

Rational Inattention with Sequential Information Sampling*

Benjamin Hébert[†]
Stanford University

Michael Woodford[‡]
Columbia University

September 5, 2017

Abstract

We propose a new principle for measuring the cost of information structures in rational inattention problems, based on the cost of generating the information used to make a decision through a dynamic evidence accumulation process. We introduce a continuous-time model of sequential information sampling, and provide assumptions under which the choice frequencies resulting from optimal information accumulation are the same as those implied by a static rational inattention problem with a particular cost function. Among the cost functions that can be justified in this way is the mutual information cost proposed by Sims (2010), but this is only one possibility. We introduce a class of “neighborhood-based” cost functions, which make it more costly to undertake experiments that differentiate between similar states. With this alternative cost function, optimal information accumulation implies choice frequencies that vary continuously with the state, even when payoffs are discontinuous, as observed in perceptual discrimination experiments.

*The authors would like to thank Mark Dean, Sebastian Di Tella, Mira Frick, Matthew Gentzkow, Emir Kamenica, Divya Kirti, Jacob Leshno, Stephen Morris, Pietro Ortoleva, Doron Ravid, Ilya Segal, Ran Shorrer, Ming Yang and participants at the Cowles Theory conference, Barcelona GSE Summer Conference on Stochastic Choice, and Stanford GSB research lunch for helpful discussions on this topic, and the NSF for research support. Portions of this paper appeared in Benjamin Hébert’s Ph.D. dissertation at Harvard University. All remaining errors are our own.

[†]Hébert: Stanford University. Email: bhebert@stanford.edu.

[‡]Woodford: Columbia University. Email: mw2230@columbia.edu.

1 Introduction

The theory of rational inattention, proposed by Christopher Sims and surveyed in Sims (2010), endogenizes the imperfect awareness that decision makers have about the circumstances under which they must choose their actions. According to the theory, a decision maker (DM) chooses her action on the basis of a subjective representation of the decision situation which provides only an imperfect indication of the true state. The information structure is assumed to be optimal, in the sense of allowing the best possible state-contingent action choice, net of a cost of information. In Sims' theory, the cost of an arbitrary information structure is proportional to the Shannon mutual information between the true state of the world and the information state available to the DM.

It is not obvious, though, that the theorems that justify the use of mutual information in communications engineering (Cover and Thomas (2012)) provide any warrant for using it as a cost function in a theory of attention allocation, either in the case of economic decisions or that of perceptual judgments.¹ Here we propose an alternative criterion for determining the cost of an information structure, that follows from assuming that the DM's subjective representation of her situation must be built up through a process of sequential evidence accumulation, in which each successive increment to the cumulatively available evidence is only very minimally informative.

In addition, the mutual-information cost function has implications that are unappealing on their face, and that seem inconsistent with evidence on the nature of sensory processing, as discussed in Woodford (2012). For example, the mutual-information cost function imposes a type of symmetry across different states of nature, so that it is equally easy or difficult to distinguish between any two states that are equally probably *ex ante*.

¹As explained in Cover and Thomas (2012), these theorems rely upon the possibility of "block coding" of a large number of independent instances of a given type of message, that can be jointly transmitted before any of the messages have to be decoded by the recipient. In our situation, an action must be taken in an individual decision problem, without waiting to learn about a large number of problems of the same form.

In the experimental task discussed by Caplin and Dean (2015), in which subjects are presented with an array of 100 red and blue balls, and must determine whether there are more red balls or more blue on a given trial, Sims' theory of rational inattention implies that, because the reward from any action (e.g., declaring that there are more red balls) is the same for all states with the property that there are more red balls than blue, the probability of a subject's choosing that response will be the same in each of those states.² In fact, it is much easier to quickly and reliably determine that there are more red balls for some arrays in this class (e.g. one with 98 red balls and only two blue balls) than others (e.g. one with 51 red balls and 49 blue balls, relatively uniformly dispersed), and subjects make more correct responses in the former case.³

Our alternative theoretical foundations exploit the special structure implied by an assumption that information sampling occurs through a sequential process, in which each additional signal that is received determines whether additional information will be sampled, and if so, the kind of experiment to be performed next. We emphasize the limiting case in which each individual experiment is only minimally informative, but a very large number of independent experiments can be performed. In this continuous-time limit, we obtain strong and relatively simple characterizations of the implications of rational inattention, owing to the fact that only local properties of the assumed cost function for individual experiments matter in this limiting case.

We believe that it is often quite realistic to assume that information is acquired through a sequential sampling process. As discussed in Fehr and Rangel (2011) and Woodford (2014), an extensive literature in psychology and neuroscience has argued that data on both the frequency of perceptual errors and the frequency distribution of response times can be explained by models of perceptual classification based on sequential sampling. More

²This is an implication of Lemma 6 in section 5.

³Dewan and Neligh (2017) present similar evidence against Sims' theory, from a related experiment in which subjects must estimate the number of dots in a visual array.

recently, some authors have proposed that data on stochastic choice and response time in economic contexts can be similarly modeled.⁴

Our paper is not the first that seeks to derive at least some features of such models from a theory of optimal information sampling. In particular, Moscarini and Smith (2001) consider both the optimal intensity of information sampling per unit of time and the optimal stopping problem, when the only possible kind of information is given by the sample path of a Brownian motion with a drift that depends on the unknown state of the world.⁵ (Fudenberg et al. (2015) consider a variant of this problem with a continuum of possible states, and an exogenously fixed sampling intensity.⁶) Woodford (2014) instead takes as given a stopping rule (motivated by the empirical psychology and neuroscience literatures), but allows a very flexible choice of the information sampling process, as in theories of rational inattention. Our approach differs from these earlier efforts in seeking to endogenize *both* the nature of the information that is sampled at each stage of the evidence accumulation process and the stopping rule that determines how much evidence is collected before a decision is made.

We also consider decision problems with an arbitrary finite number of choice alternatives, rather than restricting attention to binary choice problems, as in both Fudenberg et al. (2015) and Woodford (2014). In the sequential information sampling problem considered here, we allow the information sampled at each stage to be chosen very flexibly, as in Woodford (2014), subject only to a “flow” information-cost function; but we also allow the decision when to stop sampling and make a decision to be made optimally, on the basis of the entire history of information sampled to that point, as in Moscarini and Smith (2001) and Fudenberg et al. (2015). Among other results, we describe a class of information-cost

⁴In addition to the references in Fehr and Rangel (2011), recent examples include Krajbich et al. (2014) and Clithero (2016). Shadlen and Shohamy (2016) provide a neural-process interpretation of sequential-sampling models of choice.

⁵Moscarini and Smith (2001) allow the instantaneous variance of the observation process to be freely chosen (subject to a cost), but this is equivalent to changing how much of the sample path of a given Brownian motion can be observed by the DM within a given amount of clock time.

⁶See also Tajima et al. (2016) for analysis of a related class of models.

functions such that in the case of a binary decision, the DM’s beliefs evolve according to a diffusion along a one-dimensional line segment, with a decision being made when either of the two endpoints is reached, as postulated by Woodford (2014).

In our continuous-time model, the optimal information-sampling problem is presented as a problem of optimal control of a diffusion process on the probability simplex (the set of possible posterior beliefs), with sampling stopping when certain (endogenously determined) boundaries are reached. For a relatively flexible family of possible cost functions for individual experiments, the continuous time model’s predictions with regard to state-dependent choice frequencies are the same as those of a static rational-inattention model, with an appropriately chosen information-cost function. The finite set of possible signals in the equivalent static model corresponds to the set of different possible *terminal* information states in the dynamic model, each of which corresponds to one of the possible actions.

For a particular family of flow information-cost functions, the cost function for the equivalent static model is just the mutual information between the action chosen and the true state of the world; we thus provide foundations for the kind of rational inattention problem proposed by Sims (2010),⁷ that do not rely on any analogy with rate-distortion theory in communications engineering. But while our dynamic model makes predictions that are equivalent to those of the rational inattention theory of Sims (2010) (and more particularly, its application to stochastic choice by Matějka et al. (2015)) for this particular family of flow information-cost functions, we show that different predictions can be obtained under other, very plausible specifications of the flow cost function.

