^N$ *be the allocation of* t *observations across sources in* $\mathcal{S}$ *that minimizes posterior variance about* ω. *Then,*

$$V\left(q^{\mathcal{S}}(t)\right) \sim \phi(\mathcal{S})^2/t$$

where the notation "$F(t) \sim G(t)$*" means* $\lim_{t\to\infty} \frac{F(t)}{G(t)} = 1$. *Moreover, this long-run speed of learning is achievable by a (stationary) sampling rule that assigns a frequency of* $|\beta_i^{\mathcal{S}}| / \left(\sum_{j \in \mathcal{S}} |\beta_j^{\mathcal{S}}|\right)$ *to each source* i.[9]

Thus, the asymptotic speed of learning is faster from sets $\mathcal{S}$ with smaller ϕ-values. Throughout we assume the following, which guarantees the existence of a unique "best" minimal spanning set:

Assumption 1 (Unique Minimizer)**.** $\phi(\mathcal{S})$ *has a unique minimizer* $\mathcal{S}^*$ *among minimal spanning sets* $\mathcal{S} \subset [N]$.

[9]To understand the coefficients $\beta_i^{\mathcal{S}}$, suppose that the set $\mathcal{S}$ has size K, so that the coefficient vectors associated with signals in $\mathcal{S}$ have full rank, and let $C_{\mathcal{S}}$ denote the matrix of these coefficient vectors. The (random) vector of realizations corresponding to one observation of each signal in this set can be written as

$$Y = (y_1, \ldots, y_K)' = C_{\mathcal{S}}\theta + \varepsilon$$

where ε is a $K \times 1$ vector of error terms. Given these realizations, the best linear unbiased estimate for ω is $\hat{\omega} = \left[C_{\mathcal{S}}^{-1} Y\right]_{11}$. Perturbing the realization of signal i by δ_i changes this estimate by $[C_{\mathcal{S}}^{-1}]_{1i} \cdot \delta_i$. One can show that the coefficients $\beta_i^{\mathcal{S}} = |[C_{\mathcal{S}}^{-1}]_{1i}|$, so the larger $\beta_i^{\mathcal{S}}$ is, the more $\hat{\omega}$ responds to changes in the realization of signal i. Thus, fact 1 thus says that agents should observe more frequently those signals whose realizations more strongly influence the best linear estimate of ω.

Among the class of sampling rules that eventually concentrate on a single minimal spanning set, it directly follows from Fact 1 that the following sampling rule maximizes speed of learning:

$$\lambda_i^* = \begin{cases} \frac{|\beta_i^{\mathcal{S}^*}|}{\sum_{j\in\mathcal{S}^*}|\beta_j^{\mathcal{S}^*}|} & \forall\ i \in \mathcal{S}^* \\ 0 & \forall\ i \notin \mathcal{S}^* \end{cases} \tag{2}$$

This sampling rule assigns zero frequency to signals outside of the best set $\mathcal{S}^*$, and samples signals within $\mathcal{S}^*$ according to the frequencies given in Fact 1.

3.3 Optimal Long-Run Acquisitions

In principle, a social planner may improve on λ^* by sampling from multiple spanning sets. Our first theorem shows to the contrary that λ^* remains optimal when arbitrary sampling procedures are permitted: So long as C satisfies Unique Minimizer, the best long-run strategy is to concentrate acquisitions on the best minimal spanning set, and sample optimally from that set. This is true whether we consider the objective of maximizing speed of learning, or the objective of maximizing a discounted-aggregate of payoffs (in the patient limit).

Theorem 1. *Assume Unique Minimizer and let λ^* be given by (2).*

(a) Learning-Optimality: *The optimal long-run frequency vector $\lambda^{OPT} = \lambda^*$.*[10,11]

(b) Payoff-Optimality: *For fixed δ, let $d_\delta(t)$ be the vector of signal counts (up to period t) associated with any strategy that maximizes U_δ (see (1)). Then for every $\varepsilon > 0$, there exists $\underline{\delta} < 1$ such that for any $\delta \geq \underline{\delta}$ it holds that*

$$\limsup_{t\to\infty} \left\| \frac{d_\delta(t)}{t} - \lambda^* \right\| \leq \varepsilon$$

where $||\cdot||$ represents the Euclidean norm.

Part (a) of Theorem 1 says that sampling according to λ^* maximizes the asymptotic speed of learning. This conclusion is related to Chaloner (1984), a result in the experimental design literature.[12] Part (b) of Theorem 1 says that λ^* approximates the long-run frequency vector

[10]The conclusion can be loosely interpreted as stating that λ^* is the "most efficient linear representation" of the payoff-relevant state in terms of the signals. Specifically, consider the following constrained minimization problem: $\min \sum_{i=1}^N |\beta_i|$ subject to $\sum_{i=1}^N \beta_i \cdot c_i = e_1$. It can be shown by linear programming that the minimum is attained exactly when $\beta_i = \beta_i^{\mathcal{S}^*}$ (that is, when focusing on a single minimal spanning set).

[11]We show in Appendix C.1 that Assumption 1 is necessary for this result: We describe an environment where Assumption 1 fails, and it is strictly optimal to observe more than K signals with positive frequency.

[12]Chaloner (1984) showed that a c-optimal design exists on at most K points. Part (a) additionally supplies a characterization of the optimal design itself and demonstrates uniqueness. Another difference is that Chaloner (1984) studies the optimal continuous design, while we impose an integer constraint on signal counts.

under the social planner's optimal information acquisition strategy (in the patient limit).[13] Together, these results justify our use of λ^* as the optimal benchmark against which to compare society's long-run information acquisitions.

4 Main Results

We now ask whether society's acquisitions converge to the optimal long-run frequencies λ^* derived previously. We show that informational environments can be classified into two kinds—those for which efficient information aggregation is guaranteed (starting from all prior beliefs), and those for which "learning traps" are possible (depending on the prior belief). Separation of these two classes depends critically on how many signals are required to identify ω.

4.1 Learning Traps

The following example demonstrates that sequential information acquisition need not lead to efficient information aggregation. Indeed, the set of signals that are observed in the long run can be disjoint from the optimal set $\mathcal{S}^*$.

Example 1. There are three available signals:

$$X_1 = \omega + \epsilon_1$$
$$X_2 = 3\omega + b_1 + \epsilon_2$$
$$X_3 = b_1 + \epsilon_3$$

Both $\{X_1\}$ and $\{X_2, X_3\}$ are minimal spanning sets, but optimal information acquisitions (as defined in Section 3) eventually concentrate on $\{X_2, X_3\}$.[14]

Consider a prior where ω and b_1 are independent, and the prior variance of b_1 is large (exceeds 8). In the first period, observation of X_1 is most informative about ω, since X_2 is perceived as a noisier signal *about* ω than X_1, and observations of X_3 provide information only about the confounding term b_1 (which is uncorrelated with ω). Thus the best choice is to observe X_1. This observation does not affect the variance of b_1, so the same argument shows that every agent observes signal X_1. We refer to $\{X_1\}$ in this example as a *learning trap.*[15]

[13] We note that our proof does not guarantee existence of limiting frequencies $\lim_{t\to\infty} d_\delta(t)/t$ for fixed δ.

[14] It is straightforward to verify that $\phi(\{X_1\}) = 1 > 2/3 = \phi(\{X_2, X_3\})$. Note also that $X_2 - X_3$ is an unbiased signal about ω, and it is more informative than two realizations of X_1; this demonstrates that $\{X_2, X_3\}$ is the best minimal spanning set without direct computation of ϕ.

[15] The existence of learning traps is not special to the assumption of normality. We report a related example with non-normal signals in Appendix C.3.

Generalizing this example, the result below (stated as a corollary, since it will follow from the subsequent Theorem 2) gives a sufficient condition for learning traps. We impose the following (generic) assumption on the informational environment, which requires that every set of $k \leq K$ signals is linearly independent:

Assumption 2 (Strong Linear Independence)**.** *$N \geq K$ and every $K \times K$ submatrix of C is of full rank.*[16]

Corollary 1. *Assume Strong Linear Independence. For every minimal spanning set $\mathcal{S}$ with $|\mathcal{S}| < K$, there exists an open set of prior beliefs given which agents exclusively observe signals from $\mathcal{S}$.*

Thus, every small set (fewer than K signals) that identifies ω is a candidate learning trap.

4.2 Efficient Information Aggregation

Suppose in contrast to the previous section that repeated observation of K sources is required to recover ω. Our next result shows that a very different long-run outcome obtains: Starting from *any* prior, information acquisition eventually approximates the optimal frequency. Thus, even though agents are short-lived ("myopic"), they end up acquiring information in a way that is socially best.

Corollary 2. *Under Unique Minimizer, if every minimal spanning set has size K, then starting from any prior belief, it holds that $m_i(t) \sim \lambda_i^* \cdot t$ for every signal i.*

This result (like Corollary 1 above) follows from the subsequent Theorem 2, and hence is stated as a corollary.

The condition that all minimal spanning sets have size K is generically satisfied.[17] However, if we expect that sources are endogenous to design or strategic motivations, the relevant informational environments may not fall under this condition. For example, the existence of an unbiased signal about ω (that is, $X = c\omega + \epsilon$) is non-generic, but plausible in practice. Sets of signals that partition into different groups with group-specific confounding terms (as studied in Sethi and Yildiz (2017)) are also economically interesting but non-generic. The previous Corollary 1 shows that inefficiency is a possible outcome in these cases.

[16] Besides trivial cases with redundant signals, Strong Linear Independence also rules out settings such as the following: $X_1 = \omega + b_1 + \epsilon_1$, $X_2 = b_1 + \epsilon_2$, $X_3 = 2\omega + b_2 + \epsilon_3$, $X_4 = b_2 + \epsilon_4$, $X_5 = 3\omega + b_3 + \epsilon_5$, and $X_6 = b_3 + \epsilon_6$. Then $K = 4$ but the four signals X_1, X_2, X_3, X_4 are *not* linearly independent.

[17] We point out that the set of coefficient matrices satisfying Unique Minimizer is "generic" in the following stronger sense: Fix the *directions* of coefficient vectors, and suppose that the *precisions* are drawn at random; then, generically different minimal spanning sets correspond to different speed of learning. In contrast, whether every minimal spanning set has size K is a condition on the *directions* themselves.

4.3 General Characterization

We now provide a complete characterization of the possible long-run observation sets for any environment. A few additional definitions are required for this characterization. For any spanning set of signals $\mathcal{S}$, define the *subspace spanned by* $\mathcal{S}$ to be the subspace (of $\mathbb{R}^K$) spanned by the coefficient vectors associated with sources in $\mathcal{S}$. We then collect all available signals whose coefficient vectors belong to that subspace, and call this $\overline{\mathcal{S}} \subseteq [N]$. A minimal spanning set $\mathcal{S}$ is *subspace-optimal* if it uniquely maximizes the speed of learning among all feasible sets of signals within its subspace.

Definition 2. *A minimal spanning set* $\mathcal{S}$ *is* subspace-optimal *if it uniquely minimizes* ϕ *among all subsets of* $\overline{\mathcal{S}}$ *that are minimally spanning.*

Example 2. Suppose the available signals are $X_1 = \omega + \epsilon_1$ and $X_2 = 2\omega + \epsilon_2$, and consider $\mathcal{S} = \{X_1\}$. Then, the coefficient vector associated with X_1 is $(1, 0)$, and the subspace it spans is $\{(\alpha, 0) : \alpha \in \mathbb{R}\}$. Since this subspace also includes the coefficient vector $(2, 0)$ associated with X_2, we have $\overline{\mathcal{S}} = \{X_1, X_2\}$. But $\{X_2\}$ permits a faster asymptotic speed of learning than $\{X_1\}$; thus, the set $\mathcal{S}$ is not subspace-optimal.

We introduce one final assumption, which strengthens Unique Minimizer to require the existence of a best minimal spanning set $\mathcal{S}$ within every subspace.

Assumption 3 (Unique Minimizer in Every Subspace)**.** *For every* $\mathcal{S} \subset [N]$*, there exists a unique minimal spanning set* $\tilde{\mathcal{S}}$ *that minimizes* ϕ *among subsets of* $\overline{\mathcal{S}}$*.*

This assumption is guaranteed if different minimal spanning sets correspond to different ϕ-values, and thus holds generically.

Our next result generalizes both the learning traps result and also the efficient information aggregation result from the previous sections. Theorem 2 says that long-run information acquisitions eventually concentrate on a set $\mathcal{S}$ (starting from some prior belief) *if and only if* $\mathcal{S}$ is a subspace-optimal minimal spanning set.

Theorem 2. *(a) Suppose* $\mathcal{S}$ *is a subspace-optimal minimal spanning set. Then, there exists an open set of prior beliefs given which agents exclusively observe signals from* $\mathcal{S}$*. Long-run frequencies are positive precisely for those signals in* $\mathcal{S}$*, and they are given by Proposition 1.*

(b) Under Assumption 3, long-run frequencies exist given any prior belief. Moreover, if $\mathcal{S}$ *denotes the signals viewed with positive long-run frequency, then* $\mathcal{S}$ *is a minimal spanning set that is subspace-optimal, and long-run frequencies are given by Proposition 1.*

This theorem directly implies our previous Corollaries 1 and 2. To see this, observe that under Strong Linear Independence, $\overline{\mathcal{S}} = \mathcal{S}$ for every minimal spanning set $\mathcal{S}$ with fewer than K signals.[18] This implies that every minimal spanning set with fewer than K signals is (trivially) optimal in its subspace, producing Corollary 1 from part (a) of the theorem.

On the other hand, if every minimal spanning set has size K, then all minimal spanning sets belong to the same subspace. Under Unique Minimizer, there can only be one minimal spanning set that is optimal in this subspace, and this must also be the best set overall (in the sense of Section 3). This yields Corollary 2 from part (b) of the theorem above.

We provide a brief intuition for this (second) part of the theorem. Since agents choose signals by comparing the marginal value of different observations, signal acquisitions eventually concentrate on a set $\mathcal{S}$ if and only if the marginal values of signals in that set become persistently higher than those of other signals. The key challenge is that the marginal value of any given signal depends critically on which signals have been observed previously. Thus, multiple long-run observation outcomes are possible when there are different minimal spanning sets that are self-reinforcing.

We show that as agents repeatedly acquire signals from any fixed subspace of signals, they eventually discover the "objective" asymptotic marginal values of each signal *in that subspace*. Thus, only those sets of signals that are best in their subspace are potentially "self-sustaining." And if all sets of signals that reveal ω span the entire space, agents will identify the best set of signals overall and achieve efficient information aggregation.

We collect below a few additional sufficient conditions precluding learning traps, which follow as a corollary of Theorem 2:

Corollary 3. *Under Unique Minimizer:*

(a) If the best set $\mathcal{S}^*$ is of size 1 (equivalently, if $\mathcal{S}^*$ consists of a single unbiased signal $\alpha\omega + \epsilon$), then learning traps cannot emerge.

(b) Learning traps of size K cannot emerge.

The first part holds because an unbiased signal belongs to every subspace spanned by a minimal spanning set. The second part follows from part (b) of Theorem 2, which shows that any long-run observation set of size K must be the overall best set $\mathcal{S}^*$.

5 Welfare and Auto-Correlation

The previous sections described conditions under which society's acquisitions fail to converge to the optimal long-run frequency. In this section we present a more detailed analysis of the welfare loss under learning traps.

[18] Suppose $|\mathcal{S}| < K$, then by assumption of Strong Linear Independence, every signal not in $\mathcal{S}$ is linearly independent from the signals in $\mathcal{S}$. Hence $\overline{\mathcal{S}}$ cannot contain any other signal.

Throughout this section, we use

$$U_\delta^M = \mathbb{E}_M\left[-\sum_{t=1}^{\infty}(1-\delta)\delta^{t-1}(a_t-\omega)^2\right]$$

to denote the δ-discounted average payoffs following the myopic signal acquisition strategy (with optimal actions a_t). We use U_δ^{SP} to denote the maximum δ-discounted average payoff across all signal acquisition strategies. Note that both payoff sums are negative, since flow payoffs are quadratic loss at every period.

In Section 5.1, we show that payoff inefficiency under learning traps can be arbitrarily large when evaluated as a (limiting) *payoff ratio* $\lim_{\delta\to1} U_\delta^M/U_\delta^{SP}$. However, the implied payoff loss is zero if we consider the (limiting) *payoff difference* $\lim_{\delta\to1}(U_\delta^{SP}-U_\delta^M)$. This second conclusion depends critically on our assumption that ω is persistent across time, which implies that agents learn ω even while in a learning trap, albeit slowly. We demonstrate next that in nearby models in which the state is not fully persistent, average payoff losses can again be arbitrarily large.

