Dynamic Information Acquisition from Multiple Sources*

Annie Liang† Xiaosheng Mu‡ Vasilis Syrgkanis§

August 14, 2017

Abstract

Consider a decision-maker who dynamically acquires Gaussian signals that are related by a completely flexible correlation structure. Such a setting describes information acquisition from news sources with *correlated biases*, as well as aggregation of *complementary information* from specialized sources. We study the optimal sequence of information acquisitions. Generically, myopic signal acquisitions turn out to be optimal at sufficiently late periods, and in classes of informational environments that we describe, they are optimal from period 1. These results hold independently of the decision problem and its (endogenous or exogenous) timing. We apply these results to characterize dynamic information acquisition in games.

Decision-makers often lack access to information that is directly revealing about what they care about; instead, they aggregate information across many sources, each of which provides relevant information. Consider for example an individual deciding whether to purchase a home. His realtor provides information about the quality of the house, but this information may be biased. Online reviews of the realtor provide a second source of information, which help the decision-maker learn how much the realtor inflates on average. These reviews also require interpretation, and the decision-maker can cross-check reviews of this realtor against reviews of *other* realtors.

This example, already involved, is simplified relative to what information acquisition frequently looks like in practice: individuals routinely acquire and aggregate information across

*We would like to thank Yash Deshpande, Drew Fudenberg, Johannes Hörner, Elliot Lipnowski, Michihiro Kandori, Eric Maskin, Tomasz Strzalecki, Juuso Toikka, Anton Tsoy, and Muhamet Yildiz, as well as seminar audiences at Columbia University and the Pennsylvania Economic Theory Conference (2017).

†Department of Economics, University of Pennsylvania

‡Department of Economics, Harvard University

§Microsoft Research

varied sources in order to learn about unknowns such as the payoff to a new policy, the value of an asset, the inflation rate next year. When sources provide correlated information, the value of information from a given source depends on what kinds of information are available from other sources, and also on what information has already been acquired. Thus, correlation across sources complicates the problem of optimal dynamic acquisition of information: a decision-maker contemplating which information source to acquire today should take into account its impact on the value of information collected in the future.

We model the problem of dynamic acquisition of information from correlated information sources as follows: there are K unknown states $\theta_1, \ldots, \theta_K \in \mathbb{R}$, which follow a multivariate Gaussian distribution. A Bayesian decision-maker (DM) has access to K different signals, each of which is a linear combination of the K unknown states and an independent Gaussian noise term. Only the state θ_1 is payoff-relevant, while other states represent *unknown parameters* of the signal generating distributions or *correlated biases* of the information sources. Information acquisition requires physical time and effort, which we model as a capacity constraint: the DM chooses one signal to observe in each period (see Section 6 for an extension to $B > 1$ signals). At a random final period t, the DM faces a decision problem, in which he takes an action $a \in A$ and receives a payoff that depends on his action a, the state θ_1 and possibly the date t. The time of the final period is determined according to a full-support distribution over periods[1]—in the main text, we suppose that this distribution is exogenous, but later demonstrate an extension to endogenous stopping.

We study the following rules for information acquisition:

(a) *myopic*: in each period, the DM chooses the signal that (combined with past acquisitions) maximizes the expected payoff he would receive if he were to immediately face the decision problem.

(b) *dynamically optimal*: in each period, the DM chooses the signal that maximizes his overall expected payoff, taking into account the randomness of the decision period and of future signal realizations.

Additionally, we say that a dynamic information acquisition strategy is *totally optimal at time t* if the DM would follow this strategy given knowledge that period t is final.

The main contribution of our paper is to show that in this problem, the myopic and dynamically optimal solutions are (generically) equivalent after finitely many periods, and in special cases they are *immediately* equivalent. Moreover, these solutions achieve total optimality at every (large) time t. Thus, despite correlation across information sources, the myopic choice is always (eventually) the best signal to acquire. This equivalence holds for

[1]A familiar special case is one in which each period (conditional on being reached) is final with a constant probability $1 - \delta$, so that δ is the DM's discount factor.

all payoff functions that the decision-maker might face, and for all possible timings of the decision period.

Towards these results, we first demonstrate "invariance" of the solution to the decision problem. Specifically, we show that the myopically optimal strategy, as well as total optimality at period t, can be determined independently of the decision problem. We show this by first presenting the decision problem of *prediction*: the DM chooses an action a to match the state θ_1, receiving payoff $-(a-\theta_1)^2$. For this problem, the myopic decision-maker deterministically acquires in each period the signal that achieves the greatest decrease in posterior variance (since his expected payoff is simply the negative of his posterior variance). We show that the signal that satisfies this criterion *Blackwell-dominates* the remaining signals, so that it is best not only for the prediction of θ_1, but for all decision problems. A similar argument obtains for total optimality, and a more complex version of this result for dynamic optimality is presented in the appendix.

This reduction means that we can suppose without loss of generality that the decision problem is prediction, and work with deterministic strategies that do not condition on signal realizations. We provide in Section 5 sufficient conditions under which the myopic and dynamically optimal signal paths are identical *from period* 1, and moreover totally optimal at every period t. In these environments, the myopic decision-maker and the forward-looking decision-maker acquire the same signals in the same ("best") order. Intuitively, our sufficient conditions align the information acquisition goal of the current period with those of all future periods: under these conditions, the signal that achieves the greatest decrease in posterior variance in any given period turns out to also allow for maximal reduction of uncertainty in all subsequent periods.

Equivalence between the myopic and forward-looking solutions does not hold for all informational environments. In Section 6, we provide a simple counterexample, showing that the signal that is most informative in the current period may not be part of any *pair* of signals that maximizes learning across two periods. This illustrates how complementarity across signals (due to correlation) can render the myopic choice sub-optimal. Nevertheless, we show that in general, the information acquisition problem becomes "approximately separable" over time: the strength of complementarity across different signals vanishes an order of magnitude faster than the strength of substitution between realizations of the same signal. With sufficiently many observations, the signals become approximately independent.

This insight allows us to derive the main results of the paper. We show that the myopic and dynamically optimal signal paths are eventually approximately the same, and approximately totally optimal: the number of signals of each type acquired under the myopic, dynamically optimal and totally optimal criteria will eventually differ by *no more than one* from each another. Moreover, this "eventual gap of one" vanishes in generic informational environments, so that at sufficiently late periods the myopic path is identical to the dynamically optimal path, and is totally optimal.[2]

As discussed above, all of our results hold independently of discounting and of the decision problem. Specifically, our sufficient conditions for the myopic signal path to be immediately optimal are stated only in terms of the informational environment: the DM's prior belief, the signals' linear coefficients and the signal variances. More generally, we demonstrate a time T for each informational environment, such that the myopic, dynamically optimal and totally optimal signal paths differ by at most one (in each signal count) after T periods. In particular, when we consider geometric arrival of the final period with parameter $1-\delta$, the time it takes to achieve approximate equivalence remains bounded as the "discount factor" δ approaches 1.

In Section 7, we show that our results extend to *intertemporal decision problems*, where the DM takes an action at each period and receives an arbitrary state-dependent payoff depending on all of his actions. This more general framework covers endogenous stopping problems, which have been extensively studied in the previous literature, and it also allows for further applications such as dynamic investment and pricing. We note that our results extend to characterize the optimal sequence of signal choices in these richer environments, but we do not characterize the optimal sequence of actions (or stopping time).

Additional extensions of the model are presented in Section 8: our results extend to i.i.d. states drawn each period, to multiple payoff-relevant states whenever the decision problem is that of prediction, and to the possibility of "free signals" that need not be acquired. We also discuss a continuous-time analogue of our setup in which at every time, the DM chooses attention levels (subject to a capacity constraint) that influence the signals he observes, in the form of diffusion processes. In a detailed appendix, we extend and strengthen many of our previous results to this setting. For instance, we show that eventual exact equivalence holds always (instead of generically), and that immediate exact equivalence occurs under a milder condition of almost independence of prior beliefs.

Sections 9 and 10 discuss interpretations and applications of our main results. For example, the robustness of our results to the decision problem suggests that a decision-maker who faces *uncertainty* or *ambiguity* over the final decision problem can act in a way that is (in special cases, immediately, and generically, eventually) best across *all* such problems. Our setting is thus one in which robust information acquisition is both possible and simple enough for decision-makers to use in practice. In another interpretation, we consider information acquisition by a sequence of decision-makers—each acquiring (public) information to

[2]The sense of *generic* is the following: for fixed prior belief and linear combinations defining the signals, eventual equivalence holds for generic signal variances. Likewise, for fixed prior and signal variances, the result holds for generic linear coefficients.

maximize a private objective. Here, our results imply that a social planner cannot improve on the amount of information aggregated by a long sequence of myopic decision-makers.

Finally, our immediate equivalence results reveal classes of environments in which the optimal signal path can be characterized independently of the payoff function. This makes it tractable to analyze dynamic information acquisition not only in single-agent decision problems, but also in strategic settings. Our main application in Section 10 leverages this to extend two results in the literature for information acquisition in games (see Hellwig and Veldkamp (2009) and Lambert, Ostrovsky and Panov (2017)) by allowing the players to acquire information over time.

Our work most directly builds on a literature about optimal information acquisition in dynamic environments.[3] The key new feature in our problem is the choice between information sources that are related by a flexible correlation structure. This emphasis on how a DM chooses between different kinds of information appears in recent work by Fudenberg, Strack and Strzalecki (2017) (section 3.5), who consider choice between two Gaussian signals, as well as in Che and Mierendorff (2017) and Mayskaya (2017), who consider choice between two Poisson signals. Relative to these papers, our model is distinguished by permitting many Gaussian signals that exhibit arbitrary correlation. See Section 11 for extended discussion of the related literature.

1 Benchmark Case: Learning the Bias of a Source

Consider a decision-maker who wants to learn an unknown state of the world $\theta \sim \mathcal{N}(0,1)$. This state is realized at $t = 0$ and persists across all subsequent periods. The DM has access to two signals: first, he can observe realizations of a *biased signal* $X^t = \theta + b + \epsilon_X^t$, where $b \sim \mathcal{N}(0,1)$ is an unknown persistent bias and $\epsilon_X^t \sim \mathcal{N}(0, \sigma_X^2)$ is an independent noise term.

Second, he can observe a *signal about the bias*, $B^t = b + \epsilon_B^t$, where $\epsilon_B^t \sim \mathcal{N}(0, \sigma_B^2)$ is an independent noise term. The noise terms ϵ_X^t and ϵ_B^t are i.i.d. over time. To save on notation, we suppress the time indices on signals throughout, referring to them simply as X and B. Notice that although signal B does not directly contain any information about the state θ, it helps the decision-maker to interpret realizations of signal X.

Time $t = 1, 2, \ldots$ is discrete, and each period (conditional on being reached) is final with probability $1 - \delta > 0$. Below, we refer to δ as the DM's discount factor. In each period up to and including the final period, the DM chooses to observe a realization of either signal X or signal B. At the final period, he provides a prediction a for the unknown state, and

[3] For example: Wald (1947), Arrow, Blackwell and Girshick (1949), Moscarini and Smith (2001), Steiner, Stewart and Matejka (2017), Ke, Shen and Villas-Boas (2016), Fudenberg, Strack and Strzalecki (2017), Che and Mierendorff (2017), Hebert and Woodford (2017), and Mayskaya (2017).

receives the payoff $-(a-\theta)^2$. We assume that past signal realizations are known at the start of every period, so that in the final period, the DM bases his prediction on all the signals acquired so far. Which signal should the DM choose to observe in each period?

Let us first consider the choices of the myopic decision-maker, who acquires information as if he were to face the prediction problem in the current period (corresponding to $\delta = 0$). When asked to predict the state, the DM's expected payoff is maximized by predicting the posterior mean of θ, and his payoff equals the negative of his posterior variance. Because the DM's prior and the available signals are Gaussian, his posterior belief about θ is also Gaussian. Crucially, the DM's posterior variance can be expressed as the following function of q_X, the number of times he has observed signal X, and q_B, the number of times he has observed signal B:

$$f(q_X, q_B) := 1 - 1 \Bigg/ \left(1 + \frac{\sigma_B^2}{\sigma_B^2 + q_B} + \frac{\sigma_X^2}{q_X}\right). \tag{1}$$

To derive this posterior variance function, we (without loss) re-order the signal acquisitions so that the q_B realizations of signal B are observed first. Following these observations, the DM's posterior belief about b has variance $1/\left(1 + \frac{q_B}{\sigma_B^2}\right)$. Let $\overline{X}$ be the random variable that is the sample mean of q_X realizations of signal X. Then, the DM's belief over $(\theta, \overline{X})$, conditional on the first q_B observations of B, is jointly Gaussian with covariance matrix[4]

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 + \frac{1}{1+\frac{q_B}{\sigma_B^2}} + \frac{1}{\frac{q_X}{\sigma_X^2}} \end{pmatrix}$$

and (1) follows from the standard formula for Gaussian conditional variance.

Given any history of observations summarized by the pair (q_X, q_B), the myopic decision-maker will choose to observe signal X if and only if $f(q_X+1, q_B) < f(q_X, q_B+1)$. Using (1), this is equivalent to the condition that

$$(\sigma_B^2 + q_B)(1 + \sigma_B^2 + q_B)\sigma_X^2 > \sigma_B^2 q_X (1 + q_X). \tag{2}$$

Thus, on the myopic signal path, the DM alternates between observing strings of X's and strings of B's. From (2) it can be further shown that over many periods, the number of X-acquisitions divided by the number of B-acquisitions converges to the ratio σ_X/σ_B.[5,6]

[4] Observe that prior to observing any realizations of signal X, the DM believes θ and b to be independent.

[5] In the special case that σ_X and σ_B are positive integers, the myopic signal path is eventually *periodic*, and the limiting ratio σ_X/σ_B is fulfilled with the shortest period possible. For example, if $\sigma_X = \sigma_B$, then after sufficiently many periods, the DM will observe $XBXB\cdots$ and not $XXBBXXBB$. Formally, let d denote the greatest common divisor of σ_X and σ_B. Then the period length is $(\sigma_X + \sigma_B)/d$.

[6] Notice that the smaller the variance σ_X^2 is, the *less* often signal X is observed (asymptotically) relative to signal B, and vice versa. This is a general feature of the solution.

The signal acquisition path described above turns out to be not only myopically optimal, but also dynamically optimal for any discount factor δ. Lemma 1 below will be key to showing this equivalence, and it says the following: fix an arbitrary period $\underline{t}$ and a signal path $h = (s_1, s_2, \dots) \in \{X, B\}^\infty$, where the sequence follows the myopic strategy in (2) beginning in period $\underline{t}$. Suppose the DM deviates at that period to some other signal and subsequently follows the myopic strategy. Call the deviation path $\tilde{h} = (\tilde{s}_1, \tilde{s}_2, \dots)$. The lemma below states that the DM's posterior variance is smaller *at every period* along signal path h than along $\tilde{h}$.

Lemma 1. $\mathrm{Var}(\theta \mid h^t) \leq \mathrm{Var}(\theta \mid \tilde{h}^t)$ *holds at every* t.

Proof. We suppose $s_{\underline{t}} = X$, so that the deviation is to $\tilde{s}_{\underline{t}} = B$; the other case where $s_{\underline{t}} = B$ follows along identical arguments. Write $\bar{t}$ for the first period after $\underline{t}$ at which $s_{\bar{t}} = B$. Observe that if (2) holds at some history (q_X, q_B), then it continues to hold for all larger q_B. This means that the incentive to choose X at any period $t \in (\underline{t}, \bar{t}]$ along signal path $\tilde{h}$ is greater than the incentive to play X at period $t-1$ along the path h. It follows that

$$\begin{aligned}(s_{\underline{t}}, \dots, s_{\bar{t}}) &= XXX \cdots XB \\ (\tilde{s}_{\underline{t}}, \dots, \tilde{s}_{\bar{t}}) &= BXX \cdots XX\end{aligned} \tag{3}$$

After $\bar{t}$ periods, the two signal paths coincide in the number of X signals and B signals that have been acquired so far. Under myopic behavior, the same signal is acquired along either path at every period $t > \bar{t}$.

