Dynamic Information Acquisition from Multiple Sources*

Xiaosheng Mu† (Job Market Paper)
with Annie Liang‡ and Vasilis Syrgkanis§

January 27, 2018

Abstract

Decision-makers often aggregate information across many sources, each of which provides relevant information. We introduce a dynamic learning model where a decision-maker learns about unknown states by sequentially sampling from a finite set of Gaussian signals with arbitrary correlation. Such a setting describes sequential search between similar products, as well as reading news articles with correlated biases. We study the optimal sequence of information acquisitions. Assuming the final decision depends linearly on the states, we show that *myopic* signal acquisitions are nonetheless optimal at sufficiently late periods. For classes of informational environments that we describe, the myopic rule is 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.

JEL codes: C44, D81, D83

Keywords: Information Acquisition, Correlation, Endogenous Attention, Myopic Choice, Robustness, Value of Information

*We thank Yash Deshpande, Johannes Hörner, Michihiro Kandori, Elliot Lipnowski, Juuso Toikka, Anton Tsoy, Muhamet Yildiz and seminar audiences at Harvard University, Massachusetts Institute of Technology, Columbia University and the Pennsylvania Economic Theory Conference (2017) for helpful comments. We are especially grateful to Drew Fudenberg, Eric Maskin and Tomasz Strzalecki for their constant guidance and support.

†Department of Economics, Harvard University. xiaoshengmu@fas.harvard.edu.

‡Department of Economics, University of Pennsylvania. liang.annie.h@gmail.com

§Microsoft Research. vasy@microsoft.com.

1 Introduction

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 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 $\theta_1, \dots, \theta_K \in \mathbb{R}$, 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: in each discrete period, the DM chooses one signal to observe (see Section 8 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:

[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. Such a setup appears also in Wilson (2014).

(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 *t-optimal* 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 t-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 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 t-optimality, 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 t-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 7 sufficient conditions under which the myopic and dynamically optimal signal paths are identical *from period* 1, and moreover t-optimal for every 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 8, 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 t-optimal: the number of signals of each type acquired under the myopic, dynamically optimal and t-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 t-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 t-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.

A number of classical papers, including Easley and Kiefer (1988) and Aghion et al. (1991), consider more general "learning by experimentation" problems. These authors find that the DM's beliefs converge over time, and experimentation motives eventually vanish. Consequently, the incentive to "exploit" eventually becomes first-order important. Our result that myopic signal choices are eventually optimal is spiritually similar to theirs. However,

[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.

the mechanisms for these results are quite different. Easley and Kiefer (1988) and Aghion et al. (1991) demonstrate the following: if there is a unique myopically optimal policy at the limiting beliefs, then the optimal policies must converge to this policy. In our setting, every policy (signal choice) is trivially myopic at the limiting beliefs (a point mass at the true parameter), so we do not have uniqueness and cannot use this argument to identify long-run behavior.[3] We take the different approach of demonstrating "approximate separability" of the dynamic problem, which leads to myopic long-run behavior.

In Section 9, 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).

Section 10 relaxes the exogenous capacity constraint of (observing) the same number of signals per period. We allow the decision-maker to costly choose how many signal realizations to acquire at each moment. Our results extend and show that myopic signal choices remain (eventually) optimal in this more general setting.

Additional extensions of the model are presented in Section 11: our results extend to i.i.d. states drawn each period and to multiple payoff-relevant states whenever the decision problem is that of prediction. We also discuss a continuous-time analogue of our setup in which at every time, the DM chooses attention levels (subject to 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 12 and 13 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

[3]The distinction arises because in our pure learning environment, signal choice by itself does not generate any payoff. This distinction also applies to other dynamic learning papers, see Section 2 for a list. As far as we are aware, previous work in the learning literature do not discuss the (eventual) optimality of myopic information acquisition.

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 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 13 leverages this to extend two results in the literature for information acquisition in games (specifically, Hellwig and Veldkamp (2009) and Lambert, Ostrovsky and Panov (2017)) by allowing the players to acquire information over time.

2 Related Literature

2.1 Dynamic Information Acquisition

Our work builds on a literature about optimal information acquisition in dynamic environments: see e.g. Moscarini and Smith (2001), Fudenberg, Strack and Strzalecki (2017), Che and Mierendorff (2017), Mayskaya (2017), Steiner, Stewart and Matějka (2017), Hébert and Woodford (2017), Zhong (2017). Also closely related is a large literature on sequential search, including Wald (1947), Arrow, Blackwell and Girshick (1949), Weitzman (1979), Morgan and Manning (1985), Callander (2011), Klabjan, Olszwski and Wolinsky (2014), Olszwski and Weber (2015), Chick and Frazier (2012), Branco, Sun and Villas-Boas (2012), Ke, Shen and Villas-Boas (2016), Ke and Villas-Boas (2017), Sanjurjo (2017) and Doval (2017).

The key new feature in our problem is the introduction of flexible correlation structures across information sources.[4] In our main model, the 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 we

[4]Among the above papers, Callander (2011) also emphasizes the correlation between different signals (or search alternatives in his model). But the signals in Callander (2011) are related by a Brownian motion path, which yields a special correlation structure. His model is further studied in Garfagnini and Strulovici (2016) and in Bardhi (2017) under different assumptions on agent behavior and different objectives.

discuss in Section 11, 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 or rejecting a particular state. 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.

The extension of our model to intertemporal decisions (Section 9) bears resemblance to the problem considered in Steiner, Stewart and Matějka (2017). These authors also study the interaction between signal choices and actions, and the solution to their dynamic problem reduces to a series of static optimizations, similar to the role of the myopic strategy in our framework. Despite these analogies, a fundamental distinction is that Steiner, Stewart and Matějka (2017) follow the rational inattention literature (Sims (2003)) and allow arbitrary information to be acquired at entropic information costs. In similar spirit, Hébert and Woodford (2017) and Zhong (2017) adopt more general "posterior-separable" information costs to study optimal information acquisition in continuous time.[5] Compared to these papers, our DM has access to a prescribed set of Gaussian signals. In Section 10, we do allow the DM to also control the intensity of information acquisition by endogenously choosing how many signals to acquire in each period. But even in that extension, we assume that the incurred information cost is a function of the number of observations. This is analogous to Moscarini and Smith (2001) and is distinguished from the above papers that measure information cost based on belief changes.

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). In particular, Sethi and Yildiz (2016) characterizes the long-run behavior of a DM who myopically acquires information from experts with independent biases. Their work inspired our benchmark model in Section 3. Additionally, Hellwig and Veldkamp (2009), Myatt and Wallace (2012), Colombo, Femminis and Pavan (2014) and Lambert, Ostrovsky

[5] Like us, Hébert and Woodford (2017) restricts to Gaussian information. Zhong (2017) does not make this restriction and demonstrates the optimality of Poisson signals for binary state space.

and Panov (2017) consider games in which players receive Gaussian information only once.[6] As we demonstrate in Section 13, our results can be used to add certain kinds of sequential information acquisition to these strategic models.

A major contribution of our paper, relative to this past work, is that we allow for general payoff functions (depending on a one-dimensional state), and demonstrate that the optimality of myopic information acquisition is independent of the objective. In concurrent work, Bardhi (2017) considers a similar setting of dynamic information acquisition from correlated normal signals. Relative to Bardhi (2017), we mostly focus the analysis on characterizing the optimal strategy given a random or endogenous decision time, while Bardhi (2017) extends in a different direction, analyzing a principal-agent version of optimal search.

2.2 The Value of Information

Since our decision-maker compares signals in every period, our paper relates also to a literature 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 extends this partial ordering by restricting to certain classes of decision problems: for example, Lehmann (1988), Persico (2000), Cabrales, Gossner and Serrano (2013) and Athey and Levin (2017).

We focus on the family of normal signals and show that the notion of myopic optimality is invariant to any utility function that depends on θ_1 (Lemma 2). This insight that one-dimensional normal signals admit a complete Blackwell-ordering in the static setting appeared in Hansen and Torgersen (1974). But we take this insight further and establish dynamic Blackwell comparison for sequences of normal signals. Specifically, Lemma 3 shows that *for a fixed future decision time*, the sequence of signals that leads to the lowest posterior variance Blackwell-dominates any other sequence. Lemma 7 shows that an information acquisition strategy Blackwell-dominates another *for any intertemporal decision problem* if and only if it leads to lower posterior variance at every time.

These results on the comparison of sequential normal experiments generalize the main result in Greenshtein (1996). To explain the connection, Greenshtein (1996) compares two *deterministic* sequences of signals, where each signal is θ_1 plus *independent* normal noise. His Theorem 3.1 implies that the former sequence is Blackwell-dominant if and only if its cumulative precision is higher at every time. Note that this statement does not refer to the

[6] For games of information acquisition beyond the Gaussian setting, see e.g. Persico (2000), Bergemann and Välimäki (2002), Yang (2015) and Denti (2017). All of these papers restrict to a single signal choice.

prior beliefs, but if we impose a normal prior on θ_1, higher cumulative precision is equivalent to lower posterior variance. Thus, the result of Greenshtein (1996) would coincide with ours if θ_1 is the only persistent state, and if all signals are independent conditional on θ_1. Nonetheless, our setting features additional correlation (across different signals) through the persistent states $\theta_2, \ldots, \theta_K$. Consequently, dynamic Blackwell comparison in our model necessarily depends on the prior beliefs.[7] This feature, together with the endogenous choice of signals (which may not be deterministic), complicates our problem relative to Greenshtein (1996).[8] Nonetheless, our proof of Lemma 7 bears similarity to Greenshtein (1996) in that we also construct "sequential Markov kernels." The difference is that we work with posterior beliefs rather than signal realizations.[9]

The comparison of sequential experiments allows us to simplify the analysis for a general decision problem to the prediction problem. However, even with this reduction, the question remains as to whether a sequence of signals exists that is best in terms of induced posterior variances. Our results show that dynamics need not alter the static ordering of signals, and such a signal sequence does (eventually) exist.

2.3 Statistics

Multi-armed Bandits. Our setting does not fall into the classic Multi-armed Bandit (MAB) framework, see Gittins (1979) and the survey of Bergemann and Välimäki (2008). The primary distinction is that in a MAB problem, learning takes place through realized flow payoffs.[10] Our model shuts down this feedback channel and assumes that payoff is only realized at the end. In this sense, what we study is a sequential search problem with *endogenous choice of information.*

Starting with Bubeck, Munos and Stoltz (2009), a recent literature studies so-called "best-arm identification," where a DM samples for a number of periods before selecting an arm and receiving its payoff.[11] This setup falls under our t-optimal problem, and our results

[7] This is already the case for static comparisons: as the prior beliefs vary, it is not always the same signal that leads to the lowest posterior variance about θ_1.

[8] We show that myopic and t-optimal strategies are indeed deterministic, but dynamically optimal strategies could in general depend on signal realizations.

[9] As explained, this modification is essential due to the dependence on prior beliefs in our setting.

[10] Another distinction, as discussed in the Introduction, is that the exploitation incentive in bandit problems directly implies the eventual optimality of myopic strategies. This result does not *a priori* carry over to our learning environment. Moreover, the myopic strategy is *immediately* optimal in MAB only under restrictive assumptions. See Berry and Fristedt (1988) and Banks and Sundaram (1992).

[11] Section 1.2 of Russo (2016) provides an excellent discussion of the relevant papers.

for two states ($K = 2$) exactly apply to the "identification" between two correlated normal arms. However, we are not able to handle more than two arms, due to our assumption of an one-dimensional payoff-relevant state.[12]

Correlation is the key feature of our setting, but we are not aware of many papers that study correlated bandits, either under the classical framework or under best-arm identification. Rothschild (1974), Keener (1985) and Mersereau, Rusmevichientong and Tsitsiklis (2009) provide a few stylized cases.

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). Lemma 3 shows that t-optimality is equivalent to *simultaneously* choosing t observations to achieve the most accurate beliefs. This can be viewed as a Bayesian optimal design problem with respect to the "c-optimality criterion", which seeks to minimize the variance of an unknown parameter. Our analysis is however focused on dynamics, and we utilize the notion of t-optimality only to establish the equivalence between myopic and dynamic optimality. While most work in the optimal design literature assumes a prescribed number of observations (corresponding to a fixed exogenous final period), we have been able to demonstrate the optimality of "greedy design" for a broad class of objectives.

3 Benchmark Case: Learning from a Biased Signal

Consider a decision-maker who wants to learn an unknown parameter $x \sim \mathcal{N}(0, 1)$. This parameter 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* $X^t = x + b + \epsilon_X^t$, 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 x, it helps

[12] With two arms, the DM only cares about the difference in their expected payoffs. The case with more than two arms involves *multiple payoff-relevant states* and *a decision problem that is not prediction*. Since multi-dimensional normal signals do not admit a complete Blackwell ordering (Hansen and Torgersen (1974)), we can no longer prove results for general payoff functions by studying the simpler posterior variance function. This technical difficulty limits the generalization of our results. We would like to mention that in settings related to ours, Sanjurjo (2017), Ke and Villas-Boas (2017) and Chick and Frazier (2012) also highlight the challenge of characterizing the optimal search strategy once there are at least three alternatives.

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 receives the payoff $-(a-x)^2$. 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 x, 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 x 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 $(x, \overline{X})$, conditional on the first q_B observations of B, is jointly Gaussian with covariance matrix[13]

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

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

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

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.[14,15]

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. $\mathrm{Var}(x \,|\, h^t) \leq \mathrm{Var}(x \,|\, \tilde{h}^t)$ *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 $t \in \left(\underline{t}, \bar{t}\right]$ 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 $\mathrm{Var}(x \,|\, h^t) = \mathrm{Var}(x \,|\, \tilde{h}^t)$ at every $t < \underline{t}$ and $t \geq \bar{t}$.[16] Now consider any period t with $\underline{t} \leq t < \bar{t}$. Then,

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

[14]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 $(\sigma_X + \sigma_B)/d$.

[15]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.

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

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

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 *t-optimality for each 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 and arbitrary correlation patterns, and second, we allow for arbitrary payoff functions.

4 General Model

4.1 Setup

The benchmark model considered in Section 3 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,[17] 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.[18] 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

[17] 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 Liang and Mu (2017).

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

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 11 discusses the case of multiple payoff-relevant states.

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

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.[19]

Other than these assumptions we have made, our results are robust to the specifics of the decision problem. In Section 9, we show that our main results generalize to *endogenous stopping*, where the DM chooses an optimal time to stop acquiring information,[20] 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'.*[21]

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

[19]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.

[20]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.

[21]That is, the inverse matrix C^{-1} exists, and its first row consists of non-zero entries.

4.2 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.

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.[22] 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 benchmark case considered in Section 3.

[22] This interpretation is equivalent to the model we presented, under suitable linear transformations.

5 Notation and Definitions

5.1 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 period t following history h^{t-1}.[23] 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.[24]

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.

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

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

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 $q^S(t)$.*

5.2 Optimality Criteria

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

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.[25] 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.[26] Thus, we will (without loss) restrict 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

[25]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.

[26]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.

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.[27]

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

Definition 2. *A strategy is* t-optimal *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 distribution of posterior beliefs $\theta \sim \mathcal{N}(\mu^t, V^t)$ at the end of period t, induced by the strategy.*[28]

Intuitively, a strategy achieves t-optimality 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 (i.e. fixed final period).

6 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 3.[29]

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

We will show now that the development of myopic optimality and t-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 t-optimal for a general decision problem if and only if it is t-optimal for prediction.

[27] 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.

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

[29] 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.

Intuitively, a myopic DM concerned about immediate payoffs only seeks to reduce the uncertainty about θ_1, but does not care about how correlated θ_1 may be with the other states (which helps future updating). By normality, when we restrict to static decision problems that depend on θ_1, the signal that leads to the lowest posterior variance about θ_1 is in fact best in Blackwell's order (Blackwell (1951)). Thus we have

Lemma 2. *Fix an arbitrary decision problem satisfying Assumption 2. A strategy is myopic if and only if it is myopic for prediction (minimizing current posterior variance about θ_1).*

The insight for the "if" direction is discussed above and has appeared for instance in Hansen and Torgersen (1974). For the "only if" part, 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. Henceforth, when working with myopic strategies, we can restrict to deterministic strategies that simply minimize the variance about θ_1 at each step.

The same equivalence turns out to hold for t-optimality (see Definition 2). To see this, observe that our previous argument regarding Blackwell ordering implies that given a fixed number of observations t, the optimal *set* of t 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 not to be advantageous.[30] 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 t-optimal 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 let $f(q_1, \dots, q_K)$ denote the DM's posterior variance function given q_i observations of each signal i. This posterior variance function plays a key role in our analysis, and Appendix A describes its key properties.

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

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

with probability 1.[31]

[30] This is roughly because posterior variance at the final period does not depend on signal realizations.

[31] 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 minimizes posterior variance.

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

Using Lemma 2 and 3, we can deduce that for the informational environment discussed in Section 3, any myopic strategy achieves t-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.

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 t-) optimality. We will nevertheless show in Section 8 that the dynamically optimal strategies are *eventually* deterministic and myopic in generic environments.

7 Sufficient Conditions for Immediate Equivalence

In Section 3, we saw an environment in which the myopic, dynamically optimal, and t-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: *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.[32] The benchmark case in Section 3 falls into this class of environments, as does its generalization below:

Example 1 (Multiple Biases). 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 $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$.

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

This environment is separable.[33]

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

Example 2 (Orthogonal Signals). Suppose the DM's prior is standard Gaussian ($V^0 = \mathbf{I}_K$), and the row vectors of C are orthogonal to one another. This environment is separable.[34]

By the same argument as in Section 3, in a separable environment, any strategy that is myopic for prediction is t-optimal for prediction at every time. The reduction lemmata in Section 6 enable us to extend the equivalence to arbitrary decision problems.[35]

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

An alternative sufficient condition for the myopic strategy to be immediately optimal is *symmetry* across the signals: *The informational environment $(V^0, C, \{\sigma_i^2\})$ is symmetric if the posterior variance function $f(q_1, \dots, q_K)$ 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 are symmetric.

In a symmetric environment, the natural strategy of observing the signal that has been *least observed* turns out to be myopically, dynamically and t-optimal. We have

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

[33] The DM's posterior variance about θ is given by

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

[34] The posterior variance is $\left[I_K - C' \left(CC' + \operatorname{diag}\left(\frac{\sigma_1^2}{q_1}, \dots, \frac{\sigma_K^2}{q_K} \right) \right)^{-1} C \right]_{11}$ By orthogonality, CC' is a diagonal matrix. Thus $\left(CC' + \operatorname{diag}\left(\frac{\sigma_1^2}{q_1}, \dots, \frac{\sigma_K^2}{q_K} \right) \right)^{-1}$ is also a diagonal matrix, and it is separable in $q_1, \dots, q_K$.

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

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.

8 Eventual Equivalence

8.1 A Counterexample to Immediate Optimality

The previous section provided conditions under which the myopic and dynamically optimal signal paths coincide at every period, and are moreover t-optimal for every t. Why might these equivalences fail? We provide a simple example below, building on the benchmark case considered in Section 3.

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 3, 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) t-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 t-optimal division for $t = 6$ is $(3, 2, 1)$, as $f(3, 2, 1) < f(4, 2, 0)$.[36]

The myopic strategy fails to achieve t-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.[37] Nonetheless, we will

[36] We point out that in period 6, 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 t-optimal.

[37] 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.

show that the strength of such complementarities vanishes at late time periods, so that the myopic signal path eventually achieves approximate dynamic and t-optimality. Moreover, in generic environments the myopic path is eventually exactly optimal.

8.2 Main Results

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 t-optimal division at time t, according to the definition in Lemma 3.[38] Thus $m_i(t)$ (resp. $d_i(t), n_i(t)$) counts the number of times a myopic (resp. forward-looking, t-optimal) DM observes signal i in the first t periods.

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 t-optimality can differ by at most 1.

Before stating the result, we impose a weak regularity 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.[39] Additionally, it is not needed for the equivalence between myopic and t-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 $(q_1, \ldots, q_K)$ 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*

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

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

any decision problem and any time $t \geq T$,[40]

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

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

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

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 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.

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.[41] 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.[42]

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

[41]By Lemma 5, this is true along any signal path in which the signal counts go to infinity *proportionally*. In Appendix D, we show such proportionality for the myopic, dynamically optimal and t-optimal strategies.

[42]We mention that eventual equivalence generalizes to a setting where the DM has access to "free signals" that need not be acquired. The proof based on Lemma 5 is unchanged.

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 t-optimality at every period (see Appendix G for the proof):[43]

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, even 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 fail to achieve t-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.[44]

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 (Immediate Equivalence under Many Observations)**.** *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 t-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.

8.3 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

[43] The key new technical tool is use of results on the Diophantine approximation (approximating real numbers by rationals). See Appendix G for details.

[44] 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.

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 R denote the operator norm of the matrix V^{-1}.*[45] *Suppose $w = (1, \ldots, 1)'$, then $|m_i(t) - n_i(t)| \leq 1$ whenever $t \geq 24(R+1)K^2$.*[46]

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.

Theorem 4 shows that the time to eventual equivalence depends on how long it takes for (transformed) states to "de-correlate," at which point potential complementarity or substitution across signals is sufficiently weak.[47] This depends on two primitives:

First, the time to eventual equivalence 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. 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 t-optimality is attained later. This is because 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 states, and de-correlation takes longer.

In contrast, suppose we hold prior precision fixed and increase the degree of prior 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 gets closer to being singular, the norm of V^{-1} increases and the time to equivalence is longer. That is, greater correlation in the prior requires more time to de-correlate.

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

[46] 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.

[47] Note that Theorem 4 only provides an upper bound on the exact number of periods it takes for myopic to become optimal, which would in general depend on the utility function and/or discount factor. Nonetheless, the comparative statics results in the subsequent paragraphs hold not just for the upper bound we derive, but also for the exact time. For instance, in the continuous-time limit of our model, doubling the prior precision also doubles the time to equivalence.

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.