To do this, we outline a sequence of autocorrelated models that converge to our main model (with perfect state persistence). Section 5.2 describes this generalization of our framework, and Section 5.3 shows that the qualitative insights from the previous sections extend—specifically, the characterization of learning traps and efficiency (based on subspace-optimality) is robust to consideration of slowly changing states. We then quantify the potential payoff losses in these nearby models: At near perfect persistence, welfare losses under learning traps can be large according to both of the above measures.

5.1 Welfare Loss Under Learning Traps

Recall from Fact 1 that the long-run speed of learning from any minimal spanning set $\mathcal{S}$ is given by $(\phi(\mathcal{S}))^2/t$, so the ratio $\phi(\mathcal{S})/\phi(\mathcal{S}^*)$ compares the long-run speed of learning from $\mathcal{S}$ versus the optimal set $\mathcal{S}^*$ (where larger ratios imply greater inefficiencies).

A small modification of the learning traps example in Section 4.1 shows that this ratio can be made arbitrarily large.

Example 3. There are three available signals:

$$X_1 = \frac{1}{L}\omega + \epsilon_1$$
$$X_2 = \omega + b_1 + \epsilon_2$$
$$X_3 = b_1 + \epsilon_3$$

where $L > 0$ is a constant. In this example, the ratio $\phi(\{X_1\})/\phi(\{X_2, X_3\}) = L/2$, but for every choice of L, there is a set of priors given which X_1 is exclusively observed.[19]

Thus, society's long-run speed of learning can be arbitrarily slower than the optimal speed. Observing that aggregated payoffs are a sum of (negative) posterior variances, the above construction also implies the ratio of flow payoffs (under optimal sampling from $\mathcal{S}$ vs. the efficient set $\mathcal{S}^*$) at late periods can be made arbitrarily large. As $\delta \to 1$, these later payoffs dominate the total payoffs from the initial periods.[20] Thus for every constant $c > 0$, there is an informational environment and prior such that the limiting payoff ratio satisfies

$$\lim_{\delta \to 1} U_\delta^M / U_\delta^{SP} > c.$$

(Note that because payoffs are negative, larger values of $U_\delta^M / U_\delta^{SP}$ correspond to greater payoff inefficiencies.)

On the other hand, the payoff difference vanishes in the patient limit; that is,

$$\lim_{\delta \to 1} \left(U_\delta^{SP} - U_\delta^M \right) = 0$$

in all informational environments and for all priors. This is because even under learning traps, agents eventually learn ω, implying that $\lim_{\delta \to 1} U_\delta^{SP} = \lim_{\delta \to 1} U_\delta^M = 0$. However, we show below that zero payoff difference is special to the environment in which ω is perfectly persistent across time. Once states (slightly) evolve over time, payoff difference between the myopic rule and the optimum is bounded away from 0 even in the patient limit.

5.2 Autocorrelated Model

Recall that in our main model, the state vector $\theta = (\omega, b_1, \dots, b_{K-1})'$ is persistent across time. We now consider a state vector θ^t that evolves according to the following law:

$$\theta^1 \sim \mathcal{N}(0, \Sigma^0); \quad \theta^{t+1} = \sqrt{\alpha} \cdot \theta^t + \sqrt{1-\alpha} \cdot \eta^t, \quad \text{where } \eta^t \sim \mathcal{N}(0, M).$$

Above, means are normalized to zero, and the prior covariance matrix of the state vector at time $t = 1$ is Σ^0. We restrict the autocorrelation coefficient $\sqrt{\alpha}$ to belong to $(0, 1)$. (Choice of $\alpha = 1$ returns our main model, and we will be interested in approximations where α is close to but strictly less than 1.) The *innovation* $\eta^t \sim \mathcal{N}(0, M)$ captures the additional noise terms that emerge under state evolution, which we assume to be i.i.d. across time. Fixing signal coefficients $\{c_i\}$, every autocorrelated model is indexed by the triple (M, Σ^0, α).

[19] The region of inefficient priors (that result in suboptimal learning) does decrease in size as the level of inefficiency increases. Specifically, as L increases, the prior variance of b_1 has to increase correspondingly in order for the first agent to choose X_1.

[20] This follows from the fact that the harmonic series diverges.

In each period, the available signals are

$$X_i^t = \langle c_i, \theta^t \rangle + \epsilon_i^t, \quad \epsilon_i^t \sim \mathcal{N}(0,1).$$

The signal noises are i.i.d. across time and further independent from the innovations in state evolution. The agent in period t myopically chooses the signal that minimizes posterior variance of ω^t, while the social planner seeks to minimize a discounted sum of such posterior variances.

Note that while Σ^0 is analogous to the prior in our main model, the covariance matrix M is a new degree of freedom, without an obvious parallel in the persistent model. It turns out that many choices of M are inappropriate if we want to interpret the autocorrelated model with α close to 1 as approximating our original model (as we do). To see why, consider the following example:

Example 4. The two available signals are:

$$X_1^t = \omega^t - b^t + \epsilon_1^t$$
$$X_2^t = b^t + \epsilon_2^t$$

where (ω^t, b^t) that evolve according to

$$(\omega^1 - b^1, b^1) \sim \mathcal{N}(0, I); \quad (\omega^{t+1}, b^{t+1}) = \sqrt{\alpha} \cdot (\omega^t, b^t) + \sqrt{1-\alpha} \cdot \mathcal{N}(0, M).$$

Suppose $\alpha < 1$ and $M = \text{diag}(1,0)$. Then, even without learning about b^t, society's posterior beliefs about this second state decreases to zero geometrically. It is then easy to show that society will eventually only sample X_1. On the other hand, if $M = \text{diag}(0,1)$, then by symmetry society eventually samples only the second signal. Both of these learning dynamics are far from the one under perfect persistence ($\alpha = 1$), in which case Fact 1 tells us that optimal sampling eventually acquires each signal in equal proportion.

The example highlights that beliefs evolve for two reasons in the autocorrelated model: First, beliefs respond to signal acquisitions (as before); second, beliefs adapt to the evolution of the states, as determined by the covariance matrix of the innovations M.[21] Clearly M then shapes the optimal sampling strategy. We would like to pick M so that the optimal sampling rule when α is near 1 is related to the optimal sampling strategy at $\alpha = 1$, and this will require the innovations to "align with" the optimal sampling strategy.

[21] Specifically observe that M represents the stationary covariance matrix in the absence of any signal acquisition, since

$$\Sigma^t = \alpha \cdot \Sigma^{t-1} + (1-\alpha) \cdot M$$

implies $\Sigma^t \to M$.

This issue can be resolved as follows: Given any minimal spanning set $\mathcal{S}$, linearly transform the states to be $\tilde{\theta}_1, \ldots, \tilde{\theta}_K$ so that the $k \leq K$ signals in $\mathcal{S}$ can be rewritten as $\tilde{\theta}_i + \mathcal{N}(0,1)$ for $1 \leq i \leq k$. The payoff-relevant state ω can then be expressed as the linear combination $\sum_{i=1}^k w_i \tilde{\theta}_i$ for weights $w_1, \ldots, w_k > 0$. In particular, $\phi(\mathcal{S})$ is the sum of these weights.

When agents optimally sample from $\mathcal{S}$ in the persistent state model, the posterior covariance matrix of the first k transformed states is eventually (approximately) proportional to $\operatorname{diag}\left(\frac{1}{w_1}, \ldots, \frac{1}{w_k}\right)$. M is chosen in line with this long-run outcome; that is, we require that

$$(\tilde{\theta}_1^{t+1}, \ldots, \tilde{\theta}_k^{t+1}) = \sqrt{\alpha} \cdot (\tilde{\theta}_1^t, \ldots, \tilde{\theta}_k^t) + \mathcal{N}\left(0, \operatorname{diag}\left(\frac{x}{w_1}, \ldots, \frac{x}{w_k}\right)\right) \tag{3}$$

where x can be any positive number. As we explain below, this choice of M guarantees that the optimal sampling rule for $\alpha = 1$ remains approximately optimal for α close to 1.[22] Note that the condition above restricts only the first k states (associated with $\mathcal{S}$).

5.3 Robustness of Main Results

Here we show that the autocorrelated model described above preserves our main results regarding long-run outcomes.

For any $\mathcal{S}$, denote by $\mathcal{M}(\mathcal{S})$ the set of all positive-definite matrices M that satisfy the criterion in (3). The following result presents the existence of learning traps.

Theorem 3. *Suppose $\mathcal{S}$ is a subspace-optimal minimal spanning set. Then there exists $M \in \mathcal{M}(\mathcal{S})$, Σ^0 such that for every $\epsilon > 0$, there is an $\underline{\alpha}(\epsilon) < 1$ where for each autocorrelated model (M, Σ^0, α) with $\alpha > \underline{\alpha}(\epsilon)$:*

1. *Every agent in the autocorrelated model myopically samples from $\mathcal{S}$.*

2. *The resulting discounted average payoff satisfies*

$$\limsup_{\delta \to 1} U_\delta^M \leq -(1-\epsilon)\phi(\mathcal{S}) \cdot \sqrt{(1-\alpha)M_{11}}\ ,$$

while it is feasible to achieve a patient payoff of

$$\liminf_{\delta \to 1} U_\delta^{SP} \geq -(1+\epsilon)\phi(\mathcal{S}^*) \cdot \sqrt{(1-\alpha)M_{11}}$$

by sampling from $\mathcal{S}^*$.

[22] The diagonal matrix means that we focus on autocorrelated models where the transformed states $\tilde{\theta}_1, \ldots, \tilde{\theta}_k$ evolve independently from each other. This corresponds to the fact that in our original model with perfect persistence, the transformed states are eventually believed to be approximately independent after sufficient learning.

Part (1) generalizes Theorem 2, showing that every subspace-optimal minimal spanning set remains a potential long-run absorbing set from some prior. This demonstrates that the property of subspace-optimality is more general than our original setting with perfectly persistent states.

Part (2) shows that the payoff difference in the patient limit satisfies

$$\liminf_{\delta \to 1} \left(U_\delta^{SP} - U_\delta^M \right) > (\phi(\mathcal{S})(1-\epsilon) - \phi(\mathcal{S}^*)(1+\epsilon)) \cdot \sqrt{(1-\alpha)M_{11}}.$$

If $\phi(\mathcal{S}) > \phi(\mathcal{S}^*)$, this payoff difference is strictly positive for α close to (but not equal to) 1, since we can set ϵ sufficiently small that $\phi(\mathcal{S})(1-\epsilon) > \phi(\mathcal{S}^*)(1+\epsilon)$. Thus, whenever $\mathcal{S}$ is not the overall efficient minimal spanning set, social acquisitions result in "first-order" payoff inefficiency for α close to 1.[23]

Moreover, since $\phi(\mathcal{S}) - \phi(\mathcal{S}^*)$ can be made arbitrarily large (independently of α), the payoff loss can be arbitrarily large across environments. Formally we show the following:

Proposition 1. *For every $\epsilon > 0$, there exists an environment as in Example 3 and an autocorrelated model (M, Σ^0, α) such that $\liminf_{\delta \to 1} U_\delta^{SP} \geq -\epsilon$ but $\limsup_{\delta \to 1} U_\delta^M \leq -\frac{1}{\epsilon}$.*

Our main result for efficient learning (Corollary 2) extends as well. The following result demonstrates that near-efficient information aggregation obtains from every prior if all minimal spanning sets span the same space.

Theorem 4. *Suppose every minimal spanning set has size K, and fix any $M \in \mathcal{M}(\mathcal{S}^*)$. For every $\epsilon > 0$, there exists $\underline{\alpha}(\epsilon) < 1$ such that in any autocorrelated model (M, Σ^0, α) with $\alpha \geq \alpha(\epsilon)$, society achieves a discounted average payoff of*

$$\liminf_{\delta \to 1} U_\delta^M \geq -(1+\epsilon) \cdot \phi(\mathcal{S}^*) \cdot \sqrt{(1-\alpha)M_{11}}$$

while it is not feasible to do better than

$$\limsup_{\delta \to 1} U_\delta^{SP} \leq -(1-\epsilon) \cdot \phi(\mathcal{S}^*) \cdot \sqrt{(1-\alpha)M_{11}}.$$

(Note that no restriction is placed on the prior covariance matrix Σ_0)

Collectively, these results demonstrate that the main insights from our perfectly persistent model extend to nearby settings with dynamically evolving states.

[23] By first-order payoff inefficiency we mean that the payoff difference is on the size of $\sqrt{1-\alpha}$. One reason we think this is a meaningful criterion is because it captures whether the *limit payoff ratio* is bounded away from 1. Moreover, in our discrete-time setting, posterior variances generally do not converge and $\lim_{\delta \to 1} U_\delta^{SP}$ may not exist. Thus it may be unreasonable to expect *any* sampling rule to exactly achieve the limsup of optimal payoffs in the patient limit. Under our criterion, however, myopic sampling *can* achieve first-order payoff efficiency in some informational environments (see Theorem 4).

6 Interventions

The previous sections detail the possibility, and welfare consequences, of learning traps. This naturally suggests a question of what kinds of policies might "free" agents from these traps. We compare several possible policy interventions: Increasing the *quality* of information acquisition (so that each signal acquisition is more informative); restructuring incentives so that agents' payoffs are based on information obtained over several periods (equivalent to acquisition of *multiple signals* each period); and providing a one-shot release of *free information*, which can then guide subsequent acquisitions.

6.1 More Precise Information

Consider first an intervention in which the precision of signal draws is increased uniformly over signals. For example, if different signals correspond to measurement of different neurochemicals in a group of lab subjects, a government agency can provide researchers with funding that permits recruitment of more subjects. This improves the quality of the estimate regardless of which neurochemical the researcher chooses to measure.

We model this intervention by supposing that each signal acquisition produces B independent observations from that source (with the main model corresponding to $B = 1$). The result below (a corollary of part (a) of Theorem 2) shows that providing more informative signals is of limited effectiveness: Any set of signals that is a potential learning trap given $B = 1$ remains a potential learning trap under arbitrary improvements to signal precision.

Corollary 4. *Suppose that for $B = 1$, there is a set of priors given which signals in $\mathcal{S}$ are (exclusively) viewed in the long run. Then, for every $B \in \mathbb{Z}_+$, there is a set of priors given which $\mathcal{S}$ is exclusively viewed in the long run.*

However, the sets of prior beliefs corresponding to different values of B need not be the same. For a *fixed* prior belief, subsidizing higher quality acquisitions may or may not move the community out of a learning trap. To see this, consider first the informational environment and prior belief from Example 1. Increasing the precision of signals is ineffective here: The first agent chooses X_1 regardless of the value of B, and our previous logic again implies that each subsequent agent also chooses signal X_1. Thus, the set $\{X_1\}$ remains a learning trap. In Appendix C.4, we provide a contrasting example in which increasing the precision of signals can indeed break agents out of a learning trap from a specified prior belief. Which of these examples is relevant depends on fine details of the informational environment as well as the prior, which the policymaker may not know in practice.

6.2 Batches of Signals

Another possibility is to restructure the incentive scheme so that agents' payoffs are based on information obtained over several periods, equivalent to acquisition of a batch of signals each period. For example, evaluation of researchers can be based on a set of papers, or researchers can be given labs and permitted to direct the work of multiple individuals simultaneously.

Both of these approaches for restructuring the environment can be modeled by permitting each agent to allocate B observations across the sources (where $B = 1$ returns the main model). Note the key difference from the previous intervention: Here it is possible for the B observations to be allocated across *different* signals. We show that it is possible to guarantee efficient information aggregation in this case:

Proposition 2. *Under Unique Minimizer, there is a B such that given acquisition of B signals every period, long-run frequency is λ^* starting from every prior belief.*

Thus, given sufficiently many observations each period, agents will allocate observations in a way that eventually approximates the optimal frequency.

The number of observations needed, however, depends on subtle details of the informational environment. In particular, the required B cannot be uniformly bounded over all environments of fixed size (number of states K and number of signals N).[24] See Appendix A.5 for further details.

6.3 Free Information

Finally, we consider provision of free information to the community. We can think of this as releasing information that a policymaker knows, or as a reduced form for funding specific kinds of research, the results of which are made public.

Formally, the policymaker chooses M signals $X_j = \langle p_j, \theta\rangle + \mathcal{N}(0, 1)$, where each $\|p_j\|_2 \leq \gamma$, so that signal precisions are bounded by γ^2. At time $t = 0$, independent realizations of these signals are made public. All subsequent agents update their prior beliefs based on this free information in addition to the history of signal acquisitions thus far.

We show that given a sufficient number of (kinds of) signals, of sufficiently high precision, efficient learning can be guaranteed. Specifically, if $k \leq K$ is the size of the optimal set $\mathcal{S}^*$, then $k - 1$ precise signals are sufficient to guarantee efficient learning:

[24] The required B depends on two properties: first, on how well the optimal frequency λ^* can be approximated via allocation of B observations—for example, $\lambda^* = (1/2, 1/2)$ can be achieved exactly using two observations, while $\lambda^* = (3/8, 5/8)$ cannot; second, on the difference in learning speed between the best set and the next best minimal spanning set, which determines the "slack" that is permitted in the approximation of λ^*. Thus, small batch sizes B are sufficient when the optimal frequency λ^* can be well-approximated using a small number of observations, or when there are large efficiency gains from observing the best set.