Thus it is clear that $\mathrm{Var}(\theta \mid h^t) = \mathrm{Var}(\theta \mid \tilde{h}^t)$ at every $t < \underline{t}$ and $t \geq \bar{t}$.[7] Now consider any period t with $\underline{t} \leq t < \bar{t}$. Then,

$$\begin{aligned}\mathrm{Var}(\theta|\tilde{h}^t) &= \mathrm{Var}(\theta|h^{\underline{t}-1}BX\cdots XX) \\ &= \mathrm{Var}(\theta|h^{\underline{t}-1}XX\cdots XB) \\ &\geq \mathrm{Var}(\theta|h^{\underline{t}-1}XX\cdots XX) = \mathrm{Var}(\theta \mid h^t),\end{aligned}$$

using exchangeability of signals in the second equality, and myopic optimality along signal path h in the final inequality. This completes the argument. □

From this lemma and the one-shot deviation principle, it follows that the myopic strategy is also dynamically optimal. Moreover, since every history is reachable by a sequence of one-shot deviations from the myopic signal path, repeated application of Lemma 1 yields that the myopic strategy achieves the lowest posterior variance at every time among all possible strategies. As mentioned previously, we call this stronger property *total optimality*.

[7] The history h^t includes all signals acquired up to and *including* period t.

The subsequent section defines a general class of information acquisition problems which takes this example as a special case. We move beyond this example in two main directions: first, we allow for more than two sources that are arbitrarily correlated, and second, we allow for arbitrary payoff functions.

2 General Model

2.1 Setup

The benchmark model considered in Section 1 can be seen as a special case of the following model. There are K persistent states $\theta_1, \dots, \theta_K \sim \mathcal{N}(\mu^0, V^0)$, where $\mu^0 \in \mathbb{R}^K$ denotes the vector of prior means and V^0 is a $K \times K$ positive-definite prior covariance matrix.

The DM has access to K different signals,[8] each of which is a linear combination of the unknown states and a Gaussian noise term

$$X_i^t = \langle c_i, \theta \rangle + \epsilon_i^t, \quad \epsilon_i^t \sim \mathcal{N}\left(0, \sigma_i^2\right).$$

where each $c_i = (c_{i1}, \dots, c_{iK})'$ is a constant $K \times 1$ vector and $\theta = (\theta_1, \dots, \theta_K)'$ is the vector of unknown states.[9] Throughout, let C be the matrix of coefficients whose i-th row is c_i'.

Time $t = 1, 2, \dots$ is discrete, and in each period the DM chooses one of the K signals to observe. At some unknown final period t, he will face a decision problem, in which he chooses an action a from a set A and receives payoff $u_t(a, \theta)$. Each u_t is an arbitrary state-dependent and time-dependent utility function. The time of decision is governed (exogenously) by an arbitrary full-support distribution. Special cases include *geometric discounting*, in which every period (conditional on being reached) has a constant probability of being final, as well as *Poisson arrival* of the final period.

We assume in the main model that there is a single payoff-relevant state.

Assumption 1 (Single Payoff-Relevant State)**.** *At every period t,*

$$u_t(a, \theta_1, \theta_{-1}) = u_t(a, \theta_1) \quad \textit{does not depend on } \theta_{-1}.$$

Section 8 discusses the case of multiple payoff-relevant states.

We also assume that the decision problem is non-trivial in the following way.

[8] Our analysis directly extends to situations where the number of signals is less than the number of states, but the case where there are more signals than states presents additional challenges, since the DM need not observe all the signals in order to learn (the payoff-relevant state θ_1). The question of how many (and which) signals are acquired in that case is addressed in a follow-up paper.

[9] Here and later, we exclusively use the apostrophe to denote vector or matrix transpose.

Assumption 2 (Payoff Sensitivity to Mean)**.** *For any period t, any variance $\sigma^2 > 0$ and any action $a^* \in A$, there exists a positive measure of μ_1 for which a^* does not maximize $\mathbb{E}[u_t(a, \theta_1) \mid \theta_1 \sim \mathcal{N}(\mu_1, \sigma^2)]$.*

In words, holding fixed the DM's belief variance, his expected value of θ_1 affects the optimal action to take at time t.

A sufficient condition for Assumption 2 is that for every t and every a^*, there exists some other action $\hat{a}$ such that $u_t(\hat{a}, \theta_1) > u_t(a^*, \theta_1)$ as $\theta_1 \to +\infty$ or as $\theta_1 \to -\infty$. That is, we require that the two limiting states $\theta_1 \to +\infty$ and $\theta_1 \to -\infty$ disagree about the optimal action. This is true for all natural applications of the model.

Other than these assumptions we have made, our results are robust to the specifics of the decision problem. In Section 7, we show that our main results generalize to *endogenous stopping*, where the DM chooses an optimal time to stop acquiring information,[10] and *intertemporal decision problems*, in which the DM both acquires a signal and also takes an action a_t in each period. For expositional clarity, we work with the simpler payoff function $u_t(a, \theta)$ with exogenous stopping in the main model.

Finally, we impose a mild assumption on the informational environment.

Assumption 3 (Full Rank and Exact Identifiability)**.** *The matrix C has full rank, and no proper subset of row vectors of C spans the coordinate vector e_1'.*

This assumption requires that no subset of signals fully reveals the payoff-relevant state. Heuristically, the DM has to observe each signal infinitely often to recover the value of θ_1.[11]

We discuss below several interpretations for this model of information acquisition.

2.2 Interpretations

Learning from signals with correlated biases. The first interpretation takes the payoff-irrelevant states $\theta_2, \ldots, \theta_K$ to be unknown biases, so that the problem is one of dynamically learning from *biased news sources* or *biased experts*. The coefficient matrix C allows for general correlation structures across the biases of the sources; thus, the decision-maker can learn about the bias of one source by observing signal realizations from another. We assume that the DM knows the correlation structure, so that he can use observations across the sources to de-bias his beliefs.

[10] Whether the time of decision is exogenous or endogenous depends on the specific setting. For example, a politician acquiring information about a policy may want to best cast her vote at a future meeting, but her objective could also be to propose an alternative policy whenever she feels sufficiently informed. Under the assumptions of our model, however, either timing yields (approximately) the same optimal signal path.

[11] Equivalently, we assume that the inverse matrix C^{-1} exists, and its first row consists of non-zero entries.

Learning a composite of unknowns. A second interpretation takes the payoff-relevant state θ_1 to be a linear combination of unknowns $\tilde{\theta}_1, \ldots, \tilde{\theta}_K$ about which the decision-maker can learn independently.[12] Such a structure emerges in a variety of settings: for example, the DM may care about the value of a conglomerate that consists of several companies, where each company i is valued at $\tilde{\theta}_i$. The DM wants to learn $\theta_1 := \tilde{\theta}_1 + \cdots + \tilde{\theta}_K$ but has access to information about each company i's value $\tilde{\theta}_i$ separately.

As another example, suppose a political group wants to learn the average perspective in the population towards an issue. There are K demographics, where the proportion of the population in the k-th demographic is p_k. The distribution of perspectives in the k-th demographic is normal with unknown mean μ_k and known variance σ_k^2, so that the average perspective is $\theta_1 := \sum_k p_k \mu_k$. The group can learn about θ_1 by sampling individuals from the population, but it is not feasible to sample individuals directly according to p. The available polling technologies are modeled as K distributions $q^{(1)}, q^{(2)}, \ldots, q^{(K)} \in \Delta^{K-1}$ over the demographics, and individuals can be sampled from any of these distributions.

Precision-accuracy tradeoff. In a final interpretation, suppose that the decision-maker observes signals from some distribution, but does not know certain parameters governing this distribution. For example, a scientist needs to measure the acidity of a substance using a potentially biased instrument. There are two sources of error: *measurement error* due to natural (idiosyncratic) fluctuations in the environment, and *systematic error* due to the (persistent) bias of the instrument he uses. He can increase the precision of his estimate by repeatedly measuring the substance, or he can make his estimates more accurate by learning about the bias of the instrument (for example, by testing the instrument on a substance with known acidity). The scientist's information acquisition problem can be abstracted into the model considered in Section 1.

3 Notation and Definitions

3.1 Strategies, Payoffs, and Beliefs

Let $[K] = \{1, 2, \ldots, K\}$ denote the set of all signals. At the beginning of any period t, the DM faces a *history* $h^{t-1} \in ([K] \times \mathbb{R})^{t-1} = H^{t-1}$ consisting of his previous signal choices as well as the realized signal values. A *strategy* is a measurable map from all finite histories to signals—that is, $S : \cup_{t \geq 1} H^{t-1} \to [K]$, where $S(h^{t-1})$ represents the signal choice in

[12] This interpretation is in fact equivalent to the model we presented: In one direction, by taking $\tilde{\theta}_i = \langle c_i, \theta \rangle$, we can rewrite the payoff-relevant state θ_1 as a linear combination of $\tilde{\theta}_1, \ldots, \tilde{\theta}_K$. In the opposite direction, starting with a model in which the DM cares about $\langle w, \tilde{\theta} \rangle$, we define $\theta_1 = \langle w, \tilde{\theta} \rangle$ and $\theta_i = \tilde{\theta}_i, \forall i > 1$. Then each signal $\tilde{\theta}_i + \epsilon_i$ is a linear combination of the states $\theta_1, \ldots, \theta_K$ plus noise, returning our main model.

period t following history h^{t-1}.[13] Together with the prior belief that $\theta \sim \mathcal{N}(\mu^0, V^0)$, each strategy induces a joint distribution over possible states and infinite histories: $\Theta \times H^\infty = \mathbb{R} \times ([K] \times \mathbb{R})^\infty$.

Since the DM's prior and available signals are Gaussian, his posterior belief about θ is also Gaussian at every history. Specifically, if the DM's belief at the beginning of a period is $\theta \sim \mathcal{N}(\mu, V)$, then a single observation of signal i updates his belief to $\theta \sim \mathcal{N}(\hat{\mu}, \hat{V})$, where the posterior covariance matrix

$$\hat{V} = \phi_i(V)$$

is a deterministic function of the prior covariance matrix V (indexed to the signal i). On the other hand, the posterior mean $\hat{\mu}$ depends on the signal realization, and it is the following random variable:

$$\hat{\mu} \sim \mathcal{N}(\mu, V - \hat{V}).$$

Note that the distribution of the posterior mean has variance $V - \hat{V}$, so its degree of dispersion exactly equals the amount of uncertainty reduction from prior to posterior beliefs.[14]

At the final period, the DM's posterior mean and variance about θ_1 are sufficient to determine his optimal action. Let $\theta_1 \sim \mathcal{N}(\mu_1, V_{11})$ be his (marginal) belief about θ_1, where μ_1 is the first coordinate of the vector μ, and V_{11} is the $(1,1)$ entry of the matrix V. Let

$$r_t(\mu, V) = r_t(\mu_1, V_{11}) = \max_{a \in A} \mathbb{E}[u_t(a, \theta_1) \mid \theta_1 \sim \mathcal{N}(\mu_1, V_{11})] \tag{4}$$

be the (maximum) expected *flow* payoff of a DM with arbitrary belief $\theta \sim \mathcal{N}(\mu, V)$, conditional on period t being final. For notational simplicity, we write the DM's belief about θ given history h^t as $\theta \sim \mathcal{N}(\mu^t, V^t)$.

We represent the on-path behavior of any strategy S in the following way, tracking the number of acquired signals of each type up to and including a given period.

Definition 1. *Fix a strategy S and any infinite history h realized under S. The* division *over signals at time t along history h is denoted by the vector $q^S(t) = (q_1^S(t), \ldots, q_K^S(t))$, where $q_i^S(0) \equiv 0$ and $q_i^S(t) = q_i^S(t-1) + \mathbb{1}\left(S(h^{t-1}) = i\right)$. That is, $q_i^S(t)$ counts the number of periods in which signal i is observed, up to and including period t. Without referencing the history h, the division at time t is just the random vector $q^S(t)$.*

3.2 Optimality Criteria

We now define the notions of *myopic*, *dynamically optimal*, and *totally optimal* behavior that are the focus of this paper.

[13] Throughout, we assume without loss that the DM uses a pure strategy.

[14] This can be shown using the formula for conditional Gaussian variance. We omit the computation.

As usual, call a strategy *myopic, or myopically optimal* if the DM's signal choice in each period maximizes the expected flow payoff in the current period.[15] For the dynamically optimal problem, write $\pi(t)$ for the ex-ante probability that period t is final. Then, the forward-looking DM faces a (time-inhomogenous) Markov decision problem with value function given by the following Bellman equation:

$$U^{t-1}\left(\mu^{t-1}, V^{t-1}\right) = \max_{i \in [K]} \mathbb{E}\left[\pi(t) \cdot r_t\left(\mu^t, \phi_i\left(V^{t-1}\right)\right) + U^t\left(\mu^t, \phi_i\left(V^{t-1}\right)\right) \right. \\ \left. \mid \mu^t \sim \mathcal{N}\left(\mu^{t-1}, V^{t-1} - \phi_i\left(V^{t-1}\right)\right)\right]. \tag{5}$$

To interpret, a DM who observes signal i in period t updates his belief about the state vector to $\theta \sim \mathcal{N}(\mu^t, \phi_i(V^{t-1}))$. With probability $\pi(t)$, this is the final period and he receives $r_t(\mu^t, \phi_i(V^{t-1}))$ by taking the optimal action. Otherwise he continues to the next period and expects to receive $U^t(\mu^t, \phi_i(V^{t-1}))$ based on the updated belief. This value function $U^{t-1}(\mu^{t-1}, V^{t-1})$ is the total payoff to be gained following history h^{t-1} from the ex-ante perspective; it is not discounted.

We note that due to the assumption of a single payoff-relevant state, the utility function $U^{t-1}(\mu^{t-1}, V^{t-1})$ depends on μ_1^{t-1} and V^{t-1}, but not on the expected value of the remaining states.[16] Thus, we will (without loss) restrict attention to Markovian strategies that depend only on the simpler tuple $(t, \mu_1^{t-1}, V^{t-1})$ rather than (t, μ^{t-1}, V^{t-1}). We will further say that a Markovian strategy is *deterministic* if it does not condition on signal realizations. Such a strategy depends only on the calendar time t and the current covariance matrix V^{t-1}. If a Markovian strategy also conditions on the expected value of θ_1, it is called *stochastic*.

Whether it be deterministic or stochastic, a Markovian strategy S is *dynamically optimal* if at every history h^{t-1} with associated belief $\theta \sim \mathcal{N}(\mu^{t-1}, V^{t-1})$, the signal choice $S(h^{t-1})$ is a maximizer on the RHS of the Bellman equation (5). This definition requires optimality even at off-path histories.[17]

Finally, we define a property of strategies that we will refer to as *total optimality*.

Definition 2. *A strategy is* totally optimal at time t *if it maximizes* $\mathbb{E}\left[r_t(\mu^t, V^t)\right]$, *the expected flow payoff in period* t, *among all strategies. The expectation is taken over the*

[15] That is, S is myopic if at every history h^{t-1}, the signal choice $S(h^{t-1})$ is the signal i that maximizes the expected flow payoff $\mathbb{E}[r_t(\mu^t, \phi_i(V^{t-1})]$ given the posterior belief $\mu^t \sim \mathcal{N}(\mu^{t-1}, V^{t-1} - \phi_i(V^{t-1}))$. Note that by this definition, there can be multiple myopic strategies.

[16] Knowing μ_1^{t-1} and V^{t-1} is sufficient to determine the evolution of μ_1^t and V^t. It is however not enough for the DM to remember only the variance of θ_1, because processing future signals requires knowing how θ_1 is correlated with the other states.