9 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.[48]

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 $U(a_1, a_2, \dots; \theta_1)$, which can exhibit arbitrary intertemporal dependence.[49] 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,[50] myopic information acquisition remains (approximately) optimal. Formally, we prove in Appendix L the following result:

Theorem 5. *Suppose myopic signal choices minimize the posterior variance about θ_1 after any number of observations. Then, for any intertemporal decision problem, there is an optimal strategy in which the DM acquires information myopically.*

Thus, whenever the informational environment satisfies the sufficient conditions in Section 7, the DM cannot do better than acquiring information myopically even if his actions may have

[48]Constant waiting cost per unit of time appears for instance in Fudenberg, Strack and Strzalecki (2017) and Che and Mierendorff (2017).

[49]Endogenous stopping arises if we take each a_t to specify both the stopping decision and the action to be taken when stopped.

[50]This assumption distinguishes our model from Multi-armed Bandit problems, see Section 2.

intertemporal consequences. 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.[51] 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.

More generally, the *eventual* optimality of myopic information acquisition also holds for intertemporal decision problems. For example, Theorem 2 remains true as stated and shows that for any intertemporal decision problem, every optimal strategy acquires information myopically after T^* periods.[52,53]

Our proof of Theorem 5 (and of its "eventual" analogue) is based on a dynamic Blackwell-dominance lemma that generalizes the static reduction results in Section 6. This lemma states that a sequence of normal signals yields higher expected payoff than another sequence (in every intertemporal decision problem depending on θ_1) if it leads to lower posterior variance about θ_1 at every period. We note that the direct extension of Blackwell-dominance to the dynamic setting says that a 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.* This turns out to be technically non-trivial and dependent on normality. Our argument in Appendix L generalizes Greenshtein (1996), see the detailed discussion in Section 2.

10 Endogenous Learning Intensities

Our results also extend beyond the exogenous capacity constraint of one signal acquisition per period. In this section, we assume instead that in each period t, the DM can choose to observe any number $N_t \in \mathbb{Z}^+$ of signal realizations, where each realization is $\sum_{k=1}^{K} c_{ik}\theta_k$ (for some i) plus independent Gaussian noise. In so doing, the DM incurs a flow cost of information acquisition, modeled as $\kappa(N_t)$ for some increasing cost function $\kappa(\cdot)$ with

[51] 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 signals. We discuss this connection in Section 11, when we introduce the continuous time version of our model.

[52] Moreover, the division over signals is t-optimal for $t \geq T^*$.

[53] Whether (and how much) this "eventual" result helps with characterizing optimal actions depends on discounting. In endogenous stopping problems, for example, signal choices after T^* periods are only relevant if the DM does not stop before then. This would be the case with high δ.

$\kappa(0) = 0$.[54] Note that this framework embeds our main model if we define $\kappa(0) = \kappa(1) = 0$ and $\kappa(N) = \infty$ for $N > 1$.

We assume that the DM faces an intertemporal decision-problem as described in the previous section. Specifically, he takes an action a_t at the end of each period t and receives total payoff $U(a_1, a_2, \ldots; \theta_1) - \sum_t \delta^{t-1} \cdot \kappa(N_t)$ for some discount factor δ.[55] For the special case of endogenous stopping, payoff simplifies to

$$\delta^\tau \cdot u(a_\tau; \theta_1) - \sum_{t=1}^{\tau} \delta^{t-1} \cdot \kappa(N_t)$$

whenever the DM stops after τ periods.

This formulation of the endogenous stopping problem can be seen as a (discrete-time) generalization of Moscarini and Smith (2001), who answer the question of how to optimally choose the precision of information over time. Their model has a single state and a single signal, corresponding to a special case of our framework where $K = 1$ and the DM only chooses the "learning intensity" N_t. Nonetheless, with many signals available in our general setting, the question is not just about how much information to acquire at each moment, but also *which* information to acquire. Despite this difficulty, our immediate equivalence results imply that optimal signal choices are often independent of optimal intensity levels:[56]

Theorem 5'. *Suppose myopic signal choices minimize the posterior variance about θ_1 after any number of observations. Then, for any intertemporal decision problem with endogenous learning intensities, there is an optimal strategy in which the DM chooses signals myopically.*

This differs from the statement of Theorem 5 since we can only conclude now that signal choices are myopic. Intensity choices, on the other hand, are not generally myopic. More concretely, suppose the informational environment satisfies the sufficient conditions in Section 7. Let h denote the myopic signal path. Then Theorem 5' implies that an optimal strategy under endogenous intensities is to "follow" h: observe the first N_1 signals in h in the first period, the next N_2 signals in the second period, so on and so forth.

Our characterization of which information to acquire is a first step toward the analysis of how much information to acquire (intensity levels) and when to stop acquiring information

[54] It is also natural to assume the convexity of $\kappa(\cdot)$, so that acquiring extra signals within a single period is increasingly costly. However, the result below does not rely on this.

[55] For convenience of exposition, this assumes that information costs are separable over time and from the actual decision. Our results can accommodate more general payoff functions of the form $U(N_1, a_1, N_2, a_2, \ldots; \theta_1)$.

[56] Eventual equivalence also generalizes, but we omit the details.

(stopping time). In future work, we hope to pursue the latter questions and re-evaluate the insights of Moscarini and Smith (2001) in our framework with multiple sources of information.

11 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.[57] 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).

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

[57] 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.

$\sum_i \beta_i(t) \le 1$), which influence the diffusion processes $X_1, \ldots, X_K$ that he observes:[58]

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

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.[59]

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.[60] This result is a special case of our Theorem 8 in Appendix M, which generalizes to arbitrary prior beliefs and also characterizes optimal attention strategy off-path.[61] Our analysis also reveals that with two signals in general, optimal attention levels are eventually constant.

For $K > 2$, we extend and strengthen many of our previous results to this continuous-time setting. Specifically, Theorem 9 proves eventual *exact* equivalence in all informational environments satisfying identifiability, thus improving upon the conclusion of Theorem 1 and 2. We also provide more permissive sufficient conditions for the myopic strategy to be immediately optimal. Specifically, Theorem 7 shows that immediate optimality obtains whenever the DM's prior beliefs over different (transformed) states are "almost independent." This corroborates the intuition we provided for Theorem 1.

12 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 "t-optimal." We discuss now certain conceptual implications of these results.

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*

[58] 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.

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

[60] While the decision problem in Fudenberg, Strack and Strzalecki (2017) involves endogenous stopping, this difference does not affect our analysis as discussed in Section 9.

[61] 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 $w_1(V_{11} + V_{12}) + w_2(V_{21} + V_{22}) \ge 0$ is satisfied for every prior covariance matrix V.

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.[62]

13 Games with Dynamic Information Acquisition

13.1 A General Framework

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

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

information acquisition in games—with application to Hellwig and Veldkamp (2009) and 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.[63] 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.[64] Signal choices and realizations are *privately observed*. It is worth mentioning that Reinganum (1983) considered a similar multi-agent model with this type of private information acquisition (specifically, firms engaging in R&D before competing in oligopoly). Her model is based on the single-agent "Pandora's box" search framework presented in Weitzman (1979), and it is further developed by Taylor (1995) within the context of research tournaments. However, these papers assume perfectly revealing signals and are thus distinguished from our setting.

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.[65] 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.[66]

For concreteness, we provide examples (adapted from Lambert, Ostrovsky and Panov (2017)) that do and do not satisfy conditional independence.[67]

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

[64] 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.

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

[66] Conditional independence is imposed on players' beliefs at $t = 0$. However, this assumption is sufficient to guarantee conditional independence for subsequent 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.

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

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.

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."[68] 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 7.

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

on their Example 1.

[68]From player i's perspective, the strategy of player j can be viewed as mappings from player j's states $(\theta_1^j, \ldots, \theta_K^j)$ to player j's mixed actions. Thus, given opponent strategies, player i's expected payoff conditional on his own states $(\theta_1^i, \ldots, \theta_K^i)$ 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.

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.*[69]

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 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.

13.2 Application: Beauty Contest

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.[70]

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

[69] 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.

[70] 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.

(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.[71] 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 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

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

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

[71] 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), Myatt and Wallace (2012) and Colombo, Femminis and Pavan (2014) 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).

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).

13.3 Application: 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 before, 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 our dynamic setting with endogenous information.

14 Conclusion

Characterization of the optimal strategy for dynamic information acquisition is challenging in many settings. Common restrictions include: parametric assumptions about the decision problem and discounting structure; separation/asynchronism of information acquisition from actions; and lack of correlation across information sources (or the assumption of specific correlation structures). 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 an arbitrary correlation pattern; 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 the formal 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, while we have presented our model and results assuming a known signal structure, it is of interest to study how to cope with potential uncertainty about the learning environment. In continuing work, we are also pursuing the question of optimal design of signals by self-interested sources that seek to maximize the long-run frequency with which their signals are chosen.

Appendix 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 $s_i = \frac{q_i + r_i}{2}$.[72] 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). $\square$

[72] 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.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.

A.2 Order Difference Lemma

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

Lemma 5. *As $q_1, \dots, q_K \to \infty$, $\frac{\partial^2 f}{\partial q_i^2}$ 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 $\frac{\partial^2 f}{\partial q_i^2} \geq \frac{1}{Lq_i^3}$ 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 $q_1, \dots, q_K$ go to infinity proportionally, which is guaranteed by Proposition 3 below.

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

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

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),

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

As $q_1, \dots, q_k \to \infty$, $\Sigma \to CV^0C'$ which is symmetric and non-singular. Thus the matrix $V^0C'\Sigma^{-1}\Delta_{ii}\Sigma^{-1}CV^0$ 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 $q_1 = \dots = q_K = +\infty$.*

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

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

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 $[V^0C'\Sigma^{-1}\Delta_{jj}\Sigma^{-1}CV^0]_{11} = 0$. Let v denote the first row vector of $V^0C'\Sigma^{-1}$, 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 $v = e_1V^0C'\Sigma^{-1}$. 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.[73] The content of Lemma 6 is to show at every history, there is some signal that provides positive marginal value.[74] In contrast, the stronger Assumption 4 described in the main text requires the same signal to have positive value at every history.

Appendix B Proofs in Section 6 (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:

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

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

[73] 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).

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

(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 t-optimal if and only if the realized division $q^S(t)$ is always a t-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 t-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 $t = 1$, t-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 $(\hat{q}_1, \ldots, \hat{q}_K)$ is a t-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[75] $\hat{S}'$ 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 t-optimal division at time t.[76] 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 t-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.

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

[76] Because S induces t-optimal divisions, its contingent strategies must induce t-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.

An analogous inductive argument proves the "only if" part; that is, S is t-optimal only if the realized division $q^S(t)$ is always a t-optimal division. $\square$

Appendix C Proofs in Section 7 (Immediate Equivalence)

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

Definition 3. *Fix a history h with length H, where each signal i has been observed H_i times. A division over signals is constrained t-optimal for some $t \geq H$ (following history h), if*

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

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) t-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 t-optimal division.[77]

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 t-optimality following any given history, for the prediction problem. By our reduction results, myopic is constrained t-optimal for any decision problem and it is thus dynamically optimal.

Take a signal path $h = (s_1, s_2, \ldots)$ that follows the myopic rule starting at time $\underline{t}$. Consider a deviation path $\tilde{h} = (\tilde{s}_1, \tilde{s}_2, \ldots)$ 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.

[77] 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.

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

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

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

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

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

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

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

Proof of Proposition 2. Suppose the informational environment is symmetric. We claim that at any time t, a division $n(t)$ is t-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* t-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.[78] The claim follows.

[78]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

From this characterization of t-optimal divisions, we see that for every t-optimal division $n(t)$, there exists some t-optimal division $n(t+1)$ with $n_i(t+1) \geq n_i(t), \forall i$. Hence a myopic DM can achieve t-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 t-optimal divisions. Specifically, a division $n^h(t)$ is constrained t-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 t-optimal divisions are again monotonic over time, proving that any myopic strategy is constrained t-optimal and dynamically optimal. $\square$

Appendix 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

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

---

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 t-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 $n_1(t), \ldots, n_K(t) \to +\infty$ 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.[79] Let q_j be the limit of $n_j(t)$ along this subsequence, then a proper subset of $q_1, \ldots, q_K$ is equal to $+\infty$. By Lemma 6, there is some signal j with $\partial_j f(q_1, q_2, \ldots, q_K) < 0$; 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 $\partial_2 f(q_1, q_2, \ldots, q_K) = 0$. By continuity, the *discrete partial derivatives* of f satisfy the following inequality:[80]

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

for sufficiently large t along this subsequence. But this implies

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

contradicting the assumption that $(n_1(t), n_2(t), \ldots, n_K(t))$ is a t-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 t-optimal division must satisfy $\partial_i f \sim \partial_j f$,[81] 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 $\Sigma = CV^0C' + D^{-1} = CV^0C' + O(\frac{1}{t})$. 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 $\frac{\lambda_i^2 + O(\frac{1}{t})}{n_i^2} = \frac{\lambda_j^2 + O(\frac{1}{t})}{n_j^2}$.[82] 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

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

[80] 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.

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

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

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.[83]

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.

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. $\square$

Appendix E Proof of Theorem 1 (Eventual Gap of One)

We present the proof for equivalence between myopic and t-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 t-optimality of the division $(n_1, \dots, n_K)$ we have

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

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

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

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

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

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

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

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

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.

Appendix F Omitted Proofs for Dynamically Optimal Strategies

In this appendix, we extend the previous proof of "eventual gap of one" to dynamically optimal strategies. This extension is technically difficult in part because dynamically optimal strategies often condition on signal realizations, so that the induced division $d(t)$ can be stochastic. Still, we will show that every realized division $d(t)$ differs by at most one (in each signal) from the deterministic division $m(t)$ under a myopic strategy or from a t-optimal division $n(t)$. To that end, we first establish a dynamic Blackwell-dominance lemma for comparing sequences of normal signals.

F.1 A Dynamic Blackwell-dominance Lemma

Recall that the DM only needs to remember the expected value about θ_1 and the covariance matrix of all the states. Thus, a history of (payoff-relevant) beliefs is $h^T = (\mu_1^0, V^0; \dots; \mu_1^T, V^T)$. For a fixed prior covariance matrix V^0, an alternative way to summarize the history is the divisions over signals. That is, a history up to and including period T can also be written as $h^T = (\mu_1^0, d(0); \dots; \mu_1^T, d(T))$, where each $d(t) = (d_1(t), \dots, d_K(t))$ counts the number of each signal acquired by time t. In this section, we will use this alternative definition of history and view a strategy S as a mapping from such sequences of divisions to signal choices. Each strategy induces a distribution of histories and determines the DM's expected payoff (assuming that he takes action optimally in the final period).

Consider a mapping $\tilde{G}$ from possible sequences of divisions to these sequences themselves: For each sequence of divisions $(d(0), \dots, d(T))$, $\tilde{G}$ maps to another sequence $(\tilde{d}(0), \dots, \tilde{d}(T))$, subject to "consistency." There are three consistency requirements: First, $\sum_i \tilde{d}_i(t) = t$, meaning that each $\tilde{d}(t)$ must be a possible division at time t. Second, $\tilde{d}_i(t) \geq \tilde{d}_i(t-1)$, meaning that the sequence $\tilde{d}$ can be realized under some strategy. Lastly, we require

$$(\tilde{d}(0), \dots, \tilde{d}(T-1)) = \tilde{G}(d(0), \dots, d(T-1))$$

so that nesting sequences are mapped to nesting sequences. For simplicity, we will often denote $d(0)$ by the vector $\mathbf{0}$.

We will use any such mapping $\tilde{G}$ to construct a deviation strategy $\tilde{S}$ from a given strategy S. Loosely speaking, whenever $d(T)$ is the realized division under strategy S, we will let $\tilde{S}$ induce the division $\tilde{d}(T) = \tilde{G}(d(T))$. Thus if $f(\tilde{d}(T)) \leq f(d(T))$ always holds, the DM's posterior variance about θ_1 is always smaller under strategy $\tilde{S}$ than under strategy S, which would imply via Blackwell dominance that $\tilde{S}$ yields higher payoff than S.

Lemma 7. *Fix any strategy S and any consistent mapping $\tilde{G}$ defined above. Suppose that for every sequence of divisions $(d(0), \ldots, d(T))$ realized under S, it holds that*

$$f(\tilde{d}(T)) \leq f(d(T))$$

with strict inequality occurring with positive probability (under S). Then S is not a dynamically optimal strategy.

We first give an intuitive discussion of the proof before supplying the formal details. As mentioned, the basic idea is to construct a strategy $\tilde{S}$ that achieves the division $\tilde{d}(T)$ whenever S would achieve $d(T)$. That is, suppose the sequence of divisions under S has been $d(0), \ldots, d(T-1)$, and S dictates observing some signal to reach the division $d(T)$ in the next period, then we let $\tilde{S}$ observe some other signal to reach the division $\tilde{d}(T)$ from the current division $\tilde{d}(T-1)$.

If the strategy S were deterministic (not depending on signals realizations), then this construction would be sufficient. However, difficulty arises whenever S conditions on the expected value about θ_1. To explain, consider a history $h^{T-1} = (\mu_1^0, d(0); \ldots; \mu_1^{T-1}, d(T-1))$. At this history, strategy S is going to achieve the division $d(T)$ in the next period, and we want strategy $\tilde{S}$ to achieve the division $\tilde{d}(T)$. But we cannot simply let $\tilde{S}$ reach $\tilde{d}(T)$ from the history $\tilde{h}^{T-1} = (\mu_1^0, \tilde{d}(0); \ldots; \mu_1^{T-1}, \tilde{d}(T-1))$. This is because the same sequence of posterior expected values is realized with different probabilities under $\tilde{S}$ than under S, which makes it impossible to compare expected payoffs.

To address this difficulty, we instead let $\tilde{S}$ achieve the division $\tilde{d}(T)$ from *a distribution of histories* $\tilde{h}^{T-1} = (\nu_1^0, \tilde{d}(0); \ldots; \nu_1^{T-1}, \tilde{d}(T-1))$, such that these histories occur with the same probability (under $\tilde{S}$) as the probability that h^{T-1} occurs (under S). This ensures that we can compare the DM's average payoff at these histories to his payoff at history h^{T-1}. Furthermore, we require these histories to induce a distribution of beliefs about θ_1 that Blackwell-dominates the belief $\theta_1 \sim \mathcal{N}(\mu_1^{T-1}, f(d(T-1)))$ at h^{T-1}, so that the DM's average payoff is indeed higher under the deviation $\tilde{S}$.

Because $f(\tilde{d}(1)) \leq f(d(1))$, the DM's distribution of posterior beliefs about θ_1 after one observation is Blackwell more informative under $\tilde{S}$ than under S. Thus the two conditions identified in the preceding paragraph are satisfied at initial times T. Our key argument is to extend these conditions inductively. Specifically, we show that if a distribution of beliefs $\theta_1 \sim \mathcal{N}(\nu_1^{T-1}, f(\tilde{d}(T-1)))$ Blackwell-dominates *a single belief* $\theta_1 \sim \mathcal{N}(\mu_1^{T-1}, f(d(T-1)))$, then after observing a signal in period T (possibly different signal under $\tilde{S}$ than under S), the posterior distribution of beliefs $\theta_1 \sim \mathcal{N}(\nu_1^T, f(\tilde{d}(T)))$ Blackwell-dominates the posterior *distribution of beliefs* $\theta_1 \sim \mathcal{N}(\mu_1^T, f(d(T)))$. This enables us to select a distribution of expected values ν_1^T for each μ_1^T such that Blackwell-dominance is preserved.

A non-technical version of the above argument is as follows: In the first period, the signal under $\tilde{S}$ is better than the signal under S about θ_1. Disregarding the effect on the other states, we can equivalently think of the signal under S as $\theta_1 + \xi_1$ and the signal under $\tilde{S}$ as two signals $\theta_1 + \xi_1$ and $\theta_1 + \eta_1$; the Gaussian noise terms ξ_1 and η_1 have appropriate variances and are independent from the state θ_1 and from each other.[84] What the deviation strategy $\tilde{S}$ does after the first period is to "temporarily forget" the extra signal $\theta_1 + \eta_1$ and modify the signal choice of S (according to the mapping $\tilde{G}$) as if it had only observed $\theta_1 + \xi_1$. Forgetting is temporary because we require the DM to remember the extra signal for future updating purposes (as will be clear below).

In the second period, we again view the signals under $\tilde{S}$ and S as θ_1 plus independent noise. The precision of these noise terms are so that the signals in the first two periods combine to yield the correct posterior variance about θ_1.[85] Now if the precision of the second-period signal is higher under $\tilde{S}$ than under S, we can again think of $\tilde{S}$ as generating an extra signal in the second period than S, and in that case $\tilde{S}$ can continue to "temporarily forget" this extra signal when deciding future information acquisition. However, we are only guaranteed that the *cumulative* precision of signals under $\tilde{S}$ is higher than under S. Thus, it is possible that the precision of the second-period signal is in fact lower under the deviation strategy $\tilde{S}$.

Then, it will be the case that strategy S observes $\theta_1 + \xi_1$ in the first period and two signals $\theta_1 + \xi_2, \theta_1 + \eta_2$ in the second period. On the other hand, the deviation strategy $\tilde{S}$ observes two signals $\theta_1 + \xi_1, \theta_1 + \eta_1$ in the first period and only one signal $\theta_1 + \xi_2$ in the second period. We note that by $f(\tilde{d}(2)) \leq f(d(2))$, the precision of η_1 must be higher than η_2. Thus, the signal $\theta_1 + \eta_1$ is equivalent to two further signals $\theta_1 + \eta_2$ and $\theta_1 + \hat{\eta}_1$. Under strategy $\tilde{S}$, the DM observes these signals in the first period, "temporarily forgets" them in choosing the second-period signal, but later "retrieves" the signal value of $\theta_1 + \eta_2$ so as to maintain the same amount of information obtained under $\tilde{S}$ as under S.[86] He then modify

[84] The variance of ξ is such that the DM's posterior variance about θ_1 given $\theta_1 + \xi_1$ is $f(d(1))$; thus the *precision* of ξ is $\frac{1}{f(d(1))} - \frac{1}{f(\mathbf{0})}$. Similarly, the precision of η_1 is $\frac{1}{f(\tilde{d}(1))} - \frac{1}{f(d(1))}$.