Proposition 3. *Let $k := |\mathcal{S}^*|$. Under Unique Minimizer, there exists a $\gamma < \infty$, and $k-1$ signals $X_j = \langle p_j, \theta \rangle + \mathcal{N}(0,1)$ with $\|p_j\|_2 \leq \gamma$, such that with these free signals provided at $t = 0$, society's long-run frequency is λ^* starting from every prior belief.*

The proof is by construction. We show that as long as agents understand those confounding terms that appear in the best set of signals (these parameters have dimension $k-1$), they will come to evaluate the signals in the best set according to their asymptotic marginal values.[25]

This intervention is most relevant in settings in which a technological advance could greatly speed up progress, but development of the technology is slow and tedious. For example, suppose that high-resolution brain scans would allow for rapid understanding of depression, but the current imaging technology is very poor. Researchers working to understand depression may prefer to exploit existing technologies, rather than contribute to development of this new technology. The government can intervene by funding preliminary development of brain imaging, which then encourages researchers to begin using brain scans. Once low-resolution brain scans are common, the payoff to advancing the imaging technology increases, and even myopic researchers may contribute to this agenda. In this way, provision of free information can nudge agents onto the right path of learning.

7 Extensions

General Payoff Functions. Our main results extend when each agent t chooses an action to maximize an arbitrary individual payoff function $u_t(a_t, \omega)$ (recall that previously we restricted to $u_t(a_t, \omega) = -(a_t - \omega)^2$). We require only that these payoff functions are nontrivial in the following sense:

Assumption 4 (Payoff Sensitivity to Mean)**.** *For every t, any variance $\sigma^2 > 0$ and any action $a^* \in A$, there exists a positive Lebesgue measure of μ for which a^* does* not *maximize $\mathbb{E}[u_t(a, \omega) \mid \omega \sim \mathcal{N}(\mu, \sigma^2)]$.*

That is, for every belief variance, the expected value of ω affects the optimal action to take. This rules out cases with a "dominant" action and ensures that each agent *strictly* prefers to choose the most informative signal. Since the signal that minimizes the posterior variance about ω Blackwell-dominates every other signal,[26] each agent's signal acquisition remains unchanged.

[25] This intervention requires knowledge of the full correlation structure, and also which set $\mathcal{S}^*$ is best. An alternative intervention, with higher demands on information provision but lower demands on knowledge of the environment, is to provide $K-1$ (sufficiently precise) signals about all of the confounding terms.

[26] See, e.g., Hansen and Torgersen (1974).

However, the interpretation of the optimal benchmark (that we defined in Section 3) is more limited. Specifically, while the optimal frequency can still be interpreted as maximizing information revelation, the relationship to the social planner problem (part (b) of Theorem 1) may fail. A detailed discussion is relegated to Appendix C.5.1.

Low Altruism. So far we have assumed that agents care only to maximize the accuracy of their own prediction of the payoff-relevant state. Consider a generalization in which agents are slightly altruistic; that is, each agent t chooses a signal as well as an action a_t to maximize discounted payoffs $\mathbb{E}\left[\sum_{t' \geq t} \delta^{t'-t} \cdot u(a_t, \omega)\right]$, assuming that future agents will behave similarly. (Note that $\delta = 0$ returns our main model.) We show in Appendix C.5.2 that for δ sufficiently small, part (a) of Theorem 2 continues to hold (in every equilibrium of this game). So the existence of learning traps is robust to a small degree of altruism.

Multiple Payoff-Relevant States. In our main model, only one of the K persistent states is payoff-relevant. Consider a generalization in which each agent predicts (the same) $r \leq K$ unknown states and his payoff is determined via a weighted sum of quadratic losses. We show in Appendix C.5.3 that our main results extend to this setting. The possibility for agents to have payoffs that depend on *heterogeneous* states is also interesting, and we leave this for future work.

8 Conclusion

We study a model of sequential learning, where short-lived agents choose what kind of information to acquire from a large set of available information sources. Because agents do not internalize the impact of their information acquisitions on later decision-makers, they may acquire information inefficiently (from a social perspective). Inefficiency is not guaranteed, however: Depending on the informational environment, myopic concerns can endogenously push agents to identify and observe only the most informative sources.

Our main results separate these possibilities, and reveal that the extent of *learning spillovers* is essential to determining which outcome emerges. Specifically, does information about unknowns of immediate societal interest (i.e., the payoff-relevant state) also teach about unknowns that are only of indirect value (i.e., the confounding terms)?

When such spillovers are present, simple incentive schemes for information acquisition—in which agents care only about immediate contributions to knowledge—are sufficient for efficient long-run learning. When these spillovers are not built into the environment, other incentives are needed. For example, forward-looking funding agencies can encourage investment in the confounding terms (our "free information" intervention). Alternatively, agents can be evaluated on the basis of a body of work (our "multiple signal" intervention). These observations are consistent with practices that have arisen in academic research, including the establishment of third-party funding agencies (e.g. the NSF) to support basic science and methodological research, and the evaluation of researchers based on advancements developed across several papers (e.g. tenure and various prizes).

We conclude below with brief mention of additional directions. So far we have considered the demand for information given an exogenous set of information sources. Another natural model would have the information sources choose the information they provide in order to maximize demand (see a related problem in Perego and Yuksel (2018).) Our Theorem 1 implies the following comparative static: If signal i is viewed with positive frequency in the optimal benchmark, then this frequency is (locally) decreasing in its precision. Thus, if demand is interpreted as λ_i^* (the long-run frequency with which source i is optimally viewed), sources face conflicting incentives: They want to provide information sufficiently precise to be included in the best set and receive viewership at all, but subject to this, they want to provide signals as imprecise as possible. These conflicting forces suggest that characterization of the equilibrium provisions of information precision is subtle.

Finally, while we have described our setting as choice between information sources, our model may apply more generally to choice between actions with complementarities. For example, suppose a sequence of managers take actions that have externalities for future managers, and each manager seeks to maximize performance of the company during his tenure. The concepts we have developed here of *efficient information aggregation* and *learning traps* have natural analogues in that setting (actions that maximize the company's long-term welfare, versus those that do not). Relative to the general setting, we study here a class of complementarities that are micro-founded in correlated signals. It is an interesting question of whether and how the forces we find here generalize to other kinds of complementarities.

References

Ali, Nageeb. 2018. "Herding with Costly Information." *Journal of Economic Theory.*

Angeletos, George-Marios, and Alessandro Pavan. 2007. "Efficient Use of Information and the Social Value of Information." *Econometrica*, 75(4): 1103–1142.

Banerjee, Abhijit. 1992. "A Simple Model of Herd Behavior." *Quaterly Journal of Economics*, 107(3): 797–817.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1992. "A Theory of Fads, Fashion, Custom, and Cultural Change as Information Cascades." *Journal of Political Economy*, 100(5): 992–1026.

Borgers, Tilman, Angel Hernando-Veciana, and Daniel Krahmer. 2013. "When Are Signals Complements Or Substitutes." *Journal of Economic Theory*, 148(1): 165–195.

Burguet, Roberto, and Xavier Vives. 2000. "Social Learning and Costly Information." *Economic Theory*.

Chade, Hector, and Edward E. Schlee. 2002. "Another Look at the Radner-Stiglitz Nonconcavity in the Value of Information." *Journal of Economic Theory*, 107: 421–452.

Chade, Hector, and Jan Eeckhout. 2018. "Matching Information." *Theoretical Economics*.

Chaloner, Kathryn. 1984. "Optimal Bayesian Experimental Design for Linear Models." *The Annals of Statistics*, 12(1): 283–300.

Chen, Yiling, and Bo Waggoner. 2016. "Informational Substitues."

Che, Yeon-Koo, and Konrad Mierendorff. 2017. "Optimal Sequential Decision with Limited Attention." Working Paper.

Dasaratha, Krishna, Ben Golub, and Nir Hak. 2018. "Social Learning in a Dynamic Environment." Working Paper.

Frongillo, Rafael, Grant Schoenebeck, and Omer Tamuz. 2011. "Social Learning in a Changing World."

Fudenberg, Drew, Philip Strack, and Tomasz Strzalecki. 2017. "Stochastic Choice and Optimal Sequential Sampling." Working Paper.

Golub, Benjamin, and Matthew Jackson. 2012. "How Homophily Affects the Speed of Learning and Best-Response Dynamics." *The Quarterly Journal of Economics*.

Hann-Caruthers, Wade, Vadim Martynov, and Omer Tamuz. 2017. "The Speed of Sequential Asymptotic Learning." Working Paper.

Hansen, Ole Havard, and Eric N. Torgersen. 1974. "Comparison of Linear Normal Experiments." *The Annals of Statistics*, 2: 367–373.

Harel, Matan, Elchanan Mossel, Philipp Strack, and Omer Tamuz. 2018. "Groupthink and the Failure of Information Aggregation in Large Groups." Working Paper.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis. 2017. "Optimal and Myopic Information Acquisition." Working Paper.

Mayskaya, Tatiana. 2017. "Dynamic Choice of Information Sources." Working Paper.

McLean, Richard, and Andrew Postlewaite. 2002. "Informational Size and Incentive Compatibility." *Econometrica*.

Milgrom, Paul, and Rober J. Weber. 1982*a*. "A Theory of Auctions and Competitive Bidding." *Econometrica*.

Milgrom, Paul, and Robert J. Weber. 1982*b*. "The Value of Information in a Sealed-Bid Auction." *Journal of Mathematical Economics*.

Mueller-Frank, Manuel, and Mallesh Pai. 2016. "Social Learning with Costly Search."

American Economic Journal: Microeconomics.

Perego, Jacopo, and Sevgi Yuksel. 2018. "Media Competition and Social Disagreement." Working Paper.

Sethi, Rajiv, and Muhamet Yildiz. 2016. "Communication with Unknown Perspectives." *Econometrica*, 84(6): 2029–2069.

Sethi, Rajiv, and Muhamet Yildiz. 2017. "Culture and Communication." Working Paper.

Smith, Lones, and Peter Sorenson. 2000. "Pathological Outcomes of Observational Learning." *Econometrica*.

Vives, Xavier. 1992. "How Fast do Rational Agents Learn?" *Review of Economic Studies*.

Vivi Alatas, Abhijit Banerjee, Arun Chandrasekhar Rema Hanna, and Ben Olken. 2016. "Network Structure and the Aggregation of Information: Theory and Evidence from Indonesia." *American Economic Review*.

A Appendix

The structure of the appendix follows that of the paper. We provide proofs for the results listed in the main text, in the order in which they appeared; the only exception is that the proof of part (b) of Theorem 1 relies on tools we develop in the other proofs, and so it is given at the end. Other results and examples are deferred to a separate Online Appendix.

A.1 Posterior Variance Function

A.1.1 Monotonicity and Convexity

Here we review and extend a basic result from Liang, Mu and Syrgkanis (2017). Specifically, we show that the posterior variance about ω weakly decreases over time, and the marginal value of any signal decreases in its signal count.

Lemma 1. *Given prior covariance matrix Σ^0 and q_i observations of each signal i, society's posterior variance about ω is*

$$V(q_1, \dots, q_N) = \left[((\Sigma^0)^{-1} + C'QC)^{-1}\right]_{11} \tag{4}$$

where $Q = \operatorname{diag}(q_1, \dots, q_N)$. The function V is decreasing and convex in each q_i whenever these arguments take non-negative real values.

Proof. Note that $(\Sigma^0)^{-1}$ is the prior precision matrix and $C'QC = \sum_{i=1}^{N} q_i \cdot [c_i c_i']$ is the total precision from the observed signals. Thus (4) simply represents the fact that for

Gaussian prior and signals, the posterior precision matrix is the sum of the prior and signal precision matrices. To prove the monotonicity of V, consider the partial order $\succeq$ on positive semi-definite matrices where $A \succeq B$ if and only if $A - B$ is positive semi-definite. As q_i increases, the matrix Q and $C'QC$ increase in this order. Thus the posterior covariance matrix $((\Sigma^0)^{-1} + C'QC)^{-1}$ decreases in this order, which implies that the posterior variance about ω decreases.

To prove that V is convex, it suffices to prove that V is *midpoint-convex* since the function is clearly continuous.[27] Take $q_1, \ldots, q_N$, $r_1, \ldots, r_N \in \mathbb{R}_+$ and let $s_i = \frac{q_i + r_i}{2}$. Define the corresponding diagonal matrices to be Q, R, S. Observe that $Q + R = 2S$. Thus by the AM-HM inequality for positive-definite matrices, we have

$$((\Sigma^0)^{-1} + C'QC)^{-1} + ((\Sigma^0)^{-1} + C'RC)^{-1} \succeq 2((\Sigma^0)^{-1} + C'SC)^{-1}.$$

Using (4), we conclude that

$$V(q_1, \ldots, q_N) + V(r_1, \ldots, r_N) \geq 2V(s_1, \ldots, s_N).$$

This proves the (midpoint) convexity of V. □

A.1.2 Inverse of Positive Semi-definite Matrices

For future use, we provide a definition of $[X^{-1}]_{11}$ for positive *semi-definite* matrices X. When X is positive definite, its eigenvalues are strictly positive, and its inverse matrix is defined as usual. In general, we can apply the Spectral Theorem to write

$$X = UDU',$$

where U is a $K \times K$ orthogonal matrix whose columns are eigenvectors of X, and $D = \operatorname{diag}(d_1, \ldots, d_K)$ is a diagonal matrix consisting of non-negative eigenvalues. Even if some of these eigenvalues are zero, we can think of X^{-1} as

$$X^{-1} = (UDU')^{-1} = UD^{-1}U' = \sum_{j=1}^{K} \frac{1}{d_j} \cdot [u_j u_j']$$

where u_j is the j-th column vector of U. We thus define

$$[X^{-1}]_{11} := \sum_{j=1}^{K} \frac{(\langle u_j, e_1 \rangle)^2}{d_j}, \tag{5}$$

[27] A function V is midpoint-convex if the inequality $V(a) + V(b) \geq 2V(\frac{a+b}{2})$ always holds. Every continuous function that is midpoint-convex is also convex.

with the convention that $\frac{0}{0} = 0$ and $\frac{z}{0} = \infty$ for any $z > 0$. Note that by this definition,

$$[X^{-1}]_{11} = \lim_{\epsilon \to 0_+} \left(\sum_{j=1}^{K} \frac{(\langle u_j, e_1 \rangle)^2}{d_j + \epsilon} \right) = [(X + \epsilon I_K)^{-1}]_{11},$$

since the matrix $X + \epsilon I_K$ has the same set of eigenvectors as X, with eigenvalues increased by ϵ. Hence our definition of $[X^{-1}]_{11}$ is a continuous extension of the usual definition to positive semi-definite matrices.

A.2 Proof of Theorem 1 Part (a)

A.2.1 Asymptotic Posterior Variance Function

We first approximate the posterior variance as a function of the frequencies with which each signal is observed. Specifically, for any $\lambda \in \mathbb{R}_+^N$, define

$$V^*(\lambda) := \lim_{t \to \infty} t \cdot V(\lambda t).$$

The following result shows V^* to be well-defined and computes its value:

Lemma 2. *Let* $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_N)$. *Then*

$$V^*(\lambda) = [(C'\Lambda C)^{-1}]_{11} \tag{6}$$

The value of $[(C'\Lambda C)^{-1}]_{11}$ *is well-defined, see (5).*

Proof. Recall that $V(q_1, \ldots, q_N) = [((\Sigma^0)^{-1} + C'QC)^{-1}]_{11}$ with $Q = \text{diag}(q_1, \ldots, q_N)$. Thus

$$t \cdot V(\lambda_1 t, \ldots, \lambda_N t) = \left[\left(\frac{1}{t} (\Sigma^0)^{-1} + C'\Lambda C \right)^{-1} \right]_{11}.$$

Hence the lemma follows from the continuity of $[X^{-1}]_{11}$ in the matrix X. □

We note that $C'\Lambda C$ is the Fisher Information Matrix when signals are observed according to frequencies λ. Thus the above lemma can also be seen as an application of the Bayesian Central Limit Theorem.

A.2.2 Reduction to the Study of V^*

The development of the function V^* is useful for the following reason:

Lemma 3. *Suppose* $\tilde{\lambda}$ *uniquely minimizes* $V^*(\cdot)$ *subject to* $\lambda \in \Delta^{N-1}$ *(the* $(N-1)$*-dimensional simplex). Then, the* t*-optimal divisions satisfy* $q_i(t) \sim \tilde{\lambda}_i \cdot t$ *for all* i.