[17] A slightly weaker definition of dynamic optimality is that S simply maximizes the ex-ante payoff $U^0(\mu^0, V^0)$. Because we are concerned in this paper with the optimality of the myopic strategy, we will work with the stronger definition that imposes optimality at all histories.

distribution of posterior beliefs $\theta \sim \mathcal{N}(\mu^t, V^t)$ at the end of period t, induced by the strategy.[18]

Intuitively, a strategy achieves total optimality at time t if a DM who knows period t to be final will follow the strategy. This can be understood as a limiting case of dynamic optimality, where the distribution over periods is degenerate.

4 Invariance to the Decision Problem

Our first set of results delivers a key simplification of the analysis.

Consider the special case of the main model in which the DM's expected flow payoff is

$$r_t(\mu, V) = -V_{11},$$

where V_{11} is the posterior variance about θ_1. This would arise, for example, if the decision problem were prediction of θ_1 and the payoff were quadratic loss, as in Section 1.[19]

In this case, because signal realizations do not affect the posterior covariance matrix, we can restrict attention to *deterministic strategies* with the property that $S(h^{t-1})$ only depends on t and V^{t-1}.

We will show now that the development of myopic optimality and total optimality for the prediction problem is without loss: any myopic strategy in the prediction problem is myopic in any decision problem, and a strategy is totally optimal for a general decision problem if and only if it is totally optimal for prediction.

Consider a DM with belief $\theta_1 \sim \mathcal{N}(\mu_1^{t-1}, [V^{t-1}]_{11})$ at the beginning of period t. His belief about θ_1 after a single observation of signal i becomes $\theta_1 \sim \mathcal{N}(\mu_1^t, [V^t]_{11})$. The posterior variance $[V^t]_{11} = [\phi_i(V^{t-1})]_{11}$ is deterministic, and the posterior mean follows the distribution

$$\mu_1^t \sim \mathcal{N}(\mu_1^{t-1}, [V^{t-1}]_{11} - [V^t]_{11}).$$

We observe that as the posterior variance $[V^t]_{11}$ decreases, this distribution of posterior means becomes less concentrated; by normality, this implies that the distribution of posterior beliefs increases in the Blackwell order (Blackwell (1951)). Thus, if acquisition of signal i induces a lower posterior variance than acquisition of signal j, then signal i Blackwell-dominates signal j. Since any myopic strategy for the prediction problem minimizes $[V^t]_{11}$, it also maximizes the immediate flow payoff in any decision problem.[20]

[18]The more general notion of total optimality following a given history is defined and used in the appendix.

[19]To recall, the DM chooses $a \in \mathbb{R}$ and receives $u(a, \theta_1) = -(a - \theta_1)^2$. Then he optimally chooses a to be the posterior mean of θ_1, and his expected payoff equals the negative of his posterior variance about θ_1.

[20]Intuitively, a myopic DM concerned about immediate payoffs only seeks to reduce the uncertainty about θ_1, but does not care how correlated θ_1 may be with the other states (which helps future updating and decisions). In his *static* signal acquisition problem, each signal can be viewed as θ_1 plus an independent noise term (see Appendix B for details). It then follows that smaller posterior variance means less noise in the signal and thus a more informative signal in the Blackwell ordering.

We record this result in the following lemma, with a formal proof in Appendix B:[21]

Lemma 2. *Fix an arbitrary decision problem satisfying Assumption 2. A strategy is myopic if and only if it is myopic for prediction.*

Henceforth, when working with myopic strategies, we can restrict attention to deterministic strategies that simply minimize the variance about θ_1 at each step.

The same equivalence turns out to hold for total optimality (see Definition 2). To see this, observe that our previous argument regarding Blackwell ordering implies that given a fixed number of observations k, the optimal *set* of k signals is independent of the decision problem. While our DM faces sequential acquisition and may condition later signal choices on earlier signal realizations, this additional flexibility turns out to not be advantageous because the variance at the final period does not depend on realized signal values. Thus, for a fixed final period, the optimal *signal sequence* is also invariant to the decision problem.

Our next lemma formalizes this, and states that a strategy is totally optimal at time t if and only if with probability 1, there is no way to reduce posterior variance by redistributing the total number of past observations across different signals. To state the lemma, we introduce a posterior variance function that will be frequently referred to in this paper:

Definition 3. *Given prior covariance matrix V^0, matrix of linear coefficients C and signal variances $(\sigma_i^2)_{i=1}^K$. For non-negative integer values $q_1, \ldots, q_K$, the function $f(q_1, \ldots, q_K)$ is the DM's* posterior variance about θ_1 *after q_i observations of each signal i (see Appendix A for details).*

Lemma 3. *Fix an arbitrary decision problem satisfying Assumption 2. A strategy S is totally optimal at time t if and only if the induced (random) division $q^S(t)$ (see Definition 1) satisfies*

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

with probability 1.[22]

Any division $(q_1, \ldots, q_K)$ in the argmin above is called a *totally optimal division at time t*.

Using Lemma 2 and 3, we can deduce that for the informational environment discussed in Section 1, any myopic strategy achieves total optimality and dynamic optimality for arbitrary time-dependent payoff functions. We show in the following section that this result extends to several classes of informational environments.

[21]Below, Assumption 2 ensures that the relation between posterior variance and Blackwell ordering is strict: a signal with strictly less noise leads to strictly higher flow payoff.

[22]Our proof in Appendix B shows, somewhat surprisingly, that the DM's expected flow payoff in period t is unchanged even if his signal choices lead to different divisions along different histories, so long as each realized division is totally optimal.

Outside of such environments, the dynamically optimal strategy for general decision problems does not admit a similar reduction (to the prediction problem). For example, compare two signal sequences X_1X_2 and X_3X_4 where X_1 is myopically better than X_3 but the pair X_3X_4 yields lower posterior variance than X_1X_2. Which sequence gives rise to higher overall expected payoff will in general depend on the DM's decision problem and on discounting. This difficulty makes it harder to establish the equivalence between dynamic and myopic (or total) optimality. We will nevertheless show in Section 6 that the dynamically optimal strategies are *eventually* deterministic and myopic (in generic environments).

5 Sufficient Conditions for Immediate Equivalence

In Section 1, we saw an environment in which the myopic, dynamically optimal, and totally optimal strategies agree from period 1. Here we generalize this equivalence to several classes of environments. One sufficient condition is that the signals are *separable* in the following sense:

Definition 4. *The informational environment* $(V^0, C, \{\sigma_i^2\})$ *is* separable *if there exist convex functions* $g_1, \ldots, g_K$ *and a strictly increasing function* F *such that the posterior variance function satisfies*

$$f(q_1, \ldots, q_K) = F(g_1(q_1) + \cdots + g_K(q_K)).$$

Intuitively, separability ensures that observing signal i does not change the relative value of other signals, but strictly decreases the marginal value of signal i relative to every other signal.[23] The benchmark case in Section 1 falls into this class of environments, as does its generalization below:

Example 1. There is a single payoff-relevant state $\theta \sim \mathcal{N}(0, v_0)$. The DM has access to observations of $X = \theta + b_1 + \cdots + b_{K-1} + \epsilon_X$, where each b_i is a persistent bias independently drawn from $\mathcal{N}(0, v_i)$, and $\epsilon_X \sim \mathcal{N}(0, \sigma_X^2)$ is a noise term i.i.d. over time. Additionally, he can learn about each bias b_i by observing $B_i = b_i + \epsilon_i$, where $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$. Then, the DM's posterior variance about θ is given by

$$f(q_1, \ldots, q_{K-1}, q_X) = v_0 - \frac{v_0^2}{v_0 + \frac{\sigma_X^2}{q_X} + \sum_{i=1}^{K-1} \left(v_i - \frac{v_i^2}{v_i + \sigma_i^2 / q_i} \right)}$$

which is indeed separable.

Another class of separable environments occurs when the signals provide orthogonal (thus independent) information:

[23] While we can write f in terms of V^0, C and $\{\sigma_i^2\}$, Definition 4 is strictly speaking not a condition on the primitives.

Example 2. Suppose the DM's prior is standard Gaussian ($V^0 = \mathbf{I}_K$), and the row vectors of C are orthogonal to one another. Then we have

$$f(q_1, \ldots, q_K) = \left[I_K - C' \left(CC' + \operatorname{diag}\left(\frac{\sigma_1^2}{q_1}, \ldots, \frac{\sigma_K^2}{q_K} \right) \right)^{-1} C \right]_{11}$$

By orthogonality, CC' is a diagonal matrix. Thus $\left(CC' + \operatorname{diag}\left(\frac{\sigma_1^2}{q_1}, \ldots, \frac{\sigma_K^2}{q_K} \right) \right)^{-1}$ is also a diagonal matrix, and it is separable in $q_1, \ldots, q_K$.

By the same argument as in Section 1, in a separable environment, any strategy that is myopic for prediction is totally optimal for prediction at every time. The reduction lemmata in Section 4 enable us to extend the equivalence to arbitrary decision problems.[24]

Proposition 1. *Suppose the informational environment is separable. Then any myopic strategy is totally optimal at every time, and it is dynamically optimal. Conversely, any dynamically optimal strategy is myopic and totally optimal at every time.*

An alternative sufficient condition for immediate equivalence (among the different optimality criteria) is that the posterior variance about θ_1 depends on each signal count in a *symmetric* way:

Definition 5. *The informational environment* $(V^0, C, \{\sigma_i^2\})$ *is* symmetric *if the posterior variance function* $f(q_1, \ldots, q_K)$ *is symmetric in its arguments.*

An example of a symmetric environment is the following.

Example 3. There are three states $\theta_1, \theta_2, \theta_3$ independently drawn from $\mathcal{N}(0, 1)$ and three signals $X_1 = \theta_2 + \theta_3 + \epsilon_1$, $X_2 = \theta_1 + \theta_3 + \epsilon_2$ and $X_3 = \theta_1 + \theta_2 + \epsilon_3$. The noise terms $\epsilon_1, \epsilon_2, \epsilon_3$ have the same variance. Suppose the DM cares about $\theta_1 + \theta_2 + \theta_3$, then the signals contribute symmetrically to the posterior variance about this payoff-relevant state.

In a symmetric environment, the natural strategy of observing the signal that has been *least observed* turns out to be myopically, dynamically and totally optimal.[25] We have the following result:

Proposition 2. *Suppose the informational environment is symmetric. Then any myopic strategy is totally optimal at every time and dynamically optimal. Any dynamically optimal strategy is myopic and totally optimal at every time.*

Finally, we also show that for the case of two signals ($K = 2$), even without the separability or symmetry assumption, myopic information acquisition achieves optimality whenever the vectors defining the signals are not too close to collinear—see Appendix I for details.

[24] To derive the equivalence with dynamic optimality, we also use the observation that if a strategy S maximizes the flow payoff at every period, then it also maximizes the ex-ante expected payoff.

[25] Starting from the null history, the (on-path) signal acquisitions divide into blocks of K, each of which consists of exactly one observation of each signal, in any order.

6 Eventual Equivalence

The previous section provided conditions under which the myopic and dynamically optimal signal paths coincide at every period, and are moreover totally optimal for every t. Why might these equivalences fail? We provide a simple example below, building on the benchmark case considered in Section 1.

Example 4. There are three states θ, b_1, b_2 independently drawn from $\mathcal{N}(0,1)$, where only θ is payoff-relevant. The DM chooses from the three signals $X = \theta + b_1 + \epsilon_X$, $B_1 = b_1 + b_2 + \epsilon_1$, and $B_2 = b_2 + \epsilon_2$, where all signal variances are equal to 1. Given a history (q_X, q_1, q_2), the DM's posterior variance about θ is

$$f(q_X, q_1, q_2) = 1 - \frac{1}{2 + \frac{1}{q_X} - \frac{1}{1+\frac{1}{q_1}+\frac{1}{1+q_2}}}. \tag{6}$$

The derivation is similar to (1) in Section 1, so we omit it. From this formula, it can be shown that the myopic decision-maker's initial signal path is XXB_1XX, which achieves the (unique) totally optimal division $(4,1,0)$ at period 5. However, the myopic DM's next signal acquisition is B_1, so that the myopic division becomes $(4,2,0)$, while the unique totally optimal division at period 6 is $(3,2,1)$, as $f(3,2,1) < f(4,2,0)$.[26]

The myopic strategy fails to achieve total optimality at $t = 6$ for the following reason: after the initial history XXB_1X, the acquisition of signal X is myopically better than either B_1 or B_2. But looking forward two periods, the *pair* of signals B_1B_2 is better than any pair that includes signal X (in particular the myopic choices XB_1).

From this example, we see how complementarities between pairs of signals (such as between B_1 and B_2 above) can render the myopic choices sub-optimal.[27] Nonetheless, we will show that the strength of such complementarities vanishes at late time periods, so that the myopic signal path eventually achieves approximate dynamic and total optimality. Moreover, in generic environments the myopic path is eventually exactly optimal.

Below, we use $m(t) = (m_1(t), \ldots, m_k(t))$ to denote the division over signals under a *deterministic* myopic strategy, $d(t)$ to denote the (random) division under an arbitrary dynamically optimal strategy, and finally $n(t)$ to denote a totally optimal division at time t, according to the definition in Lemma 3.[28] Thus $m_i(t)$ (resp. $d_i(t), n_i(t)$) counts the number

[26] We point out that in the sixth period, the myopic DM is in fact indifferent between observing B_1 or B_2; if he observes B_2 instead, his division would be $(4,1,1)$, which is also not totally optimal.

[27] However, if signals are fully divisible, then even a myopic DM can take advantage of the complementarity between B_1 and B_2 by devoting equal attention to them. In continuous time, the myopic strategy in Example 4 is in fact optimal from the beginning. See Appendix M for details.

[28] Although non-deterministic myopic strategies may exist due to tie-breaking, any realized division over

of times a myopic (resp. forward-looking, totally optimal) DM observes signal i in the first t periods. Our results will be stated in terms of these signal counts.

Our first result in this section says that the differences (in terms of signal counts) across the three optimality criteria become minimal after sufficiently many periods. Specifically, at every late period t, the number of times any signal has been observed under myopic, dynamic and total optimality can differ by at most 1.

Before stating the result, we impose a weak technical condition. Fix any proper subset of signals I. We require that some signal j outside of I *strictly decreases posterior variance about* θ_1 whenever each signal in I has been observed sufficiently many times. This assumption guarantees that the dynamically optimal strategy observes each signal infinitely often along every history of signal realizations. We comment that the assumption is satisfied for generic informational environments[29]; additionally, it is not needed for the equivalence between myopic and total optimality, or in the continuous-time variant of our model (Appendix M).

Assumption 4 (Strict Variance Decrease)**.** *For any proper subset of signals I, there exists $j \notin I$ and $\epsilon > 0$ such that: for any division $(q_1, \dots, q_K)$ with $q_j = 0$ and q_i sufficiently large ($\forall i \in I$), it holds that $f(q_j + 1, q_{-j}) < f(q_j, q_{-j}) - \epsilon$.*

Our eventual (approximate) equivalence result is now stated.

Theorem 1 (Eventual Gap of One)**.** *Suppose the informational environment $(V^0, C, \{\sigma_i^2\})$ satisfies Assumption 3 and 4. There exists a large finite T such that the following holds: for any decision problem and any time $t \geq T$,*[30]

(a) $|m_i(t) - n_i(t)| \leq 1, \forall i$.

(b) $|d_i(t) - n_i(t)| \leq 1,\ \forall i$ *for every realized* $d(t)$.

(c) $|d_i(t) - m_i(t)| \leq 1, \forall i$ *for every realized* $d(t)$.