We focus on the implications of an attractive family of flow information-cost functions, which we call “neighborhood-based” cost functions. The idea of this class of information-cost specifications is that information structures are more costly the greater the extent

⁷Morris and Strack (2017) provide a related foundation for the mutual-information cost function, but for the special case in which there are only two possible states.

to which they allow intrinsically similar states of the world (states that share a “neighborhood”) to be discriminated; the dependence on a concept of intrinsic similarity between states (the “neighborhood structure”) distinguishes these cost functions from the mutual-information cost function. We show that versions of our theory that assume a flow information-cost function in this family can explain the kind of continuous variation of response frequencies with changes in the characteristics of the alternatives presented that is commonly observed in perceptual discrimination experiments (but that would not be predicted by the standard theory of rational inattention).⁸

Section 2 begins directly with a description of our continuous-time model, and introduces the information-cost matrix function as a way of parameterizing information costs in this model. Section 3 then presents one of our main results (Theorem 1), that in a large set of cases, the solution to the continuous-time model is equivalent, in terms of the joint distribution of choices and states, to the implications of a static rational inattention model with a suitable static information-cost function.

In section 4, we discuss the connection between the information-cost matrix function of the continuous-time model and the flow information-cost function for an individual signal, and state a set of general assumptions that flow information-cost functions are assumed to satisfy, in the spirit of the treatment of static rational inattention problems by De Oliveira et al. (2017). We show that all flow cost functions satisfying these conditions must have a particular type of local structure when individual “experiments” are nearly uninformative.⁹ Section 5 then introduces a specific class of flow cost functions that satisfy these general conditions, our “neighborhood-based” cost functions.

Finally, section 6 provides a justification for the continuous-time model proposed in

⁸In section A.2 of the appendix, we describe a specific case of this class of cost functions, based on the Fisher information as a measure of the informativeness of an information structure, that can be applied to rational inattention models with a continuum of states.

⁹This method, based on Chentsov (1982), is also used by Hébert (2014) in a different context.

section 2, and for the connection between flow cost function specifications and information-cost matrix function of the continuous-time model asserted in section 4. Here we show that a discrete-time dynamic evidence accumulation problem, in which the cost of each individual signal is given by a flow cost function satisfying the assumptions stated in section 4, leads to the continuous-time problem discussed in section 2, in the limit as the number of successive signals per time period is made large, while the informativeness of each individual signal is made small at a corresponding rate. Section 7 concludes.

2 Continuous-Time Sequential Evidence Accumulation

We begin by directly introducing our continuous-time model of sequential evidence accumulation, leaving for later (section 6) the demonstration that it arises as a limiting case of an explicit discrete-time dynamic evidence accumulation problem. Let $x \in X$ be the underlying state of the nature, and $a \in A$ be the action taken by the DM. For simplicity, we assume that A and X are finite sets. We also assume that the number of states is weakly larger than the number of actions, $|X| \geq |A|$. The DM's utility from taking action a in state x at time t is $u_{a,x} - \kappa t$. The parameter $\kappa > 0$ governs the penalty for delaying making a decision; the DM does not discount the future. We assume a penalty of this kind, rather than time discounting, for reasons of tractability.

The DM does not perfectly observe the state $x \in X$. At each time t , the DM holds beliefs $q_t \in \mathcal{P}(X)$, where $\mathcal{P}(X) \subset \mathbb{R}^{|X|}$ denotes the probability simplex over X . That is, q_t is a vector of length $|X|$, whose elements, denoted $q_{x,t}$, are the probability, under the DM's beliefs at time t , of state x . Time begins at $t = 0$, when the DM holds prior beliefs q_0 . At each moment in time, the DM faces two decisions: whether to gather information about the state $x \in X$, and whether to stop and make a decision. When stopping with beliefs q_τ at time τ , the DM will simply choose a to maximize $u_a^T \cdot q_\tau$, where u_a is the vector of utilities

associated with action a , resulting in payoff $\hat{u}(q_\tau) - \kappa\tau$.

When the DM gathers information, she chooses the variance-covariance matrix of possible changes in her beliefs, subject to certain constraints. In our model, the DM's beliefs evolve as

$$dq_{x,t} = q_{x,t} \sigma_{x,t} \cdot dB_t, \quad (1)$$

where dB_t is an $|X| - 1$ -dimensional Brownian motion,¹⁰ σ_t is a matrix that can be chosen by the DM, and $\sigma_{x,t}$ is a particular row of that matrix.

The DM's choice of σ_t is subject to restrictions — a trivial one to ensure that the beliefs stay in the simplex, and an economic restriction that limits the amount of information the DM can acquire. The trivial restriction is that

$$1^T \cdot dq_t = 0$$

always, where 1 is a vector of ones. This restriction is equivalent to requiring that

$$\sigma_t^T q_t = \vec{0}.$$

We will use $M(q_t)$ to denote the set of $|X| \times |X|$ matrices satisfying this condition. Our notation enforces the requirement that $dq_{x,t} = 0$ if $q_{x,t} = 0$.

The non-trivial restriction, which limits the quantity of information the DM can acquire at each moment, is

$$\frac{1}{2} tr[\sigma_t \sigma_t^T k(q_t)] \leq \chi, \quad (2)$$

where $k(q_t)$ is an $|X| \times |X|$ dimensional matrix-valued function we will refer to as the “information-cost matrix function”, $tr[\cdot]$ is the trace, and χ is a positive constant that in-

¹⁰Note that this is largest possible number of independent Brownian motions of which dq_t may be a linear combination.

dexes the tightness of the constraint. We discuss this constraint, and the information-cost matrix function, in more detail below. For now, we note simply that the information-cost matrix function satisfies certain properties: for any q_t , $k(q_t)$ is symmetric and positive semi-definite, and its null space is the space of vectors that are constant for all $x \in X$ in the support of q_t .¹¹

Using her control of the volatility of her beliefs, and subject to the constraints imposed by the information-cost matrix function, our DM attempts to maximize her expected payoff. Her sequence problem can be written, given beliefs q_t at time t ,

$$V(q_t) = \sup_{\{\sigma_s \in M(q_s)\}, \tau \geq t} E_t[\hat{u}(q_\tau) - \kappa(\tau - t)],$$

where τ is the DM's endogenous stopping time, subject to the constraints listed previously.

Wherever this value function is twice-differentiable and the DM does not choose to stop, the problem can be given a simple recursive representation:

$$\sup_{\sigma_t \in M(q_t)} \frac{1}{2} \text{tr}[\sigma_t^T D(q_t) V_{qq}(q_t) D(q_t) \sigma_t] = \kappa,$$

subject to the information constraint (2), where $D(q_t)$ is a diagonal matrix with the elements of q_t on its diagonal, and $V_{qq}(q_t)$ is the Hessian of $V(q)$ evaluated at $q = q_t$.¹²

¹¹Actually, because we require that $\sigma_t \in M(q_t)$, constraint (2) only involves the quadratic form $v^T k(q_t) w$ defined for vectors v and w such that $v^T q_t = w^T q_t = 0$. We extend the definition of the quadratic form to all vectors $v, w \in \mathbb{R}^{|X|}$, in order to obtain a unique representation in terms of a matrix $k(q_t)$, by adding the requirement that $k(q)v = 0$ for any vector $v \in \mathbb{R}^{|X|}$ with the property that v_x is equal to a constant for all x in the support of q .

¹²In the case of a differentiable function $V(q)$ defined on the probability simplex $\mathcal{P}(X)$, in order to write the Hessian of the function as a matrix, we must adopt a coordinate system for the tangent space to the probability simplex. Throughout this paper, we do this by extending the function to the domain $\mathbb{R}_+^{|X|}$ by defining the function to be homogeneous of degree one on this larger domain (an assumption that does not restrict the function's values on the simplex). Vectors in the tangent space are then simply vectors in $\mathbb{R}^{|X|}$, which we express using the natural set of basis vectors corresponding to each element of X . The Hessian matrices appearing in equations such as (3), (11), (30), and (16) below should also all be understood in this way.