Proof. Fix any increasing sequence of times $t_1, t_2, \ldots$. It suffices to show that whenever the limit $\lambda_i := \lim_{m\to\infty} \frac{q_i(t_m)}{t_m}$ exists for each i, this limit λ must be $\tilde{\lambda}$. Suppose not, then by assumption $V^*(\lambda) > V^*(\tilde{\lambda})$. For $\epsilon > 0$, define another vector $\hat{\lambda} \in \mathbb{R}_+^N$ with $\hat{\lambda}_i = \lambda_i + \epsilon, \forall i$. By the continuity of V^*, it holds that $V^*(\hat{\lambda}) > V^*(\tilde{\lambda})$ for sufficiently small ϵ.

Since $\lambda_i = \lim_{m\to\infty} \frac{q_i(t_m)}{t_m}$, there exists M sufficiently large such that $q_i(t_m) \leq (\lambda_i + \epsilon) \cdot t_m$ for each i and $m \geq M$. Hence, for $m \geq M$,

$$t_m \cdot V(q_1(t_m), \ldots, q_N(t_m)) \geq t_m \cdot V(\hat{\lambda}_1 \cdot t_m, \ldots, \hat{\lambda}_N \cdot t_m) \to V^*(\hat{\lambda})$$

where the inequality uses the monotonicity of V. On the other hand,

$$t_m \cdot V(\tilde{\lambda}_1 \cdot t_m, \ldots, \tilde{\lambda}_N \cdot t_m) \to V^*(\tilde{\lambda}).$$

Comparing the above two displays, we see that for sufficiently large m,

$$V(q_1(t_m), \ldots, q_N(t_m)) > V(\tilde{\lambda}_1 \cdot t_m, \ldots, \tilde{\lambda}_N \cdot t_m).$$

But this contradicts the t-optimality of the division vector $q(t_m)$, as society could do better by following frequencies $\tilde{\lambda}$. The lemma is thus proved. □

A.2.3 Crucial Lemma

We pause to demonstrate the following technical lemma:

Lemma 4. *Suppose $\mathcal{S}^* = \{1, \ldots, K\}$ uniquely minimizes $\phi(\cdot)$ and let C^* be the $K \times K$ submatrix of C corresponding to the first K signals. Further suppose $\beta_j^{\mathcal{S}^*} = [(C^*)^{-1}]_{1j}$ is positive for $1 \leq j \leq K$. Then for any signal $i > K$, if we write $c_i = \sum_{j=1}^K \alpha_j \cdot c_j$ (which is a unique representation), then $|\sum_{j=1}^K \alpha_j| < 1$.*

Proof. By assumption, we have the vector identity

$$e_1 = \sum_{j=1}^K \beta_j \cdot c_j \quad \text{with } \beta_j = [(C^*)^{-1}]_{1j} > 0.$$

Suppose for contradiction that $\sum_{j=1}^K \alpha_j \geq 1$ (the opposite case where the sum is ≤ -1 can be similarly treated). Then some α_j must be positive. Without loss of generality, we assume $\frac{\alpha_1}{\beta_1}$ is the largest among such ratios. Then $\alpha_1 > 0$ and

$$e_1 = \sum_{j=1}^K \beta_j \cdot c_j = \left(\sum_{j=2}^K \left(\beta_j - \frac{\beta_1}{\alpha_1} \cdot \alpha_j\right) \cdot c_j\right) + \frac{\beta_1}{\alpha_1} \cdot \left(\sum_{j=1}^K \alpha_j \cdot c_j\right)$$

This represents e_1 as a linear combination of the vectors $c_2, \ldots, c_K$ and c_i, with coefficients $\beta_2 - \frac{\beta_1}{\alpha_1} \cdot \alpha_2, \ldots, \beta_K - \frac{\beta_1}{\alpha_1} \cdot \alpha_K$ and $\frac{\beta_1}{\alpha_1}$. Observe that these coefficients are non-negative: For each $2 \leq j \leq K$, $\beta_j - \frac{\beta_1}{\alpha_1} \cdot \alpha_j$ is clearly positive if $\alpha_j \leq 0$ (since $\beta_j > 0$). And if $\alpha_j > 0$, then by assumption $\frac{\alpha_j}{\beta_j} \leq \frac{\alpha_1}{\beta_1}$ and $\beta_j - \frac{\beta_1}{\alpha_1} \cdot \alpha_j$ is again non-negative.

By definition, $\phi(\{2, \dots, K, i\})$ is the sum of the absolute value of these coefficients. This sum is

$$\sum_{j=2}^{K} \left(\beta_j - \frac{\beta_1}{\alpha_1} \cdot \alpha_j\right) + \frac{\beta_1}{\alpha_1} = \sum_{j=1}^{K} \beta_j + \frac{\beta_1}{\alpha_1} \cdot \left(1 - \sum_{j=1}^{K} \alpha_j\right) \leq \sum_{j=1}^{K} \beta_j.$$

But then $\phi(\{2, \dots, K, i\}) \leq \phi(\{1, 2, \dots, K\})$, leading to a contradiction. Hence the lemma must be true. □

A.2.4 Proof of Theorem 1 Part (a) when $|\mathcal{S}^*| = K$

Given Lemma 3, part (a) of Theorem 1 will follow once we show that λ^* uniquely minimizes $V^*(\cdot)$ over the simplex; recall that λ^* denotes the optimal frequencies for the minimal spanning set $\mathcal{S}^*$ that minimizes ϕ. In this section, we prove that λ^* is indeed the unique minimizer whenever this best subset $\mathcal{S}^*$ contains exactly K signals. Later on we will prove the same result even when $|\mathcal{S}^*| < K$, but that proof will require additional techniques.

Lemma 5. *For $\lambda \in \Delta^{N-1}$, the function $V^*(\lambda)$ is uniquely minimized at $\lambda = \lambda^*$.*

Proof. First, we assume that $[(C^*)^{-1}]_{1i}$ is positive for $1 \leq i \leq K$. This is without loss because we can always work with the "negative" of any signal (replace c_i with $-c_i$), which does not affect agents' behavior.

Since $V(q_1, \dots, q_N)$ is convex in its arguments, $V^*(\lambda) = \lim_{t \to \infty} t \cdot V(\lambda_1 t, \dots, \lambda_N t)$ is also convex in λ. To show λ^* uniquely minimizes V^*, we only need to show λ^* is a *local minimum*. In other words, it suffices to show $V^*(\lambda^*) < V^*(\lambda)$ for any λ that belongs to an ϵ-neighborhood of λ^*. By assumption, $\mathcal{S}^*$ is minimally-spanning and so its signals are linearly independent. Thus its signals must span all of the K states. From this it follows that the $K \times K$ matrix $C'\Lambda^* C$ is positive definite, and by (6) the function V^* is differentiable near λ^* (see Remark 1 below).

We claim that the partial derivatives of V^* satisfy the following inequality:

$$\partial_K V^*(\lambda^*) < \partial_i V^*(\lambda^*) \leq 0, \forall i > K. \tag{*}$$

Once this is proved, we will have, for λ close to λ^*,

$$V^*(\lambda_1, \dots, \lambda_K, \lambda_{K+1}, \dots, \lambda_N) \geq V^*\left(\lambda_1, \dots, \lambda_{K-1}, \sum_{k=K}^{N} \lambda_k, 0, \dots, 0\right) \geq V^*(\lambda^*). \tag{7}$$

The first inequality is based on (*) and differentiability of V^*, while the second inequality is because λ^* uniquely minimizes V^* when only the first K signals are observed. Moreover, when $\lambda \neq \lambda^*$, one of these inequalities is strict so that $V^*(\lambda) > V^*(\lambda^*)$ holds strictly.

To prove (*), we recall that

$$V^*(\lambda) = e_1'(C'\Lambda C)^{-1}e_1.$$

Since $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_N)$, its derivative is $\partial_i \Lambda = \Delta_{ii}$, which is an $N \times N$ matrix whose (i,i)-th entry is 1 with all other entries equal to zero. Using properties of matrix derivatives, we obtain

$$\partial_i V^*(\lambda) = -e_1'(C'\Lambda C)^{-1}C'\Delta_{ii}C(C'\Lambda C)^{-1}e_1.$$

As the i-th row vector of C is c_i', $C'\Delta_{ii}C$ is the $K \times K$ matrix $c_i c_i'$. The above simplifies to

$$\partial_i V^*(\lambda) = -[e_1'(C'\Lambda C)^{-1}c_i]^2.$$

At $\lambda = \lambda^*$, the matrix $C'\Lambda C$ further simplifies to $(C^*)' \cdot \operatorname{diag}(\lambda_1^*, \ldots, \lambda_K^*) \cdot (C^*)$, which is a product of $K \times K$ invertible matrices. We thus deduce that

$$\partial_i V^*(\lambda^*) = -\left[e_1' \cdot (C^*)^{-1} \cdot \operatorname{diag}\left(\frac{1}{\lambda_1^*}, \ldots, \frac{1}{\lambda_K^*}\right) \cdot ((C^*)')^{-1} \cdot c_i\right]^2.$$

Crucially, note that the term in the brackets is a linear function of c_i. To ease notation, we write $v' = e_1' \cdot (C^*)^{-1} \cdot \operatorname{diag}\left(\frac{1}{\lambda_1^*}, \ldots, \frac{1}{\lambda_K^*}\right) \cdot ((C^*)')^{-1}$ and $\gamma_i = \langle v, c_i \rangle$. Then

$$\partial_i V^*(\lambda^*) = -\gamma_i^2, \; 1 \leq i \leq N. \tag{8}$$

For $1 \leq i \leq K$, $((C^*)')^{-1} \cdot c_i$ is just e_i. Thus, using the assumption $[(C^*)^{-1}]_{1j} > 0, \forall j$, we have

$$\gamma_i = e_1' \cdot (C^*)^{-1} \cdot \operatorname{diag}\left(\frac{1}{\lambda_1^*}, \ldots, \frac{1}{\lambda_K^*}\right) \cdot e_i = \frac{[(C^*)^{-1}]_{1i}}{\lambda_i^*} = \sum_{j=1}^{K} [(C^*)^{-1}]_{1j} = \phi(\mathcal{S}^*) \tag{9}$$

On the other hand, choosing any $i > K$, we can uniquely write the vector c_i as a linear combination of $c_1, \ldots, c_K$. By Lemma 4, for any $i > K$ we have

$$\gamma_i = \langle v, c_i \rangle = \sum_{j=1}^{K} \alpha_j \cdot \langle v, c_j \rangle = \sum_{j=1}^{K} \alpha_j \cdot \gamma_j = \phi(\mathcal{S}^*) \cdot \sum_{j=1}^{K} \alpha_j, \tag{10}$$

where the last equality uses (9). Since $|\sum_{j=1}^{K} \alpha_j| < 1$, the absolute value of γ_i for any $i > K$ is strictly smaller than the absolute value of γ_K. This together with (8) proves the desired inequality (*), and Lemma 5 follows. □

Remark 1. *The essence of this proof is the following nontrivial property: The subset* $\{1, \ldots, K\}$ *uniquely minimizes* ϕ *among all subsets of size* K *if and only if*

$$\phi(\{1, \ldots, K\}) < \phi(\{1, \ldots, K\} \cup \{i\} \backslash \{j\}), \; \forall 1 \leq j \leq K < i \leq N.$$

That is, if a set of K signals does not minimize ϕ, then we can improve the speed of learning by adding one *signal to replace* one *existing signal. This property enables us to reduce the general problem with N signals to a much simpler problem with $K+1$ signals. We are then able to use calculus to resolve the latter problem, see (*).*

This argument breaks down if we start with a set of less than K signals, since V^* can fail to be differentiable at the frequency vector of interest. It is for this reason that we need a different proof of Lemma 5 when $|\mathcal{S}^*| < K$, which we present next.

A.2.5 A Perturbation Argument

To handle the case of $|\mathcal{S}^*| < K$, we first extend the definition of $\phi(\cdot)$ to arbitrary sets of siganls (not necessarily minimally-spanning) as follows. For any set $\mathcal{A}$ that contains a minimal spanning set, define $\phi(\mathcal{A}) = \min_{\mathcal{S}\subset\mathcal{A}} \phi(\mathcal{S})$, where the minimum is taken over all minimal spanning sets $\mathcal{S}$ contained in $\mathcal{A}$. If such $\mathcal{S}$ does not exist (i.e., $\mathcal{A}$ is not itself spanning), we let $\phi(\mathcal{A}) = \infty$. In particular,

$$\phi([N]) = \min_{\mathcal{S}\subset[N]} \phi(\mathcal{S})$$

represents the minimum asymptotic standard deviation achievable by only observing the signals in *some* minimal spanning set.

Our previous analysis shows that whenever $\phi(\mathcal{S})$ is uniquely minimized by a set $\mathcal{S}$ containing exactly K signals,

$$\min_{\lambda\in\Delta^{N-1}} V^*(\lambda) = V^*(\lambda^*) = \min_{\mathcal{S}\subset[N]} \phi(\mathcal{S})^2 = \phi([N])^2$$

We now show this equality holds more generally.

Lemma 6. *For any coefficient matrix C,*

$$\min_{\lambda\in\Delta^{N-1}} V^*(\lambda) = \phi([N])^2. \tag{11}$$

Proof. Because society can choose to focus on any minimal spanning set, it is clear that $\min_\lambda V^*(\lambda) \leq \phi([N])^2 = \min_{\mathcal{S}}(\phi(\mathcal{S}))^2$. It remains to prove $V^*(\lambda) \geq \phi([N])^2$ for any $\lambda \in \Delta^{N-1}$. By Lemma 2, we need to show $[(C'\Lambda C)^{-1}]_{11} \geq \phi([N])^2$.

This was already proved for *generic* coefficient matrices C (specifically, those for which $\phi(\mathcal{S})$ is minimized by a set of K signals). But even if C is "non-generic", we can approximate it by a sequence of "generic" matrices C_m.[28] Along this sequence, we have

$$[(C_m'\Lambda C_m)^{-1}]_{11} \geq \phi_m([N])^2$$

[28] First, we may add repetitive signals to ensure $N \geq K$. This does not affect the value of $\min V^*(\lambda)$ or $\phi([N])$. Whenever $N \geq K$, it is generically true that every minimal spanning set contains exactly K signals. Moreover, the equality $\phi(\mathcal{S}) = \phi(\tilde{\mathcal{S}})$ for $\mathcal{S} \neq \tilde{\mathcal{S}}$ induces a non-trivial polynomial equation over the entries in C. This means we can always find $C^{(m)}$ close to C such that for each coefficient matrix $C^{(m)}$, different subsets $\mathcal{S}$ (of size K) attain different values of ϕ.

where ϕ_m is the speed of learning from the N signals given by coefficient matrix C_m. As $m \to \infty$, the LHS above approaches $[(C'\Lambda C)^{-1}]_{11}$. Thus the lemma will follow once we show that $\limsup_{m\to\infty} \phi_m([N]) \geq \phi([N])$.

For this we invoke the following characterization

$$\phi([N]) = \min_{\beta \in \mathbb{R}^N} \sum_{i=1}^{N} |\beta_i| \quad \text{s.t.} \quad e_1 = \sum_{i=1}^{N} \beta_i \cdot c_i.$$

If $e_1 = \sum_i \beta_i^{(m)} \cdot c_i^{(m)}$ along the convergent sequence, then $e_1 = \sum_i \beta_i \cdot c_i$ for any limit point β of $\beta^{(m)}$. This enables us to conclude $\liminf_{m\to\infty} \phi_m([N]) \geq \phi([N])$, which is more than what we need. □

A.2.6 Proof of Theorem 1 Part (a) when $|\mathcal{S}^*| < K$

Let $\mathcal{S}^* = \{1, \ldots, k\}$. We will now show that even if $k < K$, λ^* is still the unique minimizer of $V^*(\cdot)$. This will imply part (a) of Theorem 1 via Lemma 3. Since $V^*(\lambda^*) = \phi(\mathcal{S}^*)^2 = \phi([N])^2$ by definition, we know from Lemma 6 that λ^* does minimize V^*. It remains to show that λ^* is the *unique* minimizer.

To do this, we will consider a perturbed informational environment in which signals $k+1, \ldots, N$ are made slightly more precise. Specifically, let $\eta > 0$ be a small positive number. Consider an alternative signal coefficient matrix $\tilde{C}$ with $\tilde{c_i} = c_i$ for $i \leq k$ and $\tilde{c_i} = (1+\eta)c_i$ for $i > k$. Let $\tilde{\phi}(\mathcal{S})$ represent the speed of learning function for this alternative problem. Then, it is clear that $\tilde{\phi}(\mathcal{S}^*) = \phi(\mathcal{S}^*)$, while $\tilde{\phi}(\mathcal{S})$ is slightly smaller than $\phi(\mathcal{S})$ for $\mathcal{S} \neq \mathcal{S}^*$. Thus with sufficiently small η, the subset $\mathcal{S}^*$ remains the best set in this perturbed environment, and λ^* remains the optimal frequency vector.