[85] The precisions are $\frac{1}{f(d(2))} - \frac{1}{f(d(1))}$ under S and $\frac{1}{f(\tilde{d}(2))} - \frac{1}{f(\tilde{d}(1))}$ under $\tilde{S}$. Their difference is the precision of η_2 (in the next paragraph), and it is lower than the precision of η_1 derived previously.

[86] At the end of two periods, the DM has not yet "retrieved" $\theta_1 + \hat{\eta}_1$.

the signal choice S in the third period, so on and so forth.

In summary, the assumption that $f(\tilde{d}(T)) \leq f(d(T))$ tells us that the deviation strategy $\tilde{S}$ generates *cumulatively* more information (about θ_1) than S. This is weaker than the statement that $\tilde{S}$ generates more information than S *in every period*, hence we cannot directly extend Blackwell's insight by letting the DM "permanently forget" the extra information he receives under $\tilde{S}$. Nonetheless, our assumption on cumulative information does ensure that the same information is received "earlier" under $\tilde{S}$ than under S. We use this observation to let the DM "temporarily forget" the extra information under $\tilde{S}$, but later "retrieve" it whenever necessary. This turns out to be sufficient to show that $\tilde{S}$ is a profitable deviation. We now present the formal proof.

Proof of Lemma 7. We follow the above proof outline and construct a deviation strategy $\tilde{S}$ as follows:

In the first period, consider the signal choice under S. This signal leads to a division $d(1)$. We let $(\mathbf{0}, \tilde{d}(1)) = G(\mathbf{0}, d(1))$. The deviation strategy $\tilde{S}$ observes the unique signal that would achieve the division $\tilde{d}(1)$ after the first period.

After the first observation, the DM's distribution of posterior beliefs about θ_1 under strategy S is $\theta_1 \sim \mathcal{N}(\mu_1^1, f(d(1)))$ with μ_1^1 a normal random variable with mean μ_1^0 and variance $f(\mathbf{0}) - f(d(1))$. By comparison, the distribution of beliefs under $\tilde{S}$ is $\theta_1 \sim \mathcal{N}(\nu_1^1, f(\tilde{d}(1)))$ with ν_1^1 drawn from $\mathcal{N}(\mu_1^0, f(\mathbf{0}) - f(\tilde{d}(1)))$. Since $f(\tilde{d}(1)) \leq f(d(1))$, the latter distribution of beliefs (under $\tilde{S}$) is Blackwell more informative. Thus in fact, we can associate *each belief* $\theta_1 \sim \mathcal{N}(\mu_1^1, f(d(1)))$ under S with a more informative distribution of beliefs $\mathcal{N}(\nu_1^1, f(\tilde{d}(1)))$ under $\tilde{S}$. Specifically, for fixed μ_1^1, the distribution of associated ν_1^1 should be $\nu_1^1 \sim \mathcal{N}(\mu_1^1, f(d(1)) - f(\tilde{d}(1)))$. We say this distribution of (payoff-relevant) beliefs under $\tilde{S}$ "imitates" the belief $(\mu_1^1, f(d(1)))$ under S. By construction, this distribution of ν_1^1 occurs under $\tilde{S}$ with the same probability as μ_1^1 occurs under S.

The likelihood of ν_1^1 given μ_1^1, the probability of μ_1^1 occurring under S and the probability of ν_1^1 occurring under $\tilde{S}$ together determine the likelihood of μ_1^1 given ν_1^1, which is a Gaussian probability kernel $p(\mu_1^1 \mid \nu_1^1)$. In the second period, the deviation strategy $\tilde{S}$ takes the current belief $(\nu_1^1, f(\tilde{d}(1)))$ and randomly selects some μ_1^1 to "imitate", with probability given by p. To be specific, given any selection of μ_1^1, we find the signal that S is going to observe in the second period given belief $(\mu_1^1, f(d(1)))$. This signal choice under S leads to a sequence of divisions $(\mathbf{0}, d(1), d(2))$, which is mapped to $(\mathbf{0}, \tilde{d}(1), \tilde{d}(2))$ under $\tilde{G}$. We let $\tilde{S}$ observe a signal in the second period that would achieve the division $\tilde{d}(2)$. By the consistency of $\tilde{G}$, this signal is in fact unique.

Let us now fix μ_1^1 and study the distribution of posterior beliefs about θ_1 after two observations. Under S, the posterior belief is $\theta_1 \sim \mathcal{N}(\mu_1^2, f(d(2)))$ with μ_1^2 normally distributed with mean μ_1^1 and variance $f(d(1)) - f(d(2))$. Under the deviation strategy $\tilde{S}$, the distribution of posterior beliefs is $\theta_1 \sim (\nu_1^2, f(\tilde{d}(2)))$ with ν_1^2 drawn from $\mathcal{N}(\mu_1^1, f(d(1)) - f(\tilde{d}(2)))$.[87]

Since $f(\tilde{d}(2)) \leq f(d(2))$, the distribution of beliefs under $\tilde{S}$ Blackwell-dominates the distribution under S, for each μ_1^1. We can thus associate each history $(\mu_1^1, d(1); \mu_1^2, d(2))$ under S with a distribution of histories $(\nu_1^1, \tilde{d}(1); \nu_1^2, \tilde{d}(2))$ under $\tilde{S}$ such that the corresponding beliefs under $\tilde{S}$ are more informative at both periods. This procedure can be repeated in the third period, ad infinitum. The upshot is that each infinite history h under S is associated with a distribution of histories $\tilde{h}$ under $\tilde{S}$ (occurring with the same probability). At every period t, the DM's belief about θ_1 along the history h is his average belief across the distribution of histories $\tilde{h}$, so that his average payoff across these histories is weakly higher (and strictly so with positive probability). Integrating over different histories h shows that $\tilde{S}$ improves upon S, proving the Lemma. $\square$

F.2 Switch Deviations

In what follows, we will apply Lemma 7 with a particular class of mappings $\tilde{G}$. Consider any sequence of divisions $(d^*(0), d^*(1), \ldots, d^*(t_0))$. Let i be the signal observed in period t_0 and j be any other signal. An "(i,j)-switch" mapping $\tilde{G}$ specifies the following:

1. Suppose $T < t_0$ or $d(t) \neq d^*(t)$ for some $t \leq t_0$, then let $\tilde{G}(d(0), \ldots, d(T))$ be itself.

2. Suppose $T \geq t_0$, $d(t) = d^*(t), \forall t \leq t_0$. If $d_j(T) = d_j(t_0)$, then let $\tilde{d}(T) = (d_i(T) - 1, d_j(T) + 1, d_{-ij}(T))$. If $d_j(T) > d_j(t_0)$, then let $\tilde{d}(T) = d(T)$.

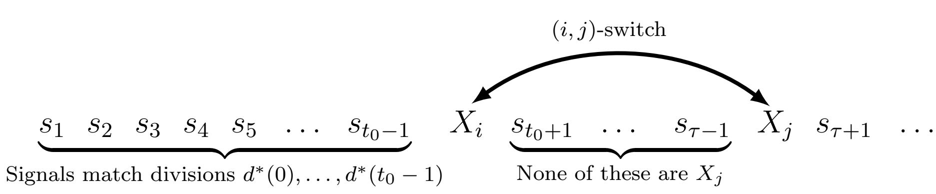

Figure 1: *Pictorial representation of an (i,j)-switch based on a sequence of divisions $d^*(0), \ldots, d^*(t_0)$.*

[87] Here we use the following technical result: suppose the DM is endowed with *a distribution of prior beliefs* $\theta \sim \mathcal{N}(\mu, V)$, with μ_1 normally distributed with mean y and variance σ^2, then upon observing signal i and performing Bayesian updating, his distribution of posterior beliefs is $\theta \sim \mathcal{N}(\hat{\mu}, \phi_i(V))$, with $\hat{\mu}_1$ normally distributed with mean y and variance $\sigma^2 + [V]_{11} - [\phi_i(V)]_{11}$. This is proved by observing that the DM's distribution of beliefs about θ_1 must integrate to the same ex-ante distribution of θ_1.

This definition is easier to interpret if we think about the resulting deviation strategy $\tilde{S}$ constructed in the proof of Lemma 7. The first case above says that $\tilde{S}$ only deviates when the history of divisions is $d^*(0), \dots, d^*(t_0 - 1)$, and S is about to observe signal i in period t_0. The second case says that $\tilde{S}$ dictates observing signal j instead at that history. Subsequently, the DM under $\tilde{S}$ observes the same signal as S (at the belief under S that is being imitated) until the first period at which S is about to observe signal j. If that period exists, the deviation strategy $\tilde{S}$ switches back to observing signal i and agrees with S afterwards.

The benefit of "switch deviations" is that their posterior variances are easily compared with the original strategy: $\tilde{d}(t) = d(t)$ except at a history that extends $(d^*(0), d^*(1), \dots, d^*(t_0 - 1))$ (and before signal j is observed again under S). And at such histories, $f(\tilde{d}(t)) < f(d(t))$ iff the *discrete partial derivatives* satisfy:

$$|\partial_i f(d_i(t) - 1, d_j(t), d_{-ij}(t))| < |\partial_j f(d_i(t) - 1, d_j(t), d_{-ij}(t))|. \tag{21}$$

Thus S is not dynamically optimal whenever we can find a realized division $d(t_0)$ under S, such that $d_i(t_0) = d_i(t_0 - 1) + 1$ (so that signal i is observed under S in period t_0) and the inequality (21) holds for all $t \geq t_0$ and $d_j(t) = d_j(t_0)$.

F.3 Proof of Proposition 3 Part (c)

We now derive the asymptotic ratios for any dynamically optimal strategy S. We will show that any realized division $d(t)$ satisfies $|d_i(t) - \lambda_i t| \leq N$, where the constant N depends only on the informational environment but not on the decision problem or the particular strategy S.

We first show that as $T \to \infty$, $d_1(T), \dots, d_K(T) \to \infty$, and the speed of divergence depends only on the informational environment. For contradiction, suppose the result is not true. Then we can find a sequence of histories $\{h^{T_m}\}$ such that $T_m \to \infty$ but $d_1(T_m)$ remains bounded. These histories need not nest one another. By passing to a subsequence, we may assume $q_i = \lim_{m \to \infty} d_i(T_m)$ exists for every signal i, and $q_i = \infty$ for a proper subset of signals $i \in I$. We further assume that the signal i^* observed in the last period of each history h^{T_m} is the same, and it belongs to I—otherwise, we simply truncate the history by finitely many periods.

Choose $j^* \notin I$ to be a signal that has positive marginal value (satisfying Assumption 4). We claim that for sufficiently large T_m, the (i^*, j^*)-switch deviation $\tilde{S}$ following history $h^{T_m - 1}$ (the first $T_m - 1$ periods of h^{T_m}) satisfies (21) with $t_0 = T_m$. This will contradict the optimality of S. To prove this claim, observe that as $T_m \to \infty$, $d_{i^*}(T_m) \to \infty$ because $i^* \in I$. Since we are only concerned with the inequality (21) for $d_{i^*}(t) \geq d_{i^*}(T_m)$, the LHS of (21) becomes vanishingly small. In contrast, the RHS is bounded away from zero by Assumption 4.[88] Hence we have verified (21) and derived a contradiction to the optimality of S.

Next, as $d_1(t), \ldots, d_K(t) \to \infty$, the following approximations hold (see (16)):

$$|\partial_i f(d_i(t)-1, d_j(t), d_{-ij}(t))| \sim \frac{\sigma_i^2 \cdot [Q_i]_{11}}{d_i(t)^2} \qquad |\partial_j f(d_i(t)-1, d_j(t), d_{-ij}(t))| \sim \frac{\sigma_j^2 \cdot [Q_j]_{11}}{d_j(t)^2}$$

If $\limsup_{t_0 \to \infty} \frac{d_i(t_0)}{d_j(t_0)} > \frac{\lambda_i}{\lambda_j}$ (recall that λ_i is proportional to $\sigma_i \cdot \sqrt{[Q_i]_{11}}$), then the above estimates imply (21) whenever $d_i(t) \geq d_i(t_0)$ (because $t \geq t_0$) and $d_j(t) = d_j(t_0)$, contradicting the optimality of S. Hence, $\limsup_{t_0 \to \infty} \frac{d_i(t_0)}{d_j(t_0)} \leq \frac{\lambda_i}{\lambda_j}$ for any pair of signals i, j, so that $\lim_{t_0 \to \infty} \frac{d_i(t_0)}{t_0} = \lambda_i, \forall i$.

Once this is established, it follows that the matrix $\Sigma = CV^0C' + D^{-1}$ converges to CV^0C' at the rate of $\frac{1}{t}$. By (16), we can deduce more precise approximations:

$$|\partial_i f(d_i(t)-1, \cdots)| = \frac{\sigma_i^2 \cdot [Q_i]_{11} + O(\frac{1}{t})}{d_i(t)^2} \qquad |\partial_j f(d_i(t)-1, \cdots)| = \frac{\sigma_j^2 \cdot [Q_j]_{11} + O(\frac{1}{t})}{d_j(t)^2}$$

If $\frac{d_i(t_0)}{d_j(t_0)} > \frac{\lambda_i}{\lambda_j} + O(\frac{1}{t_0})$, then these refined estimates would again imply (21) whenever $d_i(t) \geq d_i(t_0)$ and $d_j(t) = d_j(t_0)$. We conclude that $\frac{d_i(t_0)}{d_j(t_0)} \leq \frac{\lambda_i}{\lambda_j} + O(\frac{1}{t_0})$. Since this holds for every pair of signals, we deduce $d_i(t_0) = \lambda_i \cdot t_0 + O(1)$ as desired.

F.4 Proof of Theorem 1 Parts (b) and (c)

(b) Here we establish the "eventual gap of one" between dynamic and t-optimality. Let us suppose for contradiction that at a large time T, $d_j(T) \leq n_j(T) - 2$ and $d_i(T) \geq n_i(T) + 1$ holds for some realized division $d(T)$ under a dynamically optimal strategy S. Let $t_0 \leq T$ be the last period along this history (leading to the division $d(T)$) in which signal i was observed. Consider the (i,j)-switch deviation $\tilde{S}$ that begins in that period. We will verify the inequality (21), which will contradict the optimality of S.

Getting rid of absolute values, we need to show for $t \geq t_0$, $d_i(t) \geq d_i(t_0)$ and $d_j(t) = d_j(t_0)$,

$$\partial_i(d_i(t)-1, d_j(t), d_{-ij}(t)) > \partial_j(d_i(t)-1, d_j(t), d_{-ij}(t)). \tag{22}$$

By t-optimality of $n(T)$, we have

$$\partial_i(n_i(T), n_j(T)-1, n_{-ij}(T)) \geq \partial_j(n_i(T), n_j(T)-1, n_{-ij}(T)). \tag{23}$$

[88] While the statement of Assumption 4 in the main text requires $q_{j^*} = 0$, it actually implies that signal j^* has positive marginal value for any given number of past observations $q_{j^*} = d_{j^*}(T_m)$.

Our goal is to show (23) implies (22).

From (23) to (22), the LHS at most decreases by some cross partials. This is because $d_i(t) \geq d_i(t_0) = d_i(T) \geq n_i(T) + 1$, where the middle equality is due to the definition of t_0. Moreover, we claim that the difference between T and t is bounded above by a constant. In one direction, since $d_i(t_0) = d_i(T)$, the asymptotic characterization implies $T - t_0$ is bounded above and so is $T - t$ whenever $t \geq t_0$. In the other direction, since $d_j(t) = d_j(t_0)$, we have $t - t_0$ is bounded above. Thus $t - T \leq t - t_0$ is also bounded above. It follows that the LHS of (22) decreases from the LHS of (23) by a bounded number of cross partials.

In contrast, the RHS of (22) is smaller than the RHS of (23) by at least a second derivative $\partial_{jj} f$ less some cross partials, because $d_j(t) = d_j(t_0) \leq d_j(T) \leq n_j(T) - 2$. Because the second derivative eventually dominates the combined effects of cross partials, we are able to deduce (22) from (23) for sufficiently large T and any $t \geq t_0$.[89] We have thus proved (21) assuming $d_j(T) \leq n_j(T) - 2$, contradicting the optimality of S. A symmetric argument rules out $d_j(T) \geq n_j(T) + 2$ (and $d_i(T) \leq n_i(T) - 1$).

(c) We follow the preceding proof and consider a history h^T where $d_j(T) \leq m_j(T) - 2$ and $d_i(T) \geq m_i(T) + 1$. Define t_0 and the (i, j)-switch deviation $\tilde{S}$ as before. To obtain a contradiction, we will verify (22) for $t \geq t_0, d_i(t) \geq d_i(t_0)$ and $d_j(t) = d_j(t_0)$.

Let $T^* \leq T$ be the last period in which the myopic DM observed signal j. Then it holds that

$$\partial_i(m_i(T^*), m_j(T^*) - 1, m_{-ij}(T^*)) \geq \partial_j(m_i(T^*), m_j(T^*) - 1, m_{-ij}(T^*)) \tag{24}$$

because signal j was myopically better than signal i in period T^*.

Note that $d_i(t) \geq d_i(t_0) = d_i(T) \geq m_i(T) + 1 \geq m_i(T^*) + 1$, while $d_j(t) = d_j(t_0) \leq d_j(T) \leq m_j(T) - 2 = m_j(T^*) - 2$. Thus essentially the same argument as in the proof of Part (b) enables us to deduce (22) from (24), contradicting the optimality of S. A symmetric argument rules out $d_j(T) \geq n_j(T) + 2$ and $d_i(T) \leq n_i(T) - 1$. Theorem 1 is now completely proved.

Appendix G Proof of Theorem 2 (Generic Eventual Equivalence)

To guide the reader through this appendix, we begin by outlining the proof of the theorem, which is broken down into several steps. We first show a simpler result that in generic

[89] Since the constants in Proposition 3 and in the order difference lemma only depend on the informational environment, so does T.

informational environments, the number of periods in which the three optimality criteria coincide has natural density 1. While this is weaker than the desired Theorem 2, it already rules out situations such as Example 4, where any strategy fails to achieve t-optimality in one-third of the periods. Our proof of this result is based on the observation that if equivalence does not hold at some time t, there must be *two different divisions* over signals for which the resulting posterior variances about θ_1 are within $O(\frac{1}{t^4})$ from each other. This leads to a Diophantine approximation inequality, which we can show only occurs at a vanishing fraction of periods t.

To improve the result and demonstrate equivalence at "all large periods," we show that the number of "exceptional periods" t is generically finite if there are *three different divisions* over signals whose posterior variances are within $O(\frac{1}{t^4})$ from each other. This allows us to conclude that in generic environments, the t-optimal divisions eventually monotonically increase (in each coordinate) over time.

In such environments, t-optimality can be achieved at every late period. Thus, whenever t-optimality obtains in *some* late period, it will be sustained in all future periods. Since we have already established that a myopic or dynamically optimal strategy achieves t-optimality infinitely often, we conclude equivalence at all large t.

G.1 Equivalence at Almost All Times

We begin by proving a weaker result. Proposition 4 below shows equivalence at almost all times.

Proposition 4. *Suppose the informational environment $(V^0, C, \{\sigma_i^2\})$ is such that for any $i \neq j$, the ratio $\frac{\lambda_i}{\lambda_j}$ is an irrational number. Then, at a set of times with natural density 1,*[90] *$m(t) = n(t) = d(t)$ holds for any decision problem. In particular, the dynamically optimal division $d(t)$ is deterministic and independent of the decision problem at these times.*

Proof of Proposition 4. Suppose that $m_1(t) \leq n_1(t) - 1, m_2(t) \geq n_2(t) + 1$. Consider the last period $\tilde{t}$ in which the myopic strategy observes signal 2. Write $\tilde{m}_i = m_i(\tilde{t})$ and $n_i = n_i(t)$. Then $\tilde{m}_1 \leq m_1(t) \leq n_1 - 1$ and $\tilde{m}_2 = m_2(t) \geq n_2 + 1$.[91]

From t-optimality of $n(t)$ and myopic optimality of $m(\tilde{t})$, we have the following inequalities regarding the *discrete partial derivatives*:

$$\partial_2 f(n_1 - 1, n_2, \ldots, n_K) \geq \partial_1 f(n_1 - 1, n_2, \ldots, n_K). \tag{25}$$

[90] Formally, for any set of positive integers A, let $A(N)$ count the number of integers in A no greater than N. Then we deinfe the natural density of A to be $\lim_{N\to\infty} \frac{A(N)}{N}$, when this limit exists.

[91] This analysis resembles the proof of Theorem 1 Part (a), in Appendix E.

$$\partial_2 f(\tilde{m}_1, \tilde{m}_2 - 1, \dots, \tilde{m}_K) \le \partial_1 f(\tilde{m}_1, \tilde{m}_2 - 1, \dots, \tilde{m}_K). \tag{26}$$

Since $\tilde{m}_2 - 1 \le n_2$, the LHS of (26) is smaller than the LHS of (25) by (at most) a number of cross partials. Similarly, the RHS of (26) is at most bigger than the RHS of by a number of cross partials. These observations together with the order difference lemma imply that the two sides of (25) differ by at most $O\left(\frac{1}{t^4}\right)$.