The following lemma describes the Hamilton-Jacobi-Bellman (HJB) equation associated with this dynamic optimization problem. It is derived by showing that the information constraint binds.¹³ The maximum eigenvalue appears in place of a maximization over σ_t , but this is just a compact way of expressing the idea that the DM is choosing in which direction(s) to update her beliefs.

Lemma 1. *Anywhere the value function $V(q_t)$ is twice-differentiable, it satisfies*

$$\max\{\lambda_1(D(q_t)V_{qq}(q_t)D(q_t) - \theta k(q_t)), \hat{u}(q_t) - V(q_t)\} = 0, \quad (3)$$

where $\theta = \chi^{-1}\kappa$, and for any $|X| \times |X|$ matrix K , $\lambda_1(K)$ denotes the largest eigenvalue of K associated with an eigenvector v such that $\iota^T v = 0$.¹⁴

Proof. See the appendix, section D.1. □

This equation has the standard form of an optimal stopping problem, with the twist that it is a ‘‘Hessian equation’’ in the continuation region. The parameter θ describes the race between information acquisition and time in this model. The larger the penalty for delay, and the tighter the information constraint, the larger the parameter θ . The caveat about twice-differentiability plays several roles. First, as is common in optimal stopping problems, the value function may not be twice differentiable on the stopping boundary. Second, the Hessian equation in the continuation region is ‘‘degenerate elliptic’’, and therefore a solution that is twice-differentiable everywhere in the continuation region may not exist. A third complication is that the beliefs q_t may come to place zero weight on a certain state —

¹³The derivation depends on an additional property of the $k(q_t)$ matrix that will be discussed below.

¹⁴Here we are interested in the eigenvectors of the matrix corresponding to elements of the tangent space to the probability simplex. Note that under our notation for writing quadratic forms over the probability simplex as matrices, explained in footnotes 11 and 12 above, ι is a null eigenvector of both $D(q)V_{qq}D(q)$ and $k(q)$, for any q ; but we do not wish to count this as one of the eigenvectors of the linear operator for purposes of defining the maximum eigenvalue, as our first-order condition actually involves a linear operator defined on the tangent space of the probability simplex.

that is, the beliefs may hit the boundary of the simplex, at which point the value function $V(q_t)$ is not twice-differentiable in all directions. Fortunately, in what follows, these issues will be a nuisance, rather than a serious obstacle.

The DM's optimal stopping rule is characterized by the standard value-matching and smooth-pasting conditions. Let $\Omega \subset \mathcal{P}(X)$ be the open subset of the simplex on which the DM continues to search for information, and let $\partial\Omega$ denote its boundary. For all $q \in \partial\Omega$, the value matching condition, $V(q) = \hat{u}(q)$, and smooth pasting condition, $V_q(q) = \hat{u}_q(q)$, will hold. Note, however, that the derivative $\hat{u}_q(q)$ does not exist everywhere — at beliefs where the DM is just indifferent between two actions with distinct state-contingent payoffs, the stopping payoff is non-differentiable.¹⁵ However, it will never be optimal for the DM to stop at one of these indifference points.

Before we describe the value function, we will provide some intuition for the volatility constraint and describe in more detail the information-cost matrix function. The volatility constraint is a limit on the information the DM can acquire, because it limits the volatility of her beliefs. Our DM is a Bayesian, meaning that she can never expect to revise her beliefs in a particular direction — her beliefs must be a martingale; this is why there can be no drift term in equation (1). If she receives a mostly uninformative signal at a particular moment, her beliefs have a small amount of volatility at that moment. In contrast, if she receives an informative signal, her beliefs will be very volatile.

Our specification assumes that her beliefs are driven by a Brownian motion, which generates continuous sample paths and does not have jumps.¹⁶ This embeds the idea that, as one looks at smaller and smaller time intervals, the informativeness of the signals the

¹⁵At this point, we have also not shown that $V(q)$ is differentiable everywhere, but this is proven in the proof of Theorem 1.

¹⁶Che and Mierendorff (2016) and Zhong (2017) explore related models with jumps in beliefs. These are assumed to represent the only possible form of information arrival in the former paper, and demonstrated to represent an optimal form of experimentation in the latter paper, under assumptions different from those made here.

DM is observing scales down. In section 6, we discuss more primitive assumptions about the cost of alternative dynamic information sampling strategies that lead the DM to want to smooth the quantity of information gathered across time, so that the continuity assumed in this section is a feature of the optimal strategy, in a continuous-time limiting case of the model presented in that section.

We derive the information constraint (equation (2)) from a model in which the DM can choose any information structure she desires at each time period, as in standard rational inattention models. One result of our derivation is the observation that the DM can choose any volatility matrix σ_t . This is, in a sense, a familiar idea — Kamenica and Gentzkow (2011), for example, emphasize the idea of choosing a distribution of posteriors, subject to the constraint that the mean posterior is equal to the prior. Our DM appears to choose only the volatility, and not the higher cumulants of the distribution of posteriors, but this is because she finds it optimal to smooth her information gathering over time, and the instantaneous volatility is sufficient to characterize the resulting process for beliefs. This result permits both a relatively parsimonious specification of the information sampling strategies available to the DM, and a relatively parsimonious specification of possible forms for the information constraint.

In modeling the evolution of the DM’s beliefs as a diffusion process, our model resembles those proposed by authors such as Krajbich et al. (2014) and Fudenberg et al. (2015), though unlike those authors we endogenize the diffusion process through which additional information arrives while sampling continues. Additionally, our model emphasizes the “unconditional” dynamics of beliefs (that is, not conditional on any particular state being the true state), whereas the models discussed by those authors are described in terms of their “conditional” dynamics (that is, conditional on some particular state being the true state).

The information-cost matrix function $k(q_t)$ is more than simply a way of obtaining a single (scalar) measure of the “size” of the elements of σ_t . The relative size of different

elements of the matrix also allows us to specify the degree to which it is more costly to obtain more precise information of some kinds rather than others. Larger (positive) diagonal elements k_{xx} for certain states x imply that it is relatively more costly to obtain signals that reveal much about the likelihood of those states; larger negative off-diagonal elements $k_{xx'}$ (relative to the size of the diagonal elements k_{xx} and $k_{x'x'}$) for pairs of states x, x' imply that it is relatively more costly to obtain signals that allow one to differentiate sharply between states x and x' .

An example of an information-cost matrix function that satisfies our general assumptions (and will be important for the discussion below) is the inverse Fisher information matrix $(g^+(q))$,¹⁷

$$k(q) = g^+(q) = \begin{bmatrix} q_1(1-q_1) & -q_1q_2 & \dots & -q_1q_{|X|} \\ -q_1q_2 & q_2(1-q_2) & \dots & -q_2q_{|X|} \\ \vdots & \vdots & \ddots & \vdots \\ -q_1q_{|X|} & -q_2q_{|X|} & \dots & q_{|X|}(1-q_{|X|}) \end{bmatrix}. \quad (4)$$

In this case, the off-diagonal element $k_{xx'}(q)$ is equal to $-q(x)q(x')$ for any pair of states x, x' ; thus it depends only on the prior probabilities of the two states, and is otherwise the same regardless of the states selected. Thus any two states are assumed to be equally easy or difficult to tell apart: it only matters whether two states are the same or not, and how likely they are to occur.

While this kind of symmetry might seem appealing on a priori grounds for some applications (where the different possible states are a set of alternatives, each equally unrelated to all of the others), we view it as quite implausible for many cases of economic relevance. For example, one is often interested in states that represent different possible values of

¹⁷The Fisher information matrix, of which this can be viewed as a pseudo-inverse, is described in section 4.

some quantity (a “state variable”), and hence can be ordered on a line.¹⁸ One might well suppose that possible methods of learning about the value of that variable will all have the property that nearby values of the state variable result in similar probabilities of receiving particular signals, and hence that it is particularly costly to arrange an information structure that makes the conditional probabilities of signals very different for states that are near one another in the ordering of states.