We provide a brief intuition for this result. For each k, define a new state $\tilde{\theta}_k = \langle c_k, \theta \rangle$, so that the signal X_k is simply $\tilde{\theta}_k$ plus independent Gaussian noise. The payoff-relevant state θ_1 can be rewritten as a linear combination of the linearly transformed states $\tilde{\theta}_1, \dots, \tilde{\theta}_K$. Our

signals under such strategies also occurs under a deterministic strategy. Since our results below are stated in terms of these divisions, they apply to all myopic strategies, deterministic or stochastic.

[29]Zero marginal values occur only if $\partial_j f(q_1, \dots, q_K) = 0$. Fixing V^0 and C, such an equation (for any $q_1, \dots, q_K$) induces a non-trivial polynomial relation among the signal variances (σ_i^2). Since the number of possible tuples $(q_1, \dots, q_K)$ is countable, zero marginal values only happen in non-generic situations.

[30]We remark that the results stated here only compare the *on-path* behavior of myopic, dynamically optimal and totally optimal strategies. However, a slight modification of our proof shows that for some large T, following any given history, the myopic, dynamically optimal and totally optimal divisions differ by at most 1 after T periods (a similar statement holds for Theorem 2 below). We omit the details.

key observation is that the DM's posterior belief over these transformed states $(\tilde{\theta}_1, \dots, \tilde{\theta}_K)$ becomes *almost independent* after sufficiently many observations of each signal. Formally, the posterior covariance between any pair of transformed states is small relative to the posterior variance about either state, and the ratio converges to zero as the number of observations grows to infinity. We note that this property of "de-correlation" is a consequence of the Bayesian Central Limit Theorem, and it extends to states and signals that need not be normally-distributed.[31]

Based on the insight that different signals are approximately independent from each other, we show that any complementarity or substitution effect across the signals is eventually weak. The property of "de-correlation" alone is not sufficient to drive this latter conclusion. While de-correlation gives the magnitude of the posterior covariance matrix, the complementarity or substitution between two signals is essentially a function of how the posterior covariance matrix varies with the acquisition of extra signals (i.e. its derivatives and second derivatives). To this end, we develop a key technical lemma (Lemma 5 in Appendix A) regarding the second derivatives of the posterior covariance matrix. This lemma says that the effect of observing a signal on the marginal value of other signals is eventually second-order to its effect on the marginal value of (further realizations of) the same signal.[32] Thus, observing a particular signal may change the ordering of other signals, but the cardinal extent of such change is limited. We conclude that the dynamic information acquisition problem is "near-separable" at sufficiently late periods, and it is approximately without loss to treat the dynamic problem as a series of static problems, for which the myopic solution is optimal. See Appendix E for the formal proof.

One may wonder whether the "gap of one" stated in the theorem can be dropped. Indeed, in *generic* informational environments, it turns out that any myopic strategy eventually coincides (exactly) with a dynamically optimal strategy and also achieves total optimality at every period (see Appendix G for the proof):[33]

Theorem 2 (Generic Eventual Equivalence). *Fix prior covariance matrix V^0 and linear coefficients C. For generic signal variances $\{\sigma_i^2\}_{i=1}^K$ (with Lebesgue measure 1), there exists T^* such that $m(t) = d(t) = n(t)$ at every time $t \geq T^*$, for any decision problem.*

In fact, the "generic" qualifier can be dropped when $K = 2$ (see Appendix I), but it cannot when $K > 2$: in Appendix J, we show that in the previous Example 4, all myopic strategies

[31] While de-correlation suggests the possibility for extending our eventual equivalence results beyond Gaussian information, our proof of Theorem 1 relies on normality to prove the stronger property of eventual "near-separability," as discussed in the next paragraph.

[32] This is true along any signal path in which the signal counts go to infinity proportionally, which we endogenously justify in Appendix D.

[33] The key new technical tool is use of results on the approximation of real numbers by rationals; see Appendix G for details.

fail to achieve total optimality infinitely often. We also provide another example in which the myopic division differs from the dynamically optimal division infinitely often, for arbitrarily low discounting. These examples suggest that our results are best possible.[34]

Finally, if we allow the DM to observe B signals (including repetitions) each period, then with sufficiently large B, we return the immediate equivalence results.

Theorem 3. *Fix any informational environment satisfying Assumption 3, and suppose that the DM acquires B signals each period. Then, if B is sufficiently large, any myopic strategy achieves total optimality at every time and is dynamically optimal.*

Intuitively, this is because the myopic strategy with B observations per period is equivalent to a strategy that plans B periods forward in our main model. The question of how large B needs to be for immediate equivalence to obtain is related to the question of how large the period T needs to be for the signal paths to be approximately equivalent (in Theorem 1). We provide now a bound for this T.

6.1 Time to Eventual Equivalence

Using the linear transformation described above, we may consider states $\tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_K \sim \mathcal{N}(\mu, V)$ such that the payoff-relevant state is $\theta^* = \langle w, \tilde{\theta} \rangle$ for some fixed vector w, and the available signals are $X_i^t = \tilde{\theta}_i + \epsilon_i^t$ with ϵ_i^t standard Gaussian noises independent from each other and over time. The primitives of this transformed informational environment are the weight vector w and the prior covariance matrix V. We are interested in how large the period T has to be for Theorem 1 to apply. The following upper bound assumes that w is the vector of all 1's, although the method of proof easily generalizes to arbitrary w:

Theorem 4. *Let S denote the operator norm of the matrix V^{-1}.*[35] *Suppose $w = (1, \ldots, 1)'$, then $|m_i(t) - n_i(t)| \leq 1$ whenever $t \geq 24(S+1)K^2$.*[36]

We can derive similar bounds for the dynamically optimal division by incorporating Assumption 4. The statement is somewhat cumbersome, so we will not give the details here.

[34] Nonetheless, these counterexamples (to eventual *exact* equivalence) rely on the discreteness of our main model. The eventual gap of one vanishes in the continuous-time limit, as we show in Appendix M.

[35] The operator norm of a matrix M is defined $\|M\|_{op} = \sup\left\{\frac{\|Mv\|}{\|v\|} : v \in \mathbb{R}^K \text{ with } v \neq \mathbf{0}\right\}$.

[36] Our bound is of order K^3 for almost all covariance matrices V, for the following reason: the positive-definite matrix V can be written as $U \cdot U'$ for some matrix U. Imagine that the entries of U are drawn i.i.d. from a fixed distribution with finite mean and variance. Then a result in random matrix theory states that $\|U^{-1}\|_{op}$ has order $\sqrt{K}$ with probability approaching 1 as $K \to \infty$ (Rudelson and Vershynin (2008), Tao and Vu (2010)). Thus $\|V^{-1}\|_{op}$ has order at most K.

Intuitively, the time to (approximate) equivalence depends on how long it takes for signals to "de-correlate," at which point potential complementarities across signals are weak. How long this takes depends on two primitives:

First, it is increasing in the number of signals K. Intuitively, the greater the number of signals K, the more pairs of signals there are that need to de-correlate; this increases the time to equivalence. Second, the bound is increasing in the norm of V^{-1}. To interpret this, suppose first that we adjust the precision of the DM's prior but fix the degree of correlation, for example by scaling V by a factor less than 1. Then, the norm of V^{-1} increases, and equivalence between myopic and total optimality is attained later. Intuitively, holding fixed the correlation among the states, a more precise prior can be understood as "re-scaling" the state space by shrinking all states towards zero. Since signal noise is not correspondingly rescaled, each signal now reveals less about the state, and de-correlation takes longer.

In contrast, suppose we hold precision fixed and increase the degree of correlation. This would correspond to fixing the diagonal entries of V and increasing the off-diagonal entries, so that the variances about individual states are unchanged but their covariances become larger in magnitude. Then, the entire matrix V becomes closer to being singular, the norm of V^{-1} increases and the time to equivalence is longer. Intuitively, greater correlation in the prior requires more time to de-correlate.

With sufficiently many observations, the DM's belief simultaneously becomes more precise and less correlated, and these two effects are confounded in our previous result that equivalence "eventually" occurs. It is tempting to think that eventual equivalence follows from the (eventual) precision of the DM's belief. However, our discussion above shows that the important feature is not precision but correlation: having an arbitrarily precise belief does not guarantee (immediate) equivalence; on the contrary, equivalence takes longer under a precise correlated prior belief.[37]

7 Endogenous Stopping and Intertemporal Decisions

While we have assumed an exogenous (random) final period so far, our results extend to endogenous stopping problems in which the DM decides the final period t and takes an action $a \in A$. His payoff is given by an arbitrary time-dependent payoff function $u_t(a, \theta_1)$, which takes into account discounting and/or a constant cost of signals.[38]

[37] It is not completely rigorous to base this conclusion on the "upper bound" we derive for the time to equivalence. Nonetheless, the intuition that a more precise prior implies a longer time to de-correlate does hold. There is an exact relationship in the continuous-time limit of our model (see Appendix M): scaling (each entry of) the prior covariance matrix V by half means the time at which equivalence occurs doubles.

[38] Constant waiting cost per unit of time appears in the work of Fudenberg, Strack and Strzalecki (2017) and Che and Mierendorff (2017). It is different from the approaches taken in (for example) Moscarini and

In fact, we will consider a more general class of intertemporal decision problems described as follows: in each period t, the DM observes one of the K signals and then chooses an action a_t; his total payoff from these actions is $U(a_1, a_2, \ldots; \theta_1)$, which can exhibit arbitrary intertemporal dependence.[39] We provide in Appendix L an example that takes this more general form: a DM chooses how to divide resources between investment in an asset with known return and an asset with unknown return; in each period, he simultaneously makes investment decisions, and also acquires information about the unknown return.

We show that as long as the DM's action choices do not affect how much he can learn about the states,[40] myopic information acquisition remains (approximately) optimal. Formally, we prove in Appendix L the following result:

Theorem 5. *Suppose myopic information acquisition achieves total optimality at every time for the prediction problem. Then, for any intertemporal decision problem in which the DM takes an action a_t at the end of each period t and receives overall payoff $U(a_1, a_2, \ldots; \theta_1)$, there is a dynamically optimal strategy in which the DM acquires information myopically.*

Thus, whenever the informational environment satisfies the sufficient conditions in Section 5, the DM cannot do better than acquiring information myopically even if his actions may have intertemporal consequences.[41] This property that the optimal signal path is independent from the optimal sequence of actions makes it possible to characterize the latter as if information arrives exogenously; in particular, the optimal stopping time can be solved for assuming myopically chosen signals until the stopping time.[42] We should however mention the distinction between "myopic signal choice" and "myopic action choice." While Theorem 5 suggests that myopic information acquisition is optimal under fairly permissive assumptions, myopic actions often do poorly under the forward-looking criterion.

Our proof of Theorem 5 is based on a dynamic Blackwell-dominance lemma (for comparing sequences of normal signals) that generalizes the static reduction results in Section 4. Loosely speaking, this lemma states that a DM can "imitate" the action choice, and achieve the same payoff, as any DM whose posterior variance about θ_1 is larger at every period. We note that the direct extension of Blackwell-dominance to the dynamic setting says that a

Smith (2001) and Hebert and Woodford (2017), where the cost of acquiring different signals varies depending on properties of that signal (e.g. precision or entropy reduction).

[39] This framework embeds endogenous stopping if we take each a_t to specify both the stopping decision and the action to be taken when stopped.

[40] This assumption distinguishes our model from Multi-armed Bandit problems, see Section 11 for detailed discussion of this connection.

[41] Our eventual equivalence results can be similarly generalized to this setting. We omit the details.

[42] In a stylized two-state model, Fudenberg, Strack and Strzalecki (2017) analyze the optimal stopping behavior under exogenous information and proceed to verify its optimality under endogenously chosen information. We discuss this connection in Section 8, when we introduce our model in continuous time.

DM with *better information in every period* obtains higher payoff. In contrast, our lemma is based on the weaker assumption that the DM has *better cumulative information up to every period*.[43] This turns out to be technically non-trivial, see Appendix L for further discussion.

8 Other Extensions of the Model

Multiple Payoff-Relevant States. In the main model, we assumed that the decision-maker's payoff function depends on a unidimensional state θ_1. When the DM cares about multiple states at once, our results extend for the specific problem of prediction: the DM predicts state vector $\hat{\theta} \in \mathbb{R}^K$ and receives payoff $-(\hat{\theta} - \theta)'\mathbf{W}(\hat{\theta} - \theta)$, where $\mathbf{W}$ is an arbitrary positive semi-definite matrix. Our equivalence results and their proofs apply without change to this setting.[44] However, extension to general decision problems fails because there does not exist a complete Blackwell ordering over signals about multi-dimensional states.

Non-Persistent i.i.d. States. So far we have considered persistent states $\theta_1, \dots, \theta_K$. All of our results extend if new states $\theta_1^t, \dots, \theta_K^t$ are independently drawn each period according to $\theta_k^t = \theta_k + \gamma_k^t$, and the signals are $X_i^t = \sum_{i=1}^K c_{ik}\theta_k + \epsilon_i^t$ as before. The noise terms γ_k^t and ϵ_i^t are independent from one another. We assume that the DM receives $u_t(a, \theta_1^t)$ in the final period t, which depends on the payoff-relevant state at that time. To see that our results extend, simply notice that the DM's posterior variance about θ_1^t is the sum of his posterior variance about θ_1 and the variance of γ_1^t. Because the latter cannot be controlled by the DM, his optimal information acquisition strategy is unchanged. We leave to future work the question of whether (and when) the myopic strategy is optimal under richer state dynamics (e.g. AR processes).

Free Signals. We have modeled limited information acquisition capacity by restricting the DM to observe only one signal per period. This restriction can be easily relaxed to allow multiple observations each period, see Theorem 3 above. We mention that our model can be further generalized to include "free signals"; that is, suppose a subset of signals is observed in each period, and the DM additionally chooses one signal (outside of this subset) to observe.

[43] The condition of Theorem 5 does not imply that the myopic strategy achieves the largest variance reduction in any given period (other than the first). It only ensures that the cumulative variance reduction is largest under the myopic strategy.

[44] For a diagonal matrix $\mathbf{W}$, the DM's objective function f is a weighted sum of posterior variances about multiple states. Generalizing Lemma 5 in Appendix A, we can show that any such f exhibits "eventual near-separability," which implies our eventual equivalence results. Even if $\mathbf{W}$ is not diagonal, by the spectral theorem, there exists an orthonormal matrix $\mathbf{J}$ and a diagonal matrix $\mathbf{X}$ such that $\mathbf{W} = \mathbf{JXJ}'$. Then the objective function is a weighted sum of posterior variances about multiple, linearly-transformed states. Our proofs still carry through as long as we modify Assumption 3 to require that each of these "transformed payoff-relevant states" is exactly identified by the signals.

Our results and their proofs remain valid.[45]

Continuous Time. In Appendix M, we provide a detailed analysis of a continuous-time version of our problem. We assume that the DM has one unit of attention in total at every point in time. He chooses attention levels $\beta_1(t), \dots, \beta_K(t)$ (subject to $\beta_i(t) \geq 0$ and $\sum_i \beta_i(t) \leq 1$), which influence the diffusion processes $X_1, \dots, X_K$ that he observes:[46]

$$dX_i^t = \beta_i(t) \cdot \tilde{\theta}_i dt + \sqrt{\beta_i(t)} dB_i^t,$$

where each B_i is an independent standard Brownian motion, and $\tilde{\theta}_i = \langle c_i, \theta \rangle$ is a "linearly transformed state." The decision problem is the same as in discrete time.[47]

We note that if $K = 2$, the DM has an independent symmetric prior about the states $\tilde{\theta}_1, \tilde{\theta}_2$ and if he cares about the difference $\tilde{\theta}_1 - \tilde{\theta}_2$, then this model becomes the one considered by Fudenberg, Strack and Strzalecki (2017) (section 3.5), who show it is optimal (on-path) to pay equal attention to both signals at every time.[48] This result is a special case of our Theorem 8 in Appendix M, which generalizes to arbitrary prior beliefs and characterizes the optimal strategy off-path.[49] Our analysis also reveals that with two signals in general, the optimal attention levels are eventually constant.