To summarize: a necessary condition for $m_1(t) \le n_1(t) - 1, m_2(t) \ge n_2(t) + 1$ to occur is that

$$|f(n_1 - 1, n_2 + 1, \dots, n_K) - f(n_1, n_2, \dots, n_K)| = O\left(\frac{1}{t^4}\right). \tag{27}$$

Likewise, we can deduce that (27) is also a necessary condition for $d_1(t) \le n_1(t) - 1, d_2(t) \ge n_2(t) + 1$ to occur (see the proof of Theorem 1 Part (b), in Appendix F). Hence, to prove the Proposition we only need to show that (27) only holds at a set of times with natural density 0. The following lemma proves exactly this property. □

Lemma 8. *Suppose $\frac{\lambda_1}{\lambda_2}$ is an irrational number. For positive constants c_0, c_1, define $\mathcal{A}(c_0, c_1)$ to be the following set of positive integers:*

$$\{t : \exists\ q_1, q_2, \dots, q_K \in \mathbb{Z}^+, s.t.\ |q_i - \lambda_i t| \le c_0, \forall i \\ \wedge\ |f(q_1, q_2 + 1, \dots, q_K) - f(q_1 + 1, q_2, \dots, q_K)| \le c_1/t^4\}.$$

Then $\mathcal{A}(c_0, c_1)$ has natural density zero.

Proof of Lemma 8. The proof relies on the following technical result, which gives a precise approximation of the *discrete partial derivatives* of f:

Lemma 9. *Fix the informational environment. There exists a constant a_j such that*

$$f(q_j, q_{-j}) - f(q_j + 1, q_{-j}) = \frac{\sigma_j^2 \cdot [Q_j]_{11}}{(q_j - a_j)^2} + O\left(\frac{1}{t^4}; c_0\right) \tag{28}$$

holds for all $q_1, \dots, q_K$ with $|q_i - \lambda_i t| \le c_0, \forall i$. The notation $O\left(\frac{1}{t^4}; c_0\right)$ means an upper bound of $\frac{L}{t^4}$, where the constant L may depend on the informational environment as well as on c_0.[92]

Given (28), we see that the condition

$$|f(q_1, q_2 + 1, \dots, q_K) - f(q_1 + 1, q_2, \dots, q_K)| \le \frac{c_1}{t^4}$$

[92] In our application of Lemma 8, c_0 is taken to be the constant N in Proposition 3 and thus depends also on the informational environment. The statement of Lemma 8 is however more general and allows for arbitrary c_0, c_1.

implies $\left|\frac{\sigma_1^2\cdot[Q_1]_{11}}{(q_1-a_1)^2}-\frac{\sigma_2^2\cdot[Q_2]_{11}}{(q_2-a_2)^2}\right|\leq\frac{c_2}{t^4}$ and thus $\left|(\frac{\lambda_1}{q_1-a_1})^2-(\frac{\lambda_2}{q_2-a_2})^2\right|\leq\frac{c_3}{t^4}$ for some larger positive constants c_2, c_3. This further implies $\left|\frac{\lambda_1}{q_1-a_1}-\frac{\lambda_2}{q_2-a_2}\right|\leq\frac{c_4}{t^3}$, which reduces to

$$\left|q_2-a_2-\frac{\lambda_2}{\lambda_1}(q_1-a_1)\right|\leq\frac{c_5}{t}. \tag{29}$$

This inequality says that the fractional part of $\frac{\lambda_2}{\lambda_1}q_1$ is very close to the fractional part of $\frac{\lambda_2}{\lambda_1}a_1-a_2$. But since $\frac{\lambda_2}{\lambda_1}$ is an irrational number, the fractional part of $\frac{\lambda_2}{\lambda_1}q_1$ is "equi-distributed" in (0,1) as q_1 ranges in the positive integers.[93] Thus the Diophantine approximation (29) only has solution at a set of times t with natural density 0, proving Lemma 8. Below we supply the technically involved proof of (28). □

Proof of Lemma 9. Fix $q_1,\dots,q_K$ and the signal j. Recall that the diagonal matrix D is given by $\mathrm{diag}(\frac{q_1}{\sigma_1^2},\dots,\frac{q_K}{\sigma_K^2})$. Consider any $\hat{q}_j\in[q_j,q_j+1]$ and let $\hat{D}$ be the analogue of D for the division $(\hat{q}_j,q_{-j})$. That is, $\hat{D}=D$ except that $[\hat{D}]_{jj}=\frac{\hat{q}_j}{\sigma_j^2}$. Let $\hat{\Sigma}=CV^0C'+\hat{D}^{-1}$. From (16), we have

$$\partial_j f(\hat{q}_j,q_{-j})=-\frac{\sigma_j^2}{\hat{q}_j^2}\cdot\left[V^0C'\hat{\Sigma}^{-1}\Delta_{jj}\hat{\Sigma}^{-1}CV^0\right]_{11}. \tag{30}$$

Here and later in this proof, $\partial_j f$ represents the usual continuous derivative rather than the discrete derivative.

Let $D_0=\mathrm{diag}\left(\frac{\lambda_1 t}{\sigma_1^2},\dots,\frac{\lambda_K t}{\sigma_K^2}\right)$ and $\Sigma_0=CV^0C'+D_0^{-1}$. For $|q_i-\lambda_i t|\leq c_0,\forall i$ we have $\hat{D}-D_0=O(c_0)$, where the Big O notation applies entry-wise. It follows that

$$\hat{\Sigma}=CV^0C'+\hat{D}^{-1}=CV^0C'+D^{-1}+O(\frac{1}{t^2};c_0)=\Sigma_0+O(\frac{1}{t^2};c_0).$$

Observe that the matrix inverse is a differentiable mapping at Σ_0 (which is $CV^0C'+D_0^{-1}\succeq CV^0C'$ and thus positive definite). Thus we have

$$\hat{\Sigma}^{-1}=\Sigma_0^{-1}+O\left(\frac{1}{t^2};c_0\right).$$

Plugging this into (30) and using $\hat{q}_j\sim\lambda_j t$, we obtain that

$$\partial_j f(\hat{q}_j,q_{-j})=-\frac{\sigma_j^2}{\hat{q}_j^2}\cdot\left[V^0C'\Sigma_0^{-1}\Delta_{jj}\Sigma_0^{-1}CV^0\right]_{11}+O\left(\frac{1}{t^4};c_0\right). \tag{31}$$

[93]Formally, the Equi-distribution Theorem states that for any irrational number α and any sub-interval $(a,b)\subset(0,1)$, the set of positive integers n such that the fractional part of αn belongs to (a,b) has natural density $b-a$. It is a special case of the Ergodic Theorem.

Since $\Sigma_0 = CV^0C' + \frac{1}{t}\cdot \text{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)$, we can apply Taylor expansion (to the matrix inverse map) and write

$$\Sigma_0^{-1} = (CV^0C')^{-1} - \frac{1}{t}(CV^0C')^{-1}\cdot \text{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)\cdot (CV^0C')^{-1} + O\left(\frac{1}{t^2}\right). \tag{32}$$

This implies

$$\begin{aligned} V^0C'\Sigma_0^{-1}\Delta_{jj}\Sigma_0^{-1}CV^0 &= V^0C'(CV^0C')^{-1}\Delta_{jj}(CV^0C')^{-1}CV^0 - \frac{M_j}{t} + O\left(\frac{1}{t^2}\right) \\ &= Q_j - \frac{M_j}{t} + O\left(\frac{1}{t^2}\right), \end{aligned} \tag{33}$$

where M_j is a fixed $K \times K$ matrix depending only on the informational environment. For future use, we note that

$$\begin{aligned} M_j &= V^0C'(CV^0C')^{-1}\,\text{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)(CV^0C')^{-1}\Delta_{jj}(CV^0C')^{-1}CV^0 \\ &\qquad + V^0C'(CV^0C')^{-1}\Delta_{jj}(CV^0C')^{-1}\,\text{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)(CV^0C')^{-1}CV^0 \\ &= C^{-1}\,\text{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)(CV^0C')^{-1}\Delta_{jj}C'^{-1} \\ &\qquad + C^{-1}\Delta_{jj}(CV^0C')^{-1}\,\text{diag}\left(\frac{\sigma_1^2}{\lambda_1}, \ldots, \frac{\sigma_K^2}{\lambda_K}\right)C'^{-1}. \end{aligned} \tag{34}$$

Using (33), we can simplifty (31) to

$$\partial_j f(\hat{q}_j, q_{-j}) = -\frac{\sigma_j^2}{\hat{q}_j^2}\cdot\left[Q_j - \frac{M_j}{t}\right]_{11} + O\left(\frac{1}{t^4}; c_0\right). \tag{35}$$

Integrating this over $\hat{q}_j \in [q_j, q_j+1]$, we conclude that

$$f(q_j, q_{-j}) - f(q_j+1, q_{-j}) = \frac{\sigma_j^2}{q_j(q_j+1)}\cdot\left[Q_j - \frac{M_j}{t}\right]_{11} + O\left(\frac{1}{t^4}; c_0\right). \tag{36}$$

We set $a_j = -\left(\frac{\lambda_j\cdot[M_j]_{11}}{2[Q_j]_{11}} + \frac{1}{2}\right)$. Then,

$$\frac{\sigma_j^2}{q_j(q_j+1)}\cdot\left[Q_j - \frac{M_j}{t}\right]_{11} = (\sigma_j^2\cdot[Q_j]_{11})\cdot\frac{1+\frac{2a_j+1}{\lambda_j t}}{q_j(q_j+1)} = \frac{\sigma_j^2\cdot[Q_j]_{11}}{(q_j-a_j)^2} + O\left(\frac{1}{t^4}; c_0\right),$$

implying the desired approximation (28). The last equality above uses $\frac{1+\frac{2a_j+1}{\lambda_j t}}{q_j(q_j+1)} = \frac{1}{(q_j-a_j)^2} + O\left(\frac{1}{t^4}; c_0\right)$, which is because

$$\frac{q_j(q_j+1)}{(q_j-a_j)^2} = 1 + \frac{2(a_j+1)}{q_j-a_j} + O\left(\frac{1}{(q_j-a_j)^2}\right) = 1 + \frac{2a_j+1}{\lambda_j t} + O\left(\frac{1}{t^2}; c_0\right)$$

dividing through by $q_j(q_j+1)$. □

G.2 A Simultaneous Diophantine Approximation Problem

Lemma 8 tells us that at most times t, there do not exist a *pair* of divisions (differing minimally on two signal counts) that lead to posterior variances close to each other (with a difference of $\frac{c_1}{t^4}$). We obtain a stronger result if a *triple* of such divisions were to exist.

Lemma 10. *Fix V^0 and C, and let signal variances vary. For positive constants c_0, c_1, define $\mathcal{A}^*(c_0, c_1)$ to be the following set of positive integers:*

$$\begin{aligned}\{t : \exists\ q_1, q_2, q_3, \ldots, q_K \in \mathbb{Z}^+, s.t.\ |q_i - \lambda_i t| \le c_0, \forall i \\ \wedge\ |f(q_1, q_2 + 1, q_3, \ldots, q_K) - f(q_1 + 1, q_2, q_3, \ldots, q_K)| \le c_1/t^4 \\ \wedge\ |f(q_1, q_2, q_3 + 1, \ldots, q_K) - f(q_1 + 1, q_2, q_3, \ldots, q_K)| \le c_1/t^4\}\end{aligned}$$

Then, except for signal variances belonging to a Lebesgue measure-zero set, $\mathcal{A}^*(c_0, c_1)$ has finite cardinality.

Proof. So far we have been dealing with fixed informational environments. However, a number of parameters defined above depend on the signal variances $\sigma = \{\sigma_i^2\}_{i=1}^K$. Specifically, while the matrix $Q_i = C^{-1}\Delta_{ii}C'^{-1}$ is independent of σ, the asymptotic proportions $\lambda_i \propto \sigma_i \cdot [Q_i]_{11}$ do vary with σ. In this proof, we write $\lambda_i(\sigma)$ to highlight this dependence.

Next, we recall the matrix M_j introduced earlier in (34). We note that for fixed matrices V^0 and C, each entry of $M_j(\sigma)$ is a fixed linear combination of $\frac{\sigma_1^2}{\lambda_1(\sigma)}, \ldots, \frac{\sigma_K^2}{\lambda_K(\sigma)}$.

Then, the parameter $a_j(\sigma)$ in (28) is given by (see the previous proof)

$$a_j(\sigma) = -\frac{1}{2} - \frac{\lambda_j(\sigma) \cdot [M_j(\sigma)]_{11}}{2[Q_j]_{11}} = -\frac{1}{2} + \lambda_j(\sigma) \sum_{i=1}^K \tilde{b}_{ji} \frac{\sigma_i^2}{\lambda_i(\sigma)} = -\frac{1}{2} + \sum_{i=1}^K b_{ji}\sigma_i\sigma_j \tag{37}$$

for some constants $\tilde{b}_{ji}, b_{ji}$ independent of σ. In the last equality above, we used the fact that $\frac{\lambda_j(\sigma)}{\lambda_i(\sigma)}$ equals a constant times $\frac{\sigma_j}{\sigma_i}$.

Thus Lemma 9 gives

$$f(q_j, q_{-j}) - f(q_j + 1, q_{-j}) = \frac{\sigma_j^2 \cdot [Q_j]_{11}}{(q_j - a_j(\sigma))^2} + O\left(\frac{1}{t^4}; c_0\right)$$

whenever $|q_i - \lambda_i(\sigma) \cdot t| \le c_0, \forall i$. We comment that the Big O constant here may depend on σ. However, a single constant suffices if we restrict each σ_i to be bounded above and bounded away from zero. Since measure-zero sets are closed under countable unions, this restriction does not affect the result we want to prove.

By the above approximation, a necessary condition for $t \in \mathcal{A}^*(c_0, c_1)$ is that q_1, q_2, q_3 satisfy

$$\left|(q_2 - a_2(\sigma)) - \frac{\eta \cdot \sigma_2}{\sigma_1}(q_1 - a_1(\sigma))\right| \leq \frac{c_6}{q_1} \tag{38}$$

as well as

$$\left|(q_3 - a_3(\sigma)) - \frac{\kappa \cdot \sigma_3}{\sigma_1}(q_1 - a_1(\sigma))\right| \leq \frac{c_6}{q_1} \tag{39}$$

for some constant c_6 independent of σ (c_6 may depend on c_0, c_1 stated in the lemma). The constant η is given by $\eta = \sqrt{[Q_2]_{11}/[Q_1]_{11}}$, and similarly for κ.

It remains to show that for generic σ, there are only finitely many positive integer triples (q_1, q_2, q_3) satisfying the *simultaneous Diophantine approximation* (38) and (39). To prove this, we assume that each σ_i is i.i.d. drawn from the uniform distribution on $[\frac{1}{L}, L]$, where L is a large constant.[94] Denote by $F(q_1, q_2, q_3)$ the event that (38) and (39) hold simultaneously. We claim that there exists a constant c_7 such that $\mathbb{P}(F(q_1, q_2, q_3)) \leq \frac{c_7}{q_1^4}$ holds for all q_1, q_2, q_3. Since $F(q_1, q_2, q_3)$ cannot occur for $q_2, q_3 > c_8 q_1$, this claim will imply

$$\sum_{q_1, q_2, q_3} \mathbb{P}(F(q_1, q_2, q_3)) < \sum_{q_1} \sum_{q_2, q_3 \leq c_8 q_1} \frac{c_7}{q_1^4} < \sum_{q_1} \frac{c_7 c_8^2}{q_1^2} < \infty. \tag{40}$$

Generic finiteness of tuples (q_1, q_2, q_3) will then follow from the Borel-Cantelli Lemma.[95]

To prove this claim, it suffices to show that if $\sigma = (\sigma_1, \sigma_2, \sigma_3, \sigma_4, \ldots, \sigma_K)$ and $\sigma' = (\sigma_1, \sigma_2', \sigma_3', \sigma_4, \ldots, \sigma_K)$ both satisfy (38) and (39), then $|\sigma_2 - \sigma_2'|, |\sigma_3 - \sigma_3'| \leq \frac{c}{q_1^2}$ for some constant c.[96] Without loss, we assume $|\sigma_2 - \sigma_2'| \geq |\sigma_3 - \sigma_3'|$. Using (37), we can rewrite the condition (38) as

$$\left| \underbrace{\left(q_2 + \frac{1}{2}\right) - \frac{\eta \cdot \sigma_2}{\sigma_1}\left(q_1 + \frac{1}{2}\right)}_{A} + \underbrace{\sum_i \beta_i \sigma_2 \sigma_i}_{B} \right| \leq \frac{c_6}{q_1}$$

[94] The uniform distribution most directly implies the Lemma regarding Lebesgue measure zero, but any other continuous distribution is equally fine for this argument.

[95] Because of the use of Borel-Cantelli Lemma, this proof (unlike Lemma 8 above) does not allow us to effectively determine, for given σ, whether (38) and (39) only have finitely many integer solutions. Nonetheless, a modification of this proof does imply the following finite-time probabilistic statement: when $\sigma_1, \ldots, \sigma_K$ are independently drawn, the probability that myopic, dynamic and t-optimality coincide at every period $t \geq T$ is at least $1 - O(\frac{1}{T})$, where the constant involved only depends on the distribution of σ.

[96] This implies that the probability of the event $F(q_1, q_2, q_3)$ conditional on any value of $\sigma_1, \sigma_4, \ldots, \sigma_K$ is bounded by $\frac{c_7}{q_1^4}$, which is stronger than the claim.

for some constants β_i independent of σ. A similar inequality holds at σ':

$$\left| \underbrace{\left(q_2 + \frac{1}{2}\right) - \frac{\eta \cdot \sigma_2'}{\sigma_1}\left(q_1 + \frac{1}{2}\right)}_{A'} + \underbrace{\sum_i \beta_i \sigma_2' \sigma_i'}_{B'} \right| \leq \frac{c_6}{q_1}$$

It follows from the above two inequalities that $|A + B - A' - B'| \leq \frac{2c_6}{q_1}$. Furthermore, since $|A - A'| \leq |A + B - A' - B'| + |B - B'|$ (by triangle inequality), we deduce

$$\left| \frac{\eta \cdot (\sigma_2' - \sigma_2)}{\sigma_1} \cdot \left(q_1 + \frac{1}{2}\right) \right| \leq \frac{2c_6}{q_1} + \left| \sum_i \beta_i (\sigma_2' \sigma_i' - \sigma_2 \sigma_i) \right| \tag{41}$$

Because $\sigma_i' = \sigma_i$ for $i \neq 2, 3$, we have

$$\begin{aligned} \left| \sum_i \beta_i (\sigma_2' \sigma_i' - \sigma_2 \sigma_i) \right| &= \left| \sum_i \beta_i (\sigma_2' - \sigma_2)\sigma_i + \sum_i \beta_i \sigma_2' (\sigma_i' - \sigma_i) \right| \\ &= \left| \left(\sum_i \beta_i (\sigma_2' - \sigma_2)\sigma_i \right) + \beta_2 \sigma_2' (\sigma_2' - \sigma_2) + \beta_3 \sigma_2' (\sigma_3' - \sigma_3) \right| \\ &\leq (K + 2) L \cdot \max_i |\beta_i| \cdot |\sigma_2' - \sigma_2| \end{aligned}$$

Plugging this estimate into (41), we obtain the desired result $|\sigma_2 - \sigma_2'| \leq \frac{c}{q_1^2}$. This completes the proof of the lemma. □

G.3 Monotonicity of t-Optimal Divisions

We apply Lemma 10 to prove the eventual monotonicity of t-optimal divisions in generic informational environments.

Lemma 11. *Fix V^0 and C. For generic signal variances $\{\sigma_i^2\}_{i=1}^K$, there exists T_0 such that for $t \geq T_0$, the t-optimal division $n(t)$ is unique, and it satisfies $n_i(t+1) \geq n_i(t), \forall i$.*

Proof. Uniqueness follows from the stronger fact that in generic informational environments, $f(q_1, \ldots, q_K)$ differs from $f(q_1', \ldots, q_K')$ whenever $q \neq q'$. This is because any such equality corresponds to a non-trivial polynomial equation over σ, and there are only countably many of them to be ruled out.

Using the order difference lemma, we can already deduce $|n_i(t+1) - n_i(t)| \leq 1, \forall i$ at sufficiently large times t. Suppose that $n_1(t+1) = n_1(t) - 1$. Then because $\sum_i (n_i(t+1) - n_i(t)) = 1$, we can without loss assume $n_2(t+1) = n_2(t) + 1$ and $n_3(t+1) = n_3(t) + 1$.

For notational ease, write $n_i = n_i(t), n_i' = n_i(t+1)$. By t-optimality, we have

$$f(n_1 - 1, n_2 + 1, n_3, \ldots, n_K) \geq f(n_1, n_2, n_3, \ldots, n_K)$$

$$f(n_1', n_2', n_3', \ldots, n_K') \leq f(n_1' + 1, n_2' - 1, n_3', \ldots, n_K')$$

These inequalities are equivalent to

$$\partial_2 f(n_1 - 1, n_2, n_3, \ldots, n_K) \geq \partial_1 f(n_1 - 1, n_2, n_3, \ldots, n_K) \tag{42}$$

$$\partial_2 f(n_1', n_2' - 1, n_3', \ldots, n_K') \leq \partial_1 f(n_1', n_2' - 1, n_3', \ldots, n_K') \tag{43}$$

with $\partial_i f$ representing the *discrete partial derivative*.

Since $n_2' - 1 = n_2$, the LHS of (43) is at most smaller than the LHS of (42) by a number of cross partials. Similarly, the RHS of (43) is at most bigger than the RHS of (42) by a number of cross partials. Thus the two sides of (42) cannot differ by more than $O(\frac{1}{t^4})$. That is, for some constant c_1 we have[97]

$$|f(n_1 - 1, n_2 + 1, n_3, \ldots, n_K) - f(n_1, n_2, n_3, \ldots, n_K)| \leq \frac{c_1}{t^4}. \tag{44}$$

An analogous argument using $n_1' = n_1 - 1$ and $n_3' = n_3 + 1$ yields

$$|f(n_1 - 1, n_2, n_3 + 1, \ldots, n_K) - f(n_1, n_2, n_3, \ldots, n_K)| \leq \frac{c_1}{t^4}. \tag{45}$$

Moreover, from Proposition 3, there is a constant c_0 such that

$$|n_i - \lambda_i t| \leq c_0, \forall i. \tag{46}$$

Now, Lemma 10 says that in generic environments, there are only finitely many integer solutions $(n_1, \ldots, n_K)$ to (44), (45) and (46). This proves what we want. □

G.4 Completing the Proof of Theorem 2

By Lemma 11, generically there exists T_0 such that $n(t)$ is monotonic over time in each signal for $t \geq T_0$. If a myopic or forward-looking DM achieves t-optimality at some time $t \geq T_0$, he can and thus will continue to do so at every future time (because such a strategy maximizes the payoff in every period). Thanks to Proposition 4, such a time t exists. The theorem is proved.