An alternative possible information-cost matrix function, also satisfying our general assumptions, is given by

$$k(q) = \begin{bmatrix} \frac{q_1 q_2}{q_1 + q_2} & -\frac{q_1 q_2}{q_1 + q_2} & 0 & \cdots & 0 \\ -\frac{q_1 q_2}{q_1 + q_2} & \frac{q_1 q_2}{q_1 + q_2} + \frac{q_2 q_3}{q_2 + q_3} & -\frac{q_2 q_3}{q_2 + q_3} & \ddots & \vdots \\ 0 & -\frac{q_2 q_3}{q_2 + q_3} & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \frac{q_{|X|-1} q_{|X|-2}}{q_{|X|-2} + q_{|X|-1}} + \frac{q_{|X|} q_{|X|-1}}{q_{|X|-1} + q_{|X|}} & -\frac{q_{|X|-1} q_{|X|}}{q_{|X|} + q_{|X|-1}} \\ 0 & \cdots & 0 & -\frac{q_{|X|} q_{|X|-1}}{q_{|X|} + q_{|X|-1}} & \frac{q_{|X|-1} q_{|X|}}{q_{|X|} + q_{|X|-1}} \end{bmatrix}. \quad (5)$$

In this case, the only off-diagonal elements $k_{xx'}(q)$ are negative elements in the case that x' directly follows x in the ordering of states (or vice versa). This form of matrix $k(q)$ implies that an information structure is costly only to the extent that there are pairs of “neighboring” states x, x' for which the conditional probabilities of signals are different ($p_{x'} \neq p_x$).

The differing implications of these two alternative assumptions about the form of the information-cost matrix function are explored in section 5. For now, we simply note that our model allows for different specifications in this regard, and that we regard this as desirable, as it will often be reasonable for the specification of information costs to incorporate a notion of “distance” between different possible states.

Our derivation of the continuous-time problem set out above from a more explicit evi-

¹⁸This is also often true of perceptual classification experiments, in which subjects are asked to classify stimuli that differ from one another in their intensity or magnitude along some single dimension.

dence accumulation problem places additional restrictions on the information-cost matrix function, beyond the properties already mentioned above: it will in fact be necessary that $k(q)$ be continuous, and that there exist a positive constant m such that $k(q) - mg^+(q)$ is positive semi-definite.¹⁹

For a large class of information-cost matrix functions $k(q_t)$, we can solve the sequence problem described in this section, and show that the solution is equivalent to a certain static rational inattention problem. We present these results in the next section.

3 Static and Dynamic Rational Inattention Problems

In most theories of rational inattention, including the classic formulations of Sims, only a single signal is collected for each decision that must be made. In a decision problem where an action is to be chosen once from a set of possibilities, the rational inattention problem is static; a signal is obtained (once) that depends on the state, an action is taken that depends on the signal, and that is all. The kind of dynamic optimization model proposed in the previous section seems quite different.

Nonetheless, we establish below that in a broad class of cases, it is possible to establish an equivalence between the information that is acquired through an optimal evidence accumulation process of the kind proposed in the previous section and the information acquired in a static model of rational inattention, with a particular type of cost function. Thus our dynamic model does not necessarily have different implications than a static rational inattention model; however, the dynamic optimization problem can provide a reason for interest in static information-cost functions of particular types.

¹⁹Examples (4) and (5) above are both continuous in q . The second of these examples does not strictly satisfy the second requirement stated in the text for $m > 0$, but is the limit of a sequence of examples that does. These examples are closely related to the mutual-information cost function proposed by Sims and to a “neighborhood-based” cost function that we introduce in section 5, respectively.

We begin by explaining the form of a static rational inattention problem. As in the previous section, let $x \in X$ be the underlying state of nature, and let $s \in S$ be a signal the DM can receive, which might convey information about the state. We assume that X and S are finite sets. Let $q \in \mathcal{P}(X)$ denote the DM's prior belief (before receiving a signal) about the probability of state x . Define $p_{s,x}$ as the probability of receiving signal s in state x , let $p_x \in \mathcal{P}(S)$ be the associated conditional probability distribution of the signals given state x , and let p be the $|S| \times |X|$ matrix whose elements are $p_{s,x}$. The matrix p , which is a set of conditional probability distributions for each state of nature, $\{p_x\}_{x \in X}$, defines an "information structure." After receiving signal s , the DM will hold a posterior, $q_s \in \mathcal{P}(X)$, which is a function of p and q , defined by Bayes' rule.

The maximum achievable expected payoff, given an information structure p and prior q , can be written as

$$\bar{u}(p, q) \equiv \max_{\{a(s)\}} \sum_{x \in X} \sum_{s \in S} q_x p_{s,x} u(a(s), x).$$

The standard static rational inattention problem, given the signal alphabet S ,²⁰ is then

$$\max_{\{p_x \in \mathcal{P}(S)\}_{x \in X}} \bar{u}(p, q) - \theta C(p, q; S), \quad (6)$$

where

$$C(\cdot, \cdot; S) : \mathcal{P}(S)^{|X|} \times \mathcal{P}(X) \rightarrow \mathbb{R} \quad (7)$$

is a cost function for information structures, and $\theta > 0$ is a multiplicative factor that lets us consider alternative assumptions about the tightness of the information constraint, given a measure of the informativeness of alternative information structures represented by the function C .

In the classic formulation of Sims, a problem of the form equation (6) is considered, in

²⁰The full problem includes a choice over the signal alphabet S . A standard result, which will hold for all of the cost functions we study, is that $|S| = |A|$ is sufficient.

which the cost function $C(p, q; S)$ is given by the Shannon mutual information between the signal and the state. This can be defined using Shannon's entropy,²¹

$$H^{Shannon}(q) \equiv - \sum_{x \in X} (e_x^T q) \ln(e_x^T q). \quad (8)$$

Shannon's entropy can in turn be used to define a measure of the degree to which the posterior q_s associated with any signal differs from the prior q , the Kullback-Leibler (KL) divergence,

$$D_{KL}(q_s || q) \equiv H^{Shannon}(q) - H^{Shannon}(q_s) + (q_s - q)^T H_q^{Shannon}(q). \quad (9)$$

Mutual information is then the expected value of the KL divergence over possible signals,

$$I^{Shannon}(p, q; S) \equiv \sum_{s \in S} (e_s^T p q) D_{KL}(q_s || q). \quad (10)$$

It is a measure of the informativeness of the signal, in that it provides a measure of the degree to which the signal changes what one should believe about the state, on average.

Shannon's mutual information is not, however, the only possible measure of the informativeness of an information structure, or the only plausible cost function for a static rational inattention problem. We discuss additional examples below, but first return to our discussion of the continuous-time information sampling problem introduced in section 2.

To obtain further results, we restrict our attention to information-cost matrix functions with the following property: there exists a twice-differentiable function $H : \mathbb{R}_+^{|X|} \rightarrow \mathbb{R}$ such

²¹We use the notation e_x to denote the vector (element of \mathbb{R}^X) with a one in the place corresponding to state x , and zeros elsewhere (column x of the identity matrix of dimension $|X|$).

that, for all q_t in the interior of the simplex,

$$D(q_t)^{-1}k(q_t)D(q_t)^{-1} = H_{qq}(q_t). \quad (11)$$

This class includes a number of information-cost matrix functions of interest: for example, it includes the case in which $k(q_t)$ is the inverse Fisher information matrix, which we will show corresponds to the standard rational inattention model, and the case in which $k(q_t)$ is the “neighborhood-based” function that we introduce in section 5.²² We shall refer to the function H as the “generalized entropy function,” for reasons that will become clear below. Using this convex function, we can define a Bregman divergence,

$$D_H(q_s||q) = H(q_s) - H(q) - (q_s - q)^T H_q(q).$$

The Kullback-Leibler divergence is a Bregman divergence (see equation (9)), with a generalized entropy function equal to the negative of Shannon’s entropy.