Let $\tilde{V}^*$ be the asymptotic posterior variance function here, then our previous analysis shows that $\tilde{V}^*$ has minimum value $\phi(\mathcal{S}^*)^2$ on the simplex. Taking advantage of the connection between V^* and $\tilde{V}^*$, we thus have

$$\begin{aligned} V^*(\lambda_1, \ldots, \lambda_N) &= \tilde{V}^*\left(\lambda_1, \ldots, \lambda_k, \frac{\lambda_{k+1}}{(1+\eta)^2}, \ldots, \frac{\lambda_N}{(1+\eta)^2}\right) \\ &\geq \frac{\phi(\mathcal{S}^*)^2}{\sum_{i\leq k} \lambda_i + \frac{1}{(1+\eta)^2} \sum_{i>k} \lambda_i}. \end{aligned}$$

The equality uses (6) and $C'\Lambda C = \sum_i \lambda_i c_i c_i' = \sum_{i\leq k} \lambda_i c_i c_i' + \sum_{i>k} \frac{\lambda_i}{(1+\eta)^2} \tilde{c}_i \tilde{c}_i{}'$. The inequality follows from the homogeneity of $\tilde{V}^*$.

The above display implies

$$\forall \lambda \in \Delta^{N-1}, \quad V^*(\lambda) \geq \frac{\phi(\mathcal{S}^*)^2}{1 - \frac{2\eta+\eta^2}{(1+\eta)^2} \sum_{i>k} \lambda_i} \geq \frac{\phi(\mathcal{S}^*)^2}{1 - \eta \sum_{i>k} \lambda_i} \quad \text{for some } \eta > 0. \tag{12}$$

Hence $V^*(\lambda) > \phi(\mathcal{S}^*)^2 = V^*(\lambda^*)$ whenever λ puts positive weight outside of the best set. But we already know that $V^*(\lambda)$ is uniquely minimized at λ^* when λ is restricted to the best set. Hence λ^* is the unique minimizer of V^* over the whole simplex. This completes the proof of Theorem 1 part (a).

Remark 2. *We can actually show a stronger result, namely $n_i(t) = \lambda_i^* \cdot t + O(1), \forall i$. Note that for any $(q_1, \dots, q_N)$ to be a t-optimal division, it is necessary that $tV(q_1, \dots, q_N) \leq tV(\lambda^* t)$. A straightforward refinement of Lemma 2 gives that whenever $V^*(\lambda)$ is finite, $t \cdot V(\lambda t)$ approaches $V^*(\lambda)$ at the rate of $\frac{1}{t}$. Thus we must have*

$$V^*\left(\frac{q_1}{t}, \dots, \frac{q_N}{t}\right) \leq V^*(\lambda^*) + O\left(\frac{1}{t}\right).$$

Together with (12), this implies $\sum_{i>k} \frac{q_i}{t} = O(\frac{1}{t})$, so that signals outside of $\mathcal{S}^*$ are only observed finitely many times. Conditional on these *signal counts, Proposition 1 shows that society's optimal allocation over the first k signals must satisfy $n_i(t) = \lambda_i^* \cdot t + O(1), \forall 1 \leq i \leq k$. This proves what we want.*

A.3 Proof of Theorem 2 Part (a)

Let signals $1, \dots, k$ (with $k \leq K$) be a minimally spanning set that is optimal in its subspace. We will demonstrate an open set of prior beliefs given which *all agents* observe these k signals. Since these signals are minimally spanning, they must be linearly independent. Thus we can consider linearly transformed states $\tilde{\theta}_1, \dots, \tilde{\theta}_K$ such that these k signals are simply $\tilde{\theta}_1, \dots, \tilde{\theta}_k$ plus standard Gaussian noise. This linear transformation is invertible, so any prior over the original states is bijectively mapped to a prior over the transformed states. Thus it is without loss to work with the transformed model and look for prior beliefs over the transformed states.

The payoff-relevant state ω becomes a linear combination $\lambda_1^*\tilde{\theta}_1 + \cdots + \lambda_k^*\tilde{\theta}_k$ (after scaling). Since the first k signals are optimal in their subspace, Lemma 4 implies that any other signal that belongs to this subspace can be written as

$$\sum_{i=1}^{k} \alpha_i \tilde{\theta}_i \ + \ \mathcal{N}(0,1)$$

with $|\sum_{i=1}^{k} \alpha_i| < 1$. On the other hand, if a signal does not belong to this subspace, it must take the form of

$$\sum_{i=1}^{K} \beta_i \tilde{\theta}_i \ + \ \mathcal{N}(0,1)$$

with $\beta_{k+1}, \dots, \beta_K$ not all equal to zero.

Now consider a prior belief such that $\tilde{\theta}_1, \dots, \tilde{\theta}_K$ are *independent* from each other. Given prior variances $v_1, \dots, v_K$, the reduction in the variance of $\lambda_1^* \tilde{\theta}_1 + \cdots + \lambda_k^* \tilde{\theta}_k$ by any signal $\sum_{i=1}^k \alpha_i \tilde{\theta}_i + \mathcal{N}(0,1)$ is

$$\frac{(\sum_{i=1}^k \alpha_i \lambda_i^* v_i)^2}{1 + \sum_{i=1}^k \alpha_i^2 v_i}$$

If $v_1, \dots, v_k$ are small positive numbers and *if the product $\lambda_i^* v_i$ is approximately constant across* $1 \leq i \leq k$, then the above is approximately $(\sum_{i=1}^k \alpha_i)^2 (\lambda_1^* v_1)^2$. Since $|\sum_{i=1}^k \alpha_i| < 1$, we deduce that any other signal belonging to the subspace of the first k signals is worse than signal 1 (in the first period), whose variance reduction is $\frac{(\lambda_1^* v_1)^2}{1+v_1}$.

Meanwhile, take any signal that does not belong to the subspace. The variance reduction by such a signal $\sum_{i=1}^K \beta_i \tilde{\theta}_i + \mathcal{N}(0,1)$ is

$$\frac{(\sum_{i=1}^k \beta_i \lambda_i^* v_i)^2}{1 + \sum_{i=1}^K \beta_i^2 v_i}$$

As $\beta_{k+1}, \dots, \beta_K$ are not all zero, the denominator above is arbitrarily large if $v_{k+1}, \dots, v_K$ are chosen to be large. Then, this signal is again worse than signal 1 for the first agent, similar to the situation in Example 1.

To summarize, we have shown that whenever the prior variances $v_1, \dots, v_K$ satisfy the following three conditions, the first agent chooses among the first k signals:

1. $v_1, \dots, v_k$ are close to 0;
2. $\lambda_1^* v_1, \dots, \lambda_k^* v_k$ have pairwise ratios close to 1;
3. $v_{k+1}, \dots, v_K$ are large.[29]

To show *every agent* chooses among the first k signals, it suffices to check that starting from any prior belief satisfying the above conditions, the posterior beliefs after observing a signal continue to satisfy these conditions. Since variances decrease over time, the first condition is obviously satisfied. By independence, learning about $\tilde{\theta}_1, \dots, \tilde{\theta}_k$ does not affect the variances of the remaining states. So $v_{k+1}, \dots, v_K$ are unchanged, and the third condition is verified. Finally, the second condition holds by Proposition 1: Since each signal $i \leq k$ is sampled according to λ_i^*, the variance v_i after t periods is approximately $\frac{1}{\lambda_i^* t}$. Hence part (a) of Theorem 2 is proved.[30]

[29] Formally, we require that for some fixed constant $\epsilon > 0$, it holds that $v_1, \dots, v_k < \epsilon$; $\max_{1 \leq i \leq k} \lambda_i^* v_i \leq (1+\epsilon) \cdot \min_{1 \leq i \leq k} \lambda_i^* v_i$; and $v_{k+1}, \dots, v_K > \frac{1}{\epsilon}$.

[30] While we restrict attention to priors that are independent over $\tilde{\theta}_1, \dots, \tilde{\theta}_K$, it can be shown that the argument extends to mild prior correlation across the transformed states. Specifically, we only need to impose the extra condition that $\mathrm{cov}(\tilde{\theta}_i, \tilde{\theta}_j) \leq \epsilon \cdot v_j, \forall i \neq j$.

A.4 Proof of Theorem 2 Part (b)

A.4.1 Preliminary Steps

Given any prior, let $\mathcal{A} \subset [N]$ be the set of all signals that are observed by infinitely many agents. We first show that $\mathcal{A}$ is a spanning set.

Indeed, by definition we can find some period t after which agents only observe signals in $\mathcal{A}$. Note that the variance reduction of any signal approaches zero as its signal count gets large. Thus, along society's signal path, the variance reduction is close to zero at sufficiently late periods.

If $\mathcal{A}$ is not spanning, society's posterior variance remains bounded away from zero. Thus in the limit where each signal in $\mathcal{A}$ has infinite signal counts, there still exists some signal j outside of $\mathcal{A}$ whose variance reduction is strictly positive.[31] By continuity, we deduce that at any sufficiently late period, observing signal j is better than observing any signal in $\mathcal{A}$. This contradicts our assumption that later agents only observe signals in $\mathcal{A}$.

Now that $\mathcal{A}$ is spanning, we can take $\mathcal{S}$ to be the optimal minimal spanning set in the subspace spanned by $\mathcal{A}$. To prove Theorem 2 part (b), we will show that long-run frequencies are positive precisely for the signals in $\mathcal{S}$. By ignoring the initial periods, we assume without loss that *only signals in* $\overline{\mathcal{A}}$ *are available*. It suffices to show that whenever the signals observed infinitely often span a subspace, agents eventually sample from the optimal subset $\mathcal{S}$ in that subspace. To ease notation, we assume this subspace is the entire $\mathbb{R}^K$, and prove the following result:

Theorem 2 part (b) Restated. *Suppose that the signals observed infinitely often span* $\mathbb{R}^K$*. Then society eventually observes signals in* $\mathcal{S}^*$ *with frequencies* λ^**.*

The next sections are devoted to the proof of this restatement.

A.4.2 Estimates of Derivatives

We introduce a few technical lemmata:

Lemma 7. *For any* $q_1, \ldots, q_N$*, we have*

$$\left| \frac{\partial_{jj} V(q_1, \ldots, q_N)}{\partial_j V(q_1, \ldots, q_N)} \right| \leq \frac{2}{q_j}.$$

[31] To see this, let $s_1, \ldots, s_N$ denote the limit signal counts, where $s_i = \infty$ if and only if $i \in \mathcal{A}$. We need to find some signal j such that $V(s_j + 1, s_{-j}) < V(s_j, s_{-j})$. If such a signal does not exist, then all partial derivatives of V at s are zero. Since V is differentiable, this would imply that all directional derivatives of V are also zero. By the convexity of V, V must be minimized at s. However, the minimum value of V is zero because there exists a spanning set. This contradicts $V(s) > 0$.

Proof. Recall that $V(q_1, \ldots, q_N) = e_1' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot e_1$. Thus

$$\partial_j V = -e_1' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j \cdot c_j' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot e_1,$$

and

$$\partial_{jj} V = 2e_1' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j \cdot c_j' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j \cdot c_j' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot e_1.$$

Let $\gamma_j = e_1' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j$, which is a number. Then the above becomes

$$\partial_j f = -\gamma_j^2; \qquad \partial_{jj} f = 2\gamma_j^2 \cdot c_j' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j.$$

Note that $(\Sigma^0)^{-1} + C'QC \succeq q_j \cdot c_j c_j'$ in matrix norm. Thus the number $c_j' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j$ is bounded above by $\frac{1}{q_j}$.[32] This proves the lemma. ☐

Since the second derivative is small compared to the first derivative, we deduce that the variance reduction of any *discrete* signal can be approximated by the partial derivative of f. This property is summarized in the following lemma:

Lemma 8. *For any $q_1, \ldots, q_N$, we have*[33]

$$V(q) - V(q_j + 1, q_{-j}) \geq \frac{q_j}{q_j + 1} |\partial_j V(q)|.$$

Proof. We will show the more general result:

$$V(q) - V(q_j + x, q_{-j}) \geq \frac{q_j x}{q_j + x} \cdot |\partial_j V(q)|, \forall x \geq 0.$$

This clearly holds at $x = 0$. Differentiating with respect to x, we only need to show

$$-\partial_j V(q_j + x, q_{-j}) \geq \frac{q_j^2}{(q_j + x)^2} |\partial_j V(q)|, \forall x \geq 0.$$

Equivalently, we need to show

$$-(q_j + x)^2 \cdot \partial_j V(q_j + x, q_{-j}) \geq -q_j^2 \cdot \partial_j V(q), \forall x \geq 0.$$

Again, this inequality holds at $x = 0$. Differentiating with respect to x, it becomes

$$-2(q_j + x) \cdot \partial_j V(q_j + x, q_{-j}) - (q_j + x)^2 \cdot \partial_{jj} V(q_j + x, q_{-j}) \geq 0.$$

This is exactly the result of Lemma 7. ☐

[32]Formally, we need to show that for any $\epsilon > 0$, the number $c_j'[c_j c_j' + \epsilon I_K]^{-1} c_j$ is at most 1. Using the identify $Trace(AB) = Trace(BA)$, we can rewrite this number as

$$Trace([c_j c_j' + \epsilon I_K]^{-1} c_j c_j') = Trace(I_K - [c_j c_j' + \epsilon I_K]^{-1} \epsilon I_K) = K - \epsilon \cdot Trace([c_j c_j' + \epsilon I_K]^{-1}).$$

The matrix $c_j c_j'$ has rank 1, so $K-1$ of its eigenvalues are zero. Thus the matrix $[c_j c_j' + \epsilon I_K]^{-1}$ has eigenvalue $1/\epsilon$ with multiplicity $K-1$, and the remaining eigenvalue is positive. This implies $\epsilon \cdot Trace([c_j c_j' + \epsilon I_K]^{-1}) > K-1$, and then the above display yields $c_j' \cdot [(\Sigma^0)^{-1} + C'QC]^{-1} \cdot c_j < 1$ as desired.

[33]Note that the convexity of V gives $V(q) - V(q_j + 1, q_{-j}) \leq |\partial_j V(q)|$. This lemma provides a converse that we need for the subsequent analysis.

A.4.3 Lower Bound on Variance Reduction

Our next result gives a lower bound on the directional derivative of V along the "optimal" direction λ^*:

Lemma 9. *For any $q_1, \ldots, q_N$, we have*

$$|\partial_{\lambda^*} V(q)| \geq \frac{V(q)^2}{\phi(\mathcal{S}^*)^2}.$$

Proof. To compute this directional derivative, we think of agents acquiring signals in fractional amounts, where a fraction of a signal is just the same signal with precision multiplied by that fraction. Consider an agent who draws λ_i^* realizations of each signal i. Then he essentially obtains the following signals:

$$Y_i = \langle c_i, \theta \rangle + \mathcal{N}\left(0, \frac{1}{\lambda_i^*}\right), \forall i.$$

This is equivalent to

$$\lambda_i^* Y_i = \langle \lambda_i^* c_i, \theta \rangle + \mathcal{N}(0, \lambda_i^*), \forall i.$$

Such an agent receives at least as much information as the sum of these signals:

$$\sum_i \lambda_i^* Y_i = \sum_i \langle \lambda_i^* c_i, \theta \rangle + \sum_i \mathcal{N}(0, \lambda_i^*) = \frac{\omega}{\phi(\mathcal{S}^*)} + \mathcal{N}(0, 1).$$

Hence the agent's posterior precision about ω (which is the inverse of his posterior variance V) must increase by at least $\frac{1}{\phi(\mathcal{S}^*)^2}$ along the direction λ^*. The chain rule of differentiation yields the lemma. □

We can now bound the variance reduction at late periods:

Lemma 10. *Fix any $q_1, \ldots, q_N$. Suppose M is a positive real number such that $(\Sigma^0)^{-1} + C'QC \succeq M c_j c_j'$ holds for each signal $j \in \mathcal{S}^*$. Then we have*

$$\min_{j \in \mathcal{S}^*} V(q_j + 1, q_{-j}) \leq V(q) - \frac{M}{M+1} \cdot \frac{V(q)^2}{\phi(\mathcal{S}^*)^2}.$$

Proof. Fix any signal $j \in \mathcal{S}^*$. Using the condition $(\Sigma^0)^{-1} + C'QC \succeq M c_j c_j'$, we can deduce the following variant of Lemma 8:[34]

$$V(q) - V(q_j + 1, q_{-j}) \geq \frac{M}{M+1} |\partial_j V(q)|.$$

[34] Even though we are not guaranteed $q_j \geq M$, we can modify the prior and signal counts such that the precision matrix $(\Sigma^0)^{-1} + C'QC$ is unchanged, and signal j has been observed at least M times. This is possible thanks to the condition $(\Sigma^0)^{-1} + C'QC \succeq M c_j c_j'$. Then, applying Lemma 8 to this modified problem yields the result here.