[97] As discussed in the proof of Lemma 10, we can find a single constant c_1 that works for all σ bounded above and bounded away from zero.

Appendix H Proof of Theorem 3 (Batch of Signals)

Observing B signals each period is the same as planning B periods at once in the original model. Let us use the latter interpretation and show that the t-optimal division is eventually increasing every B periods, whenever $B \geq K-1$. This is because by the order difference lemma, if t is large and $n_i(t+B) \leq n_i(t) - 1$ for some signal i, then $n_j(t+B) \leq n_j(t) + 1$ for every other signal j.[98] But these inequalities would together imply $B = (n_i(t+B) - n_i(t)) + \sum_{j \neq i}(n_j(t+B) - n_j(t)) \leq K-2$, leading to a contradiction.

For sufficiently large B, it is clear that these t-optimal divisions are monotone from the beginning. Thus any myopic strategy can and thus will achieve t-optimality and dynamic optimality.

Appendix I Equivalence Results for $K = 2$

I.1 Immediate Equivalence with Complementary Signals

When there are only two states and two signals, immediate equivalence holds for a broad class of environments:[99]

Proposition 5. *Suppose $K = 2$, the prior is standard Gaussian ($V^0 = \mathbf{I}_2$), and both signals have variance 1.*[100] *Write $C = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and assume without loss that $|ad| \geq |bc|$. Then myopic, dynamic and t-optimality coincide from period 1 whenever the following inequality holds:*

$$(1 + 2b^2) \cdot |ad - bc| \geq |ad + bc|. \tag{47}$$

In particular, this is true whenever $abcd \leq 0$.

Proof. Under the assumptions, the DM's posterior variance about θ_1 is computed to be

$$f(q_1, q_2) = \frac{1 + b^2 q_1 + d^2 q_2}{1 + (a^2 + b^2)q_1 + (c^2 + d^2)q_2 + (ad - bc)^2 q_1 q_2}$$

[98] The argument is similar to (18), which we omit.

[99] In the proposition below, if the linear coefficients a, b, c, d were picked at random, then with probability $\frac{1}{2}$ we would have $abcd \leq 0$. Immediate equivalence follows.

[100] We make these simplifying assumptions on prior and signal variances so that the condition for equivalence is easy to state and interpret. They are not essential in any way.

Given q_i observations of each signal i in the past, a myopic DM chooses signal 1 if and only if $f(q_1+1, q_2) < f(q_1, q_2+1)$, which reduces to

$$\begin{aligned}&(ad-bc)^2b^2q_1^2 + (1+b^2)(ad-bc)^2q_1 - (a^2d^2-b^2c^2)q_1 + c^2(1+b^2) \\ <&(ad-bc)^2d^2q_2^2 + (1+d^2)(ad-bc)^2q_2 + (a^2d^2-b^2c^2)q_2 + a^2(1+d^2)\end{aligned} \tag{48}$$

The condition $|ad| \geq |bc|$ ensures that the RHS is an increasing function of q_2, because the coefficients in front of q_2^2 and q_2 are both positive. Meanwhile, the condition $(1+2b^2)|ad-bc| \geq |ad+bc|$ implies the LHS is larger when $q_1 = 1$ than when $q_1 = 0$, so that the LHS is increasing in (integer values of) q_1. Thus, the proof of Lemma 1 extends without change.[101] Any myopic strategy is t-optimal and dynamically optimal. □

To interpret, (47) requires that the determinant of the matrix C, $ad-bc$, to be not too small (holding other terms constant). Put differently, the two vectors (in $\mathbb{R}^2$) defining the signals should not be close to collinear. This rules out situations where the two signals provide such similar information in the initial periods that they substitute one another.[102]

I.2 Eventual Equivalence

Even when immediate equivalence does not obtain, the signal paths under myopic, dynamic and t-optimality are eventually *exactly* identical for the case of $K=2$. This holds for arbitrary prior beliefs and signals, not just generically (c.f. Theorem 2).

Theorem 6. *Suppose $K=2$ and suppose the informational environment satisfies Assumption 3.*[103] *There exists a large time T^* such that the following holds: any myopic division $m(t)$ is t-optimal at times $t \geq T^*$, and so is any dynamically optimal division $d(t)$ for any decision problem.*

[101] Specifically, after a one-shot deviation from signal 1 to signal 2, the deviation path keeps observing signal 1 until the first time the original myopic path observes signal 2. The same exchangeability argument then shows that the posterior variances along the deviation path are uniformly larger than the original path.

[102] For example, when $a=3, b=1, c=5, d=2$, it can be shown that the myopic path begins with $BXXXXXXXXX$, while the dynamically optimal path (for the prediction problem and $\delta=0.9$) is to observe $XXXXXXXXXB$. Lemma 1 fails here because following a one-shot deviation from B to X in the first period, the DM finds it myopically better to *continue* observing X rather than B, as $f(2,0) < f(1,1)$. This suggests strong substitution between the two signals so that deviating to X in the first period decreases the marginal value of signal B *even more* than it does signal X.

[103] Recall that Assumption 4 is always satisfied when $K=2$

That is, when $K = 2$, any myopic or dynamically optimal strategy achieves t-optimality after finitely many periods.[104]

A key step of the proof is to show that when $K = 2$, the "constrained t-optimal divisions" (see Definition 3 in Appendix C) are eventually monotonic, so that payoff can (and thus will) be minimized period by period.

Lemma 12. *Fix an informational environment $(V^0, C, \{\sigma_i^2\})$ with $K = 2$. There exists H^* such that for any history h with signal counts $H_1, H_2 \geq H^*$, the constrained t-optimal divisions following h satisfy $n_i^h(t+1) \geq n_i^h(t)$, $\forall i$, $\forall t \geq H$.*

Proof of Lemma 12. We first show that the unconstrained t-optimal divisions are monotonic for $t \geq 2H^*$.[105] Let (n_1, n_2) be the t-optimal division. Then by Proposition 3, the $t+1$-optimal division is $n_1 + b + 1, n_2 - b$ for some bounded integer b. Without loss assume $b \geq 0$, then we need to show b is in fact equal to zero. This is because for fixed $b \geq 1$,

$$f(n_1 + 1, n_2 - 1) \geq f(n_1, n_2) \implies f(n_1 + b + 1, n_2 - b) > f(n_1 + b, n_2 + 1 - b) \tag{49}$$

as n_1, n_2 increase to ∞ proportionally. The proof of this uses the order difference lemma, similar to (18), so we omit it. Since the LHS of (49) holds by t-optimality, we deduce that the unconstrained t-optimal divisions are indeed increasing over time.

Let us turn to the constrained divisions. Fix a history h with length $H \geq 2H^*$. If $n^h(H)$ is unconstrained t-optimal, then by monotonicity, any future unconstrained t-optimal division satisfies the constraints. Thus $n^h(t)$ is unconstrained t-optimal at all $t \geq H$. The conclusion of the lemma holds in this case.

Otherwise, suppose without loss that $H_1 < n_1(H)$ and $H_2 > n_2(H)$. This means the DM has observed an excess of signal 2 than optimal. Denote by $T > H$ the first time at which $n_2(T) = H_2$; such a time exists because $n_2(t)$ increases by 0 or 1 each time. We claim that for $H \leq t < T$, $n^h(t) = (t - H_2, H_2)$, and for $t \geq T$, $n^h(t) = n(t)$. This will prove the lemma.

Intuitively, the claim says that the DM should keep observing signal 1 until its signal count catches up with the unconstrained optimum. To prove it formally, we note from the convexity (single-peakedness) of f that

$$f(n_1(t), n_2(t)) < f(t - H_2, H_2) \implies$$
$$f(t - H_2, H_2) < f(t - q_2, q_2), \forall q_2 > H_2 > n_2(t).$$

[104] We do not state the result as $m(t) = d(t) = n(t)$ because these divisions need not be unique.

[105] Together with Proposition 4, this monotonicity is sufficient to prove Theorem 6 under the extra assumption that λ_1/λ_2 is an irrational number. Our proof below does not rely on this assumption.

By the choice of T, we have $n_2(t) < H_2$ for $t \in [H, T)$. Thus the above display implies $n^h(t) = (t - H_2, H_2)$ is the constrained optimum at such times t. For larger t, the unconstrained t-optimal division satisfies the constraints. Hence the constrained optimum must also be an unconstrained optimum, proving the claim and the lemma. □

With Lemma 12, we can quickly prove Theorem 6 as follows. The lemma implies that following any history h with $H_i \geq H^*$, a myopic continuation strategy can, and thus will, achieve constrained t-optimality at every time. Such a strategy maximizes the flow payoff in every period following h, so it is dynamically optimal. By Proposition 3, $H_i \geq H^*$ holds along the myopic and dynamically optimal signal paths whenever the history h is sufficiently long. Moreover, any constrained optimum eventually becomes unconstrained optimum. Theorem 6 is proved.

Appendix J Examples Without Eventual Equivalence

J.1 Myopic Not t-Optimal Infinitely Often

Here we continue to study Example 4 presented in the main text. Consider the t-optimal division at time t. The payoff function in (6) suggests that the problem can be separated into two parts: choosing q_X, and allocating the remaining observations between q_1 and q_2. The latter allocation problem is identical to the benchmark case—an optimal division between signals B_1 and B_2 satisfies $q_2 = q_1 - 1$ or $q_2 = q_1$. With some extra algebra, we obtain that for $N \geq 1$:

1. If $t = 3N + 1$, then the unique t-optimal division (q_X, q_1, q_2) is $(N + 2, N, N - 1)$;
2. If $t = 3N + 2$, then the unique t-optimal division is $(N + 3, N, N - 1)$;
3. If $t = 3N + 3$, then the unique t-optimal division is $(N + 2, N + 1, N)$.

Crucially, note that when transitioning from $t = 3N + 2$ to $t = 3N + 3$, the t-optimal number of X signals is decreased. This reflects the complementarity between signals B_1 and B_2, which causes the DM to observe them in pairs. Due to this failure of monotonicity, t-optimality at all large times is unattainable.[106]

[106] We can however show that in this example, any myopic strategy maximizes the ex-ante expected payoff (in the prediction problem), whenever the DM is moderately impatient (specifically, $\delta \leq 0.91$). Moreover, for any discounting $\delta < 1$, the myopic signal path eventually coincides with the forward-looking solution.

J.2 Myopic Not Dynamically Optimal

Next, we provide another example in which, for *any* positive discount factor δ, the myopic division differs from the dynamically optimal division at infinitely many periods. The setup is almost the same as Example 4, except that the noise term ϵ_X has variance 2. The DM has posterior variance about θ given by

$$f(q_X, q_1, q_2) = 1 - \frac{1}{1 + \frac{2}{q_X} + 1 - \frac{1}{1+\frac{1}{q_1}+\frac{1}{1+q_2}}} \tag{50}$$

The change from the previous example is that $\frac{2}{q_X}$ appears in the denominator, rather than $\frac{1}{q_X}$.

As in the analysis of the previous example, there will be times $\hat{t}-1$ and $\hat{t}$ at which the t-optimal division transitions from $(M, N, N-1)$ to $(M-1, N+1, N)$, for some positive integers M, N. By t-optimality, we must have $f(M, N, N-1) \leq f(M-1, N, N)$ and $f(M-1, N+1, N) \leq f(M, N, N)$. Using (50), we deduce that any such pair (M, N) satisfies

$$2\left(N^2+5N+6\right) - 2 - \frac{4}{N+1} \leq M(M-1) \leq 2\left(N^2+5N+6\right) + 2 + \frac{4}{N}. \tag{51}$$

Observe that for any $N > 2$, we have $\frac{4}{N+1} < 2$ and $\frac{4}{N} < 2$. Since both $M(M-1)$ and $2\left(N^2+5N+6\right)$ are even numbers, the above condition reduces to

$$M(M-1) = 2(N^2+5N+6+\phi), \tag{52}$$

where $\phi \in \{-1, 0, 1\}$. If $\hat{t} = M + 2N$ for such M, N, we call it a "bad" time.

We can show by induction that the myopic strategy achieves t-optimality at any time t that is not "bad". Specifically, suppose $\hat{t}$ is a bad time and the myopic division at time $\hat{t}-2$ is the t-optimal division $(M-1, N, N-1)$.[107] Then, at time $\hat{t}-1$, the DM's myopic choice is to observe signal X, which achieves the t-optimal division $(M, N, N-1)$. t-optimality cannot be achieved at the bad time $\hat{t}$, but the myopic DM observes B_1 to reach the division $(M, N+1, N-1)$, since $f(M, N+1, N-1) = f(M, N, N) < f(M+1, N, N-1)$. Next, at time $\hat{t}+1$, the DM myopically observes signal B_2. This yields the division $(M, N+1, N)$, which is t-optimal at time $\hat{t}+1$. Afterwards, the t-optimal divisions are monotonic until the next bad time.[108] Hence the myopic strategy preserves t-optimality until then.

Thus this is an example in which myopic information acquisition eventually achieves dynamic optimality, but does not eventually achieve t-optimality.

[107] t-optimality can be verified using (50) and (52).

[108] The key is to show that $n_1(t+1) - n_1(t), n_2(t+1) - n_2(t) \in \{0, 1\}$. Given this, non-monotonicity of the t-optimal division must mean that $n_X(t+1) = n_X(t) - 1, n_1(t+1) = n_1(t) + 1, n_2(t+1) = n_2(t) + 1$.

Take any pair (M, N) satisfying (52) with $\phi = 1$. This pair determines a bad time $\hat{t} = M + 2N$. From the preceding paragraph, the myopic strategy observs XB_1B_2 at times $\hat{t}-1, \hat{t}$ and $\hat{t}+1$. Consider a deviation strategy which observes B_1B_2X in these 3 periods, and which agrees with the myopic signal path in all other periods. The flow payoffs obtained under the deviation differ from the payoffs under the myopic strategy only in two periods: $\hat{t}-1$ and $\hat{t}$.

At time $\hat{t}-1$, the deviation does worse by an amount of

$$f(M-1, N+1, N-1) - f(M, N, N-1) = \frac{\frac{2}{M(M-1)} - \frac{1}{N^2+5N+7+\frac{2}{N}}}{\left(1 + \frac{2}{M-1} + 1 - \frac{1}{1+\frac{1}{N}+\frac{1}{1+N}}\right)\left(1 + \frac{2}{M} + 1 - \frac{1}{1+\frac{1}{N}+\frac{1}{N}}\right)} \tag{53}$$

At time $\hat{t}$ however, the deviation does better by an amount of

$$f(M, N+1, N-1) - f(M-1, N+1, N) = \frac{\frac{1}{N^2+5N+5-\frac{2}{N+1}} - \frac{2}{M(M-1)}}{\left(1 + \frac{2}{M-1} + 1 - \frac{1}{1+\frac{1}{1+N}+\frac{1}{1+N}}\right)\left(1 + \frac{2}{M} + 1 - \frac{1}{1+\frac{1}{N}+\frac{1}{1+N}}\right)} \tag{54}$$

Since $M(M-1) = 2(N^2 + 5N + 7)$, the RHS of (53) is of order $\frac{1}{N^5}$, while the RHS of (54) is of order $\frac{1}{N^4}$. Thus for any positive discount factor δ, the deviation strategy achieves higher discounted total payoff than the myopic strategy whenever N is sufficiently large.[109]

Thus, for any $\delta > 0$, it cannot be the case that $m(t) = d(t)$ for every large t; otherwise, a 3-period deviation constructed above (with N sufficiently large) achieves higher ex-ante payoff.[110]

[109] The Diophantine equation (52) has infinitely many positive integer solutions (M, N) for $\phi = 1$. To see this, we rewrite (52) as $(2M-1)^2 - 2(2N+5)^2 = 7$, which is a *Pell's equation*. All integer solutions are characterized by $(2M-1) + (2N+5)\sqrt{2} = (3 \pm \sqrt{2})(3 + 2\sqrt{2})^{\ell}$, with ℓ an arbitrary integer. Thus there are infinitely many bad times, corresponding to $(M, N) = (6, 2), (14, 7), \ldots$.

[110] Our analysis leaves open the possibility that $m(t) - d(t)$ equals a nonzero constant vector at large times t, so that the myopic and forward-looking *signal paths* are eventually the same, but the *signal counts* continue to differ. We can however show it does not occur in this example: first recall that we have shown $m(t) = n(t)$ at a set of (bad) times with natural density 1. Suppose $d(t) \neq m(t)$ for all large t, then $d(t) \neq n(t)$ at almost all t. But $\lambda_X/\lambda_1 = \lambda_X/\lambda_2 = \sqrt{2}$, so the proof of Proposition 4 implies $d_X(t)$ cannot be different from $n_X(t)$ (or $m_X(t)$). The remaining case is if $d_1(t) = m_1(t)+1, d_2(t) = m_2(t)-1$ (or vice versa) for large t. Note that, holding fixed $m_X(t)$, the myopic division always minimizes the DM's posterior variance (since any division $m_1(t), m_2(t)$ with $m_1(t) - m_2(t) \in \{1, 2\}$ is t-optimal for the prediction of b_1). Thus the forward-looking DM can switch an observation of signal B_1 to signal B_2 (or vice versa) and reduce his posterior variance at every future period, leading to a contradiction.

J.3 Beyond Linear Signals

The example below highlights the significance of linear signals for our equivalence results.

Example 7. Consider a setup similar to Example 4 before, with three states θ, b_1, b_2 drawn independently. The DM has access to three signals

$$\begin{aligned} X &= \theta + \mathrm{sign}(b_1) + \epsilon_X \\ B_1 &= \mathrm{sign}(b_1 b_2) + \epsilon_1 \\ B_2 &= b_2 + \epsilon_2 \end{aligned}$$

where $\epsilon_X, \epsilon_1, \epsilon_2$ are Gaussian noise terms. The distinction from Example 4 is that the signals here are not *linear* combinations of the unknown states, but rather depend on the "sign" of b_1 and b_2. We focus on the prediction problem, in which the DM seeks to minimize squared prediction error about θ.

Note that prior to the first observation of signal B_1, signal B_2 is completely uninformative about the payoff-relevant state θ (even when combined with previous observations of signal X). Similarly, signal B_1 is individually uninformative about b_1 and thus about θ — the sign of $b_1 b_2$ does not contain any new information about b_1 when b_2 is equally likely to be positive or negative. These imply that the DM's uncertainty about θ is not reduced upon the first observation of either B_1 or B_2. Hence, the myopic strategy in this example is to always observe X. A myopic DM *gets stuck in observing* X, contrary to Lemma 6.

In so doing, a myopic DM never fully learns the value of θ. By contrast, a forward-looking DM would observe each signal infinitely often and completely identify the value of θ. Thus, eventual equivalence between the myopic and dynamically optimal signal paths fails (drastically) in this example for *all* signal variances, violating the conclusion of Theorem 2.

Appendix K Proof of Theorem 4 (Time to "Eventual Gap One")

The proof of the bound resembles the proof of Theorem 1 Part (a), except that we need sharper estimates (for the posterior variance function and its derivatives). We now turn to these estimates. Throughout, we work with the linearly-transformed model, where each signal X_i is simply $\tilde{\theta}_i$ plus standard Gaussian noise, and the DM's prior covariance matrix over the transformed states is V. Let $w = (1, \dots, 1)'$, and $\gamma = \gamma(q_1, \dots, q_K)$ represents the following $K \times 1$ vector:

$$\gamma' = w' \cdot V(V+E)^{-1}$$

with $E = \text{diag}(\frac{1}{q_1}, \ldots, \frac{1}{q_K})$. Our more precise approximations of the posterior variance function f and its derivatives are based on this vector γ. For $1 \le i \le K$, γ_i denotes the i-th coordinate of γ.