With these information-cost matrix functions, it is easy to show (using equation (3)) that the quantity $V(q) - \theta H(q)$ is a local martingale inside the continuation region, anywhere the value function is twice differentiable. Ignoring several technicalities, which are discussed in the proof, we can apply the optional stopping theorem:

$$\begin{aligned} V(q_0) &= E_0[V(q_\tau) - \theta H(q_\tau) + \theta H(q_0)] \\ &= E_0[\hat{u}(q_\tau) - \theta H(q_\tau) + \theta H(q_0)]. \end{aligned}$$

Using this idea, and the notion that, in an optimal stopping problem, the DM “chooses the boundaries,” we conjecture and verify the following result:

²²It is more restrictive, however, than the class of information-cost matrix functions defined in section 2.

Theorem 1. *There exists a unique solution to the continuous time sequential evidence accumulation problem, in which*

$$V(q_0) = \max_{\pi \in \mathcal{P}(A), \{q_a \in \mathcal{P}(X)\}_{a \in A}} \sum_{a \in A} \pi(a) (u_a^T \cdot q_a) - \theta \sum_{a \in A} \pi(a) D_H(q_a || q_0),$$

subject to the constraint that $\sum_{a \in A} \pi(a) q_a = q_0$.

There exist maximizers of this problem, π^* and q_a^* , such that π^* is the unconditional probability, in the dynamic problem, of choosing a particular action, and q_a^* , for all a such that $\pi^*(a) > 0$, is the unique belief the DM will hold when stopping and choosing that action.

Proof. See the appendix, section D.2. □

Thus the sequential evidence accumulation problem is equivalent to a static rational inattention problem of the kind described above, with a particular kind of static information-cost function,

$$\begin{aligned} C(p, q_0; S) &= \sum_{s \in S} \pi(s) D_H(q_s || q_0), \\ &= \sum_{s \in S} \pi(s) H(q_s) - H(q_0) \end{aligned} \tag{12}$$

where $\pi(s)$ now refers to the unconditional probability of receiving signal s in the static problem, and q_s to the posterior when signal s is received, with the signal space S in the static problem identified with the set of possible actions A .²³ We call a cost function that can be written in the form (12) “posterior-separable.”²⁴

The mutual-information cost function (10) proposed by Sims is one such cost func-

²³The “signal” can thus be viewed as an instruction as to which action is advisable.

²⁴This kind of cost function is instead called “uniformly posterior-separable” in Caplin et al. (2017). The class of static cost functions that can be justified by Theorem 1 is also related to the class of “GERI” cost functions defined by Fosgerau et al. (2016).

tion. In this case, the generalized entropy function H is the negative of Shannon’s entropy (8), the corresponding information-cost matrix function is the inverse Fisher information matrix (4), the Bregman divergence is the Kullback-Leibler divergence (9), and the information measure defined by (12) is then the Shannon mutual information (10). Thus Theorem 1 provides a foundation for assuming endogenous information of the same kind as the standard static rational inattention model, and hence for the same predictions regarding stochastic choice as are obtained by Matějka et al. (2015).²⁵ This assumes a particular information-cost matrix function; but below we show not only that a continuous-time model with this particular information-cost matrix function can arise as the limit of a well-behaved discrete-time model of sequential evidence accumulation, but also that any of an entire class of possible specifications of the flow information cost function for that discrete-time model will lead to this result in the continuous-time limit.

On the other hand, Theorem 1 also implies that other posterior-separable cost functions can similarly be justified. Indeed, any static information cost function (12), where D_H is the Bregman divergence derived from some convex, twice-differentiable function H , can be given such a justification.²⁶ We give additional examples in section 5. The divergence D_H associated with such a model is of interest apart from its role in defining the equivalent static information-cost function (12). In particular, the expected value of the divergence indicates the expected time cost required for the DM to reach a decision.

Theorem 1 shows that we can associate a particular posterior-separable static information-cost function with any matrix-valued function $k(q)$ satisfying certain conditions: this is the information cost function that defines a static rational-inattention problem that is equivalent in certain respects to the continuous-time problem defined by $k(q)$. But we also show

²⁵For a related foundation for this static cost function, in the special case in which there are only two possible states, see Morris and Strack (2017).

²⁶The continuous-time information sampling process that is required is simply the one in which the information-cost matrix function is given by equation (11).

below how to derive a particular information-cost matrix function $k(q)$ corresponding to any information-cost function $C(p, q; S)$ satisfying certain conditions: $k(q)$ defines a local approximation to the function $C(p, q; S)$, and defines a continuous-time problem that represents a limiting case of a discrete-time optimal evidence accumulation problem, in which $C(p, q; S)$ is the flow information-cost function specifying the cost associated with each individual signal that is received.

Thus, there is a two-way relationship between matrix-valued functions $k(q)$ and cost functions $C(p, q; S)$. Posterior-separable cost functions are the fixed points of the resulting mapping: if one uses a posterior-separable cost function to derive an information-cost matrix function, and then solves the continuous-time model defined by that function $k(q)$, one will recover that same posterior-separable cost function as the static information-cost function of the equivalent static rational inattention problem.

Another important general consequence of Theorem 1 concerns posterior beliefs at the time of an eventual decision. The probability distribution $q_a^* \in \mathcal{P}(X)$ is the DM's belief conditional on taking action $a \in A$. The vector q_a^* is unique, given a particular action a , meaning that there is only one belief the DM can reach before choosing to stop and take a particular action. (The further this belief is from the DM's prior, q_0 , as measured by the divergence D_H , the more time it will take, in expectation, for the DM to arrive at this belief before acting.) The martingale property of beliefs during the evidence accumulation process thus requires that beliefs q_t at each stage of the process are some convex combination of the finite set of posteriors $\{q_a^*\}_{a \in A}$.

Hence beliefs diffuse on a simplex of dimension $|A| - 1$ during the decision process; if there are only two possible actions (as in the binary choice problems to which the drift-diffusion model is applied by authors such as Fehr and Rangel (2011)), then the belief state must diffuse along a line segment, as assumed in the DDM, regardless of the number of possible states $|S|$. Thus an apparently arbitrary feature of the DDM (outside the two-state

case in which the DDM is known to correspond to optimal Bayesian decision making, see Fudenberg et al. (2015)) can be shown to follow from optimal sequential evidence accumulation, if the information sampling is flexible in the way that we model it here.

4 Flow Information Costs

In this section, we elaborate on the connection, suggested above, between the information-cost matrix function in our continuous-time model of information sampling and the kind of cost function for an individual signal that is assumed in static rational inattention models. The continuous-time model can be viewed as a limiting case of a discrete-time model of optimal evidence accumulation, in which an endogenously determined signal is received each period, and the choice of the signal to receive each period is subject to a cost for more informative signals, specified by an information-cost function similar to the kind assumed in static rational inattention models. We call this cost function for an individual signal in a dynamic model of evidence accumulation the “flow information-cost function.”

While we defer until section 6 a complete discussion of how the continuous-time model can be derived as a limiting case of a discrete-time dynamic model, in this section we preview certain conclusions from that analysis by explaining the connection between the flow information-cost function of the discrete-time dynamic model and the information-cost matrix function of the continuous-time model. (Here the connection is simply asserted; the connection is proven in section 6.) We preview the results before proceeding to the complete derivation, because understanding them can help to explain the assumptions about the information-cost matrix function that we have proposed above. Additionally, in the following section we discuss a particular class of information-cost matrix functions, and wish to motivate this specification in terms of the form of flow information-cost functions from which these information-cost matrix functions can be derived.

An important conclusion of this section is our demonstration that many different flow information-cost functions can give rise to the same information-cost matrix function, and hence to the same predictions about the information that will be accumulated in the continuous-time limit. This is one of the main advantages, in our view, of considering the continuous-time limit: while our conclusions still depend on assumptions about the nature of information costs, there are less ways in which our conclusions can vary once we pass to the continuous-time limit.

At each stage of the discrete-time sequential evidence accumulation problem discussed in section 6, the DM chooses an information structure. Each information structure p has a cost $C(p, q; S)$, given by a function of the form of (7), where q indicates the DM's prior in this stage (that is, the posterior beliefs following from observations prior to the current stage of the dynamic problem), and S is again the signal alphabet.²⁷ Our most general results depend only on assuming that this flow information-cost function satisfies a set of six general conditions, stated below.