For $K > 2$, we extend and strengthen many of our previous results to this continuous-time setting. Specifically, we show that eventual *exact* equivalence obtains in all informational environments satisfying identifiability, thus improving upon the conclusion of Theorem 1 and 2. We also provide more permissive sufficient conditions for immediate equivalence, which can be interpreted as whenever the DM's prior belief over different states is "almost independent," as mentioned previously.

9 Discussion of Results

The equivalence results presented earlier show that under certain conditions, a decision-maker will (eventually) acquire the same sequence of signals whether he optimizes a myopic criterion or a forward-looking criterion, and that these signal choices are "totally optimal." We discuss now certain conceptual implications of these results.

[45]Every signal that is not "free" is still observed with positive frequency, as in Proposition 3 in Appendix D. We could then apply Lemma 5 to deduce eventual equivalence.

[46]This formulation can be seen as a limit of our discrete-time model, if we take period length to zero and also "divide" the signals to maintain the same amount of information that can be gathered every second.

[47]At an exogenously determined random final time t (drawn with density $\pi(t)$), the DM takes an action a and receives payoff $u_t(a, \theta_1)$.

[48]While the decision problem in Fudenberg, Strack and Strzalecki (2017) involves endogenous stopping, as we have discussed in Section 7, this difference does not affect our analysis.

[49]To apply Theorem 8, we observe that the payoff weights are $w_1 = w_2 = 1$ (replacing $\tilde{\theta}_2$ by its negative). Thus the condition $w_1(V_{11} + V_{12}) + w_2(V_{21} + V_{22}) \geq 0$ is satisfied for every prior covariance matrix V.

Robust Information Acquisition. It is standard to assume that decision-makers know their objective function. In practice, however, decision-makers often do not know *when* or *how* acquired information will be useful. For example, students take classes to acquire knowledge without clear practical applications, and CEOs learn about their industry to inform decisions that they cannot yet anticipate. These decision-makers' beliefs over which decision problems they will ultimately face, and when they will face them, can be highly complex. In fact, decision-makers may not have well-defined beliefs, facing ambiguity over the final decision problem and its timing. In general, the information acquisition problem for such decision-makers can be difficult to describe and solve. Our results show that there are informational environments in which these challenges can be sidestepped: by behaving myopically, the DM acquires information in a way that is simultaneously best across *all* decision problems and for *all* timings of decision. This suggests a domain in which robust information acquisition is possible, and moreover simple enough for decision-makers to use in practice.

Multiple Decision-Makers. Consider a sequence of decision-makers who each acquires a signal, whose realization is public, and then chooses an action (based on all past signal realizations) to maximize a private objective. This model resembles the social learning frameworks first introduced in Banerjee (1992) and Bikhchandani, Hirshleifer and Welch (1992), without the classic friction that decision-makers only observe coarse summary statistics of past information acquisitions. In this setting, it is obvious that the decision-makers will eventually learn the payoff-relevant state. But an interesting feature of our environment is that not only does learning occur, it turns out to occur "as fast as possible." If the informational environment is separable or symmetric, a social planner cannot improve on the amount of information aggregated by a sequence of myopic decision-makers choosing information for different and private objectives.[50]

10 Application: Games with Dynamic Information Acquisition

In general, the possibility for a decision-maker to jointly acquire information and also to choose an optimal action introduces substantial technical complications—in particular, it is often the case that the optimal signal choices and the optimal action sequences need to be solved *jointly*. Our immediate equivalence results tell us that there are domains in which we can separate the concern of optimal information acquisition from other details of the problem, thus simplifying the analysis. We describe in detail below one such domain—dynamic information acquisition in games—with application to Hellwig and Veldkamp (2009) and

[50]By Theorem 2, the planner (generically) can do no better than a long sequence of decision-makers.

Lambert, Ostrovsky and Panov (2017).

Consider a normal-form game with N players and action profiles $A = \times_i A_i$. Each player i's payoff $u_i(a, \omega)$ is a function of the realized action profile $a \in A$ and an unknown real-valued state ω. This payoff function is routinely extended to mixed action profiles by linearity.

The normal-form game will be played once. Time is discrete, and a full-support distribution π determines the final period in which the game will be played.[51] In each period t up to and including that final period, each player i has access to signals from the set $(X_k^{it})_{k=1}^K$, defined as follows:

$$X_k^{it} = \langle c_k, \theta^i \rangle + \epsilon_k^{it}, \tag{7}$$

where c_k is a $K \times 1$ vector of coefficients, the vector $\theta^i = (\theta_1^i, \theta_2^i, \ldots, \theta_K^i)$ represents *persistent* unknown states pertaining to player i's observations, and ϵ_k^{it} are standard Gaussian noise terms that are independent across signals, players and time.[52]

We require that the players share a common prior over ω and the states $(\theta^i)_{1\leq i\leq N}$ with the following *conditional independence* property: for each player i, conditional on the value of θ_1^i, the payoff-relevant state ω and the *other* players' unknown states $(\theta^j)_{j\neq i}$ are conditionally independent from player i's states θ^i.[53] This ensures that no player i infers anything about ω or about any other player j's information beyond what he (player i) learns about θ_1^i, which essentially makes θ_1^i the only state of interest for player i.[54]

For concreteness, we provide examples (adapted from Lambert, Ostrovsky and Panov (2017)) that do and do not satisfy conditional independence.[55]

Example 5 (Satisfies Conditional Independence). The payoff-relevant state is $\omega \sim \mathcal{N}(0,1)$. One player has access to noisy observations of $\omega + \rho_1\xi + b_1$ and b_1, where b_1 is independent of ω, ξ and ρ_1 is a constant. The other player has access to noisy observations of $\omega + \rho_2\xi + b_2$ and b_2, where b_2 is independent of ω, ξ and b_1, and ρ_2 is a constant. Then, defining $\theta_1^1 = \omega + \rho_1\xi$ and $\theta_1^2 = \omega + \rho_2\xi$, we see that for each player i, conditional on the value of θ_1^i, the state of nature ω and also the the other player's states (θ_1^j, b_j) are independent from b_i. Thus, the best way for player i to learn about ω and about the other player's information is to learn $\theta_1^i = \omega + \rho_i\xi$ as precisely as possible.

[51] It is not important that actions are synchronous. Our subsequent observations will hold even if players take actions at different random times.

[52] We make the simplifying assumption that the coefficients c_k do not depend on the player, so that players face symmetric informational environments. Assuming unit signal variances is just a normalization.

[53] Note however that we do not impose conditional independence between ω and the other players' states.

[54] Conditional independence is imposed on players' beliefs at $t = 0$. However, this assumption is sufficient to guarantee conditional independence for all possible posterior beliefs—given the value of θ_1^i, each signal is simply a linear combination of player i's other states plus noise. Thus conditional independence is preserved.

[55] Example 5 is based on Example OA.3 in Lambert, Ostrovsky and Panov (2017), and Example 6 is based on their Example 1.

Example 6 (Fails Conditional Independence). The payoff-relevant state is ω. One player has access to noisy observations of $\omega + \xi$, where ξ is independent of ω. The other player has access to noisy observations of both ω and ξ. Because both states ω and ξ covary with $\omega + \xi$, there is no way to define the second player's "state of interest" that would satisfy conditional independence.

Throughout, we maintain Assumptions 2 and 3. In the current context, Assumption 3 requires that each θ_1^i is exactly identified by the signals available to player i. Assumption 2 is similarly modified to state that for each player i and arbitrary opponent strategies, player i's expected payoff, conditional on the value of θ_1^i, satisfies "payoff sensitivity to the mean."[56] This ensures that regardless of how opponents play, each player i always has strict incentive to acquire information about θ_1^i. Finally, we assume that the informational environment is *separable* or *symmetric*, as defined in Section 5.

In each period until the final period, each player i acquires B independent observations of his signals described above, possibly obtaining multiple (independent) realizations of the same signal. Both signal choices and their realizations are private information.

For each player, a history at the end of t periods is a sequence of (that player's) signal choices and their realizations up to and including period t. For a fixed player i, an *information acquisition strategy* specifies a mapping from every history to a (multi-)set of B signals among $(X_k^i)_{1 \le k \le K}$. A *decision strategy* specifies a mapping from every history to a mixed action $s_i \in \Delta(A_i)$. Player i's *strategy* in this model consists of an information acquisition strategy as well as a decision strategy. The players' strategies, together with the distribution π governing the final period, determine a joint distribution over states, histories and realized action profiles. Each player seeks to maximize his expected payoff with respect to this distribution, and we look for Nash equilibria of this model.

While this setup is rather general, it turns out to admit a simple solution:

Corollary 1. *Under the above assumptions, each player's signal path is myopic in every NE of this model. That is, at every realized history, each player i acquires the B signals that achieve the greatest immediate decrease in his belief variance about θ_1^i.*[57]

In fact, our equivalence results show that the myopic information acquisition strategy is *dominant* in the following sense: for arbitrary opponent strategies, player i's best response

[56]From player i's perspective, the strategy of player j can be viewed as mappings from player j's states $(\theta_1^j, \dots, \theta_K^j)$ to player j's mixed actions. Thus, given opponent strategies, player i's expected payoff conditional on his own states $(\theta_1^i, \dots, \theta_K^i)$ depends on his conditional belief about other players' states and ω. By "conditional independence", this expectation is unchanged if player i only conditions on the value of θ_1^i, according to the prior.

[57]Using the stronger solution concept of Perfect Bayesian equilibrium or Sequential equilibrium, we can further deduce that the entire information acquisition strategy is myopic. That is, each player acquires signals myopically at every history, on-path or off-path.

consists of acquiring signals myopically. Crucially, the optimal signal choices are independent of the game matrix itself (and of opponent strategies). We now illustrate the use of this corollary with two examples.

10.1 A Beauty Contest Game

Hellwig and Veldkamp (2009) introduced a beauty contest game with information acquisition. We build on this by modifying the information acquisition stage so that players *sequentially* acquire information over many periods (rather than once), and face a *capacity constraint* each period (rather than costly signals). We show that the basic insights of Hellwig and Veldkamp (2009) hold in this setting.

Specifically, suppose that at an unknown final period, a unit mass of players simultaneously chooses prices $p_i \in \mathbb{R}$ to minimize the (normalized) squared distance between their price and an unknown target price p^*, which depends on the unknown state ω and also on the average price $\overline{p} = \int p_i \, di$:

$$u_i(p_i, \overline{p}, \omega) = -\frac{1}{(1-r)^2} \cdot (p_i - p^*)^2 \quad \text{where } p^* = (1-r) \cdot \omega + r \cdot \overline{p} \tag{8}$$

The constant $r \in (-1, 1)$ determines whether pricing decisions are complements or substitutes.[58]

In every period up until the final period, each player acquires B signals from the set (X_k^i), as in the framework we have developed. To closely mirror the setup in Hellwig and Veldkamp (2009), we set each $\theta_1^i = \omega$. Assuming "conditional independence" of players' signals, we can directly apply Corollary 1 and conclude that in every equilibrium, players choose a deterministic (myopic) sequence of information acquisitions. This result echoes Hellwig and Veldkamp (2009) (section 1.3.4), who show that equilibrium is unique when players choose from *private* signals.[59] Relative to these authors, our extension is to introduce dynamics and show how the dynamic problem can be reduced into a static one, as we describe below.

Let $\Sigma(t)$ be the posterior variance of a myopic decision-maker about ω after the first t observations. Since the players in our model myopically acquire B signals per period, their (common) posterior variance at the end of t periods is given by $\Sigma(BT)$. Thus, conditional on

[58] When $r > 0$, best responses are increasing in the prices set by other players, thus decisions are complements. Conversely, $r < 0$ implies decisions are substitutes.

[59] Hellwig and Veldkamp (2009) also study a case in which players observe signals that are distorted by a common noise (which violates conditional independence). They show that multiple equilibria generally arise with such "public signals". Dewan and Myatt (2008) and Myatt and Wallace (2012) restore a unique linear symmetric equilibrium by assuming perfectly divisible signals, similar to the continuous-time variant of our model. In contrast, our equilibrium analysis relies on the informational environment (conditional independence), but not on symmetry (across the players) or linearity (of the best reply function).

period T being the final period, our game is as if the players acquire a batch of BT signals and then choose prices. This means that equilibrium prices are determined in the same way as in Hellwig and Veldkamp (2009):

$$p(\mathcal{I}^i_{\leq BT}) = \frac{1-r}{1-r+r\cdot\Sigma(BT)}\cdot\mathbb{E}(\omega|\mathcal{I}^i_{\leq BT}) \tag{9}$$

where $\mathcal{I}^i_{\leq BT}$ represents player i's information set, consisting of BT signal realizations.

We can use this characterization of equilibrium to re-evaluate the main insight in Hellwig and Veldkamp (2009): the incentive to acquire more informative signals is increasing in aggregate information acquisition if decisions are complements and decreasing if decisions are substitutes. For this purpose, we augment the model with a period 0, in which each player i invests in a capacity level B_i at some cost. Afterwards, players acquire information myopically (under possibly differential capacity constraints) and participate in the beauty contest game.

Let $\mu \in \Delta(\mathbb{Z}^+)$ be the distribution over capacity levels chosen by player i's opponents. Then, player i's expected utility from choosing capacity B_i is given by

$$EU(B_i,\mu) = -\mathbb{E}_{T\sim\pi}\left[\frac{\Sigma(B_iT)}{\left(1-r+r\cdot\int_B\Sigma(BT)\,d\mu(B)\right)^2}\right]. \tag{10}$$

Above, the expectation is taken with respect to the random final period T distributed according to π, while inside the expectation, the term $\int_B\Sigma(BT)\,d\mu(B)$ is the average posterior variance among the players. Similar to Proposition 1 in Hellwig and Veldkamp (2009), we have the following result:

Corollary 2. *Suppose $\hat{B}_i > B_i$ and $\hat{\mu} > \mu$ in the sense of first-order stochastic dominance. Then the sign of the difference $EU(B_i,\mu) + EU(\hat{B}_i,\hat{\mu}) - EU(B_i,\hat{\mu}) - EU(\hat{B}_i,\mu)$ is*

(a) zero, if there is no strategic interaction ($r = 0$);

(b) positive, if decisions are complementary ($r > 0$);

(c) negative, if decisions are substitutes ($r < 0$).

When decisions are complements, the value of additional information is increasing in the amount of *aggregate* information. Thus player i has a stronger incentive to choose a higher signal capacity if his opponents (on average) acquire more signals. This incentive goes in the opposite direction when decisions are substitutes, which confirms the main finding of Hellwig and Veldkamp (2009).

10.2 Strategic Trading

We consider the strategic trading game introduced in Lambert, Ostrovsky and Panov (2017), in which individuals trade given asymmetric information about the value of an asset. We endogenize the information available to traders by adding a pre-trading stage in which traders sequentially acquire signals. As in our main framework, we suppose that trading occurs at a final time period that is determined according to an arbitrary full-support distribution.

In more detail: at the final time period, a security with unknown value v is traded in a market, and each of n traders submits a demand d_i. There are additionally liquidity traders who generate exogenous random demand u. A market-maker privately observes a signal θ_M (possibly multi-dimensional) and the total demand $D = \sum_i d_i + u$. He sets the price $P(\theta_M, D)$, which in equilibrium equals $\mathbb{E}[v \mid \theta_M, D]$. Each strategic trader then obtains profit $\Pi_i = d_i \cdot (v - P(\theta_M, D))$.