K.1 Preliminary Estimates

In this subsection, we re-derive the posterior variance function f, its derivatives and second derivatives. The formulae below take as primitives V and w, but they are equivalent to those presented in Appendix A (for the original model). *All partial derivatives in this subsection are the usual continuous derivative.*

Fact 1 (Posterior Variance)**.**

$$f(q_1, \ldots, q_K) = w'(V - V(V+E)^{-1}V)w. \tag{55}$$

Fact 2 (Partial Derivatives of Posterior Variance)**.**

$$\partial_i f(q_1, \ldots, q_K) = -\frac{1}{q_i^2} \cdot w'V(V+E)^{-1}\Delta_{ii}(V+E)^{-1}Vw = -\frac{\gamma_i^2}{q_i^2}. \tag{56}$$

Fact 3 (Second-Order Partial Derivatives of Posterior Variance)**.**

$$\begin{aligned}
&\partial_{ii} f(q_1, \ldots, q_K) \\
&\quad= \frac{2 \cdot w'V(V+E)^{-1}\Delta_{ii}(V+E)^{-1}Vw}{q_i^3} - \frac{2 \cdot w'V(V+E)^{-1}\Delta_{ii}(V+E)^{-1}\Delta_{ii}(V+E)^{-1}Vw}{q_i^4} \\
&\quad= \frac{2\gamma_i^2}{q_i^3} \cdot \left(1 - \frac{[(V+E)^{-1}]_{ii}}{q_i}\right)
\end{aligned} \tag{57}$$

Fact 4 (Cross-Partial Derivatives of Posterior Variance)**.**

$$\begin{aligned}
\partial_{ij} f(q_1, \ldots, q_K) &= \frac{-2}{q_i^2 q_j^2} \cdot w'V(V+E)^{-1}\Delta_{ii}(V+E)^{-1}\Delta_{jj}(V+E)^{-1}Vw \\
&= \frac{-2\gamma_i\gamma_j}{q_i^2 q_j^2} \cdot [(V+E)^{-1}]_{ij}.
\end{aligned} \tag{58}$$

The above facts are proved by simple linear algebra. The upshot is the following result which quantifies the asymptotic characterization in Proposition 3 (easy to see that $\lambda_i = 1$):

Proposition 6. *For $T \ge 21.5(R+1)K^2$, it holds that $0.948 \cdot \frac{T}{K} \le m_i(T), n_i(T) \le 1.054 \cdot \frac{T}{K}$.*

Proof. We will derive the bounds for $m_i(t)$, which is more difficult. Note from the definition of γ that $\gamma' \cdot (V+E) = w' \cdot V$, thus $(V+E)\gamma = Vw$ and $V(w-\gamma) = E\gamma = (\frac{\gamma_1}{q_1}, \dots, \frac{\gamma_K}{q_K})'$. It follows that

$$w - \gamma = V^{-1} \cdot \left(\frac{\gamma_1}{q_1}, \dots, \frac{\gamma_K}{q_K}\right)'.$$

From the definition of the operator norm, we deduce

$$\sum_{i=1}^{K} (1-\gamma_i)^2 = \|w-\gamma\|^2 \leq R^2 \cdot \left(\sum_{j=1}^{K} \frac{\gamma_j^2}{q_j^2}\right). \tag{59}$$

Let γ be evaluated at $(m_1(t), \dots, m_K(t))$, and suppose signal j is the myopic choice in period $t+1$, so that $|\partial_j f| = \frac{\gamma_j^2}{(m_j(t))^2}$ is maximal.[111] Then the above inequality implies $|1-\gamma_j| \leq \frac{R\sqrt{K} \cdot \gamma_j}{m_j(t)}$. Assuming $m_j(t) > 20RK$, this gives $|1-\gamma_j| \leq \frac{\gamma_j}{20\sqrt{2}}$, so that $\gamma_j < 1.037$. Plugging back into (59) and using $m_j(t) > 20RK$, we have for each i,

$$(1-\gamma_i)^2 + (1-\gamma_j)^2 < \frac{1.037^2 R^2 K}{(m_j(t))^2} < 0.00135$$

which implies $0.963 < \gamma_i, \gamma_j < 1.037$ and $\gamma_j - \gamma_i < 0.052$.[112] Thus $\frac{\gamma_j}{\gamma_i} = 1 + \frac{\gamma_j - \gamma_i}{\gamma_i} < 1 + \frac{0.052}{0.963} < 1.054$. But by assumption $|\frac{\gamma_j}{m_j(t)}| \geq |\frac{\gamma_i}{m_i(t)}|$, so that $\frac{m_j(t)}{m_i(t)} < 1.054$.

Now since $T \geq 21.5(R+1)K^2$, the DM must have observed *some* signal j in period $t_0 + 1 \leq T$ with the property that $m_j(t_0+1) \geq \frac{T}{K}$. Thus $m_j(t_0) \geq 21.5RK$ and our preceding analysis yields $m_i(t_0) > \frac{21.5RK}{1.054} > 20RK$ for *every* signal i. It follows that when $t \geq t_0 + 1$, $m_i(t) > \frac{m_j(t)-1}{1.054}$ holds for every pair of signals i, j. Summing across i or j yields the desired bounds.[113] $\square$

[111] This claim is somewhat incorrect, since the myopic signal maximizes (the absolute value of) the *discrete partial derivative*, rather than the continuous derivative. However, there is minimal distinction between the two: the discrete partial derivative always satisfies

$$\frac{\gamma_j^2}{q_j(q_j+1)} \leq |f(q_j+1, q_{-j}) - f(q_j, q_{-j})| \leq \frac{\gamma_j^2}{q_j^2},$$

where γ is evaluated at $(q_1, \dots, q_K)$. The RHS follows from the convexity of f, while the LHS uses the property that γ_j^2 increases in q_j, because $\frac{\partial \gamma_j(q)}{\partial q_j} = \frac{\gamma_j}{q_j^2} \cdot [(V+E)^{-1}]_{jj}$.

Thus, if signal j is the myopic choice in period $t+1$, we always have $\frac{\gamma_j^2}{(m_j(t))^2} \geq \frac{\gamma_i^2}{m_i(t) \cdot (m_i(t)+1)} \geq \frac{\gamma_i^2}{2(m_i(t))^2}$. Except for having to adjust the constants, the rest of this proof is minimally affected—in fact, these constants need not change much either, because we have $\frac{\gamma_j^2}{(m_j(t))^2} \geq \frac{\gamma_i^2}{m_i(t) \cdot (m_i(t)+1)} \geq \frac{\gamma_i^2}{1.1(m_i(t))^2}$ once we have shown $m_i(t) \geq 10$. The constant 1.1 appearing here is easily absorbed into the subsequent estimates.

[112] The latter uses $(1-\gamma_i)^2 + (1-\gamma_j)^2 \geq \frac{(\gamma_j - \gamma_i)^2}{2}$.

[113] Fixing j and directly summing this inequality across i, we would obtain $m_j(t) < 1.054\frac{T}{K} + 1$. The extra "+1" can be removed by noting $m_j(t) > \frac{m_j(t)+K-1}{1.054}$ and summing this with the inequalities $m_i(t) > \frac{m_j(t)-1}{1.054}$ across i *different from* j.

A particular step of the above proof is recorded below.

Fact 5. *Suppose* $q_1, \ldots, q_K \geq 20RK$*. Then* $0.963 \leq \gamma_i(q_1, \ldots, q_K) \leq 1.037$ *and* $|\gamma_i - \gamma_j| \leq 0.052$.

As a corollary, we have a quantitative version of the order difference lemma:

Lemma 13. *Suppose* $q_1, \ldots, q_K \geq 20RK$*. Then* $\frac{1.80}{q_i^3} \leq \partial_{ii} f \leq \frac{2.16}{q_i^3}$*, while* $|\partial_{ij} f| \leq \frac{2.16R}{q_i^2 q_j^2}, \forall i \neq j$.

The proof of this is based on the previously listed facts and $0 \leq [(V+E)^{-1}]_{ii} \leq [V^{-1}]_{ii} \leq R$ (thus also $|[(V+E)^{-1}]_{ij}| \leq R$).[114]

K.2 Bound on Consecutive Observations

We now bound the number of periods between consecutive observations of signal i.

Lemma 14. *Suppose* $t \geq 21.5(R+1)K^2$*, and the myopic DM observes signal* j *in period* t*. Then* $m_j(t+3K) > m_j(t)$.

Proof. Let $T = t + 3K$. Write $m_i = m_i(t)$ and $M_i = m_i(T)$. Without loss assume $j = 1$. We need to show $M_1 > m_1$.

Suppose for contradiction that $M_1 = m_1$. Since $\sum_{i=1}^{K}(M_i - m_i) = T - t = 3K$, we can without loss assume $M_2 - m_2 \geq 4$. Since signal 1 was myopically optimal at time t, the *discrete partial derivatives* satisfy

$$\partial_1 f(m_1 - 1, m_2, \ldots, m_K) \leq \partial_2 f(m_1 - 1, m_2, \ldots, m_K).$$

We claim that this implies

$$\partial_1 f(M_1, M_2 - 1, \ldots, M_K) < \partial_2 f(M_1, M_2 - 1, \ldots, M_K),$$

so that the DM cannot myopically observe signal 2 at time T (observing signal 1 would be better).

To prove this claim, consider the difference between the preceding two displays. Since $M_1 = m_1$, the LHS of the latter display is smaller than the LHS of the former plus $\partial_{11} f$ and

[114] This estimate is rather crude: when V is diagonal, $[(V+E)^{-1}]_{ij} = 0$ but R can be arbitrarily large. This suggests potential improvements to our bound that better capture correlation in the prior.

at most $T-t$ cross partials. The RHS, however, is larger by $(M_2 - m_2 - 1)\partial_{22} f$ minus at most $T-t$ cross partials. It thus suffices to show

$$3\partial_{22} f \geq \partial_{11} f + 6K|\partial_{ij} f|. \tag{60}$$

These derivatives are evaluated at points $(q_1, \ldots, q_K)$ with $m_i \leq q_i \leq M_i$, so that $0.948\frac{t}{K} \leq q_i \leq 1.054\frac{T}{K} \leq 1.13\frac{t}{K}$.[115] Hence, by Lemma 13, we have

$$\frac{\partial_{11} f}{\partial_{22} f} \leq \frac{2.16 \cdot q_2^3}{1.8 \cdot q_1^3} \leq \frac{2.16 \cdot 1.13^3}{1.8 \cdot 0.948^3} < 2.1$$

and also

$$\frac{|\partial_{ij} f|}{\partial_{22} f} \leq \frac{2.16 \cdot q_2^3 \cdot R}{1.8 \cdot q_i^2 q_j^2} \leq \frac{2.16 \cdot 1.13^3}{1.8 \cdot 0.948^4} \cdot \frac{RK}{t} < \frac{2.15RK}{t} < \frac{1}{10K}.$$

Hence (60) holds, proving the myopic DM did not observe signal 2 at time T. Essentially the same argument can be used to derive a contradiction assuming that signal 2 was observed at some time between t and T. But then $M_2 = m_2$, again a contradiction. The Lemma is thus proved. □

K.3 Completing the Proof of Theorem 4

Fix $t \geq 24(R+1)K^2$. Suppose for contradiction that $|m_i(t) - n_i(t)| > 1$ for some i. Without loss assume $|m_1(t) - n_1(t)| > 1$ and this difference is largest among all signals. We further assume $m_1(t) \leq n_1(t) - 2$ and $m_2(t) \geq n_2(t) + 1$. A symmetric argument applies to $m_1(t) \geq n_1(t) + 2$.

Let $\tilde{t} \leq t$ be the last period when the myopic DM observed signal 2. By Lemma 14, $t - \tilde{t} \leq 3K$.[116] Write $n_i = n_i(t), m_i = m_i(t), \tilde{m}_i = m_i(\tilde{t})$. As in the proof of Theorem 1 Part (a), we will prove (18) and deduce a contradiction. We have (considering the difference between (19) and (20))

$$\begin{aligned}
&f(\tilde{m}_1, \tilde{m}_2, \ldots \tilde{m}_K) - f(\tilde{m}_1 + 1, \tilde{m}_2 - 1, \ldots \tilde{m}_K) - f(n_1 - 1, n_2 + 1, \ldots, n_K) + f(n_1, n_2, \ldots, n_K) \\
=&\partial_2(\tilde{m}_1, \tilde{m}_2 - 1, \ldots, \tilde{m}_K) - \partial_1(\tilde{m}_1, \tilde{m}_2 - 1, \ldots, \tilde{m}_K) - \partial_2(n_1 - 1, n_2, \ldots, n_K) + \partial_1(n_1 - 1, n_2, \ldots, n_K) \\
=&(\tilde{m}_2 - n_2 - 1)\partial_{22} f + (n_1 - \tilde{m}_1 - 1)\partial_{11} f - (n_1 - \tilde{m}_1 - 1 + \tilde{m}_2 - n_2 - 1)\partial_{12} f + \sum_{j>2}(n_j - \tilde{m}_j)(\partial_{1j} f - \partial_{2j} f) \\
\geq&(n_1 - m_1 - 1) \cdot \partial_{11} f - (|n_1 - \tilde{m}_1| + |n_2 - \tilde{m}_2|) \cdot |\partial_{12} f| - \sum_{j>2} |n_j - \tilde{m}_j| \cdot (|\partial_{1j} f| + |\partial_{2j} f|).
\end{aligned} \tag{61}$$

[115] The last step uses $t \geq 42K$ and so $\frac{T}{t} \leq \frac{15}{14}$.

[116] Consider $t_0 = 21.5(R+1)K^2$, then by Proposition 6, $m_2(t_0) \leq 1.054t_0 < 22.7(R+1)K^2 < 0.948t \leq m_2(t)$. Thus the myopic DM observed signal 2 between time t_0 and time t, and $\tilde{t} \geq t_0$. Applying Lemma 14 to $\tilde{t}$ yields $t - \tilde{t} \leq 3K$.

All partial derivatives in the above are discrete partial derivatives, and the last inequality holds because $\tilde{m}_1 \le m_1$ and $\tilde{m}_2 = m_2 \ge n_2 + 1$.

The total number of cross partials that appear above can be bounded by

$$\begin{aligned} 2\sum_j |n_j - \tilde{m}_j| &\le 2\sum_j |n_j - m_j| + 2\sum_j |m_j - \tilde{m}_j| \\ &\le 2K(n_1 - m_1) + 2(t - \tilde{t}) \le 10K(n_1 - m_1 - 1), \end{aligned}$$

where the first step uses the triangle inequality, the second step uses the assumption $|n_j - m_j| \le n_1 - m_1$, and the last step uses $t - \tilde{t} \le 3K$, $n_1 - m_1 \ge 2$.

However, we have shown previously that $\frac{|\partial_{ij} f|}{\partial_{11} f} < \frac{1}{10K}$. Hence the RHS of (61) is strictly positive, completing the proof of the theorem.

Appendix L Intertemporal Decisions

L.1 Example: Investment in a Risky Asset with Unknown Return

We first provide a concrete example to illustrate the class of intertemporal decision problems discussed in Section 9. Suppose a worker repeatedly decides how much to consume, how much to save in a liquid asset (bond), and how much to invest in an illiquid asset (pension fund). The return to the liquid asset is known: 1 dollar saved today is worth e^r dollars tomorrow, with $r > 0$. The return to the illiquid asset is unknown, and it is the payoff-relevant state in the worker's decision problem; for now, we assume that every dollar invested today in the pension fund deterministically yields e^{θ_1} dollar(s) tomorrow. The worker works for T periods, and in each of these periods he learns about θ_1 and then allocates his wealth across consumption, saving and investment. In period $T + 1$, the worker retires and receives all the returns from his investments (into the illiquid asset). His objective is to maximize a discounted sum of consumption utilities as well as the utility from his wealth upon retirement.

Our equivalence results applied to this example imply that in separable or symmetric informational environments, the worker's optimal information acquisition strategy is myopic, minimizing the posterior variance about θ_1 at every period. This property would then allow us to analyze the worker's optimal consumption/saving/investment behavior as if information were exogenously given. While we will not comment on the optimal sequence of actions, our results provide an important simplification of the analysis by separating the concern of optimal information acquisitions from the question of optimal decisions.

An important (hidden) assumption of the above example is that the return to investment is deterministic and only observed at the end. However, by considering the realized return as a particular signal, our model can also cover situations in which the DM (perhaps periodically) learns from past, stochastic returns. Our results would extend as long as the distribution of the realized (log) return does not depend on the amount of investment—specifically, the worker could counterfactually figure out the return to his investment, even when he did not invest.

L.2 Proof of Theorem 5

To simplify notation, we (equivalently) let the DM take action a_t at the *beginning* of each period t, prior to observing a signal in that period. The DM's strategy in this problem consists of an information acquisition strategy S as well as a decision strategy A. Specifically, at any history h^T (which is a sequence of divisions and expected values about θ_1), $A(h^T)$ specifies the action to be taken in period $T+1$, while $S(h^T)$ is the signal choice.[117] Fix any pair (S, A), we will demonstrate another pair $(\tilde{S}, \tilde{A})$ that achieves the same expected payoff, where $\tilde{S}$ acquires information myopically. This will prove the theorem.

Our argument closely follows the proof of Lemma 7 in Appendix F. Let $\tilde{G}$ maps each sequence of divisions $(d(0), \ldots, d(T))$ under S to the deterministic sequence of myopic divisions $(m(0), \ldots, m(T))$. Then, as in the proof of Lemma 7 (especially the last paragraph there), we can associate each history of beliefs $h^T = (\mu_1^0, f(\mathbf{0}); \mu_1^1, f(d(1)); \ldots; \mu_1^T, f(d(T)))$ under S to a distribution of belief histories $\tilde{h}^T = (\mu_1^0, f(\mathbf{0}); \nu_1^1, f(m(1)); \ldots, \nu_1^T, f(m(T)))$ that occurs with the same probability under $\tilde{S}$ and is more informative about θ_1 at every period. Then, at a given history $\tilde{h}^T$ under $\tilde{S}$, we let the DM randomly draw a history h^T to "imitate" and follow the action choice given by A. That is, we set $\tilde{A}(\tilde{h}^T) = A(h^T)$ with the appropriate probability that $\tilde{h}^T$ imitates h^T.[118]

With this construction of $\tilde{A}$, we see that a DM following the decision strategy A obtains the same payoff along any infinite belief history h as another DM who uses the decision strategy $\tilde{A}$ and faces the distribution of belief histories $\tilde{h}$. Integrating over h, we have shown that $(\tilde{S}, \tilde{A})$ achieves the same payoff as (S, A). Hence the Theorem follows.

[117] Since pure strategies suffice for the DM's problem, it is not necessary to condition on past actions.

[118] The appropriate probability kernel has been discussed in the proof of Lemma 7.

Appendix M The Continuous-Time Model

In this appendix, we consider a continuous-time variation of our main model. We assume that the DM has one unit of attention in total at every point in time. He chooses attention levels $\beta_1(t), \ldots, \beta_K(t)$ (subject to $\beta_i(t) \geq 0$ and $\sum_i \beta_i(t) \leq 1$), which influence the diffusion processes $X_1, \ldots, X_K$ that he observes:[119]

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

where each B_i is an independent standard Brownian motion, and $\tilde{\theta}_i = \langle c_i, \theta \rangle$ is a "transformed state." The payoff-relevant state is

$$\theta_1 = \langle w, \tilde{\theta} \rangle$$

for some fixed $K \times 1$ vector w. We assume exact identifiability, so each coordinate $w_i > 0$.

At any time t, the DM's past attention levels integrate to a division $q(t)$ over the signals. A Markovian information acquisition strategy S maps $\left(q(t), \left\{X_i^{\leq t}\right\}_{i=1}^K\right)$ to $\Delta([K])$, representing how the DM divides attention at each instant as a function of his posterior beliefs.

The decision problem is the same as in discrete time: at an exogenously determined random final time t (drawn with density $\pi(t)$), the DM takes an action a_t and receives payoff $u_t(a_t, \theta_1)$. The forward-looking DM thus maximizes $\mathbb{E}\left[\int_0^\infty \pi(t) \cdot u_t(a_t, \theta_1)\ dt\right]$. We maintain the assumptions of *a single payoff-relevant state* and *payoff sensitivity*.

As in the discrete-time model (see Appendix K), the DM's posterior variance about θ_1 is a function of his prior covariance matrix V over $\tilde{\theta}$, the division q and the weights w:

$$f(q_1, \ldots, q_K) = w'(V - V(V+E)^{-1}V)w$$

with $E = \operatorname{diag}\left(\frac{1}{q_1}, \ldots, \frac{1}{q_K}\right)$. This posterior variance together with the posterior expected value of θ_1 determines the optimal action a_t that maximizes the flow payoff $\mathbb{E}[u_t(a_t, \theta_1)]$. We will thus focus on the dynamically optimal information acquisition strategy, and take for granted the corresponding optimal actions.

The notion of constrained t-optimality following a given history h is defined in the usual way. As before, a strategy S is *constrained t-optimal* if and only if the induced division $q^S(t)$ is almost surely a constrained t-optimal division:

$$q^S(t) \in \underset{q_i \geq H_i, \sum_i q_i = t}{\operatorname{argmin}} f(q_1, \ldots, q_K).$$

[119] 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.

The only difference from discrete time is that q_i can now take non-integer values. We note that by using an argument analogous to Footnote 78, we can show there is a unique constrained t-optimal division $n_i^h(t)$ at every time t.

M.1 Sufficient Conditions for Immediate Equivalence

We first examine the question of when myopic information acquisition is dynamically and t-optimal from the beginning of time. Let us define a myopic strategy in continuous time to be any limit of dynamically optimal strategies as the DM becomes infinitely impatient. Then, we seek sufficient conditions under which dynamic optimality coincides with t-optimality. As is now routine, this equivalence holds if and only if *the constrained t-optimal divisions are monotonic over time, following any given history.*

Proposition 1 and 2 directly extend and show that immediate equivalence obtains whenever the informational environment is separable or symmetric. More surprisingly, Example 4 presented in the main text now satisfies immediate equivalence, even though this equivalence fails in the discrete-time model (see Appendix J for details).[120] We mention that the example can be generalized to allow more states $\theta_1, b_1, \ldots, b_{K-2}, b_{K-1}$ and a "hierarchy" of noisy signals about $\theta_1 + b_1, b_1 + b_2, \ldots, b_{K-2} + b_{K-1}, b_{K-1}$. Immediate equivalence still obtains.

We now show that in continuous time, immediate equivalence holds whenever the DM's prior is *almost independent*. This formally confirms the intuition we provided for Theorem 1. Recall that V is the DM's prior covariance matrix over the transformed states $\tilde{\theta}_1, \ldots, \tilde{\theta}_K$.

[120] To see immediate equivalence in continuous time, we recall that Example 4 has three states θ, b_1, b_2 and three noisy signals X, B_1, B_2 about $\theta + b_1, b_1 + b_2, b_2$ respectively. The t-optimal problem is to allocate $t = t_X + t_1 + t_2$ units of time across these three signals to minimize the posterior variance about θ. Because the DM's prior beliefs over the states are independent, his allocation problem can be broken down into two sub-problems: first, for fixed sum $t_1 + t_2$, what is the optimal way to allocate t_1, t_2 to minimize the posterior variance about state b_1? Second, given the minimal variance about b_1 as a function of $t_1 + t_2$, what is the optimal pair $(t_X, t_1 + t_2)$? Note that the first sub-problem is the same as our benchmark model. Thus, the optimal pair (t_1, t_2) is increasing in the sum $t_1 + t_2$. On the other hand, because the minimal variance about b_1 is a *convex* function of $t_1 + t_2$, the second sub-problem also exhibits separability. Hence, the optimal pair $(t_X, t_1 + t_2)$ is increasing in t, which implies that the entire triple (t_X, t_1, t_2) is monotonic.