All of these conditions are satisfied by the mutual-information cost function (10) proposed by Sims, but they are also satisfied by many other cost functions. (Additional examples are given in section 5.) They are closely related to conditions that other authors have also proposed as attractive general properties to assume about information-cost function, though in the context of static information-cost functions of the kind discussed in section §3. Here we *assume* that the flow information-cost function in our dynamic model satisfies all six of these conditions; we then *prove* that under our assumptions, the equivalent static rational inattention problem (the existence of which is guaranteed by Theorem 1) involves a static information-cost function that satisfies these conditions.

²⁷The information-cost functions that we study, like mutual information, are defined for all finite signal alphabets S . Note, however, that mutual information is also defined over alternative sets of states of nature X . We do not impose this requirement on our more general cost functions — all of our analysis takes the set of states of nature as given.

Condition 1. Information structures that convey no information ($p_x = p_{x'}$ for all x, x' in the support of q) have zero cost. All other information structures have a weakly positive cost.

This condition ensures that the least costly strategy for the DM in the standard static rational inattention problem is to acquire no information, and make her decision based on the prior. The requirement that gathering no information has zero cost is a normalization.

The next condition is called mixture feasibility by Caplin and Dean (2015). Consider two information structures, $\{p_{1,x}\}_{x \in X}$, with signal alphabet S_1 , and $\{p_{2,x}\}_{x \in X}$, with alphabet S_2 . Given a parameter $\lambda \in (0, 1)$, we define a mixed information structure, $\{p_{M,x}\}_{x \in X}$ over the signal alphabet $S_M = (S_1 \cup S_2) \times \{1, 2\}$. For each $s = (s_1, 1)$ in the alphabet S_M , $p_{M,x}(s)$ is equal to $\lambda p_{1,x}(s)$ if $s_1 \in S_1$, and equal to 0 otherwise. Likewise, for each $s = (s_2, 2)$, $p_{M,x}(s)$ is equal to $(1 - \lambda)p_{2,x}(s)$ if $s_2 \in S_2$, and equal to 0 otherwise.

That is, this information structure results, with probability λ , in a posterior associated with information structure p_1 , and with probability $1 - \lambda$ in a posterior associated with information structure p_2 . The distribution of posteriors under the mixed information structure is a convex combination of the distributions of posteriors under the two original information structures, as if the DM flipped a coin, observed the result, and then randomly chose one of the two information structures. The mixture feasibility condition requires that choosing a mixed information structure costs no more than the cost of randomizing over information structures (using a mixed strategy in the rational inattention problem).

Condition 2. Given two information structures, $\{p_{1,x}\}_{x \in X}$, with signal alphabet S_1 , and $\{p_{2,x}\}_{x \in X}$, with alphabet S_2 , the cost of the mixed information structure is weakly less than the weighted average of the cost of the separate information structures:

$$C(p_M, q; S_M) \leq \lambda C(p_1, q; S_1) + (1 - \lambda)C(p_2, q; S_2).$$

The next condition uses Blackwell's ordering. Consider two signal structures, $\{p_x\}_{x \in X}$,

with signal alphabet S , and $\{p'_x\}_{x \in X}$, with alphabet S' . The first information structure Blackwell dominates the second information structure if, for all utility functions $u(a, x)$ and all priors $q \in \mathcal{P}(X)$,

$$\bar{u}(p, q) \geq \bar{u}(p', q).$$

If one information structure Blackwell dominates another, it is weakly more useful for every decision maker, regardless of that decision maker's utility function and prior. In this sense, it conveys weakly more information. This ordering is incomplete; most information structures neither dominate nor are dominated by a given alternative information structure. However, when an information structure does Blackwell dominate another one, we assume that the dominant information structure is weakly more costly.

Condition 3. If the information structure $\{p_x\}_{x \in X}$ with signal alphabet S is more informative, in the Blackwell sense, than $\{p'_x\}_{x \in X}$, with signal alphabet S' , then, for all $q \in \mathcal{P}(X)$,

$$C(\{p_x\}_{x \in X}, q; S) \geq C(\{p'_x\}_{x \in X}, q; S').$$

The first three conditions are, from a certain perspective, almost innocuous. For any joint distribution of actions and states that could have been generated by a DM solving a rational inattention type problem, with an arbitrary information cost function, there is a cost function consistent with these three conditions that also could have generated that data (Theorem 2 of Caplin and Dean (2015)). The result arises from the possibility of the DM pursuing mixed strategies over information structures, or in the mapping between signals and actions. These conditions also characterize “canonical” rational inattention cost functions, in the terminology of De Oliveira et al. (2017).

The mixture feasibility condition (Condition 2) and Blackwell monotonicity condition (Condition 3) are equivalent to requiring that the cost function be convex over information

structures and Blackwell monotone.

Lemma 2. *Let p and p' be information structures with signal alphabet S . A cost function is convex in information structures if, for all $\lambda \in (0,1)$, all signal alphabets S , and all $q \in \mathcal{P}(X)$,*

$$C(\lambda p + (1 - \lambda)p', q; S) \leq \lambda C(p, q; S) + (1 - \lambda)C(p', q; S).$$

A cost function satisfies mixture feasibility and Blackwell monotonicity (Conditions 2 and 3) if and only if it is convex in information structures and satisfies Blackwell monotonicity.

Proof. See the appendix, section D.3. □

The fourth condition that we assume, which is not imposed by Caplin and Dean (2015), Caplin et al. (2017), or De Oliveira et al. (2017), is a differentiability condition that will allow us to characterize the local properties of our cost functions.

Condition 4. For all signal alphabets S , in a neighborhood around any uninformative information structure, the information cost function is continuously twice-differentiable in information structures $\{p_x\}_{x \in X}$, in all directions that do not change the support of the signal distribution, and directionally differentiable, with continuous directional derivatives, with respect to perturbations that increase the support of the signal distribution. The information cost function is also Lipschitz-continuous in q .

While this may seem a relatively innocuous regularity condition, it is not completely general; for example, it rules out the case in which the DM is constrained to use only signals in a parametric family of probability distributions, and the cost of other information structures is infinite. Thus it rules out information structures of the kind assumed in Fudenberg et al. (2015) or Morris and Strack (2017). Condition 4 also rules out other proposed alternatives, such as the channel-capacity constraint suggested by Woodford (2012).

The next condition that we assume, which is also not imposed by Caplin and Dean (2015), Caplin et al. (2017), or De Oliveira et al. (2017), is a sort of local strong convexity. We will assume that the cost function exhibits strong convexity, in the neighborhood of an uninformative information structure, with respect to information structures that hold fixed the unconditional distribution of signals, uniformly over the set of possible priors.

Condition 5. There exists constants $m > 0$ and $B > 0$ such that, for all priors $q \in \mathcal{P}(X)$, and all information structures that are sufficiently close to uninformative ($C(p, q; \mathcal{S}) < B$),

$$C(p, q; \mathcal{S}) \geq \frac{m}{2} \sum_{s \in \mathcal{S}} (e_s^T p q) \|q_s - q\|_X^2,$$

where q_s is the posterior given by Bayes' rule and $\|\cdot\|_X$ is an arbitrary norm on the tangent space of $\mathcal{P}(X)$.

This condition is slightly stronger than Condition 1; it is essentially an assumption of “local strong convexity” instead of merely local convexity. It implies that all informative information structures have a strictly positive cost, and that (regardless of the DMs' current beliefs) there are no informative information structures that are “almost free.”

The mutual-information cost function (10) satisfies each of these five conditions. However, it is not the only cost function to do so. For example, we can construct a family of such cost functions, using the family of “f-divergences,” defined as

$$D_f(q_s \| q) = \sum_{x \in X} (e_x^T q) f\left(\frac{e_x^T q_s}{e_x^T q}\right),$$

where f is any strictly convex, twice-differentiable function with $f(1) = f'(1) = 0$ and $f''(1) = 1$. (The KL divergence is a member of this family, corresponding to $f(u) = u \ln u -$

$u + 1$.) For any divergence in this family, we can define an information cost function

$$I_f(p, q; S) = \sum_{s \in S} (e_s^T p q) D_f(q_s || q). \quad (13)$$

(When I_f is the KL divergence, this is just mutual information.) It is relatively easy to observe that this family of information cost functions satisfies all five of the conditions described above.²⁸

As another example of a class of cost functions that satisfy the conditions, we can establish the following.