Since V is always differentiable, $\partial_{\lambda^*} V(q)$ is a convex combination of the partial derivatives of V.[35] Thus

$$\max_{j\in\mathcal{S}^*} |\partial_j V(q)| \geq |\partial_{\lambda^*} V(q)|$$

These inequalities together with Lemma 9 complete the proof. □

A.4.4 Proof of the Restatement of Theorem 2 Part (b)

We will show $V(m(t)) \sim \frac{\phi(\mathcal{S}^*)^2}{t}$, so that society eventually approximates the optimal speed of learning. Since λ^* is the unique minimizer of V^*, this will imply the desired conclusion $m(t)/t \to \lambda^*$.

To estimate $V(m(t))$, we note that for any fixed M, society's acquisitions $m(t)$ eventually satisfy the condition $(\Sigma^0)^{-1} + C'QC \succeq Mc_jc_j'$. This is due to our assumption that the signals observed infinitely often span $\mathbb{R}^K$, which implies that $C'QC$ becomes arbitrarily large in matrix norm. Hence, we can apply Lemma 10 to find that

$$V(m(t+1)) \leq V(m(t)) - \frac{M}{M+1} \cdot \frac{V(m(t))^2}{\phi(\mathcal{S}^*)^2}$$

for all $t \geq t_0$, where t_0 depends only on M.

We introduce the auxiliary function $g(t) = V(m(t)) \cdot \frac{M}{(M+1)\phi(\mathcal{S}^*)^2}$. Then the above simplifies to

$$g(t+1) \leq g(t) - g(t)^2.$$

Inverting both sides, we have

$$\frac{1}{g(t+1)} \geq \frac{1}{g(t)(1-g(t))} = \frac{1}{g(t)} + \frac{1}{1-g(t)} \geq \frac{1}{g(t)} + 1. \tag{13}$$

This holds for all $t \geq t_0$. Thus by induction, $\frac{1}{g(t)} \geq t - t_0$ and so $g(t) \leq \frac{1}{t-t_0}$. Going back to the posterior variance function V, this implies

$$V(m(t)) \leq \frac{M+1}{M} \cdot \frac{\phi(\mathcal{S}^*)^2}{t-t_0}. \tag{14}$$

Hence, by choosing M sufficiently large in the first place and then considering large t, we find that society's speed of learning is arbitrarily close to the optimal speed $\phi(\mathcal{S}^*)^2$. This completes the proof.

We comment that the above argument leaves open the possibility that some signals outside of $\mathcal{S}^*$ are observed *infinitely often*, yet with *zero long-run frequency*. In Appendix C.2, we show this does not happen.

[35] While this may be a surprising contrast with V^*, the difference arises because the formula for V always involves a full-rank prior covariance matrix, whereas its asymptotic variant V^* corresponds to a flat prior.

A.5 Proof of Proposition 2

Given any history of observations, an agent can always allocate his B observations as follows: He draws $\lfloor B \cdot \lambda_i^* \rfloor$ realizations of each signal i, and samples arbitrarily if there is any capacity remaining. Here $\lfloor \ \rfloor$ denotes the floor function.

Fix any $\epsilon > 0$. If B is sufficiently large, then the above strategy acquires at least $(1-\epsilon) \cdot B \cdot \lambda_i^*$ observations of each signal i. Adapting the proof of Lemma 9, we see that the agent's posterior precision about ω must increase by $\frac{(1-\epsilon)B}{\phi(\mathcal{S}^*)^2}$ under this strategy. Thus the same must hold for his optimal strategy, so that society's posterior precision at time t is at least $\frac{(1-\epsilon)Bt}{\phi(\mathcal{S}^*)^2}$. This implies that society's speed of learning (per signal) is at least $\frac{\phi(\mathcal{S}^*)^2}{(1-\epsilon)}$, which can be arbitrarily close to the optimal speed $\phi(\mathcal{S}^*)^2$ with appropriate choice of ϵ.

Since λ^* is the unique minimizer of V^*, society's long-run frequencies must be close to λ^*. In particular, with ϵ sufficiently small, we can ensure that each signal in $\mathcal{S}^*$ are observed with positive frequencies. The restated Theorem 2 part (b) extends to the current setting and implies that society eventually samples from $\mathcal{S}^*$. This yields the proposition.[36]

A.6 Proof of Proposition 3

Suppose without loss that the best set is $\{1, \ldots, k\}$. By taking a linear transformation, we can further assume each signal i with $1 \le i \le k$ only involves ω and the first $k-1$ confounding terms $b_1, \ldots, b_{k-1}$. We claim that whenever $k-1$ sufficiently precise signals are provided about each of these confounding terms, long-run frequencies must converge to λ^*.

Fix any positive real number M. Since the $k-1$ free signals are very precise, it is as if the prior precision matrix satisfies

$$(\Sigma^0)^{-1} \succeq M^2 \sum_{i=2}^{k} \Delta_{ii}$$

where Δ_{ii} be the $K \times K$ matrix that has one at the (i,i) entry and zero otherwise. Recall also that society eventually learns ω. Thus at some late period t_0, society's acquisitions must satisfy

$$C'QC \succeq M^2 \Delta_{11}.$$

[36]This proof also suggests that how small ϵ (and how large B) need to be depends on the distance between the optimal speed of learning and the "second-best" speed of learning from any other minimal spanning set. Intuitively, in order to achieve long-run efficient learning, agents need to allocate B observations in the best set to approximate the optimal frequencies. If another set of signals offers a speed of learning that is only slightly worse, we will need B sufficiently large for the approximately optimal frequencies in the best set to beat this other set.

Adding up the above two displays, we have

$$(\Sigma^0)^{-1} + C'QC \succeq M^2 \sum_{i=1}^{k} \Delta_{ii} \succeq M c_j c_j', \forall 1 \le j \le k.$$

The last inequality uses the fact that each c_j only involves the first k coordinates.

Now this is exactly the condition we need in order to apply Lemma 10: Whether or not the condition is met for $j \notin \mathcal{S}^*$ does not affect the argument. Thus we can follow the proof of the restated Theorem 2 part (b) to deduce (14). That is, for fixed M and corresponding free information, society's long-run speed of learning cannot be slower than $(1 + 1/M) \cdot \phi(\mathcal{S}^*)^2$. This can be made arbitrarily close to the optimal speed, in which case we use Theorem 2 part (b) to conclude that society eventually samples according to λ^*.

A.7 Proof of Theorem 1 Part (b)

Recall that λ^* uniquely minimizes the function V^*. Thus the proposition is equivalent to the following: Fix $\epsilon > 0$, then for any δ close to 1, $V(d_\delta(t)) \le \frac{(1+\epsilon)\phi(\mathcal{S}^*)^2}{t}$ holds for sufficiently large t. That is, we only need to show that as $\delta \to 1$, the achieved speed of learning is close to the optimal speed.

Suppose for contradiction that $V(d_\delta(t)) > \frac{(1+\epsilon)\phi(\mathcal{S}^*)^2}{t}$ at some large t. Let $\tau < t$ be the last period with $V(d_\delta(\tau)) \le \frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{\tau}$. Below we first assume such a period τ exists; later we will show how to modify the proof when it does not. Consider the following deviation:

1. Agents $i \le \tau$ choose signals according to d_δ (i.e., they do not deviate);

2. Starting in period $\tau + 1$, the next Mk agents sample each signal in the best set (of size k) exactly M times, in an arbitrary order;

3. Starting in period $\tau + Mk + 1$, each future agent chooses the signal that maximizes his own expected payoff, as in our main model.

In what follows we will show that for appropriately chosen M as well as sufficiently large δ and t, this deviation yields a higher δ-discounted payoff than the original strategy d_δ.

By construction, the deviation strategy achieves the same payoff as the original strategy in the first τ periods. Next we consider those periods $\tau + 1$ through t. For $1 \le j \le t - \tau$, let $\tilde{V}(\tau + j)$ denote the posterior variance at time $\tau + j$ under the deviation strategy. We can bound it from above as follows: Our previous analysis in Appendix A.6 (specifically (13)) gives that for $j > Mk$,

$$\frac{(M+1)\phi(\mathcal{S}^*)^2}{M \cdot \tilde{V}(\tau + j)} \ge \frac{(M+1)\phi(\mathcal{S}^*)^2}{M \cdot \tilde{V}(\tau + Mk)} + j - Mk, \tag{15}$$

Using $\tilde{V}(\tau+Mk) \leq \tilde{V}(\tau) \leq \frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{\tau}$, the above inequality further yields

$$\frac{(M+1)\phi(\mathcal{S}^*)^2}{M\cdot\tilde{V}(\tau+j)} \geq \frac{(M+1)\tau}{M(1+\epsilon/2)} + j - Mk. \tag{16}$$

With slight algebra, we obtain from the above

$$\frac{\tilde{V}(\tau+j)}{\phi(\mathcal{S}^*)^2} \leq \frac{1}{\frac{\tau}{1+\epsilon/2} + \frac{j-Mk}{1+1/M}}. \tag{17}$$

Fixing ϵ, we now choose M so that $\frac{1}{M} < \frac{\epsilon}{4}$. Then there exists $\underline{j}$ (depending only on ϵ, M and K) such that for $j > \underline{j}$, it holds that

$$\frac{j-Mk}{1+1/M} \geq \frac{j+1}{1+\epsilon/2}.$$

Thus, (17) implies

$$\tilde{V}(\tau+j) \leq \frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{\tau+j+1}, \quad \forall \underline{j}+1 \leq j \leq t-\tau. \tag{18}$$

On the other hand, for small j we have the following crude estimate:

$$\tilde{V}(\tau+j) \leq \tilde{V}(\tau) \leq \frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{\tau}, \quad \forall 1 \leq j \leq \underline{j}. \tag{19}$$

Now we go back to the original strategy and make payoff comparisons. Our choice of τ ensures that posterior variance under the *original* strategy is at least $\frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{\tau+j}$, for $1 \leq j \leq t-\tau$. Hence by deviating, the payoff gain in periods $\tau+1 \sim t$ is at least[37]

$$\delta^\tau \cdot \underbrace{\left(\sum_{j=\underline{j}+1}^{t-\tau} \delta^{j-1}\left[\frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{(\tau+j)(\tau+j+1)}\right] - \sum_{j=1}^{\underline{j}} \delta^{j-1}\left[\frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2 j}{\tau(\tau+j)}\right]\right)}_{(**)}.$$

Note that $\underline{j}$ has already been fixed. So as $\delta \to 1$ and $t-\tau \to \infty$,[38] the term $(**)$ above converges to

$$(1+\epsilon/2)\phi(\mathcal{S}^*)^2 \cdot \left[\frac{1}{\tau+\underline{j}+1} - \sum_{j=1}^{\underline{j}} \frac{j}{\tau(\tau+j)}\right].$$

For large τ, the above expression is larger than $\phi(\mathcal{S}^*)^2/\tau$.

[37] In this derivation we use (18), (19) as well as the identities $\frac{1}{\tau+j} - \frac{1}{\tau+j+1} = \frac{1}{(\tau+j)(\tau+j+1)}$ and $\frac{1}{\tau} - \frac{1}{\tau+j} = \frac{j}{\tau(\tau+j)}$.

[38] Since $\frac{(1+\epsilon)\phi(\mathcal{S}^*)^2}{t} \leq V(d_\delta(t)) \leq V(d_\delta(\tau)) \leq \frac{(1+\epsilon/2)\phi(\mathcal{S}^*)^2}{\tau}$, we have $\tau \leq \frac{1+\epsilon/2}{1+\epsilon}t$. So as t becomes large, the difference $t-\tau$ also necessarily becomes large.

Summarizing the above, we have shown that whenever $\tau > \underline{\tau}$, the deviation strategy achieves payoff gain in periods $\tau + 1$ through t of at least $\delta^\tau \phi(\mathcal{S}^*)^2/\tau$ (for δ close to 1). Although the deviation strategy might do worse in periods $t+1$ onwards, the potential payoff loss is at most $O(\frac{\delta^t}{1-\delta})$, which is smaller than the aforementioned payoff gain $\delta^\tau \phi(\mathcal{S}^*)^2/\tau$ as $t - \tau \to \infty$ (since $\tau \leq \frac{1+\epsilon/2}{1+\epsilon} t$). Hence whenever $\tau > \underline{\tau}$, the deviation we constructed is a profitable deviation, and the proposition holds in these situations.

Finally, we need to address the case where the previously-defined τ is weakly less than $\underline{\tau}$. This covers the case in which τ does not exist according to our earlier definition (simply let $\tau = 0$). Instead of (16), we use the following weaker inequality

$$\frac{(M+1)\phi(\mathcal{S}^*)^2}{M \cdot \tilde{V}(\tau + j)} \geq j - Mk.$$

That is, $\tilde{V}(\tau + j) \leq \frac{(1+\frac{1}{M})\phi(\mathcal{S}^*)^2}{j - Mk}$. Recall that $\frac{1}{M} < \frac{\epsilon}{4}$ and $\tau \leq \underline{\tau}$ is now bounded. Thus for $j > \bar{j}$ (where $\bar{j}$ may need to be larger than $\underline{j}$), we would have

$$\tilde{V}(\tau + j) \leq \frac{(1+\epsilon/4)\phi(\mathcal{S}^*)^2}{\tau + j}, \quad \forall \bar{j} + 1 \leq j \leq t - \tau. \tag{20}$$

And for small j we can simply bound the posterior variance by the prior:

$$\tilde{V}(\tau + j) \leq c, \quad \forall 1 \leq j \leq \bar{j}. \tag{21}$$

Using the estimates (20) and (21) in place of (18) and (19), we find that the deviation strategy achieves payoff gain in periods $\tau + 1$ through t of at least

$$\delta^\tau \cdot \left(\sum_{j=\bar{j}+1}^{t-\tau} \delta^{j-1} \left[\frac{\epsilon/4 \cdot \phi(\mathcal{S}^*)^2}{\tau + j} \right] - \sum_{j=1}^{\bar{j}} \delta^{j-1} c \right).$$

Importantly, because (20) improves upon (18), we now have a harmonic sum in (the first part of) the parenthesis, which becomes arbitrarily large for δ close to 1. Hence the above payoff gain is at least δ^τ as $\delta \to 1$ and $t \to \infty$. Once again, this payoff gain dominates any potential loss after period t, showing that the deviation strategy is profitable. This proves part (b) of Theorem 1.

B Proofs for the Autocorrelated Model (Section 5)

B.1 Proof of Theorem 3

We work with the transformed model. Choose M so that the innovations corresponding to the transformed states are independent from each other. In other words, $\tilde{M}$ (the transformed version of M) is given by $\text{diag}(\frac{x}{w_1}, \dots, \frac{x}{w_k}, y_{k+1}, \dots, y_K)$. Here x is a small positive number, while $y_{k+1}, \dots, y_K$ are large positive numbers. We further choose $\Sigma^0 = M$, which is the stable belief without learning.

With these choices, it is clear that if all agents only sample from $\mathcal{S}$, society's beliefs about the transformed states remain independent at every period. Let v_i^{t-1} denote the belief variance about $\tilde{\theta}_i^t$ at the beginning of period t (before the signal acquisition in that period). Then as long as agent t would continue to sample a signal $\tilde{\theta}_j + \mathcal{N}(0,1)$ in $\mathcal{S}$, these prior variances would evolve as follows: $v_i^0 = \frac{x}{w_i}$ for $1 \le i \le k$ and $v_i^0 = y_i$ for $i > k$. And for $t \ge 1$,

$$v_i^t = \begin{cases} \alpha \cdot v_i^{t-1} + (1-\alpha)\tilde{M}_{ii}, & \text{if } i \neq j; \\ \alpha \cdot \frac{v_i^{t-1}}{1+v_i^{t-1}} + (1-\alpha)\tilde{M}_{ii} & \text{if } i = j. \end{cases}$$

By induction, it is clear that $v_i^t \le \tilde{M}_{ii}$ holds for all pairs i, t, with equality for $i > k$ (since any signal in $\mathcal{S}$ must have $j \le k$). Thus at the beginning of each period t, assuming that all previous agents have sampled from $\mathcal{S}$, agent t's prior uncertainties about $\tilde{\theta}_1, \dots, \tilde{\theta}_k$ are small while his uncertainties about $\tilde{\theta}_{k+1}, \dots, \tilde{\theta}_K$ are large. As such, our previous proof for the existence of learning traps with persistent states carries over, and we deduce that agent t's myopic signal choice does belong to $\mathcal{S}$. This proves the first half of the theorem.