The reason this argument does not apply in discrete time is because the minimal variance about b_1 is not a convex function of $t_1 + t_2$, when t_1, t_2 are restricted to integer values. More specifically, for given $t_1 + t_2$, the DM's optimal allocation for learning b_1 involves $t_1 = t_2 + 1$. But this is not always feasible in discrete time when the desired sum $t_1 + t_2$ is an even integer. It is in those situations that a discrete-time DM does not fully take advantage of the complementarity between signals B_1 and B_2.

Theorem 7. *Suppose time is continuous. Given K, myopic, dynamic and t-optimality coincide at every time whenever the prior covariance matrix satisfies*

$$|V_{ij}| \leq \frac{1}{2K-3} \cdot V_{ii}, \forall i \neq j.$$

Importantly, this sufficient condition does not depend on the weight vector w, so that the myopic solution is optimal regardless of what the DM's payoff-relevant state is.[121] We comment that the constant $\frac{1}{2K-3}$ is best possible in the following sense: for any $\rho > \frac{1}{2K-3}$, there exists some weight vector w and some prior covariance matrix V satisfying $|V_{ij}| \leq \rho \cdot V_{ii}$, such that immediate equivalence does *not* obtain.[122]

For the special case of two states and two signals, we are able to derive a sufficient and necessary condition for immediate equivalence. Replacing each $\tilde{\theta}_i$ by its negative if necessary, we will assume $w_1, w_2 > 0$.

Theorem 8. *Suppose time is continuous and $K = 2$. Then myopic, dynamic and t-optimality coincide at every time if and only if*

$$w_1(V_{11} + V_{12}) + w_2(V_{21} + V_{22}) \geq 0.$$ [123]

The proofs of these results are deferred to later in this appendix.

M.2 Eventual Equivalence

Next, we investigate the informational environments in which immediate equivalence does not occur. Recall that in discrete time, Theorems 1 and 2 show that the number of signals acquired of each type under different optimality criteria differ by at most one, and this "eventual gap of one" vanishes in generic environments. We now show that in continuous time, the "eventual gap of one" can be dropped in all (not just generic) environments. Intuitively, the continuous-time model can be seen as a limit of our main discrete-time model as the period length shrinks to zero. Since the "gap of one" corresponds to a difference of one period, it vanishes in the limit.

[121] This condition is not necessary; for example, separability does not require that the prior covariance matrix satisfies the inequality above.

[122] For $K = 2$, this follows from the "only if" characterization in Theorem 8 below. For general K, we construct an informational environment with K signals that turns out to be equivalent to a simpler informational environment with 2 signals. Details are provided following the proof of Theorem 7.

[123] This condition can be rewritten as $Cov\,(w_1\tilde{\theta}_1 + w_2\tilde{\theta}_2, \tilde{\theta}_1 + \tilde{\theta}_2) \geq 0$.

Theorem 9. *Suppose time is continuous. There exists a large time T^* such that for any decision problem, the division $d(t)$ under the dynamically optimal strategy is deterministic when $t \geq T^*$, and $d(t) = n(t)$ is the unique t-optimal division.*

We make a few comments. First, this result does not rely on the assumption that some signal always decreases the posterior variance about θ_1 *strictly*, i.e. Assumption 4. Second, we only state above the equivalence between dynamic and t-optimality, but since a myopic strategy is the limit of dynamically optimal strategies, we also deduce the eventual optimality of myopic information acquisition. Lastly, our proof can be quantified to give an effective upper bound on T^*, analogous to what we did in Appendix K; we omit the details to save space.

M.3 Proof of Theorem 7

All partial derivatives in this appendix are the usual continuous derivative.

It suffices to show that whenever $|V_{ij}| \leq \frac{1}{2K-3} V_{ii}$ holds for every pair of signals $i \neq j$, the constrained t-optimal divisions following any given history are monotonic over time. This implies (constrained) t-optimality can and thus will be achieved at every time (following any given history). We will in fact prove this monotonicity under a weaker condition. Specifically, we will only assume that

$$V_{ii} \geq (K-2) \max_{j \neq i} |V_{ij}| + \sum_{j \neq i} |V_{ij}|. \tag{62}$$

Fix any history h with signal counts H_i. The unique constrained t-optimal division $n_i^h(t)$ minimizes $f(n_1, \dots, n_K)$ subject to $n_i \geq H_i, \forall i$ and $\sum_i n_i = t$. Thus, by the maximum theorem, $n^h(t)$ is continuous in t.

At any given time $t > \sum_i H_i$, the Kuhn-Tucker condition for optimality implies that whenever $n_i^h(t) > H_i$, the partial derivative $\partial_i f(n^h(t))$ must be a constant independent of the signal i. Fix a time t^* and let us classify the K signals into two groups: if $n_j^h(t) = H_j$ for all t in a small neighborhood around t^*, we put signal j into the first group. Every other signal i is put into the second group. That means, if signal i is in the second group, then we can find t arbitrarily close to t^* such that $n_i^h(t) > H_i$.

Since $t^* > \sum_i H_i$, we can suppose without loss that $n_1^h(t^*) > H_1$. Thus by continuity, $n_1^h(t) > H_1$ for t sufficiently close to t^*. Then, if signal i is in the second group, we can find times $t \to t^*$ such that $n_i^h(t) > H_i$ and $n_1^h(t) > H_1$. It follows that $\partial_i f(n^h(t)) = \partial_1 f(n^h(t))$. Taking the limit as $t \to t^*$ and using the continuity of $n^h(t)$, we deduce $\partial_i f(n^h(t^*)) =$

$\partial_1 f(n^h(t^*))$.[124]

Let us suppose without loss of generality that the second group consists of signals $1, \ldots, k$. Then, the preceding paragraph shows for t close to t^*:

$$\partial_i f(n_1^h(t), \ldots, n_K^h(t)) \quad \text{is constant for } 1 \le i \le k.$$

Differentiating with respect to t, we obtain[125]

$$\sum_{j=1}^{K} \frac{\partial n_j^h(t)}{\partial t} \cdot \partial_{ij} f(n_1^h(t), \ldots, n_K^h(t)) \quad \text{is constant for } 1 \le i \le k.$$

By definition of the groups, $\frac{\partial n_j^h(t)}{\partial t} = 0$ for $j > k$, we thus have

$$\sum_{j=1}^{k} \frac{\partial n_j^h(t)}{\partial t} \cdot \partial_{ij} f(n_1^h(t), \ldots, n_K^h(t)) = c, \ \forall 1 \le i \le k.$$

This is a system of k linear equations in the k linear unknowns $\frac{\partial n_j^h(t)}{\partial t}$. Thus, we deduce by matrix algebra

$$\left(\frac{\partial n_1^h(t)}{\partial t}, \ldots, \frac{\partial n_k^h(t)}{\partial t}\right)' = c \cdot Hess^{-1} \cdot (1, \ldots, 1)' \tag{63}$$

with $Hess$ being the top-left $k \times k$ sub-matrix of the Hessian matrix of f (evaluated at $n^h(t)$).

We note that

$$1 = \sum_{i=1}^{k} \frac{\partial n_i^h(t)}{\partial t} = c \cdot (1, \ldots, 1) \cdot Hess^{-1} \cdot (1, \ldots, 1)'$$

where the first equality follows from $\sum_i n_i^h(t) = t$ and the second follows from (63). Since f is convex, $Hess$ is positive-definite and so is $Hess^{-1}$. Thus $(1, \ldots, 1) \cdot Hess^{-1} \cdot (1, \ldots, 1)' > 0$, which implies $c > 0$ via the preceding display.

Recall that $\frac{\partial n_j^h(t)}{\partial t} = 0$ for $j > k$ and t in a neighborhood of t^*. Thus, proving the (local) monotonicity of $n^h(t)$ reduces to proving $\frac{\partial n_i^h(t)}{\partial t} \ge 0$ for $1 \le i \le k$. By (63), this is further equivalent to each coordinate of $Hess^{-1} \cdot (1, \ldots, 1)'$ being non-negative. This is given in the following lemma:

[124] This is obviously true if $n_i^h(t^*) > H_i$, but the second group of signals may also contain some i for which $n_i^h(t^*) = H_i$ (but $n_i^h(t) > H_i$ at times t close to t^*).

[125] It can be shown from the Implicit Function Theorem that the derivatives $\frac{\partial n_j^h(t)}{\partial t}$ almost always exists and is given by (63) below. The only exception occurs when t is the largest time at which $n_j^h(t) = H_j$—even in that case, the right-sided derivative exists and is given by (63).

Lemma 15. *Suppose the prior covariance matrix V satisfies (62). Let $1 \le k \le K$ and suppose a division $(q_1, \dots, q_K)$ is such that $\partial_i f(q_1, \dots, q_K) = \partial_j f(q_1, \dots, q_K)$ holds for $1 \le i, j \le k$. Denote by $\mathbf{H}$ the top $k \times k$ sub-matrix of the Hessian matrix of f (evaluated at q). Then $\mathbf{H}^{-1} \cdot (1, \dots, 1)' \ge \mathbf{0}$.*

Proof of Lemma 15. We will prove the Lemma assuming $k = K$. The proof for smaller k is similar (and does not use the full strength of Assumption (62)).

By Farkas' lemma, the vector $\mathbf{H}^{-1} \cdot (1, \dots, 1)'$ has non-negative coordinates if and only if for any real numbers $x_1, \dots, x_K$, $(x_1, \dots, x_K) \cdot \mathbf{H} \ge \mathbf{0}$ implies $x_1 + \cdots + x_K \ge 0$. Suppose for contradiction that $x_1 + \cdots + x_K < 0$. Without loss assume x_1 is the smallest (most negative) among $x_1, \dots, x_K$. We will show under these assumptions that

$$|x_1| \cdot \mathbf{H}_{11} > \sum_{j \ne 1} |x_j| \cdot |\mathbf{H}_{1j}|, \tag{$\star$}$$

which will contradict $(x_1, \dots, x_K) \cdot \mathbf{H} \ge \mathbf{0}$.

We recall the following computations from the beginning of Appendix K:

$$\partial_1 f = -\frac{\gamma_1^2}{q_1^2}; \quad \partial_j f = -\frac{\gamma_j^2}{q_j^2}$$

$$\partial_{11} f = \frac{2\gamma_1^2}{q_1^3} \cdot \left(1 - \frac{[(V+E)^{-1}]_{11}}{q_1}\right); \quad \partial_{1j} f = \frac{-2\gamma_1\gamma_j}{q_1^2 q_j^2} \cdot [(V+E)^{-1}]_{1j}.$$

The assumption that $\partial_1 f = \partial_j f$ tells us $|\frac{\gamma_1}{q_1}| = |\frac{\gamma_j}{q_j}|$. It follows that

$$\frac{|\partial_{1j} f|}{\partial_{11} f} = \frac{|[(V+E)^{-1}]_{1j}/q_j|}{1 - [(V+E)^{-1}]_{11}/q_1}.$$

Using this, the desired inequality ($\star$) becomes (recall $\mathbf{H}$ is the Hessian matrix)

$$1 - \frac{[(V+E)^{-1}]_{11}}{q_1} \ge \sum_{j \ne 1} \frac{|x_j|}{|x_1|} \cdot \frac{|[(V+E)^{-1}]_{1j}|}{q_j} \tag{$\star\star$}$$

We now evaluate the LHS above. To do this, we introduce a piece of notation. For subsets $I, J \subset \{1, \dots, K\}$, we write M_{-IJ} the sub-matrix of M after removing rows $i \in I$ and columns $j \in J$; $M_{-\{i\}\{j\}}$ is simplified as M_{-ij}. Then we have (by directly expanding the determinant and collecting terms)

$$\det(V+E) = \sum_{S \subset \{1,\dots,K\}} \prod_{i \in S} \frac{1}{q_i} \cdot \det(V_{-SS}).$$

$$\det((V+E)_{-11}) = \sum_{S\subset\{2,\dots,K\}} \prod_{i\in S} \frac{1}{q_i} \cdot \det(V_{-S^1S^1}) \text{ with } S^1 \text{ representing } S\cup\{1\}.$$

We use the convention that when S is the empty set, $\prod_{i\in S}\cdots = 1$.

The above two equalities and Cramer's rule for matrix inverse imply that

$$1 - \frac{[(V+E)^{-1}]_{11}}{q_1} = 1 - \frac{\det((V+E)_{-11})}{q_1 \det(V+E)} = \frac{\sum_{S\subset\{2,\dots,K\}} \prod_{i\in S} \frac{1}{q_i} \cdot \det(V_{-SS})}{\det(V+E)}. \tag{64}$$

In a similar way, we also have

$$\begin{aligned} \frac{|[(V+E)^{-1}]_{1j}|}{q_j} = \frac{|\det((V+E)_{-1j})|}{q_j \det(V+E)} &= \frac{|\frac{1}{q_j} \cdot \sum_{T\subset\{2,\dots,K\}-\{j\}} \prod_{i\in T} \frac{1}{q_i} \cdot \pm \det(V_{-T^1T^j})|}{\det(V+E)} \\ &\leq \frac{\sum_{T\subset\{2,\dots,K\}-\{j\}} \prod_{i\in T^j} \frac{1}{q_i} \cdot |\det(V_{-T^1T^j})|}{\det(V+E)} \end{aligned} \tag{65}$$

with $T^1 = T\cup\{1\}, T^j = T\cup\{j\}$. The plus or minus sign in the first line arises in expanding $\det((V+E)_{-1j})$.

Using (64) and (65), the desired inequality ($\star\star$) is further reduced to

$$\sum_{S\subset\{2,\dots,K\}} \prod_{i\in S} \frac{1}{q_i} \cdot \det(V_{-SS}) > \sum_{j=2}^{K} \frac{|x_j|}{|x_1|} \cdot \sum_{T\subset\{2,\dots,K\}-\{j\}} \prod_{i\in T^j} \frac{1}{q_i} \cdot |\det(V_{-T^1T^j})|. \tag{$\star\star\star$}$$

We organize the RHS above according to T^j and rewrite the desired inequality as

$$\sum_{S\subset\{2,\dots,K\}} \prod_{i\in S} \frac{1}{q_i} \cdot \det(V_{-SS}) > \sum_{S\subset\{2,\dots,K\}} \prod_{i\in S} \frac{1}{q_i} \cdot \sum_{j\in S} \frac{|x_j|}{|x_1|} \cdot |\det(V_{-S^1_jS})|$$

where S^1_j represents the set $S\cup\{1\}-\{j\}$. Thus, we only need to show for each non-empty set $S\subset\{2,\dots,K\}$ (the following inequality is strict when $S=\emptyset$):

$$\det(V_{-SS}) \geq \sum_{j\in S} \frac{|x_j|}{|x_1|} \cdot |\det(V_{-S^1_jS})|. \tag{$\star\star\star\star$}$$

Fix S. We claim that the RHS of ($\star\star\star\star$) is at most $(K-2)\cdot\max_{j\in S} |\det(V_{-S^1_jS})| + \sum_{j\in S}|\det(V_{-S^1_jS})|$. To see this, recall that we assume x_1 is the most negative. Thus whenever $x_j<0$, the ratio $\frac{|x_j|}{|x_1|}$ is bounded above by 1. Moreover, since $x_1+\cdots+x_K<0$, those positive x_j sum to at most $(K-1)\cdot|x_1|$. Hence our claim. Consequently, it suffices to show

$$\det(V_{-SS}) \geq (K-2)\cdot\max_{j\in S} |\det(V_{-S^1_jS})| + \sum_{j\in S}|\det(V_{-S^1_jS})|.$$

This is equivalent to

$$\det(V_{-SS}) \geq \pm(K-1)\det(V_{-S^1_{j^*}S}) + \sum_{j\in S, j\neq j^*} \pm\det(V_{-S^1_j S}) \qquad (\star\star\star\star\star)$$

for any $j^* \in S$ and any choice of plus-minuses.

Relabeling if necessary, we assume $S = \{l+1,\ldots,K\}$. For $1 \leq i \leq K$, let u_i be the l-dimensional vector corresponding to the first l coordinates of the i-th row of V. Then V_{-SS} is a $l \times l$ matrix whose row vectors are $u_1, u_2, \ldots, u_l$. By contrast, $V_{-S^1_j S}$ has row vectors $u_2, \ldots, u_l, u_j$—since $S^1_j = S \cup \{1\} - \{j\}$, the first row of V is deleted while the j-th row is not. Suppose we have the following linear equation in vectors (with $y_1, \ldots, y_l$ appropriate scalars):

$$\pm(K-1)u_{j^*} + \sum_{j>l, j\neq j^*} \pm u_j = \sum_{i=1}^{l} y_i \cdot u_i. \tag{66}$$

Then, by the multi-linearity of the determinant,

$$\pm(K-1)\det(V_{-S^1_{j^*}S}) + \sum_{j>l, j\neq j^*} \pm\det(V_{-S^1_j S}) = y_1\det(V_{-SS}).$$

Hence, to prove ($\star\star\star\star\star$), it remains to prove $y_1 \leq 1$ assuming (66). Choose $1 \leq i^* \leq l$ to maximize $|y_i|$, and suppose for contradiction that this maximum is greater than 1. The equality (66) applied to the i^*-th coordinate gives

$$\pm(K-1)V_{j^*i^*} + \sum_{j>l, j\neq j^*} \pm V_{ji^*} = \sum_{i=1}^{l} y_i V_{ii^*}.$$

The triangle inequality together with $|y_{i^*}| \geq |y_i|$, $|y_{i^*}| > 1$ thus implies

$$|y_{i^*}| \cdot V_{i^*i^*} \leq \sum_{j\neq i^*, j^*} |y_{i^*}| \cdot |V_{ji^*}| + (K-1)\cdot|V_{j^*i^*}|. \tag{67}$$

But this contradicts our assumption (62).[126] This contradiction completes the proof of the Lemma. Theorem 7 also follows. $\square$

M.3.1 An Example Showing the Optimality of $\frac{1}{2K-3}$

We provide a specific example to show that the constant $\frac{1}{2K-3}$ is best possible for Theorem 7 to be true. Let the prior covariance matrix V have diagonal entries 1 and off-diagonal

[126] That assumption implies $|y_{i^*}| \cdot V_{i^*i^*} \geq \sum_{j\neq i^*, j^*} |y_{i^*}| \cdot |V_{ji^*}| + (K-1)|y_{i^*}| \cdot |V_{j^*i^*}|$, with equality only if $|V_{j^*i^*}|$ is maximal. Together with (67), we deduce $|V_{j^*i^*}| = 0$ and all equalities hold equal. Thus every $V_{ji^*} = 0$, which contradicts (67) because $V_{i^*i^*} > 0$.

entries $-\rho$, with $\rho > \frac{1}{2K-3}$. This means the DM's prior beliefs over the states are symmetric: each state has the same prior variance, and there is constant prior correlation between any pair of states. The DM cares about $\langle w, \tilde{\theta} \rangle$, where we choose $w_1 = w_2 = \cdots = w_{K-1} = 1$ and w_K a small positive number.

For this problem, we will show that the (unconstrained) t-optimal divisions are *not* monotonic over time, so that immediate equivalence between dynamic and t-optimality does not hold. Note that signals $1, \ldots, K-1$ have symmetric prior as well as symmetric payoff weights. Thus, the posterior variance function $f(q_1, \ldots, q_{K-1}, q_K)$ is symmetric in its first $K-1$ arguments. This implies the t-optimal division $n(t)$ must satisfy $n_1(t) = \cdots = n_{K-1}(t)$; otherwise it would not be unique. The t-optimal problem simplifies to

$$(n_1, n_K) \in \operatorname*{argmin}_{q_1, q_K : (K-1)q_1 + q_K = t} f(q_1, \ldots, q_1, q_K).$$

That is, the DM optimally divides attention between signal K and the remaining signals (which are always given equal attention).

The posterior beliefs of such a DM are derived assuming he had observed the following K signals: $\tilde{\theta}_i + \mathcal{N}(0, \frac{1}{q_1})$ for $1 \le i \le K-1$ and $\tilde{\theta}_K + \mathcal{N}(0, \frac{1}{q_K})$. We claim that the (marginal) belief about $\langle w, \tilde{\theta} \rangle$ is unchanged if the DM had instead observed only two signals: $\frac{1}{K-1} \sum_{i=1}^{K-1} \tilde{\theta}_i + \mathcal{N}(0, \frac{1}{(K-1)q_1})$ and $\tilde{\theta}_K + \mathcal{N}(0, \frac{1}{q_K})$. We provide an intuitive argument for this claim (rather than going through the computations, which are also doable). First, by symmetry, the DM's belief about $\sum_{i=1}^{K-1} \tilde{\theta}_i$ is the same whether he observed the $K-1$ signals $\tilde{\theta}_i + \mathcal{N}(0, \frac{1}{q_1})$ for $1 \le i \le K-1$, or their average $\frac{1}{K-1} \sum_{i=1}^{K-1} \tilde{\theta}_i + \mathcal{N}(0, \frac{1}{(K-1)q_1})$. Next, we need to show that the $K-1$ signals do not provide more information than their average for the DM to learn about $\tilde{\theta}_K$. This is because the extra information provided by the $K-1$ signals takes the form of $\tilde{\theta}_j - \tilde{\theta}_k$ for $1 \le j < k \le K-1$, and any such difference between the states is independent from $\tilde{\theta}_K$ conditional on the sum $\sum_{i=1}^{K-1} \tilde{\theta}_i$.[127]

Under the equivalence discussed in the preceding paragraph, the t-optimal decision problem with K states and K signals is the same as if there were only two states $\theta_1^* = \frac{1}{K-1} \sum_{i=1}^{K-1} \tilde{\theta}_i, \theta_2^* = \tilde{\theta}_K$ and correspondingly two signals. In the latter, simplified problem, the DM chooses to devote $(K-1)q_1$ units of time to the first signal and q_K units to the second. The constraint $((K-1)q_1 + q_K = t)$ and the objective function (minimizing the posterior variance) are the same as in the original problem, as soon as we change the payoff weights to be $w_1^* = K-1$ and $w_2^* = w_K$.