Corollary 1. *Under the assumptions of Theorem 1, the posterior-separable cost function (12) that defines the equivalent static rational inattention problem satisfies Conditions 1-5.*

This follows from the form of the cost function and Lemma 3, described in the next section.

We are now in position to discuss our approximation of the information cost function. We use Taylor’s theorem to approximate the cost function and its gradient up to order Δ (we use Δ because in section 6, we will be looking at small time intervals). We consider perturbations that, as above, preserve the support of the signal structure. As a result, this theorem should be interpreted as applying to “frequent but not very informative” signals, as opposed to “rare but informative” signals. We will discuss the latter type of signals shortly. The theorem is derived from the results of Chentsov (1982), which are discussed in appendix section C.1.

Theorem 2. *Suppose that an information structure $\{p_x\}_{x \in X}$, with signal alphabet S , is described by the equation*

$$p_x = r + \Delta^{\frac{1}{2}} \tau_x + o(\Delta^{\frac{1}{2}}),$$

where, for any $x \in X$ and any $\Delta \geq 0$, $e_s^T p_x \neq 0 \Rightarrow e_s^T r > 0$. Let $C(\cdot)$ be an information

²⁸This follows from Lemma 3 in the next section.

cost function that satisfies Conditions 1-4. Then, for Δ sufficiently small, the cost of this information structure can be written as

$$C(\{p_x\}_{x \in X}; q; \mathcal{S}) = \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \tau_{x'}^T g(r) \tau_x + o(\Delta),$$

where the matrix $k(q)$ is positive semi-definite and symmetric, and satisfies $k(q)\mathbf{1} = \mathbf{0}$.

If in addition the cost function satisfies Condition 5, Then there exists a constant $m_g > 0$ such that the difference between $k(q)$ and the inverse Fisher information matrix, $g^+(q)$, multiplied by that constant, is positive semi-definite: $k(q) - m_g g^+(q) \succeq \mathbf{0}$.

Proof. See the appendix, section C.1 and section D.4. □

Our re-use of the notation $k(q)$ is intentional— this matrix valued function, which is defined based on a local approximation of the cost function, will be the information cost matrix function defined in the continuous time model described in section 2. In the case of the mutual-information cost function, the matrix $k(q)$ is itself the inverse Fisher information matrix. Written in terms of the coordinate system used previously in the paper,²⁹

$$k(q) = g^+(q) = D(q) - qq^T.$$

In general, however, the matrix-valued function $k(q)$ is not the inverse Fisher information matrix, but rather an arbitrary matrix-valued function satisfying certain restrictions.

There are, in effect, two ways for a signal to contain a small amount of information, and different costs associated with these different types of signals. The results of Theorem 2 characterize, for any rational inattention cost function satisfying our conditions, the cost of receiving frequently, but relatively uninformative, signals. As Corollary 3 below demon-

²⁹This corresponds to the pseudo-inverse of the standard definition of the Fisher information matrix, if derivatives of smooth functions defined on the probability simplex are written in terms of the coordinates explained in footnote 12.

strates, the posteriors associated with these signals are close to the prior (order $\Delta^{\frac{1}{2}}$). We will discuss the cost of receiving a rare but informative signal below. Previewing the results of section 6, these two types of uninformative signals correspond, in the continuous-time limit, to the diffusion and jump components of the belief process.

The theorem substantially restricts the local structure of the cost of commonly occurring, but not particularly informative, signals, relative to the most general possible alternatives (which would not satisfy our conditions). Potential information structures $\{p_x\}_{x \in X}$ can be represented as vectors of dimension $N = (|S| - 1) \times |X|$. Under the assumptions of Condition 1, convexity, and Condition 4 (but not the Blackwell ordering condition, Condition 3), the cost function must locally resemble an inner product with respect to a positive semi-definite, $N \times N$ matrix. If we impose Condition 3 as well, the results of Theorem 2 show that we can restrict this matrix to the $k(q)$ matrix, an $|X| \times |X|$ matrix. If the DM were only allowed binary signals ($|S| = 2$), this restriction would be trivial. When the DM is allowed to contemplate more general information structures, the restriction is non-trivial.

Several authors (Caplin and Dean (2015); Kamenica and Gentzkow (2011)) have observed that it is easier to study rational inattention problems by considering the space of posteriors, conditional on receiving each signal, rather than space of signals. We can re-define the cost function using the posteriors and unconditional signal probabilities, rather than the prior and the conditional probabilities of signals. The results are described in Corollary 3 in Appendix section C.1. They are based on the matrix-valued function

$$\bar{k}(q) = D^+(q)k(q)D^+(q), \tag{14}$$

where $D^+(\cdot)$ is the pseudo-inverse of the diagonal matrix. In the case of the mutual infor-

mation cost function, the matrix $\bar{k}(q)$ is the Fisher information matrix,

$$\bar{k}(q) = g(q) = D(q)^+ - \iota \iota^T. \quad (15)$$

The Theorem 2 and Corollary 3 describe the costs of receiving frequent, but relatively uninformative signals. We next discuss the cost of receiving rare, but informative signals. These types of signals, in the limit that we discuss in section 6, will lead to jumps in beliefs. After we describe the cost of these signals, we will introduce a condition that ensures that jumps in beliefs are not part of an optimal evidence accumulation strategy.

Corollary 2. *Under the assumptions of Theorem 2, define the signal structure*

$$\hat{p} = \bar{p}_\Delta + \Delta \omega,$$

where p_Δ is a signal structure of the type described in Theorem 2, with $\lim_{\Delta \rightarrow 0^+} \bar{p}_\Delta = r \iota^T$, and $\sum_{s \in S} \omega e_x = 0$ for all $x \in X$, with $e_s^T \omega e_x \geq 0$ for all $s \in S$ such that $e_s^T \bar{p}_\Delta = 0$.

The cost of this information structure can be written in the form

$$\begin{aligned} C(p_n; q; S) &= \frac{1}{2} \Delta \sum_{s \in S: e_s^T r > 0} (e_s^T r)(q_s - q)^T \bar{k}(q)(q_s - q) \\ &+ \sum_{s \in S: e_s^T r = 0} (e_s^T \phi) D^*(q_s || q) + o(\Delta), \end{aligned}$$

where the divergence D^* is finite and twice-differentiable in its first argument for q' sufficiently close to q , with

$$\frac{\partial^2 D^*(r || q)}{\partial r^i \partial r^j} \Big|_{r=q} = \bar{k}(q). \quad (16)$$

Proof. See the appendix, section D.6. □

The divergence D^* represents the cost of acquiring an infrequent, but potentially in-

formative, signal. Naturally, if the signal is in fact not very informative, this cost must be closely related to the costs of other uninformative signals, which gives rise to the condition on the Hessian of the divergence. Note that the corollary requires that the cost is additive with respect to the other signals being received (at least up to order Δ). The result follows from the directional differentiability of the cost function with respect to signals that occur with zero probability, the continuity of that directional derivative, and invariance.

We now introduce the last condition we will impose on our cost functions. This condition, which is expressed in terms of the $\bar{k}(q)$ matrix-valued function and the divergence D^* , is a sufficient condition to ensure that the discrete-time models that we study in section 6 converge to the model with continuous sample paths (no jumps) described in section 2. The condition reflects an assumption that learning gradually over time, receiving frequent but never very informative signals, is less costly than receiving rare signals that lead to large changes in beliefs when they occur.