We suppose that in each period up to and including the final time period, each trader i chooses to observe a signal from his set (X^i_k) (described above). We maintain all of the previous assumptions on the informational environment. The requirement of conditional independence is strengthened to apply to a payoff-relevant *vector* $\omega = (v, \theta_M, u)$ (instead of a real-valued unknown): that is, for each player i, conditional on the value of θ^i_1, the payoff-relevant vector ω and the other players' unknown states $(\theta^j)_{j\neq i}$ are assumed to be conditionally independent from player i's states θ^i. Relative to the fully general setting considered in Lambert, Ostrovsky and Panov (2017), this assumption allows for flexible correlation *within* a player's signals, but places a strong restriction on correlation *across* different players' signals. Applying Corollary 1, we can conclude that:

Corollary 3. *Under the above assumptions, there is an essentially unique linear NE in which the on-path signal acquisitions are myopic, and in the final period, players play the unique linear equilibrium described in Lambert, Ostrovsky and Panov (2017).*

Thus, the closed-form solutions that are a key contribution of Lambert, Ostrovsky and Panov (2017) extend to this setting with endogenous information.

11 Related Literature

11.1 Dynamic Information Acquistion

Our work builds on a literature about optimal information acquisition in dynamic environments: see e.g. Wald (1947), Arrow, Blackwell and Girshick (1949), Moscarini and Smith (2001), Steiner, Stewart and Matejka (2017), Ke, Shen and Villas-Boas (2016), Fudenberg, Strack and Strzalecki (2017), Che and Mierendorff (2017), Hebert and Woodford (2017), and Mayskaya (2017).

The main new feature in our problem is the introduction of flexible correlation across information sources. Our DM chooses between a limited number of correlated information sources, in contrast to the classic problem of choosing the precision of information (e.g. Moscarini and Smith (2001)). This question of how to optimally choose between different "kinds" of information is posed in the concurrent work of Fudenberg, Strack and Strzalecki (2017), Che and Mierendorff (2017) and Mayskaya (2017). As mentioned in Section 8, the continuous-time version of our model generalizes section 3.5 of Fudenberg, Strack and Strzalecki (2017) to allow for many normally-distributed states and signals that may be arbitrarily correlated. In contrast, Che and Mierendorff (2017) and Mayskaya (2017) consider a DM who allocates attention between Poisson sources that provide (noisy) evidence confirming one of a finite number of states. Whether Gaussian or Poisson uncertainty is more appropriate depends on the setting: the Poisson model captures lumpy information, while Gaussian learning is applicable when small amounts of information arrive frequently. Additionally, the two approaches are distinguished by the cardinality of the state space—Che and Mierendorff (2017) and Mayskaya (2017) assume (respectively) a binary and ternary state space, while we work with a continuous state space. The former approach is more suited to problems such as learning about whether a defendant is guilty or innocent, while the latter is more suited to learning about the (real-valued) return to an investment.

Acquisition of Gaussian signals whose means are (special) linear combinations of unknown states appears previously in the work of Sethi and Yildiz (2016) and also Meyer and Zwiebel (2007).[60] In particular, Sethi and Yildiz (2016) characterizes the long-run behavior of a DM who myopically acquires information from experts with independent biases. Additionally, Persico (2000), Hellwig and Veldkamp (2009), Myatt and Wallace (2012), and Lambert, Ostrovsky and Panov (2017) consider games in which players receive Gaussian information only once. As we demonstrated in Section 10, our results can be used to add certain kinds of sequential information acquisition to these models.

A key contribution of our paper, relative to this past work, is that we allow for arbitrary payoff functions, and demonstrate that the optimality of myopic information acquisition is independent of the objective. While previous dynamic models often focus on optimal stopping under geometric discounting (or linear waiting cost), our results do not rely on such specifications and extend also to situations in which the DM repeatedly takes actions at the same as he acquires information. As discussed in Section 7, this extension makes it possible to analyze a wider range of applications.

[60] Sethi and Yildiz (2016) inspired our benchmark model in Section 1.

11.2 The Value of Information

Since our decision-maker compares signals in every period, our paper connects also to a large body of work about the value of information. Blackwell (1951)'s classic work provides a partial ordering over signals corresponding to when a signal is more valuable than another in every decision problem. Subsequent work extended this partial ordering by restricting to certain classes of decision problems: for example, Lehmann (1988) and Persico (2000). Athey and Levin (2017) characterize the value of information for "monotone decision problems", and Cabrales, Gossner and Serrano (2013) characterize the value of information for a class of investment problems. We take the opposite approach of restricting the class of information structures (considering only normal-linear signals), while allowing for general and dynamic decision problems. Our equivalence results can be interpreted as showing that dynamics need not alter the static ordering of signals, and that this property is independent of the decision problem.

11.3 Statistics

Multi-armed Bandits. Our setting does not in general fall into the Multi-armed Bandit (MAB) framework for which the classic work of Gittins (1979) applies (see also the survey of Bergemann and Välimäki (2008)). The primary distinction is that in a MAB problem, learning takes place through realized flow payoffs. In our model however, the DM's action does not influence his information acquisition. Because of this, our (totally optimal) problem is close to the recent literature on *best-arm-identification*, where a DM samples for a number of periods before selecting an arm and receiving its payoff (see Bubeck, Munos and Stoltz (2009)). In fact, our results for two states ($K = 2$) exactly apply to best-arm-identification problems with two arms and correlated normal-linear signals,[61] although we cannot handle more than two arms.[62]

Optimal Design. Our work is closely related to the field of optimal design, initiated by the the early work of Robbins (1952) (see Chernoff (1972) for a survey). Under our total optimality criterion, the DM simultaneously chooses t observations to achieve the most

[61] To see this, let μ_1, μ_2 be the unknown expected payoffs of the two arms. Then the DM only cares about the difference in means: $\theta_1 = \mu_1 - \mu_2$, which fits our assumption of a *single payoff-relevant state*. We are aware of very few papers that characterize the optimal policy for correlated bandits; see Mersereau, Rusmevichientong and Tsitsiklis (2009) for the analysis of a special correlation structure.

[62] Our analysis can handle settings in which: (a) the payoff function is arbitrary, but it depends on a single state; and (b) the payoff function depends on multiple states, but it takes the special form of (weighted sum of) squared prediction errors. While the two-arm case is included into (a), the setting with more than two arms would involve *multiple payoff-relevant states* and *a decision problem that is not prediction*. Our equivalence results would not apply.

accurate belief. This can be viewed as an optimal design problem with respect to the "c-optimality criterion", which seeks to minimize the variance of an unknown parameter. A crucial difference is that our DM has access to a *pre-defined* set of coefficient vectors $(c_{i1}, \ldots, c_{iK})$. This distinguishes our model from most work in this literature, which assumes fully flexible coefficients.

12 Conclusion

Characterizing the optimal strategy for dynamic information acquisition is challenging in many settings. Restrictions commonly imposed for tractability include: parametric assumptions about the decision problem and the discounting structure; separation in time between information acquisitions and actions; and the assumption that information sources are uncorrelated (or have a specific correlation structure). Even with these restrictions, the optimal solution often cannot be explicitly characterized.

We show that many of these limitations can be lifted by considering environments with Gaussian signals. The setting that we propose and analyze is the following: a decision-maker has access to Gaussian signals that exhibit arbitrary correlation; in each period he acquires a fixed number of signals, and at a final time period he chooses an action based on the information acquired so far. We provide sufficient conditions on the informational primitives such that the myopic sequence of signal acquisitions is exactly optimal, thus permitting simple characterization of forward-looking behavior. Generically, myopic signal acquisitions are optimal at sufficiently late periods, permitting exact analysis of long-run behavior. These results require no additional parametric assumptions on the decision problem and extend also to contemporaneous action choices (including endogenous stopping problems).

Conceptually, our results demonstrate a class of environments in which myopic decision making turns out to have strong robustness and optimality properties. This challenges the conventional understanding that forward-looking information acquisition often requires great sophistication. While our proof techniques in this paper rely on the assumption of normality, we believe that qualitative features of our results extend for "approximately normal" environments, which would emerge naturally in settings where each acquisition consists of large number of non-normal signals. We leave verification of this conjecture to future work.

Methodologically, our results simplify the analysis of optimal dynamic information acquisition in an informational environment that is commonly studied in economics. We demonstrate how existing papers that consider normal-linear signals can be extended to accommodate dynamic information acquisition. Finally, several extensions are of interest: specifically, the presence of a greater number of sources than unknown states (over-identification), and the possibility of evolving states.

A Preliminary Results

We begin by presenting a number of preliminary results that are used throughout the appendix. The first two lemmas below characterize the function f mentioned in the main text, which maps signal counts to the DM's posterior variance about the payoff-relevant state θ_1.

Lemma 4. *Given prior covariance matrix V^0 and q_i observations of each signal i, the DM's posterior variance about θ_1 is given by*

$$f(q_1, \ldots, q_K) = [V^0 - V^0 C' \Sigma^{-1} C V^0]_{11} \tag{11}$$

where $\Sigma = CV^0C' + D^{-1}$ and $D = \operatorname{diag}\left(\frac{q_1}{\sigma_1^2}, \ldots, \frac{q_K}{\sigma_K^2}\right)$. The function f is decreasing and convex in each q_i whenever these arguments take non-negative extended real values: $q_i \in \overline{\mathbb{R}_+} = \mathbb{R}_+ \cup \{+\infty\}$.

Proof. The expression (11) comes directly from the conditional variance formula for multivariate Gaussian distributions. To prove $\frac{\partial f}{\partial q_i} \leq 0$, consider the partial order $\succeq$ on positive semi-definite matrices so that $A \succeq B$ if and only if $A - B$ is positive semi-definite. As q_i increases, the matrices D^{-1} and Σ decrease in this order. Thus Σ^{-1} increases in this order, which implies that $V^0 - V^0C'\Sigma^{-1}CV^0$ decreases in this order. In particular, the diagonal entries of $V^0 - V^0C'\Sigma^{-1}CV^0$ are uniformly smaller, so that f becomes smaller. Intuitively, more information always improves the decision-maker's estimates.

To prove f is convex, it suffices to prove f is *midpoint-convex* since the function is clearly continuous. Take $q_1, \ldots, q_K,\ r_1, \ldots, r_K \in \overline{\mathbb{R}_+}$ and let $s_i = \frac{q_i + r_i}{2}$.[63] Define the corresponding diagonal matrices to be D_q, D_r, D_s. We need to show $f(q_1, \ldots, q_K) + f(r_1, \ldots, r_K) \geq 2f(s_1, \ldots, s_K)$. For this, we first use the Woodbury inversion formula to write

$$\Sigma^{-1} = (CV^0C' + D^{-1})^{-1} = J - J(J + D)^{-1}J,$$

with $J = (CV^0C')^{-1}$. Plugging this back into (11), we see that it suffices to show the following matrix order:

$$\frac{(J + D_q)^{-1} + (J + D_r)^{-1}}{2} \succeq (J + D_s)^{-1}.$$

Inverting both sides, we need to show $2\left((J + D_q)^{-1} + (J + D_r)^{-1}\right)^{-1} \preceq J + D_s$. From definition, $D_q + D_r = \operatorname{diag}(\frac{q_1 + r_1}{\sigma_1^2}, \ldots, \frac{q_K + r_K}{\sigma_K^2}) = 2D_s$. Thus the above follows from the AM-HM inequality for positive definite matrices, see for instance Ando (1983). $\square$

A.1 The Matrix Q_i

Let us define for each $1 \leq i \leq K$,

$$Q_i = C^{-1} \Delta_{ii} C'^{-1} \tag{12}$$

where Δ_{ii} is the matrix with '1' in the (i, i)-th entry, and zeros elsewhere. We note that $[Q_i]_{11} = ([C^{-1}]_{1i})^2$, which is strictly positive under Assumption 3.

[63] We allow the function f to take $+\infty$ as arguments. This relaxation does not affect the properties of f, and it is convenient for our future analysis.

A.2 Order Difference Lemma

In this subsection we establish the asymptotic order for the second derivatives of f.

Lemma 5. *As $q_1, \ldots, q_K \to \infty$, $\frac{\partial^2 f}{\partial q_i^2}$ is positive with order $\frac{1}{q_i^3}$, whereas $\frac{\partial^2 f}{\partial q_i \partial q_j}$ has order at most $\frac{1}{q_i^2 q_j^2}$ for any $j \neq i$. Formally, there is a positive constant L depending on the informational environment, such that $\frac{\partial^2 f}{\partial q_i^2} \geq \frac{1}{L q_i^3}$ and $|\frac{\partial^2 f}{\partial q_i \partial q_j}| \leq \frac{L}{q_i^2 q_j^2}$.*

To interpret, the second derivative $\partial^2 f / \partial q_i^2$ is the effect of observing signal i on the marginal value of the next observation of signal i. Our lemma says that this second derivative is always eventually positive, so that each observation of signal i makes the next observation of signal i less valuable. The cross-partial $\partial^2 f / \partial q_i \partial q_j$ is the effect of observing signal i on the marginal value of the next observation of a different signal j, and its sign is ambiguous.

The key content of the lemma is that regardless of the sign of the cross partial, it is always of lower order compared to the second derivative. In words, the effect of observing a signal on the marginal value of other signals (as quantified by the cross-partial) is eventually second-order to its effect on the marginal value of further realizations of the same signal (as quantified by the second derivative). This is true for any signal path in which the signal counts $q_1, \ldots, q_K$ go to infinity proportionally, which is guaranteed by Proposition 3 below.

Proof. Recall from Lemma 4 that $f(q_1, \ldots, q_K) = [V^0 - V^0 C' \Sigma^{-1} C V^0]_{11}$ and therefore

$$\frac{\partial^2 f}{\partial q_i \partial q_j} = [\partial_{ij}(V^0 - V^0 C' \Sigma^{-1} C V^0)]_{11} \qquad \frac{\partial^2 f}{\partial q_i^2} = [\partial_{ii}(V^0 - V^0 C' \Sigma^{-1} C V^0)]_{11} \tag{13}$$

Using properties of matrix derivatives,

$$\partial_{ii}(\Sigma^{-1}) = \Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1} - \Sigma^{-1}(\partial_{ii}\Sigma)\Sigma^{-1} + \Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}. \tag{14}$$

The relevant derivatives of the covariance matrix Σ are

$$\partial_i \Sigma = -\frac{\sigma_i^2}{q_i^2}\Delta_{ii} \qquad \partial_{ii}\Sigma = \frac{2\sigma_i^2}{q_i^3}\Delta_{ii}$$

Plugging these into (14), we obtain $\partial_{ii}(\Sigma^{-1}) = -\frac{2\sigma_i^2}{q_i^3}(\Sigma^{-1}\Delta_{ii}\Sigma^{-1}) + O\left(\frac{1}{q_i^4}\right)$. Thus by (13),

$$\frac{\partial^2 f}{\partial q_i^2} = \left[-V^0 C' \cdot \frac{\partial^2(\Sigma^{-1})}{\partial q_i^2} \cdot C V^0\right]_{11} = \frac{2\sigma_i^2}{q_i^3} \cdot \left[V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0\right]_{11} + O\left(\frac{1}{q_i^4}\right). \tag{15}$$

As $q_1, \ldots, q_k \to \infty$, $\Sigma \to C V^0 C'$ which is symmetric and non-singular. Thus the matrix $V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0$ converges to the matrix Q_i defined earlier in (12). From (15) and $[Q_i]_{11} > 0$, we conclude that $\frac{\partial^2 f}{\partial q_i^2}$ is positive with order $\frac{1}{q_i^3}$. Similarly, for $i \neq j$, we have

$$\partial_{ij}(\Sigma^{-1}) = \Sigma^{-1}(\partial_j \Sigma)\Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1} - \Sigma^{-1}(\partial_{ij}\Sigma)\Sigma^{-1} + \Sigma^{-1}(\partial_i \Sigma)\Sigma^{-1}(\partial_j \Sigma)\Sigma^{-1}.$$

The relevant derivatives of the covariance matrix Σ are

$$\partial_i \Sigma = -\frac{\sigma_i^2}{q_i^2}\Delta_{ii} \qquad \partial_j \Sigma = -\frac{\sigma_j^2}{q_j^2}\Delta_{jj} \qquad \partial_{ij}\Sigma = \mathbf{0}$$

From this it follows that $\partial_{ij}(\Sigma^{-1}) = O\left(\frac{1}{q_i^2 q_j^2}\right)$. The same holds for $\frac{\partial^2 f}{\partial q_i \partial q_j}$ because of (13), completing the proof of the lemma. $\square$

A.3 The Myopic DM Never Gets Stuck

The following technical lemma will be used to show that the myopic strategy observes each signal infinitely often (see the proof of Prop 1 Part (a) below). As mentioned in the main text, proving the analogous result for the (possibly stochastic) dynamically optimal strategy requires something stronger, namely Assumption 4.