For the second half, we first show the myopic rule leads to discounted average payoff at most $-(1-\epsilon)\phi(\mathcal{S}) \cdot \sqrt{(1-\alpha)M_{11}}$. Since we are concerned with the patient limit, it suffices to show that society's *posterior* variance about ω is larger than $(1-\epsilon)\phi(\mathcal{S}) \cdot \sqrt{(1-\alpha)M_{11}}$ *in every late period*. Below we show that the *prior* variance about ω, given by $\sum_{i=1}^k w_i^2 \cdot v_i^t$, is eventually larger than $(1-\epsilon/2)\phi(\mathcal{S}) \cdot \sqrt{(1-\alpha)M_{11}}$. This will imply the result because the difference between the prior and posterior variances is $w_j^2 \cdot (v_j^t - \frac{v_j^t}{1+v_j^t}) \le (w_j \cdot v_j^t)^2 = O(1-\alpha)$ when signal j is acquired in that period.[39]

To prove the desired lower bound on $\sum_{i=1}^k w_i^2 \cdot v_i^t$, we recall that if signal j is acquired in period t, then

$$v_j^t = \frac{\alpha v_j^{t-1}}{1+v_j^{t-1}} + \frac{(1-\alpha)x}{w_j} = \alpha v_j^{t-1} + \frac{(1-\alpha)x}{w_j} - \frac{\alpha (v_j^{t-1})^2}{1+v_j^{t-1}}.$$

This gives the following inequality

$$\frac{1}{v_j^t} = \frac{1}{\alpha v_j^{t-1} + \frac{(1-\alpha)x}{w_j}} + \frac{\alpha (v_j^{t-1})^2/(1+v_j^{t-1})}{v_j^t \cdot (\alpha v_j^{t-1} + \frac{(1-\alpha)x}{w_j})} \le \frac{1}{\alpha v_j^{t-1} + \frac{(1-\alpha)x}{w_j}} + \frac{1}{\alpha},$$

where we used $v_j^t \ge \frac{\alpha v_j^{t-1}}{1+v_j^{t-1}}$ and $\alpha v_j^{t-1} + \frac{(1-\alpha)x}{w_j} \ge \alpha v_j^{t-1}$.

[39] Since the myopic rule samples signals $1 \sim k$ with positive frequencies, each of $v_1^t, \dots, v_k^t$ is eventually of order $\sqrt{1-\alpha}$.

Recall also that for every signal $i \neq j$ we have $v_i^t = \alpha v_i^{t-1} + \frac{(1-\alpha)x}{w_i}$. Thus it holds that

$$\sum_{i=1}^k \frac{1}{v_i^t} \leq \frac{1}{\alpha} + \sum_{i=1}^k \frac{1}{\alpha v_i^{t-1} + (1-\alpha)x/w_i}. \tag{22}$$

This inequality holds regardless of the signal choice in period t. Now, by the Cauchy-Schwarz inequality,

$$(\alpha v_i^{t-1} + (1-\alpha)x/w_i) \cdot \left(\frac{1}{v_i^{t-1}} + \frac{w_i}{\phi(\mathcal{S})}\right) \geq \left(\sqrt{\alpha} + \sqrt{\frac{(1-\alpha)x}{\phi(\mathcal{S})}}\right)^2 \geq 1 + (2 - \frac{\epsilon}{2})\sqrt{\frac{(1-\alpha)x}{\phi(\mathcal{S})}},$$

where the latter inequality is true for α close to 1. Plugging back to the RHS of (22), we deduce

$$\sum_{i=1}^k \frac{1}{v_i^t} \leq \frac{1}{\alpha} + \frac{1}{1 + (2-\frac{\epsilon}{2})\sqrt{\frac{(1-\alpha)x}{\phi(\mathcal{S})}}} \cdot \sum_{i=1}^k (\frac{1}{v_i^{t-1}} + \frac{w_i}{\phi(\mathcal{S})}) = 2 + O(\sqrt{1-\alpha}) + \frac{1}{1 + (2-\frac{\epsilon}{2})\sqrt{\frac{(1-\alpha)x}{\phi(\mathcal{S})}}} \cdot \sum_{i=1}^k \frac{1}{v_i^{t-1}}$$

Hence $\limsup_{t\to\infty} \sum_{i=1}^k \frac{1}{v_i^t} \leq \frac{2+O(\sqrt{1-\alpha})}{(2-\epsilon/2)\sqrt{\frac{(1-\alpha)x}{\phi(\mathcal{S})}}} < \frac{1}{1-\epsilon/2} \cdot \sqrt{\frac{\phi(\mathcal{S})}{(1-\alpha)x}}$. Applying the Cauchy-Schwarz inequality again,

$$\sum_{i=1}^k w_i^2 v_i^t \geq \frac{(\sum_{i=1}^k w_i)^2}{\sum_{i=1}^k \frac{1}{v_i^t}}.$$

Since for large t we have $\sum_{i=1}^k \frac{1}{v_i^t} \leq \frac{1}{1-\epsilon/2} \cdot \sqrt{\frac{\phi(\mathcal{S})}{(1-\alpha)x}}$, we conclude from the above that $\sum_{i=1}^k w_i^2 v_i^t \geq (1-\epsilon/2)\phi(\mathcal{S})^{1.5} \cdot \sqrt{(1-\alpha)x} = (1-\epsilon/2)\phi(\mathcal{S}) \cdot \sqrt{(1-\alpha)M_{11}}$, as $M_{11} = \phi(S) \cdot x$.

The preceding analysis yields the desired upper bound on average payoff under the myopic rule. It remains to establish the lower bound on the optimal payoff. This follows from the observation that myopically sampling from the best set $\mathcal{S}^*$ yields a patient payoff at least $-(1+\epsilon)\phi(\mathcal{S}^*) \cdot \sqrt{(1-\alpha)M_{11}}$, which is proved later as part of Theorem 4.

B.2 Proof of Proposition 1

The environment in Example 3 is equivalent to one with three signals $\frac{1}{L}\omega$, $\frac{\omega+b}{2}$, $\frac{\omega-b}{2}$, each with standard Gaussian noise (just let $b = \omega + 2b_1$). We assume L is large, so that $\mathcal{S}^*$ consists of the latter two signals. However, suppose we choose $M = \Sigma^0 = \text{diag}(x, x)$ with $x \geq L^2$. Then assuming that all previous agents have sampled the first signal, agent t's prior variance about b remains $x \geq L^2$. As such, he continues to sample the first signal myopically. Thus under the myopic rule, all agents sample the first signal. In this case the prior variance v^t about ω evolves according to

$$v^t = \alpha \cdot \frac{L^2 \cdot v^{t-1}}{L^2 + v^{t-1}} + (1-\alpha)x.$$

It is not difficult to show that v^t must converge to the (positive) fixed point of the above equation. Let us in particular take $\alpha = 1 - \frac{1}{L^3}$ and $x = L^2$, then the long-run prior variance v solves $v = \frac{(L^2 - \frac{1}{L})v}{L^2+v} + \frac{1}{L}$. This yields exactly that $v = \sqrt{L}$. Hence long-run posterior variance is $\frac{L^2 \cdot v}{L^2+v} > \sqrt{L}/2$, which implies $\limsup_{\delta \to 1} U_\delta^M \leq -\sqrt{L}/2$.

Let us turn to the optimal rule. Write $\tilde{\theta}_1 = \frac{\omega+b}{2}$ and $\tilde{\theta}_2 = \frac{\omega-b}{2}$. In this transformed model, $\tilde{M} = \tilde{\Sigma}^0 = \text{diag}(\frac{x}{2}, \frac{x}{2})$, and the payoff-relevant state is the sum of $\tilde{\theta}_1$ and $\tilde{\theta}_2$. Consider now a strategy that samples the latter two signals alternatively. Then the beliefs about $\tilde{\theta}_1$ and $\tilde{\theta}_2$ remain independent (as in $\tilde{M}$ and $\tilde{\Sigma}^0$), and their variances evolve as follows: $v_1^0 = v_2^0 = \frac{x}{2}$; in odd periods t

$$v_1^t = \alpha \cdot \frac{v_1^{t-1}}{1+v_1^{t-1}} + (1-\alpha)\frac{x}{2} \text{ and } v_2^t = \alpha \cdot v_2^{t-1} + (1-\alpha)\frac{x}{2},$$

and symmetrically for even t.

These imply that for odd t, v_1^t converges to v_1 and v_2^t converges to v_2 below (while for even t $v_1^t \to v_2$ and $v_2^t \to v_1$):

$$v_1 = \alpha \cdot \frac{\alpha v_1 + (1-\alpha)x/2}{1 + \alpha v_2 + (1-\alpha)x/2} + (1-\alpha)x/2;$$

$$v_2 = \alpha^2 \cdot \frac{v_2}{1+v_2} + (1-\alpha^2) \cdot \frac{x}{2}.$$

From the second equation, we obtain $(1-\alpha^2)(\frac{x}{2} - v_2) = \alpha^2 \cdot \frac{(v_2)^2}{1+v_2}$. With $\alpha = 1 - \frac{1}{L^3}$ and $x = L^2$, it follows that

$$v_2 = (1+o(1))\frac{1}{\sqrt{L}}.$$

Thus we also have $v_1 = \alpha\frac{v_2}{1+v_2} + (1-\alpha)\frac{x}{2} = (1+o(1))\frac{1}{\sqrt{L}}$.

Hence under this alternating rule, long-run posterior variances about $\tilde{\theta}_1$ and $\tilde{\theta}_2$ are both bounded above by $\frac{2}{\sqrt{L}}$. Since $\omega = \tilde{\theta}_1 + \tilde{\theta}_2$, we conclude that $\liminf_{\delta \to 1} U_\delta^{SP} \geq -\frac{4}{\sqrt{L}}$. Choosing L large proves the proposition.

For Online Publication

C Other Results

C.1 Example Failing Unique Minimizer

There are $K = 3$ states ω, b_1, b_2 independently drawn with prior variances $\frac{1}{\alpha}, \frac{1}{\beta}, \frac{1}{\gamma}$. $N = 4$ signals are available, and they are respectively

$$\begin{aligned} X_1 &= \omega + b_1 + \epsilon_1 \\ X_2 &= b_1 + \epsilon_2 \\ X_3 &= \omega + b_2 + \epsilon_3 \\ X_4 &= b_2 + \epsilon_4 \end{aligned}$$

with standard normal errors. Note that the former two signals and the latter two signals are both spanning, and these two sets generate the same asymptotic variance. Thus Assumption 1 is not satisfied.

The posterior variance about ω as a function of the number of observations q_1, q_2, q_3, q_4 of each signal type can be derived as follows. First, given q_2 observations of signal X_2 and q_4 observations of signal X_4, posterior variance about θ_2 and θ_3 are $1/(q_2 + \beta)$ and $1/(q_4 + \gamma)$ respectively. Consider now q_1 additional observations of X_1; this provides the same information about the payoff-relevant state ω as the signal $\omega + \epsilon'$, where ϵ' is an independent noise term with variance $\frac{1}{q_1} + \frac{1}{q_2+\beta}$. Similarly, q_3 additional observations of X_3 are equivalent to the signal $\omega + \epsilon''$, where ϵ'' is an independent noise term with variance $\frac{1}{q_3} + \frac{1}{q_4+\gamma}$. From this we deduce that posterior variance about ω is

$$V(q_1, q_2, q_3, q_4) = 1 \Bigg/ \left(\alpha + \frac{1}{\frac{1}{q_1} + \frac{1}{q_2+\beta}} + \frac{1}{\frac{1}{q_3} + \frac{1}{q_4+\gamma}} \right).$$

The optimal division vector thus seeks to *maximize*

$$\frac{1}{\frac{1}{q_1} + \frac{1}{q_2+\beta}} + \frac{1}{\frac{1}{q_3} + \frac{1}{q_4+\gamma}} \tag{23}$$

It is useful to rewrite (23) in the following way:

$$\frac{1}{4} \left(q_1 + q_2 + \beta + q_3 + q_4 + \gamma - \frac{(q_1 - q_2 - \beta)^2}{q_1 + q_2 + \beta} - \frac{(q_3 - q_4 - \gamma)^2}{q_3 + q_4 + \gamma} \right).$$

Then, since $q_1 + q_2 + \beta + q_3 + q_4 + \gamma = t + \beta + \gamma$ is fixed at any time t, it is equivalent to choose q_1, q_2, q_3, q_4 to minimize the sum of ratios

$$\frac{(q_1 - q_2 - \beta)^2}{q_1 + q_2 + \beta} + \frac{(q_3 - q_4 - \gamma)^2}{q_3 + q_4 + \gamma}.$$

Ideally, if signals were perfectly divisible, the optimum would be to choose $q_1 = q_2 + \beta$ and $q_3 = q_4 + \gamma$. But as each q_i is restricted to integer values, this continuous optimum is not feasible whenever β and γ are not integers.

The solution to this integer optimization problem is involved, and we need some additional notation. Let r be the integer that minimizes $|r - \beta|$ (the distance to β) and let s be the integer that minimizes $|s - \gamma|$. Further, let $\langle\beta\rangle$ and $\langle\gamma\rangle$ be the value of these absolute differences.

Claim 1. *When the period t has the same parity as $r + s$, the t-optimal (q_1, q_2, q_3, q_4) satisfy*

$$q_1, q_2 \approx \frac{\langle\beta\rangle}{2\langle\beta\rangle + 2\langle\gamma\rangle} \cdot t; \quad q_3, q_4 \approx \frac{\langle\gamma\rangle}{2\langle\beta\rangle + 2\langle\gamma\rangle} \cdot t.$$

Otherwise the t-optimal (q_1, q_2, q_3, q_4) satisfy

$$q_1, q_2 \approx \frac{\langle\beta\rangle}{2\langle\beta\rangle + 2 - 2\langle\gamma\rangle} \cdot t; \quad q_3, q_4 \approx \frac{1 - \langle\gamma\rangle}{2\langle\beta\rangle + 2 - 2\langle\gamma\rangle} \cdot t.$$

Thus, all four signals are observed with positive frequency in the long run according to the optimal criterion.

Proof. To solve the integer maximization problem (23), let r be the integer that minimizes $|r - \beta|$ (the distance to β) and let s be the integer that minimizes $|s - \gamma|$. Further, let $\langle\beta\rangle$ and $\langle\gamma\rangle$ be the value of these absolute differences. We assume $2\beta, 2\gamma$ are not integers, so that $0 < \langle\beta\rangle, \langle\gamma\rangle < \frac{1}{2}$. We also assume $\langle\beta\rangle \neq \langle\gamma\rangle$, and by symmetry focus on the case of $\langle\beta\rangle < \langle\gamma\rangle$.

With these assumptions, it is clear that when q_1, q_2 are integers, the minimum value of $|q_1 - q_2 - \beta|$ is $\langle\beta\rangle$, achieved if and only if $q_1 = q_2 + r$. Similarly the minimum value of $|q_3 - q_4 - \gamma|$ is $\langle\gamma\rangle$, achieved when $q_3 = q_4 + s$. Now if the total number of observations t has the *same parity* as $r + s$, it is possible to choose q_1, q_2, q_3, q_4 such that their sum is t and $q_1 = q_2 + r$, $q_3 = q_4 + s$—any pair q_2, q_4 with sum $\frac{t-r-s}{2}$ leads to such a solution. Given these constraints, then, the optimum is to choose q_2, q_4 to minimize $\frac{\langle\beta\rangle^2}{2q_2+r+\beta} + \frac{\langle\gamma\rangle^2}{2q_4+s+\gamma}$. The optimal q_2 and q_4 satisfy $q_2/q_4 \approx \langle\beta\rangle/\langle\gamma\rangle$, which together with $q_2 + q_4 = \frac{t-r-s}{2}$ implies

$$q_1, q_2 \approx \frac{\langle\beta\rangle}{2\langle\beta\rangle + 2\langle\gamma\rangle} \cdot t; \quad q_3, q_4 \approx \frac{\langle\gamma\rangle}{2\langle\beta\rangle + 2\langle\gamma\rangle} \cdot t.$$

On the other hand, suppose t has the *opposite parity* to $r + s$. In this case $q_1 = q_2 + r$ and $q_3 = q_4 + s$ cannot both hold, thus $|q_1 - q_2 - \beta|$ and $|q_3 - q_4 - \gamma|$ cannot both take their minimum values $\langle\beta\rangle$ and $\langle\gamma\rangle$. It turns out that the best one can do is choose $q_1 = q_2 + r$ and

$q_3 = q_4 + s \pm 1$ so that $|q_1 - q_2 - \beta| = \langle\beta\rangle$ and $|q_3 - q_4 - \gamma| = 1 - \langle\gamma\rangle$. Then, the optimal choice of q_2, q_4 with sum $\frac{t-r-s\mp 1}{2}$ to minimize $\frac{\langle\beta\rangle^2}{2q_2+r+\beta} + \frac{(1-\langle\gamma\rangle)^2}{2q_4+s+\gamma\pm 1}$. This yields

$$q_1, q_2 \approx \frac{\langle\beta\rangle}{2\langle\beta\rangle + 2 - 2\langle\gamma\rangle} \cdot t; \quad q_3, q_4 \approx \frac{1-\langle\gamma\rangle}{2\langle\beta\rangle + 2 - 2\langle\gamma\rangle} \cdot t$$

as desired. □

The intuition for the conclusion above is simple: We would most prefer to set $q_1 = q_2 + \beta$ and $q_3 = q_4 + \gamma$, but this is not feasible when β and γ are not integers. Thus, there is inevitably some loss from the ideal case where signals are perfectly divisible. This loss turns out to be convex in signal counts, so both groups of signals are observed infinitely often to minimize total loss.