[127] In the DM's prior, $\tilde{\theta}_j - \tilde{\theta}_k$ has zero covariance with $\sum_{i=1}^{K-1} \tilde{\theta}_i$ and with $\tilde{\theta}_K$. Thus $\tilde{\theta}_j - \tilde{\theta}_k$ is independent from $\sum_{i=1}^{K-1} \tilde{\theta}_i$ and $\tilde{\theta}_K$, *jointly*. Conditional independence between $\tilde{\theta}_j - \tilde{\theta}_k$ and $\tilde{\theta}_K$ follows.

The prior covariance matrix V^* in the simplified (two by two) problem is given by

$$V^*_{11} = Var\left(\frac{1}{K-1}\sum_{i=1}^{K-1}\tilde{\theta}_i\right) = \frac{1-(K-2)\rho}{K-1},$$

$V^*_{12} = V^*_{21} = -\rho$ and $V^*_{22} = 1$. Now $\rho > \frac{1}{2K-3}$ implies $V^*_{11} + V^*_{12} < 0$. Thus for $w^*_2 = w_K$ sufficiently small, it holds that

$$w^*_1(V^*_{11} + V^*_{12}) + w^*_2(V^*_{21} + V^*_{22}) < 0.$$

By the proof of Theorem 8 below, we conclude that the t-optimal divisions in the two-state problem are not monotonic over time. Hence the t-optimal divisions in the original K-state problem are also not monotonic.[128]

M.4 Proof of Theorem 8

Recall that $\partial_i f(q_1, q_2) = -\frac{\gamma_i^2}{q_i^2}$, with $\gamma = (V+E)^{-1}Vw$ and $E = \operatorname{diag}(\frac{1}{q_1}, \frac{1}{q_2})$. We then directly compute that

$$\begin{aligned}
\partial_1 f(q_1, q_2) &= \frac{-(w_1 \cdot \det(V) \cdot q_2 + w_1 V_{11} + w_2 V_{21})^2}{q_1^2 q_2^2 \det^2(V+E)} := \frac{-(x_1 q_2 + y_1)^2}{q_1^2 q_2^2 \det^2(V+E)}; \\
\partial_2 f(q_1, q_2) &= \frac{-(w_2 \cdot \det(V) \cdot q_1 + w_1 V_{12} + w_2 V_{22})^2}{q_1^2 q_2^2 \det^2(V+E)} := \frac{-(x_2 q_1 + y_2)^2}{q_1^2 q_2^2 \det^2(V+E)}.
\end{aligned} \tag{68}$$

For notational ease, we define $x_1 = w_1 \cdot \det(V)$, $x_2 = w_2 \cdot \det(V)$, $y_1 = w_1 V_{11} + w_2 V_{21}$ and $y_2 = w_1 V_{12} + w_2 V_{22}$. Observe that $x_1, x_2 > 0$. We need to show immediate equivalence holds if and only if $y_1 + y_2 \geq 0$.

In one direction, suppose that $y_1 + y_2 \geq 0$. Let us consider the t-optimal divisions and show they are monotonic over time. To see this, without loss assume $y_1 \geq y_2$, which means $y_1 \geq 0$ because $y_1 + y_2 \geq 0$. Take any time $t > 0$. Then from (68), $\partial_1 f(0,t) < \partial_2 f(0,t)$. Thus the first-order-condition for t-optimality implies that at the t-optimal division $(q_1, t - q_1)$, either $\partial_1 f(q_1, q_2) = \partial_2 f(q_1, q_2)$ or $q_1 = t, q_2 = 0$. As $x_1 q_2 + y_1 + x_2 q_1 + y_2 > 0$, the partial derivatives are equal if and only if $x_1 q_2 + y_1 = x_2 q_1 + y_2$. Combined with $q_1 + q_2 = t$, this yields a candidate t-optimal division $q_1 = \frac{x_1 t + y_1 - y_2}{x_1 + x_2}, q_2 = \frac{x_2 t - y_1 + y_2}{x_1 + x_2}$. For $t < \frac{y_1 - y_2}{x_2}$, this

[128] This example shows if we replace $\frac{1}{2K-3}$ with any larger constant in the statement of Theorem 7, then there exist *some* payoff weights w and some prior such that the Theorem fails. One may wonder if equivalence can be restored by constraining the payoff weights—e.g., for $K = 2$, Theorem 8 establishes immediate equivalence for equal payoff weights $w_1 = w_2$ and an arbitrary prior. However, an example similar to the one given here shows equal payoff weights do not guarantee immediate equivalence when $K > 2$.

candidate solution is not feasible because $q_2 < 0$. Thus the t-optimal division at $t < \frac{y_1 - y_2}{x_2}$ is $(t, 0)$. For $t \geq \frac{y_1-y_2}{x_2}$, it can be verified that the above candidate solution indeed minimizes f (because $\partial_1 f - \partial_2 f$ turns from negative to positive only once). Thus the t-optimal division at $t \geq \frac{y_1-y_2}{x_2}$ is $(\frac{x_1 t + y_1 - y_2}{x_1+x_2}, \frac{x_2 t - y_1 + y_2}{x_1 + x_2})$, which is seen to be monotonic over time. We can use the same argument to show that the constrained t-optimal divisions are monotonic,[129] proving the equivalence between dynamic and t-optimality.

In the opposite direction, suppose that $y_1 + y_2 < 0$. We will show that the t-optimal divisions are *not* monotonic over time. First notice that one of y_1, y_2 must be positive, because $w_1 y_1 + w_2 y_2 = w_1^2 \cdot V_{11} + 2 w_1 w_2 \cdot V_{12} + w_2^2 \cdot V_{22} > 0$. Without loss, let us assume $y_1 > 0$ and $y_2 < -y_1$. From (68), we see that for t sufficiently small ($t < \frac{-y_1-y_2}{q_1+q_2}$), $\partial_2 f(q_1, q_2) < \partial_1 f(q_1, q_2)$ holds whenever $q_1 + q_2 = t$. Thus, the t-optimal division at such times t is $(0, t)$. But consider $\hat{t} = \frac{-y_2}{x_2}$. By (68) we have $\partial_2 f(\hat{t}, 0) = 0 > \partial_1 f(\hat{t}, 0)$. Hence $f(\hat{t}, 0) < f(\hat{t} - \epsilon, \epsilon)$ for any small ϵ. Together with the convexity of f, we deduce that the t-optimal division at $\hat{t}$ is precisely $(\hat{t}, 0)$. This violates monotonicity, as we desire to show. The proof of the theorem is complete.

M.5 Proof of Theorem 9

As in Appendix I, an important step of the proof is to show that following any history in which each signal is observed sufficiently often, the constrained t-optimal divisions are monotonic.

Lemma 16. *Suppose time is continuous. There exists H^* such that for any history h with signal counts $H_i \geq H^*, \forall i$, the constrained t-optimal divisions following h are increasing (in each coordinate) over time.*

Proof. We only need to show that the DM's posterior covariance matrix exhibits "approximate independence" after sufficiently many observations of each signal. Once this is proved, we can apply Theorem 7 to the posterior beliefs and deduce the Lemma.

The posterior covariance matrix is given by

$$\hat{V} = V - V(V+E)^{-1}V = V(V + E^{-1})E = E - E(V+E)^{-1}E,$$

[129] Suppose we want to minimize $f(q_1, q_2)$ subject to $q_1 + q_2 \geq t$, $q_1 \geq H_1, q_2 \geq H_2$. Writing $q_1 = H_1 + \hat{q}_1$ and $q_2 = H_2 + \hat{q}_2$, then the partial derivatives become $\partial_1 f = -(x_1 \hat{q}_2 + \hat{y}_1)^2 / \ldots$ and $\partial_2 f = -(x_2 \hat{q}_1 + \hat{y}_2)^2 / \ldots$, with $\hat{y}_1 = y_1 + x_1 H_2$ and $\hat{y}_2 = y_2 + x_2 H_1$ taking the role of y_1, y_2. Since $\hat{y}_1 + \hat{y}_2 \geq y_1 + y_2 \geq 0$, our method of proof remains valid.

with V being the prior covariance matrix and $E = \text{diag}(\frac{1}{q_1}, \ldots, \frac{1}{q_K})$. If S denotes the operator norm of V^{-1}, then $\hat{V}_{ii} = \frac{1}{q_i} - \frac{[(V+E)^{-1}]_{ii}}{q_i^2} \geq \frac{1}{q_i} - \frac{R}{q_i^2}$. On the other hand, $|\hat{V}_{ij}| = \frac{|[(V+E)^{-1}]_{ij}|}{q_i q_j} \leq \frac{R}{q_i q_j}$. These estimates imply $\hat{V}_{ii} \geq (2K-3) \cdot |\hat{V}_{ij}|$ whenever $q_i, q_j \geq (2K-2)R$. Choosing $H^* = (2K-2)R$ proves the Lemma. □

Proof of Theorem 9. Let H^* be given by the preceding lemma. Suppose we can demonstrate a time T_0 such that any realized division $d(T_0)$ under a dynamically optimal strategy S satisfies $d_i(T_0) \geq H^*$. Then, because constrained t-optimality is attainable at every future time, the optimal strategy S must attain it. That is, at any time $t \geq T_0$, the division $d(t)$ must be constrained t-optimal subject to $d_i(t) \geq d_i(T_0), \forall i$. For sufficiently large t, this constrained t-optimal division coincides with the unconstrained t-optimal division,[130] which proves the theorem.

Hence, the remaining difficulty is to prove $d_i(t) \to \infty$. In Appendix F, we showed the same result in discrete time by considering (i,j)-switch deviations. The proof strategy here is similar but somewhat trickier, because we do not assume some signal strictly decreases posterior variance.

To start, we recall the class of deviation strategies in Appendix F and adapt them to the continuous-time setting. For each *continuous path of divisions* $(d(t))_{0 \leq t \leq T}$, $\tilde{G}$ maps to another path $(\tilde{d}(t))_{0 \leq t \leq T}$ subject to consistency. The continuous-time analogue of the dynamic Blackwell-dominance lemma (Lemma 7) is that if $f(\tilde{d}(T)) \leq f(d(T))$ always holds, then the deviation strategy $\tilde{S}$ that simulates S yields a weakly higher expected payoff.

Let $\underline{H} > H^*$ be a large constant and $\overline{H}$ be an even larger constant, to be determined later. Fix a dynamically optimal strategy S. Consider any history under S such that some signal has been observed at least $\overline{H}$ units of time, while some other signal has been observed fewer than $\underline{H}$ units of time. Relabelling the signals if necessary, we assume that (t_0 denotes the calendar time of that history)

$$d_1(t_0), \ldots d_k(t_0) < \underline{H} \leq d_{k+1}(t_0), \ldots, d_{K-1}(t_0); \quad d_K(t_0) \geq \overline{H}.$$

Below we will construct a deviation strategy $\tilde{S}$ following this history.

We define the deviation $\tilde{S}$ and its induced divisions $\tilde{d}(T)$ jointly. At any time $T \geq t_0$, there are three possibilities:

1. Suppose that, given the path of divisions $(d(t))_{0 \leq t \leq T}$, the original strategy S devotes attention to signal K. Then we let $\tilde{S}$ distribute this amount of attention *evenly* among

[130] Our asymptotic characterization of $n(t)$ shows that the unconstrained t-optimal division satisfies the constraints at large t. Thus it is also the constrained t-optimal division.

those signals $j \in \{1, \ldots, k\}$ for which $\tilde{d}_j(T) < \underline{H}$. If no such signal exists, let $\tilde{S}$ also observe signal K.

In other words, whenever the time derivative of $d_K(T)$ is positive, we set the time derivative of $\tilde{d}_K(T)$ to be zero, and compensate it by increasing the time derivatives of $d_j(T)$ for those signals j that are insufficiently observed.

2. Suppose that S devotes attention to signal $k+1, \ldots, K-1$. Then we let $\tilde{S}$ observe the same signal.

3. Suppose that S devotes attention to signal $j \in \{1, \ldots, k\}$. If $\tilde{d}_j(T) < \underline{H}$ or $\tilde{d}_j(T) = d_j(T)$, then we let $\tilde{S}$ also observe signal j. Otherwise $\tilde{d}_j(T) = \underline{H} > d_j(T)$, and we let $\tilde{S}$ observe signal K instead.

To interpret, $\tilde{S}$ deviates from S as soon as some signal (signal K) has been observed too often compared to some other "deficient signals". Following that history, $\tilde{S}$ refrains from observing signal K and instead devotes attention to the deficient signals, until all deficient signals reach a certain signal count. Moreover, the third case above says that when a deficient signal j reaches the target level, $\tilde{S}$ switches back to observing signal K until $\tilde{S}$ agrees with S on the signal count of j. In this sense, the deviation $\tilde{S}$ considered here is a natural extension of the "switch deviations" introduced in the discrete-time model (see Appendix F).

Let us verify that either $\tilde{d}(T) = d(T)$, or $f(\tilde{d}(T)) < f(d(T))$. Suppose $\tilde{d}(T) \neq d(T)$, then there must exist some signal $j \in \{1, \ldots, k\}$ such that $d_j(T) < \tilde{d}_j(T)$. By the above construction, there could be four types of signals (we omit the dependence on T to ease notation):

$$\begin{aligned} \forall 1 \leq j \leq m, &\quad d_j < \tilde{d}_j < \underline{H} \text{ and } \tilde{d}_j - d_j = \alpha; \\ \forall m+1 \leq j \leq l, &\quad d_j < \tilde{d}_j = \underline{H} \text{ and } \tilde{d}_j - d_j = \alpha_j \leq \alpha; \\ \forall l+1 \leq j \leq K-1, &\quad d_j = \tilde{d}_j \geq \underline{H}; \\ &\quad d_K > \tilde{d}_K \geq \overline{H} \text{ and } d_K - \tilde{d}_K = m\alpha + \sum_{j=m+1}^{l} \alpha_j \end{aligned} \tag{69}$$

for some indices $0 \leq m \leq l \leq k$, $l \geq 1$ and positive real number α.[131]

If $m = 0$, then (the first type does not exist and) $\tilde{d}_j \geq \underline{H}$ for every signal j. Thus for some constant L and any sufficiently large $\underline{H}$, it holds that $|\partial_j f(\tilde{d})| \geq 1/(L(\tilde{d}_j)^2)$. Without

[131]Every $\tilde{d}_j - d_j$ is equal to the same α for $1 \leq j \leq m$ because attention devoted to signal K is always distributed evenly among the deficient signals.

loss assume α_1 is largest. Then

$$f(\tilde{d}) \leq f(d_1, \tilde{d}_2, \ldots, \tilde{d}_K) + (\tilde{d}_1 - d_1) \cdot \partial_1 f(\tilde{d}) \leq f(d_1, \tilde{d}_{-1}) - \frac{\alpha_1}{L\underline{H}^2} \leq f(d_1, \ldots, d_{K-1}, \tilde{d}_K) - \frac{\alpha_1}{L\underline{H}^2}$$

where the first inequality uses the convexity of f, and the last inequality uses its monotonicity. On the other hand,

$$f(d) \geq f(d_1, \ldots, d_{K-1}, \tilde{d}_K) + (d_K - \tilde{d}_K) \cdot \partial_K f(d_{-K}, \tilde{d}_K) \geq f(d_1, \ldots, d_{K-1}, \tilde{d}_K) - (K-1)\alpha_1 \cdot \frac{L}{(\tilde{d}_K)^2}.$$

Since $\tilde{d}_K \geq \overline{H}$, we do have $f(\tilde{d}) < f(d)$ whenever $\overline{H}$ is significantly larger than $\underline{H}$.

Now consider $m > 0$. Similar to the above, we have

$$f(d) \geq f(d_1, \ldots, d_m, \tilde{d}_{m+1}, \ldots, \tilde{d}_{K-1}, d_K) \geq f(d_1, \ldots, d_m, \tilde{d}_{m+1}, \ldots, \tilde{d}_K) - (K-1)\alpha \cdot \frac{L}{(\tilde{d}_K)^2}.$$

And for any $j \in \{1, \ldots, m\}$,

$$f(\tilde{d}) \leq f(d_j, \tilde{d}_{-j}) - \alpha \cdot |\partial_j f(\tilde{d})| \leq f(d_1, \ldots, d_m, \tilde{d}_{m+1}, \ldots, \tilde{d}_K) - \alpha \cdot |\partial_j f(\tilde{d})|.$$

Since there is at least one signal j with $|\partial_j f(\tilde{d})| \geq \frac{1}{L\underline{H}^2}$,[132] we can again deduce $f(\tilde{d}) < f(d)$ whenever $\overline{H}$ is much larger than $\underline{H}$.

Hence, we have shown that $\tilde{S}$ is a profitable deviation from S unless $\tilde{d}(t)$ is always the same division as $d(t)$. By construction, this implies that whenever a signal is observed more than $\overline{H}$ units of time, the dynamically optimal strategy S stops observing it until no other signal is deficient. At time $T_0 \geq K \cdot \overline{H}$, some signal must have been observed more than $\overline{H}$. Thus $d_i(T_0) \geq \underline{H} \geq H^*$ for every signal i, completing the proof of the theorem. $\square$

[132]To see this, let q denote the division $\tilde{d}$, with $q_i \geq \underline{H}, \forall m+1 \leq i \leq K$. Then use the previously derived inequality (59) to show that either some $\gamma_j (1 \leq j \leq m)$ is at least $1/2$, or the RHS is at least $1/4$. The result $\frac{\gamma_j^2}{q_j^2} \geq \frac{1}{L\underline{H}^2}$ holds either way.

References

Aghion, Philippe, Patrick Bolton, Christopher Harris, and Bruno Jullien. 1991. "Optimal Learning by Experimentation." *Review of Economic Studies*, 58: 621–654.

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." *Research in Economics*. Forthcoming.

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

Banks, Jeffrey S., and Rangarajan K. Sundaram. 1992. "A Class of Bandit Problems Yielding Myopically Optimal Strategies." *Journal of Applied Probability*, 29: 625–632.

Bardhi, Arjada. 2017. "Optimal Discovery and Influence Through Selective Sampling." Working Paper.

Bergemann, Dirk, and Juuso Välimäki. 2002. "Information Acquisition and Efficient Mechanism Design." *Econometrica*, 70: 1007–1033.

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.

Berry, D., and B. Fristedt. 1988. "Optimality of Myopic Stopping Times for Geometric Discounting." *Journal of Applied Probability*, 25: 437–443.

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.

Branco, Fernando, Monic Sun, and J. Miguel Villas-Boas. 2012. "Optimal Search for Product Information." *Management Science*, 58(11): 2037–2056.

Bubeck, Sébastian, Rémi 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.

Callander, Steven. 2011. "Searching and Learning by Trial and Error." *American Economic Review*, 101: 2277–2308.

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.

Chick, Stephen E., and Peter Frazier. 2012. "Sequential Sampling with Economics of Selection Procedures." *Management Science*, 58(3): 550–569.

Colombo, Luca, Gianluca Femminis, and Alessandro Pavan. 2014. "Information Acquisition and Welfare." *Review of Economic Studies*, 81: 1438–1483.

Denti, Tommaso. 2017. "Unrestricted Information Acquisition." 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.

Doval, Laura. 2017. "Whether or Not to Open Pandora's Box." Working Paper.

Easley, David, and Nicolas M. Kiefer. 1988. "Controlling a Stochastic Process with Unknown Parameters." *Econometrica*, 56: 1045–1064.

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

Garfagnini, Umberto, and Bruno Strulovici. 2016. "Social Experimentation with Interdependent and Expanding Technologies." *Review of Economic Studies*, 83: 1579–1613.

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

Greenshtein, Eitan. 1996. "Comparison of Sequential Experiments." *The Annals of Statistics*, 24: 436–448.

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

Hébert, 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.

Keener, R. 1985. "Further Contributions to the Two-armed Bandit Problem." *Annals of Statistics*, 13: 418–422.

Ke, T. Tony, and J. Miguel Villas-Boas. 2017. "Optimal Learning Before Choice." Working Paper.

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.

Klabjan, Diego, Wojciech Olszwski, and Asher Wolinsky. 2014. "Attributes." *Games and Economic Behavior*, 88: 190–206.

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.

Liang, Annie, and Xiaosheng Mu. 2017. "Overabundant Information and Learning Traps." Working Paper.

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.

Morgan, Peter, and Richard Manning. 1985. "Optimal Search." *Econometrica*, 53(4): 923–944.

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.

Olszwski, Wojciech, and Richard Weber. 2015. "A More General Pandora Rule?" *Journal of Economic Theory*, 160: 429–437.

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

Reinganum, Jennifer. 1983. "Nash Equilibrium Search for the Best Alternative." *Journal of Economic Theory*, 30: 139–152.

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

Rothschild, Michael. 1974. "A Two-Armed Bandit Theory of Market Pricing." *Journal of Economic Theory*, 9: 185–202.

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.

Russo, Daniel. 2016. "Simple Bayesian Algorithms for Best-Arm Identification." Working

Paper.

Sanjurjo, Adam. 2017. "Search with Multiple Attributes: Theory and Empirics." *Games and Economic Behavior*, , (104): 535–562.

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

Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics*, 50(3): 665–690.

Steiner, Jakub, Colin Stewart, and Filip Matějka. 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.

Taylor, Curtis R. 1995. "Digging for Golden Carrots: An Analysis of Research Tournaments." *American Economic Review*, 85(4): 872–890.

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

Weitzman, Martin L. 1979. "Optimal Search for the Best Alternative." *Econometrica*, 47(3): 641–654.

Wilson, Andrea. 2014. "Bounded Memory and Biases in Information Processing." *Econometrica*, 82: 2257–2294.

Yang, Ming. 2015. "Coordination with Flexible Information Acquisition." *Journal of Economic Theory*, 158: 721–738.

Zhong, Weijie. 2017. "Optimal Dynamic Information Acquisition." Working Paper.