Lemma 6. *For $q_1, \dots, q_K \in \overline{\mathbb{R}_+}$, $\partial_i f(q_1, \dots, q_K) = 0, \forall i$ if and only if $q_1 = \cdots = q_K = +\infty$.*

Proof. From the proof of Lemma 5, we have in general

$$\partial_i f = -\frac{\sigma_i^2}{n_i^2} \cdot [V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0]_{11}. \tag{16}$$

Suppose that each $\partial_i f$ is zero, and $q_i = +\infty$ for a proper subset I of signals. Then for any $j \notin I$, it holds that $[V^0 C' \Sigma^{-1} \Delta_{jj} \Sigma^{-1} C V^0]_{11} = 0$. Let v denote the first row vector of $V^0 C' \Sigma^{-1}$, then $v_j = 0$ for any $j \notin I$. Thus

$$v\Sigma = v(CV^0C' + D^{-1}) = vCV^0C' + vD^{-1} = vCV^0C'$$

where the last equality is because $v_j = 0$ whenever $j \notin I$, while $D^{-1} = \operatorname{diag}\left(\frac{\sigma_1^2}{q_1}, \dots, \frac{\sigma_K^2}{q_K}\right)$ is zero on those rows i with $i \in I$. Recall that we defined $v = e_1 V^0 C' \Sigma^{-1}$. Hence from the preceding display,

$$e_1 = v\Sigma(V^0C')^{-1} = vCV^0C'(V^0C')^{-1} = vC.$$

That is, v is the first row of C^{-1}. By Assumption 3, all coordinates of v are non-zero. Thus $I = [K]$, proving the desired statement. $\square$

We note that $\partial_i f = 0$ could happen for some signal i, so that f is not necessarily strictly decreasing in its arguments.[64] The content of Lemma 6 is to show at every history, there is some signal that provides positive marginal value.[65] In contrast, the stronger Assumption 4 described in the main text requires the same signal to have positive value at every history.

B Proofs in Section 4 (Reduction)

Proof of Lemma 2. Following the discussion in the main text, we only need to show that the signal with greatest immediate decrease in variance dominates every other signal in the Blackwell sense. Consider any signal i that yields posterior variance $[V^t]_{11}$ about θ_1 with $V^t = \phi_i(V^{t-1})$. We recall that the DM's distribution of posterior beliefs about θ_1 is $\theta_1 \sim \mathcal{N}(\mu_1^t, [V^t]_{11}$, with the posterior mean μ_1^t randomly drawn from $\mathcal{N}(\mu_1^{t-1}, [V^{t-1}]_{11} - [V^t]_{11})$. It is easily checked that the same distribution of posterior beliefs is generated if instead the DM observes the following signal:

$$\tilde{X} = \theta_1 + \epsilon_X; \qquad \epsilon_X \sim \mathcal{N}\left(0, \frac{[V^{t-1}]_{11} \cdot [V^t]_{11}}{[V^{t-1}]_{11} - [V^t]_{11}}\right), \qquad \epsilon_X \perp \theta_1.$$

[64]Suppose $K = 2$, $V^0 = \mathbf{I}_2$, $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, and signal variances are 1. Then $\partial_1 f(q_1, q_2) = 0$ iff $(ad - bc)dq_2 + a = 0$. This occurs when $a = d = \frac{1}{3}, b = c = \frac{2}{3}$ and $q_2 = 3$. In such an environment, a DM who has observed signal 2 three times does not benefit from signal 1 (until he observes the next signal 2).

[65]This relies on normal-linear signals, see Appendix J.3.

It is then clear that a signal with larger posterior variance $[V^t]_{11}$ corresponds to the noise term ϵ_X having larger variance, which is necessarily a *garbling* according to Blackwell. This proves that the myopic signal choice for prediction is myopic for any decision problem.

To prove the converse, we need to show that a signal with *strictly* larger posterior variance leads to *strictly* lower current-period flow payoff. For this, consider a pair of signals i, j with $[\phi_i(V^{t-1})]_{11} < [\phi_j(V^{t-1})]_{11}$. These signals are equivalent (in terms of the induced distribution of beliefs about θ_1) to $\tilde{X} = \theta_1 + \epsilon_X$ and $\tilde{Y} = \theta_1 + \epsilon_Y$ as defined above, with the noise term ϵ_X having smaller variance than ϵ_Y. Then $\tilde{Y}$ is further equivalent to $\tilde{Z} = \theta_1 + \epsilon_X + \epsilon_Z$, with ϵ_Z a Gaussian noise independent from θ_1 and ϵ_X. This analysis shows, as we mentioned, that $\tilde{Z} = \tilde{X} + \epsilon_Z$ is a garbled signal of $\tilde{X}$. A DM observing any realization of $\tilde{X}$ can randomly draw ϵ_Z and take the optimal action according to the resulting value of $\tilde{Z}$. By payoff sensitivity (Assumption 2), there is no single action that is optimal for all realizations of $\tilde{Z}$. Thus, by taking appropriate *mixed* actions (with the same support) upon any signal realization, a DM receiving signal i can achieve the same expected payoff as another DM receiving signal j. But payoff sensitivity implies that pure actions do even better, completing the proof. □

Proof of Lemma 3. We will show that starting with *any* prior covariance matrix V^0, a strategy S is totally optimal at time t if and only if the realized division $q^S(t)$ is always a totally optimal division defined with respect to this prior (minimizing the posterior variance about θ_1 after t periods).

First consider the "if" part. Take any strategy S that induces totally optimal divisions. We need to show S yields a weakly higher flow payoff in period t than every other strategy S'. We prove this by induction on t. When $t = 1$, total optimality reduces to myopic optimality, and the claim follows from Lemma 2.

Suppose the result holds for $t - 1$. With t periods, we view any strategy S' following the prior (μ^0, V^0) as consisting of two parts: a signal choice i in the first period, and a family of contingent strategies following the new belief $(\mu^1, \phi_i(V^0))$. Applying the induction hypothesis to every such belief, we see that the expected payoff (in period t) of any contingent strategy is no more than a contingent strategy that observes each signal j for $\hat{q}_j$ periods, independent of signal realizations. Here $(\hat{q}_1, \dots, \hat{q}_K)$ is a totally optimal division for $t - 1$ periods (from period 2 to period t), defined with respect to the new prior covariance matrix $\phi_i(V^0)$. Equivalently, $\hat{q}$ is a division that minimizes $f(\hat{q}_i + 1, \hat{q}_{-i})$ subject to $\hat{q}_1 + \cdots + \hat{q}_K = t - 1$.

In such a way, we have found a *deterministic* strategy[66] $\hat{S}'$ that yields a weakly higher payoff in period t than the strategy S'. By similar reasoning, we can find a deterministic strategy $\hat{S}$ that yields the same period-t payoff as S, but $\hat{S}$ induces a deterministic totally optimal division at time t.[67] Now observe that $\hat{S}$ and $\hat{S}'$ are both deterministic strategies, thus a DM using either strategy is equivalently choosing a collection of t signals to observe, and sequentiality does not matter. By definition of totally optimal divisions, $\hat{S}$ induces lower variance than $\hat{S}'$ in period t. We can thus invoke the Blackwell ordering argument to conclude that $\hat{S}$ yields a weakly higher payoff than $\hat{S}'$ in period t. This implies S is better than S', completing the induction.

An analogous inductive argument proves the "only if" part; that is, S is totally optimal at time t only if the realized division $q^S(t)$ is always a totally optimal division. □

[66]The signal choice in the first period is always deterministic, and by construction later signal choices are also deterministic, not depending on signal realizations.

[67]Because S induces totally optimal divisions, its contingent strategies must induce totally optimal divisions at time $t-1$ (defined with respect to the new prior). By induction hypothesis, these contingent strategies maximize payoff in period t. This payoff is the same as if the contingent strategy does not condition on signal realizations.

C Proofs in Section 5 (Immediate Equivalence)

We first introduce a notion of "total optimality following a given history":

Definition 6. *Fix a history h with length H, where each signal i has been observed H_i times. A division over signals is constrained totally optimal at time $t \geq H$ (following history h), if*

$$(n_1, \dots, n_K) \in \underset{q_i \geq H_i, \sum_i q_i = t}{\operatorname{argmin}} f(q_1, \dots, q_K).$$

That is, the division minimizes the DM's posterior variance about θ_1 at time t among all divisions "reachable from history h". We write any such division as $n^h(t)$.

When h is the null history, this definition reduces to (unconstrained) total optimality as defined in the main text. In general, Lemma 3 implies that following history h, a continuation strategy maximizes the flow payoff in period t iff it always induces a constrained totally optimal division.[68]

Proof of Proposition 1. Suppose the informational environment is separable. Similar to Lemma 1 for our benchmark case, we claim that after a one-shot deviation from the myopic rule, the posterior variances along the deviation path are uniformly larger than along the original myopic path. Once this is proved, we can conclude that any myopic strategy achieves constrained total optimality following any given history, for the prediction problem. By our reduction results, myopic is constrained totally optimal for any decision problem and it is thus dynamically optimal.

Take a signal path $h = (s_1, s_2, \dots)$ that follows the myopic rule starting at time $\underline{t}$. Consider a deviation path $\tilde{h} = (\tilde{s}_1, \tilde{s}_2, \dots)$ that observes some signal $i \neq s_{\underline{t}}$ in period $\underline{t}$ but subsequently follows the myopic rule. Let $\bar{t}$ be the first period after $\underline{t}$ such that $s_{\bar{t}} = i$. We will show that in any period $t \in (\underline{t}, \bar{t}]$, $\tilde{s}_t = s_{t-1}$, so that the deviation path attempts to "catch up" with the original myopic path.

To see this, we use induction on t. Suppose we have shown that the deviation path up to time $t-1$ is

$$\tilde{h}^{t-1} = (s_1, \dots, s_{\underline{t}-1}, i, s_{\underline{t}}, \dots, s_{t-2})$$

while the original myopic path up to time $t-1$ satisfies

$$h^{t-1} = (s_1, \dots, s_{\underline{t}-1}, s_{\underline{t}}, \dots, s_{t-1}).$$

Let $j = s_{t-1}$ and $(q_1, \dots, q_K)$ be the myopic division at time $t-2$. Then myopic optimality at time $t-1$ implies

$$f(q_j + 1, q_{-j}) \leq f(q_k + 1, q_{-k}), \forall k \in \{1, \dots, K\}.$$

Using $f(q_1, \dots, q_K) = F(g_1(q_1) + \dots + g_K(q_K))$ and the monotonicity of F, we can rewrite the above as

$$g_j(q_j + 1) - g_j(q_j) \leq g_k(q_k + 1) - g_k(q_k), \forall k.$$

This implies $g_j(q_j + 1) - g_j(q_j) \leq g_k(q_k + 1) - g_k(q_k), \forall k \neq i$ and $g_j(q_j + 1) - g_j(q_j) \leq g_i(q_i + 2) - g_i(q_i + 1)$ by the convexity of g_i. Now observe that the deviation path has division $(q_i + 1, q_{-i})$ at time $t-1$. Thus, the previous inequalities imply that signal j is the myopic choice at history $\tilde{h}^{t-1}$, completing the characterization of the deviation path. With this, we can apply the same exchangeability argument as in the proof of Lemma 1. Hence the Proposition follows. □

[68] We use the convention that the induced division of a continuation strategy includes the signals observed in the initial history. This simplifies notation in the sequel.

Proof of Proposition 2. Suppose the informational environment is symmetric. We claim that at any time t, a division $n(t)$ is totally optimal if and only if $|n_i(t) - n_j(t)| \leq 1$ holds for every pair of signals i, j. Obviously, the divisions that have this property (and $\sum_i n_i(t) = t$) are symmetric (as tuples) to one another. Hence they achieve the same payoff, and it suffices to prove the "only if" part of the claim.

Consider any division $(q_1, \ldots, q_K)$ with $q_1 \geq \cdots \geq q_K$. We will prove that if $q_1 - q_K \geq 2$, then this division is *not* totally optimal. By symmetry and convexity of f, we have

$$f(q_1, q_2, \ldots, q_{K-1}, q_K) = f(q_K, q_2, \ldots, q_{K-1}, q_1) \geq f(q_1 - 1, q_2, \ldots, q_{K-1}, q_K + 1).$$

because the vector $(q_1 - 1, \ldots, q_K + 1)$ is a convex combination of the vectors $(q_1, \ldots, q_K)$ and $(q_K, \ldots, q_1)$. Using Lemma 6, we can show the inequality here must be strict.[69] The claim follows.

From this characterization of totally optimal divisions, we see that for every totally optimal division $n(t)$, there exists some totally optimal division $n(t+1)$ with $n_i(t+1) \geq n_i(t), \forall i$. Hence a myopic DM can achieve total optimality at every time, and so he will. This proves that the myopic strategy maximizes the ex-ante payoff.

We can further prove the optimality of the myopic strategy following any history. The argument is essentially the same: given any history h consisting of H_i observations of each signal i, we can characterize the constrained totally optimal divisions. Specifically, a division $n^h(t)$ is constrained totally optimal if and only if it has the following property: for any pair of signals i, j, if $n_i^h(t) > H_i$, then $n_i^h(t) - n_j^h(t) \leq 1$. These constrained totally optimal divisions are again monotonic over time, proving that any myopic strategy is constrained totally optimal and dynamically optimal. □

D Asymptotic Characterization

An important step toward proving our equivalence results is to show that the signal counts grow to infinity proportionally, under any of the three optimality criteria.

Proposition 3. *Suppose the informational environment $(V^0, C, \{\sigma_i^2\})$ satisfies Assumption 3 and 4. Then there exist constants $\lambda_1, \ldots, \lambda_K > 0$ with $\sum_i \lambda_i = 1$ and a large constant N such that*

(a) $|n_i(t) - \lambda_i t| \leq N, \forall i$.

(b) $|m_i(t) - \lambda_i t| \leq N, \forall i$.

(c) $|d_i(t) - \lambda_i t| \leq N, \forall i$ *for every realized division* $d(t)$.

The constant N only depends on the informational environment but not on the decision problem. The asymptotic proportions $\lambda_1, \ldots, \lambda_K$ are given by

$$\lambda_i = \frac{|[C^{-1}]_{1i}| \cdot \sigma_i}{\sum_{j=1}^{K} |[C^{-1}]_{1j}| \cdot \sigma_j}. \tag{17}$$

[69] Suppose equality holds. By convexity, $f(q_1 - \epsilon, q_2, \ldots, q_{K-1}, q_K + \epsilon)$ is a constant c for $\epsilon \in [0, 1]$. Because f is a rational function (quotient of polynomials), this constant value extends to all $\epsilon \in \mathbb{R}$. Letting $\epsilon \to +\infty$, we deduce $f(-\infty, q_2, \ldots, q_{K-1}, +\infty) = c$. Hence $f(+\infty, q_2, \ldots, q_{K-1}, +\infty) = c$ also holds, because the Σ matrix for $q_1 = +\infty$ is the same as for $q_1 = -\infty$. Thus $f(q_1, q_2, \ldots, q_{K-1}, q_K) = f(+\infty, q_2, \ldots, q_{K-1}, +\infty)$. By the monotonicity of f, this implies $f(\hat{q}_1, q_2, \ldots, q_{K-1}, \hat{q}_K) = c$ whenever $\hat{q}_1 \geq q_1$ and $\hat{q}_K \geq q_K$. By the rational function argument again, this constant value extends to all $\hat{q}_1$ and $\hat{q}_K$. Thus $f(q_1, q_2, \ldots, q_{K-1}, q_K) = f(0, q_2, \ldots, q_{K-1}, 0)$. By Lemma 6, there exists a signal $i \in \{2, \ldots, K-1\}$ such that $\partial_i f(0, q_2, \ldots, q_{K-1}, 0)$ is strictly negative. Without loss assume $i = 2$, then $f(0, q_2 + q_1 + q_K, q_3, \ldots, q_{K-1}, 0) < f(0, q_2, \ldots, q_{K-1}, 0) = f(q_1, q_2, \ldots, q_{K-1}, q_K)$, contradicting total optimality.