C.2 Strengthening of Theorem 2 part (b)

Here we show the following result, which strengthens the restated Theorem 2 part (b) (see Appendix A.4). It implies that under Assumption 3, any signal that is observed with zero long-run frequency must in fact be observed only finitely often.

Stronger Version of Theorem 2 part (b). *Suppose that the signals observed infinitely often span $\mathbb{R}^K$. Then $m_i(t) = \lambda_i^* \cdot t + O(1), \forall i$.*

The proof is divided into two subsections below.

C.2.1 A Weaker Result

Recall that we have previously shown $m_i(t) \sim \lambda_i^* \cdot t$. We can first improve the estimate of the residual term to $m_i(t) = \lambda_i^* \cdot t + O(\ln t)$. Indeed, Lemma 10 yields that for some constant L and every $t \geq L$,

$$V(m(t+1)) \leq V(m(t)) - \left(1 - \frac{L}{t}\right) \cdot \frac{V(m(t))^2}{\phi(\mathcal{S}^*)^2}. \tag{24}$$

This is because we may apply Lemma 10 with $M = \min_{j \in \mathcal{S}^*} m_j(t)$, which is at least $\frac{t}{L}$.

Let $g(t) = \frac{V(m(t))}{\phi(\mathcal{S}^*)^2}$. Then the above simplifies to

$$g(t+1) \leq g(t) - \left(1 - \frac{L}{t}\right) g(t)^2.$$

Inverting both sides, we have

$$\frac{1}{g(t+1)} \geq \frac{1}{g(t)} + \frac{1 - L/t}{1 - (1 - L/t)g(t)} \geq \frac{1}{g(t)} + 1 - \frac{L}{t}. \tag{25}$$

This enables us to deduce

$$\frac{1}{g(t)} \geq \frac{1}{g(L)} + \sum_{x=L}^{t-1} \left(1 - \frac{L}{x}\right) \geq t - O(\ln t).$$

Thus $g(t) \leq \frac{1}{t - O(\ln t)} \leq \frac{1}{t} + O(\frac{\ln t}{t^2})$. That is,

$$V(m(t)) \leq \frac{\phi(\mathcal{S}^*)^2}{t} + O\left(\frac{\ln t}{t^2}\right).$$

Since $t \cdot V(\lambda t)$ approaches $V^*(\lambda)$ at the rate of $\frac{1}{t}$, we have

$$V^*\left(\frac{m(t)}{t}\right) \leq t \cdot V(m(t)) + O\left(\frac{1}{t}\right) \leq \phi(\mathcal{S}^*)^2 + O\left(\frac{\ln t}{t}\right). \tag{26}$$

Suppose $\mathcal{S}^* = \{1, \dots, k\}$. Then the above estimate together with (12) implies $\sum_{j>k} \frac{m_j(t)}{t} = O(\frac{\ln t}{t})$. Hence $m_j(t) = O(\ln t)$ for each signal j outside of the best set.

Now we turn attention to those signals in the best set. If these were the only available signals, then the analysis in Liang, Mu and Syrgkanis (2017) gives $\partial_i V(m(t)) = -\left(\frac{\beta_i^{\mathcal{S}^*}}{m_i(t)}\right)^2$. In our current setting, signals $j > k$ affect this marginal value of signal i, but the influence is limited because $m_j(t) = O(\ln t)$. Specifically, we can show that

$$\partial_i V(m(t)) = -\left(\frac{\beta_i^{\mathcal{S}^*}}{m_i(t)}\right)^2 \cdot \left(1 + O\left(\frac{\ln t}{t}\right)\right).$$

This then implies $m_i(t) \leq \lambda_i^* \cdot t + O(\ln t)$.[40] Using $\sum_{i \leq k} m_i(t) = t - O(\ln t)$, we deduce that $m_i(t) \geq \lambda_i^* \cdot t - O(\ln t)$ must also hold. Hence $m_i(t) = \lambda_i^* \cdot t + O(\ln t)$ for each signal i.

C.2.2 Getting Rid of the Log

In order to remove the $\ln t$ residual term, we need a refined analysis. The reason we ended up with $\ln t$ is because we used (24) and (25) at *each* period t; the "$\frac{L}{t}$" term in those equations adds up to $\ln t$. In what follows, instead of quantifying the variance reduction in each period (as we did), we will lower-bound the variance reduction over multiple periods. This will lead to better estimates and enable us to prove $m_i(t) = \lambda_i^* \cdot t + O(1)$.

To give more detail, let $t_1 < t_2 < \dots$ denote the periods in which some signal $j > k$ is chosen. Since $m_j(t) = O(\ln t)$ for each such signal j, $t_l \geq 2^{\epsilon \cdot l}$ holds for some positive constant ϵ and each positive integer l. Continuing to let $g(t) = \frac{V(m(t))}{\phi(\mathcal{S}^*)^2}$, our goal is to estimate the difference between $\frac{1}{g(t_{l+1})}$ and $\frac{1}{g(t_l)}$.

[40] Otherwise, consider $\tau + 1 \leq t$ to be the last period in which signal i was observed. Then $m_i(\tau)$ is larger than $\lambda_i^* \cdot \tau$ by several $\ln(\tau)$, while there exists some other signal $\hat{i}$ in the best set with $m_{\hat{i}}(\tau) < \lambda_i^* \cdot \tau$. But then $|\partial_i V(m(\tau))| < |\partial_{\hat{i}} V(m(\tau))|$, meaning that the agent in period $\tau + 1$ should not have chosen signal i.

Ignoring period t_{l+1} for the moment, we are interested in $\frac{\phi(\mathcal{S}^*)^2}{V(m(t_{l+1}-1))} - \frac{\phi(\mathcal{S}^*)^2}{V(m(t_l))}$, which is just the difference in the *precision* about ω when the division vector changes from $m(t_l)$ to $m(t_{l+1}-1)$. From the proof of Lemma 9, we can estimate this difference if the change were along the direction λ^*:

$$\frac{\phi(\mathcal{S}^*)^2}{V(m(t_l)+\lambda^*(t_{l+1}-1-t_l))} - \frac{\phi(\mathcal{S}^*)^2}{V(m(t_l))} \geq t_{l+1}-1-t_l. \tag{27}$$

Now, the vector $m(t_{l+1}-1)$ is not exactly equal to $m(t_l)+\lambda^*(t_{l+1}-1-t_l)$, so the above estimate is not directly applicable. However, by our definition of t_l and t_{l+1}, any difference between these vectors must be in the first k signals. In addition, the difference is bounded by $O(\ln t_{l+1})$ by what we have shown. This implies[41]

$$V(m(t_{l+1}-1)) - V(m(t_l)+\lambda^*(t_{l+1}-1-t_l)) = O\left(\frac{\ln^2 t_{l+1}}{t_{l+1}^3}\right).$$

Since $V(m(t_{l+1}-1))$ is on the oder of $\frac{1}{t_{l+1}}$, we thus have (if the constant L is large)

$$\frac{\phi(\mathcal{S}^*)^2}{V(m(t_{l+1}-1))} - \frac{\phi(\mathcal{S}^*)^2}{V(m(t_l)+\lambda^*(t_{l+1}-1-t_l))} \geq -\frac{L\ln^2 t_{l+1}}{t_{l+1}}. \tag{28}$$

(27) and (28) together imply

$$\frac{1}{g(t_{l+1}-1)} \geq \frac{1}{g(t_l)} + (t_{l+1}-1-t_l) - \frac{L\ln^2 t_{l+1}}{t_{l+1}}.$$

Finally, we can apply (25) to $t = t_{l+1}-1$. Altogether we deduce

$$\frac{1}{g(t_{l+1})} \geq \frac{1}{g(t_l)} + (t_{l+1}-t_l) - \frac{2L\ln^2 t_{l+1}}{t_{l+1}}.$$

Now observe that $\sum_l \frac{2L\ln^2 t_{l+1}}{t_{l+1}}$ converges (this is the sense in which our estimates here improve upon (25), where $\frac{L}{t}$ leads to a divergent sum). Thus we are able to conclude

$$\frac{1}{g(t_l)} \geq t_l - O(1), \quad \forall l.$$

In fact, this holds also at periods $t \neq t_l$. Therefore $V(m(t)) \leq \frac{\phi(\mathcal{S}^*)^2}{t} + O(\frac{1}{t^2})$, and

$$V^*\left(\frac{m(t)}{t}\right) \leq t\cdot V(m(t)) + O\left(\frac{1}{t}\right) \leq \phi\left(\mathcal{S}^*\right)^2 + O\left(\frac{1}{t}\right). \tag{29}$$

This equation (29) improves upon the previously-derived (26). Hence by (12) again, $m_j(t) = O(1)$ for each signal $j > k$. And once these signal counts are fixed, Proposition 1 implies $m_i(t) = \lambda_i^* \cdot t + O(1)$ also holds for signal $i \leq k$. This completes the proof.

[41] By the mean-value theorem, the difference can be written as $O(\ln t_{l+1})$ multiplied by a certain directional derivative. Since the coordinates of $m(t_{l+1}-1)$ and of $m(t_l)+\lambda^*(t_{l+1}-1-t_l)$ both sum to $t_{l+1}-1$, this directional derivative has a direction vector whose coordinates sum to zero. Combined with $\partial_i V(m(t)) = -(\frac{\phi(\mathcal{S}^*)^2}{t})\cdot(1+O(\frac{\ln t}{t}))$ (which we showed before), this directional derivative has size $O(\frac{\ln t}{t^3})$.

C.3 Example of a Learning Trap with Non-Normal Signals

The payoff-relevant state $\theta \in \{\theta_1, \theta_2\}$ is binary and agents have a uniform prior. There are three available information sources. The first, X_1, is described by the information structure

	θ_1	θ_2
s_1	p	$1-p$
s_2	$1-p$	p

with $p > 1/2$. Information sources 2 and 3 provide perfectly correlated signals (conditional on θ) taking values in $\{a, b\}$: In state θ_1, there is an equal probability that $X_2 = a$ and $X_3 = b$ or $X_2 = b$ and $X_3 = a$. In state θ_2, there is an equal probability that $X_2 = X_3 = a$ and $X_2 = X_3 = b$.

In this environment, every agent chooses to acquire the noisy signal X_1, even though one observation of each of X_2 and X_3 would perfectly reveal the state.[42]

C.4 Example Mentioned in Section 6.1

Suppose the available signals are

$$\begin{aligned} X_1 &= 10x + \epsilon_1 \\ X_2 &= 10y + \epsilon_2 \\ X_3 &= 4x + 5y + 10b \\ X_4 &= 8x + 6y - 20b \end{aligned}$$

where $\omega = x + y$ and b is a payoff-irrelevant unknown. Set the prior to be

$$(x, y, b)' \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.039 \end{pmatrix} \right).$$

It can be computed that agents observe only the signals X_1 and X_2, although the set $\{X_3, X_4\}$ is optimal with $\phi(\{X_1, X_2\}) = 1/5 > 3/16 = \phi(\{X_3, X_4\})$. Thus, the set $\{X_1, X_2\}$ constitutes a learning trap for this problem. But if each signal choice were to produce ten independent realizations, agents starting from the above prior would observe only the signals X_3 and X_4. This breaks the learning trap.

[42]We thank Andrew Postlewaite for this example.

C.5 Supplementary Material to Section 7

C.5.1 General Payoff Functions

We comment here on the possibilities for (and limitations to) generalizing Theorem 1 beyond the quadratic loss payoff function. As discussed in the main text, Part (a) of the theorem extends to general payoff functions. On the other hand, part (b) extends to *some* other "prediction problems," in which every agent's payoff function $u(a, \omega)$ is the same and depends only on $|a - \omega|$. For example, the proposition holds for $u(a, \omega) = |a - \omega|^\gamma$ with any exponent $\gamma \in (0, 2]$; such an extension only requires minor changes to our proof.

Nonetheless, even restricting to prediction problems, part (b) of Theorem 1 does *not* hold in general. For a counterexample, consider $u(a, \omega) = -\mathbf{1}_{\{|a-\omega|>1\}}$, which punishes the agent for any prediction that differs from the true state by more than 1.[43] Intuitively, the payoff gain from further information decreases sharply (indeed, exponentially) with the amount of information that has already been acquired. Thus, even with a forward-looking objective function, the range of future payoffs is limited and each agent cares mostly to maximize his own payoff. This results in an optimal sampling rule that resembles myopic behavior (and differs from the rule that would maximize speed of learning).

The above counterexample illustrates the difficulty in estimating the value of information when working with an arbitrary payoff function. In order to make intertemporal payoff comparisons, we need to know how much payoff is gained/lost when the posterior variance is decreased/increased by a certain amount. This can be challenging in general, see Chade and Schlee (2002) for a related discussion.[44]

Finally, while it is more than necessary to assume that agents have the same payoff function, the truth of part (b) of Theorem 1 does require some restrictions on how the payoff functions differ. Otherwise, suppose for example that payoffs take the form $-\alpha_t(a_t - \omega)^2$, where α_t decreases exponentially fast. Then even with the δ-discounted objective, the social planner puts most of the weight on earlier agents, resembling the myopic behavior of individual agents.

C.5.2 Low Altruism

Here we argue that part (a) of Theorem 2 generalizes to agents who are not completely myopic, but are sufficiently impatient. That is, we will show that if signals $1, \ldots, k$ are subspace-optimal, then there exist priors given which agents with low δ always observe these signals in equilibrium.

[43] We thank Alex Wolitzky for this example.

[44] This difficulty becomes more salient if we try to go beyond prediction problems: The value of information in that case will depend on signal realizations.

We follow the construction in Appendix A.3. The added difficulty here is to show that if any agent ever chooses a signal $j > k$, the payoff loss in that period (relative to myopically choosing among the first k signals) is of the same magnitude as possible payoff gains in future periods. Then, for sufficiently small δ, such a deviation is not profitable.

Suppose that agents sample only from the first k signals in the first $t-1$ periods, with frequencies close to λ^*. Then, the posterior variances $v_1, \ldots, v_k$ (which are also the prior for period t) are on the order of $\frac{1}{t}$. Thus any signal acquisition in period t leads to a variance reduction on the order of $\frac{1}{t^2}$. However, using the computation in Appendix A.3, we can show that for some positive constant ξ (independent of t), the variance reduction of any signal $j > k$ is smaller than the variance reduction of any of the first k signals by $\frac{\xi}{t^2}$. This is the amount of payoff loss in period t under a deviation to signal j.

Of course, this deviation could improve the posterior variance in future periods. But even for the best continuation strategy, the posterior variance in period $t+m$ can at most be reduced by $O(\frac{m}{t^2})$.[45] Thus if we choose ξ to be small enough, the payoff gain in period $t+m$ is bounded above by $\frac{m}{\xi t^2}$. Note that for δ sufficiently small,

$$-\frac{\xi}{t^2} + \sum_{m \geq 1} \delta^m \cdot \frac{m}{\xi t^2} < 0.$$

Hence the deviation is not profitable and the proof is complete.

C.5.3 Multiple Payoff-Relevant States

Let $V(q_1, \ldots, q_N)$ be a weighted sum of posterior variances about the r payoff-relevant states. As before, define V^* to be a normalized, asymptotic version of V. Let $q(t)$ continue to represent any allocation of t observations that minimizes V. Then, under a modification of the Unique Minimizer assumption—we require V^* to be *uniquely* minimized—the optimal frequency vector $\lambda^{OPT} := \lim_{t\to\infty} q(t)/t$ is well-defined. Nonetheless, we emphasize that with $r > 1$, these optimal allocations generally involve more than K signals. A theorem of Chaloner (1984) shows that λ^{OPT} is supported on at most $\frac{r(2K+1-r)}{2}$ signals.

We can generalize the notion of "minimal spanning sets" as follows: A set of signals $\mathcal{S}$ is minimally-spanning if optimal sampling from $\mathcal{S}$ puts positive frequency on *every* signal in $\mathcal{S}$. When $r = 1$, this definition agrees with the definition in our main model.

Similarly, we say that a minimal spanning set $\mathcal{S}$ is "subspace-optimal" if, when agents are constrained to choose from $\overline{\mathcal{S}}$, the optimal frequency vector is supported on $\mathcal{S}$. With these definitions, Theorem 2 and its proof generalize without change.

[45]This is because over m periods, the increase in the *precision* matrix is $O(m)$.