Below we present the proofs for the first two parts of this proposition. The corresponding result for dynamically optimal strategies is proved later, in Appendix F.

Proof of Proposition 3 Part (a). Let us first show $n_1(t), \dots, n_K(t) \to +\infty$ as $t \to +\infty$. Suppose this is not true, then we can find a subsequence of times, such that for *each* signal j, $n_j(t)$ either remains constant or diverges to infinity as t increases along this subsequence, and the former ($n_j(t)$ is constant) occurs for some signal j.[70] Let q_j be the limit of $n_j(t)$ along this subsequence, then a proper subset of $q_1, \dots, q_K$ is equal to $+\infty$. By Lemma 6, there is some signal j with $\partial_j f(q_1, q_2, \dots, q_K) < 0$; thus q_j is finite in particular. Relabeling the signals if necessary, we assume $j = 1$. Let us further assume $q_2 = +\infty$ so that $\partial_2 f(q_1, q_2, \dots, q_K) = 0$. By continuity, the *discrete partial derivatives* of f satisfy the following inequality:[71]

$$\partial_1 f(n_1(t), n_2(t) - 1, \dots, n_K(t)) < \partial_2 f(n_1(t), n_2(t) - 1, \dots, n_K(t))$$

for sufficiently large t along this subsequence. But this implies

$$f(n_1(t) + 1, n_2(t) - 1, \dots, n_K(t)) < f(n_1(t), n_2(t), \dots, n_K(t)),$$

contradicting the assumption that $(n_1(t), n_2(t), \dots, n_K(t))$ is a totally optimal division.

Next, as each $n_i \to +\infty$, the matrix $V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0$ converges to the matrix Q_i defined above in (12). It follows from (16) that $\partial_i f \sim \frac{-\sigma_i^2}{n_i^2} \cdot [Q_i]_{11}$ (ratio converges to 1). Since a totally optimal division must satisfy $\partial_i f \sim \partial_j f$,[72] we deduce that n_i, n_j must grow proportionally. Using $[Q_i]_{11} = ([C^{-1}]_{1i})^2$, we deduce $n_i(t) \sim \lambda_i t$.

Finally, from $n_i \sim \lambda_i t$ we have $\Sigma = C V^0 C' + D^{-1} = C V^0 C' + O(\frac{1}{t})$. Thus the matrix $V^0 C' \Sigma^{-1} \Delta_{ii} \Sigma^{-1} C V^0$ converges to Q_i at the rate of $\frac{1}{t}$. From (16), we obtain $\partial_i f = \frac{-\sigma_i^2 \cdot [Q_i]_{11} + O(\frac{1}{t})}{n_i^2}$. Thus the first-order condition $\partial_i f = \partial_j f$ implies $\frac{\lambda_i^2 + O(\frac{1}{t})}{n_i^2} = \frac{\lambda_j^2 + O(\frac{1}{t})}{n_j^2}$.[73] This is equivalent to $\lambda_i^2 n_j^2 - \lambda_j^2 n_i^2 = O(t)$, which yields $\lambda_i n_j - \lambda_j n_i = O(1)$ after factorization. Hence $n_i = \lambda_i t + O(1)$. □

Proof of Proposition 3 Part (b). We turn to a myopic decision-maker. The first-order condition $\partial_i f = \partial_j f$ need not hold, but we do know that if signal i maximizes $|\partial_i f|$ at time t, it will be observed in the next period. This property allows us to use the same argument as in the preceding proof to show $m_i(t) \to +\infty, \forall i$. Furthermore, the approximation $\partial_i f \sim \frac{-\sigma_i^2}{m_i^2} \cdot [Q_i]_{11}$ implies that if $m_i(t) > (\lambda_i + \epsilon)t$, then $\partial_i f(m(t)) > \partial_j f(m(t))$ for the signal j with $m_j(t) < \lambda_j t$. Thus this signal i cannot be myopically observed in period $t+1$. This shows $m_i(t) \sim \lambda_i t$ for sufficiently large t.[74]

With $m_i(t) \sim \lambda_i(t)$, we have the better approximation $\partial_i f = \frac{-\sigma_i^2 \cdot [Q_i]_{11} + O(\frac{1}{t})}{m_i^2}$. Pick L to be a sufficiently large constant. Whenever $m_i(t) > \lambda_i t - 1$ and $m_j(t) < \lambda_j t - L$, we have $\partial_i f(m(t)) > \partial_j f(m(t))$ so that signal i is not observed in period $t+1$. The "-1" is chosen for later convenience.

[70] Either $n_j(t)$ remains bounded, or there is a subsequence that diverges to infinity. Moreover, a bounded sequence necessarily has a constant subsequence.

[71] Here and later, we will often abuse notation and let $\partial_i f$ also denote the discrete partial derivative of f: $\partial_i f(q_i, q_{-i}) = f(q_i + 1, q_{-i}) - f(q_i, q_{-i})$, which equals an integral of the usual continuous derivative of f over the interval $[q_i, q_i + 1]$. We will similarly abuse the second derivatives $\partial_{ii} f$ and $\partial_{ij} f$. Whether the discrete or the continuous derivative is used will be clarified in the context.

[72] Because we are doing discrete optimization, $\partial_i f$ and $\partial_j f$ need not be exactly equal. But they must be approximately equal.

[73] These partials need not be exactly equal, but the error terms that arise are on the order of $O(\frac{1}{t})$.

[74] We omit the detailed argument, which is similar to what we do in the next two paragraphs.

Define $z_k(t) = m_k(t) - \lambda_k t, \forall k$ and $Z(t) = \sum_k \frac{z_k^2(t)}{\lambda_k}$. These functions measure the discrepancy between the myopic division and its linear asymptote. Note that $\sum_k z_k(t) = 0$. We claim that $Z(t)$ is a bounded function. This is trivially true if each $z_k(t) \geq -L$. Suppose instead that $z_j(t) < -L$ for some j. Then by the analysis in the previous paragraph, the signal i that is observed in period $t+1$ must satisfy $m_i(t) \leq \lambda_i t - 1$. Under the myopic strategy, $m_i(t+1) = m_i(t)+1$ and $m_k(t+1) = m_k(t)$ for every $k \neq i$. Thus $z_i(t+1) = z_i(t) - \lambda_i + 1$, and $z_k(t+1) = z_k(t) - \lambda_k$ for $k \neq i$. From this, and using $\sum_k \lambda_k = 1, \sum_k z_k(t) = 0$, we can deduce after simplifications that

$$Z(t+1) = Z(t) + \frac{2z_i(t) - \lambda_i + 1}{\lambda_i}$$

Since $z_i(t) \leq -1$, $\lambda_i \in (0,1)$, we have $Z(t+1) < Z(t) - 1$. Hence $Z(t)$ and each $z_i(t)$ remains bounded, proving the proposition. □

E Proof of Theorem 1 (Eventual Gap of One)

We present the proof for equivalence between myopic and total optimality. The comparison with dynamic optimality follows similar arguments, with extra technicalities addressed in Appendix F.

Suppose for contradiction that $m_1(t) \leq n_1(t) - 2$ (the opposite case will be treated later). Since $\sum_{i=1}^K m_i(t) = t = \sum_{i=1}^K n_i(t)$, we can assume without loss $m_2(t) \geq n_2(t) + 1$. For notational ease, write $n_i = n_i(t), \forall\ 1 \leq i \leq K$. By total optimality of the division $(n_1, \dots, n_K)$ we have

$$f(n_1 - 1, n_2 + 1, \dots, n_K) \geq f(n_1, n_2, \dots, n_K).$$

Consider the last period $\tilde{t} \leq t$ in which a myopic decision-maker observed signal 2. Write $\tilde{m}_i = m_i(\tilde{t}), \forall i$. By assumption we have $\tilde{m}_1 \leq m_1(t) \leq n_1 - 2$, and $\tilde{m}_2 = m_2(t) \geq n_2 + 1$. Moreover, from Proposition 3, we know that $t - \tilde{t}$, and thus also each $|\tilde{m}_i - n_i|$, is bounded above by a constant independent of t. Let us show under these conditions that

$$\begin{aligned} &f(n_1 - 1, n_2 + 1, \dots, n_K) \geq f(n_1, n_2, \dots, n_K) \\ \Longrightarrow\ &f(\tilde{m}_1, \tilde{m}_2, \dots, \tilde{m}_K) > f(\tilde{m}_1 + 1, \tilde{m}_2 - 1, \dots, \tilde{m}_K). \end{aligned} \tag{18}$$

This will imply that the myopic decision-maker could have deviated to observing signal 1 rather than signal 2 at time $\tilde{t}$ and achieve a smaller posterior variance (at time $\tilde{t}$), yielding a contradiction.

To show (18), we first rewrite the (first-half) assumption as

$$\partial_2 f(n_1 - 1, n_2, \dots, n_K) \geq \partial_1 f(n_1 - 1, n_2, \dots, n_K), \tag{19}$$

where $\partial_i f$ denotes the *discrete partial derivative* with respect to signal i. We also rewrite the (second-half) conclusion in (18) as

$$\partial_2 f(\tilde{m}_1, \tilde{m}_2 - 1, \dots, \tilde{m}_K) > \partial_1 f(\tilde{m}_1, \tilde{m}_2 - 1, \dots, \tilde{m}_K). \tag{20}$$

Our goal is to show (19) implies (20). To do this, let us first compare the LHS of (20) to the LHS of (19). The difference can be rewritten as a sum of second derivatives:

$$(\tilde{m}_1 - n_1 + 1)\partial_{21} f + (\tilde{m}_2 - 1 - n_2)\partial_{22} f + \sum_{j>2} (\tilde{m}_j - n_j)\partial_{2j} f.$$

Since $\tilde{m}_2 \geq n_2 + 1$, the contribution of the second summand $(\tilde{m}_2 - 1 - n_2)\partial_{22} f$ is non-negative. Thus we deduce that that LHS of (20) is at least the LHS of (19) minus $O(\frac{1}{t^4})$, which captures the combined effects of cross partials (by Lemma 5 and Proposition 3).

On the other hand, the RHS of (20) differs from the RHS of (19) by

$$(\tilde{m}_1 - n_1 + 1)\partial_{11} f + (\tilde{m}_2 - 1 - n_2)\partial_{12} f + \sum_{j>2} (\tilde{m}_j - n_j)\partial_{1j} f$$

which is negative with order $O(\frac{1}{t^3})$ because of the first summand (recall $\tilde{m}_1 \leq n_1 - 2$). Hence we have shown that from (19) to (20), the RHS decreases by more than the LHS for t sufficiently large. Thus (19) implies (20) as desired, and we have ruled out $m_1(t) \leq n_1(t) - 2$.

Suppose instead that $m_1(t) \geq n_1(t) + 2$ and $m_2(t) \leq n_2(t) - 1$. Then we can take $\tilde{t}$ to be the last period in which a myopic decision-maker observed signal 1. A symmetric argument shows that the DM could have profitably deviated to observing signal 2 at time $\tilde{t}$. The theorem follows.

References

Ando, Tsuyoshi. 1983. "On the Arithmetic-Geometric-Harmonic-Mean Inequalities for Positive Definite Matrices." *Linear Algebra and its Applications*, 52: 31–37.

Arrow, K. J., D. Blackwell, and M. A. Girshick. 1949. "Bayes and Minimax Solutions of Sequential Decision Problems." *Econometrica*, 17(3): 213–244.

Athey, Susan, and Jonathan Levin. 2017. "The Value of Information in Monotone Decision Problems." Working Paper.

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

Bergemann, Dirk, and Juuso Välimäki. 2008. "Bandit Problems." In *The New Palgrave Dictionary of Economics.* , ed. Steven N. Durlauf and Lawrence E. Blume. Basingstoke:Palgrave Macmillan.

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.

Blackwell, David. 1951. "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability." Chapter Comparison of Experiments, 93–102. Berkeley and Los Angeles: University of California Press.

Bubeck, Sebastian, Remi Munos, and Gilles Stoltz. 2009. "Pure Exploration in Multi-armed Bandits Problems." In *Algorithmic Learning Theory. ALT 2009. Lecture Notes in Computer Science, vol 5809.* Springer, Berlin, Heidelberg.

Cabrales, Antonio, Olivier Gossner, and Roberto Serrano. 2013. "Entropy and the Value of Information for Investors." *American Economic Review*, 103(1): 360–377.

Chernoff, Herman. 1972. *Sequential Analysis and Optimal Design.* Society for Industrial and Applied Mathematics.

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

Dewan, Torun, and David P. Myatt. 2008. "The Qualities of Leadership: Direction, Communication and Obfuscation." *American Political Science Review*, 102(3): 351–368.

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

Gittins, J. C. 1979. "Bandit processes and dynamic allocation indices." *Journal of the Royal Statistical Society, Series B*, 148–177.

Hebert, Ben, and Michael Woodford. 2017. "Rational Inattention with Sequential Information Sampling." Working Paper.

Hellwig, Christian, and Laura Veldkamp. 2009. "Knowing What Others Know: Coordination Motives in Information Acquisition." *The Review of Economic Studies*, 76: 223–251.

Ke, T. Tony, Zuo-Jun Max Shen, and J. Miguel Villas-Boas. 2016. "Search for Information on Multiple Products." *Management Science*, 62(12): 3576–3603.

Lambert, Nicolas, Michael Ostrovsky, and Mikhail Panov. 2017. "Strategic Trading in Informationally Complex Environments." Working Paper.

Lehmann, E. L. 1988. "Comparing Location Experiments." *The Annals of Statistics*, 16(2): 521–533.

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

Mersereau, Adam J, Paat Rusmevichientong, and John N Tsitsiklis. 2009. "A Structured Multiarmed Bandit Problem and the Greedy Policy." *IEEE Transactions On Automatic Control*, 54(12): 2787–2802.

Meyer, Margaret, and Jeffrey Zwiebel. 2007. "Learning and Self-Reinforcing Behavior." Working Paper.

Moscarini, Guiseppe, and Lones Smith. 2001. "The Optimal Level of Experimentation." *Econometrica*, 69(6): 1629–1644.

Myatt, David P., and Chris Wallace. 2012. "Endogenous Information Acquisition in Coordination Games." *The Review of Economic Studies*, 79: 340–374.

Persico, Nicola. 2000. "Information Acquisition in Auctions." *Econometrica*, 68(1): 135–148.

Robbins, Herbert. 1952. "Some aspects of the sequential design of experiments." *Bull. Amer. Math. Soc.*, 58(5): 527–535.

Rudelson, Mark, and Roman Vershynin. 2008. "The Least Singular Value of a Random Matrix is $O(n^{-1/2})$." *C. R. Math. Acad. Sci. Paris*, 346: 893–896.

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

Steiner, Jakub, Colin Stewart, and Filip Matejka. 2017. "Rational Inattention Dynamics: Inertia and Delay in Decision-Making." *Econometrica*, 85(2): 521–553.

Tao, Terence, and Van Vu. 2010. "Random Matrices: the Distribution of the Smallest Singular Values." *Geometric and Functional Analysis*, 20(1): 260–297.

Wald, Abraham. 1947. "Foundations of a General Theory of Sequential Decision Functions." *Econometrica*, 15(4): 279–313.