Condition 6. The matrix-valued function $\bar{k}(q)$ and divergence D^* associated with the cost function $C(p, q; S)$ satisfy, for all $q, q' \in \mathcal{P}(X)$ with $q' \ll q$,

$$D^*(q' || q) \geq \frac{1}{2}(q' - q)^T \left(\int_0^1 \bar{k}(sq' + (1-s)q) ds \right) (q' - q).$$

We will say that a cost function satisfying this condition exhibits a preference for gradual learning. We will call this preference “strict” if the inequality is strict for all $q' \neq q$. If the $\bar{k}(q)$ function is the Hessian of some generalized entropy function (see equation (11)), this condition is equivalent to requiring that

$$D^*(q' || q) \geq D_H(q' || q), \tag{17}$$

where D_H is the associated Bregman divergence. In the particular case of mutual informa-

tion, both D^* and D_H are the KL divergence, and the condition is (weakly) satisfied. It is also easy to construct cases in which it is strictly satisfied, as the example below shows.³⁰

Consider the family of information cost functions built from f-divergences defined in equation (13) above. All of the cost functions in this family resemble mutual information, to second order, in the sense defined by Corollary 3. Assuming that the posteriors induced by the information structure p and prior q , $\{q_s\}_{s \in S}$, are close to the prior q , and that the prior q is on the interior of the simplex,

$$I_f(p, q; S) \approx \frac{1}{2} \sum_{s \in S} (e_s^T p q) (q_s - q)^T D(q)^+ g^+(q) D(q)^+ (q_s - q). \quad (18)$$

In other words, in a sense that we show formally in section 6, all of these flow cost functions induce the same information-cost matrix function in the continuous-time problem.

However, all such functions do not induce the same divergence D^* . (Note that for this family, $D^* = D_f$.) Nonetheless, if $D_f \geq D_{KL}$ (which holds strictly, for example, in the case of the χ^2 -divergence), these cost functions will generate the same solution in the continuous-time problem: the solution to a static rational-inattention problem with the mutual-information cost function. Regardless of whether the f-divergence used to construct the flow cost function is the KL divergence or not, the KL divergence will appear in the solution to the continuous-time problem.³¹

We argued in section 2 that the inverse Fisher information matrix, when used as the

³⁰It is the assumption that the flow cost function in our dynamic evidence accumulation problem satisfies Condition 6 that allows us to avoid considering the possibility of Poisson jumps in the posterior belief state of the kind assumed by Che and Mierendorff (2016) and Zhong (2017) in the continuous-time model presented in section 2. Zhong (2017) presents conditions under which information accumulation with Poisson jumps can be optimal, but considers only posterior-separable flow cost functions of the form (12) based on a Bregman divergence, so that (17) holds with equality rather than an inequality. In this special case, in our framework jumps can also be among the optimal policies, but an equally good outcome can always be achieved by an information sampling strategy that involves no jumps, as we establish in section 6. When the inequality is instead strict, jumps cannot be optimal in the continuous-time limit of the kind of dynamic evidence accumulation problem considered in this paper.

³¹In fact, this result applies to the larger class of invariant divergences, which includes the f-divergences, and follows from Chentsov's theorem (see appendix section C.1).

information-cost matrix function, lacks certain desirable properties related to the distance between different states of the world. In the next section, we introduce a new family of cost functions which can induce information cost matrix functions that capture these notions. These information-cost matrix functions also satisfy equation (11), and therefore Theorem 1 applies. In the appendix, sections A.1 and A.2, we solve examples of the static model implied by Theorem 1 and compare it to the same static model with mutual information, illustrating why notions of the distance between states matter in economic applications.

5 Neighborhood-Based Cost Functions

Suppose that the state space X can be written as the union of a finite collection of “neighborhoods” $\{X_i\}$, and suppose furthermore that the state space is connected, in the sense that any two states can be connected by a sequence of overlapping neighborhoods. That is, for any two states $x, x' \in X$, there exists a sequence of states $\{x_0, \dots, x_n\}$ with $x_0 = x, x_n = x'$, and the property that for any $1 \leq m \leq n$, states x_m and x_{m-1} belong to a common neighborhood. Define the selection matrices E_i as the $|X_i| \times |X|$ matrices that select each of the elements of X_i from a vector of length $|X|$.

For any prior $q \in \mathcal{P}(X)$, let $\mathcal{I}(q)$ be the (necessarily non-empty) set of neighborhoods X_i such that some state belonging to X_i has positive probability under the prior, and let $\bar{q}_i \equiv \sum_{x \in X_i} e_x^T q$ be the prior probability that some state belonging to neighborhood X_i occurs. Let $q_i \in \mathcal{P}(X_i)$ be the conditional probability distribution over states in neighborhood X_i , given the prior q and conditional on the state being in neighborhood X_i . That is, for all $x \in X_i$, $q_i \equiv \frac{1}{\bar{q}_i} E_i q$.

Similarly, let $q_s \in \mathcal{P}(X)$ be the posterior after receiving signal $s \in S$, and let $q_{i,s} \in \mathcal{P}(X_i)$ be the posterior over states in neighborhood X_i , conditional on receiving signal s and having the state be part of neighborhood X_i . That is, for all $x \in X_i$, $q_{i,s} \equiv \frac{1}{\bar{q}_{i,s}} E_i q_s$, with

$\bar{q}_{i,s} \equiv \sum_{x \in X_i} e_x^T q_s$. We adopt the convention that $q_{i,s} = q_i$ if $\bar{q}_{i,s} = 0$. Finally, let $\bar{p}_i \in \mathcal{P}(S)$ be the conditional distribution of signals under the information structure p and prior q , $\bar{p}_i = \frac{1}{\bar{q}_i} \sum_{x \in X_i} p e_x e_x^T q$.

We will say that a cost function has a ‘‘neighborhood structure’’ if it can be written in the form

$$C_N(p, q; S) = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S} e_s^T \bar{p}_i D_i(q_{i,s} || q_i), \quad (19)$$

where for each $i \in \mathcal{I}(q)$, $D_i(\cdot || \cdot)$ is a divergence (not necessarily the same for all i) defined over probability distributions in $\mathcal{P}(X_i)$ that is a twice-differentiable and strongly convex in its first argument.³² Mutual information is an example of a flow cost function in this family, corresponding to the case in which there is only a single neighborhood, consisting of the entire state space X , and the divergence is the KL divergence, so that

$$C(p, q; S) = \sum_{s \in S} (e_s^T p q) D_{KL}(q_s || q) = I^{Shannon}(\{p_x\}, q; S).$$

The information cost functions based on f-divergences, defined by (13), are also single-neighborhood examples of neighborhood-based cost functions.

The following lemma shows that all cost functions with a neighborhood structure satisfy the conditions defined in section 4.

Lemma 3. *All cost functions with a neighborhood structure (19) satisfy Conditions 1-4 stated in section 4. If the neighborhood structure includes a neighborhood containing all of the states $x \in X$, the cost function also satisfies Condition 5.*

Proof. See the appendix, section D.7. □

An implication of this lemma is that any posterior-separable cost function (12) based on a strongly convex generalized entropy function H satisfies Conditions 1-5. Below, we give a

³²The f-divergences defined previously satisfy these conditions (Amari and Nagaoka (2007)).

sufficient condition for Condition 6 to be satisfied as well.

We will study a particular family of cost functions with a neighborhood structure, the “neighborhood-based cost functions.” This family is defined by the additional requirements that (i) the divergences D_i be invariant (a term defined in appendix section C.1, which applies to all of the f-divergences defined previously), and (ii) each of the D_i is bounded below by some positive multiple of D_{KL} , the Kullback-Leibler divergence.³³ As an example of the possibility of satisfying these latter requirements, the D_i may be α -divergences (or Rényi divergences, van Erven and Harremoës (2014)) of order $\alpha \geq 1$:³⁴

$$D_\alpha(p_i||q_i) \equiv \frac{1}{\alpha - 1} \log \sum_{x \in X_i} \frac{p_i(x)^\alpha}{q_i(x)^{\alpha-1}}.$$

This family can have complex neighborhood structures, for which the requirement that each of the individual divergences D_i be invariant is a less restrictive requirement. The idea of this class of cost functions is that information structures are costly only to the extent that they result in different signal distributions for states that are “similar” to one another, in the sense of belonging to the same neighborhood. If there is only one neighborhood that includes all of the states (the mutual-information case), all states are equally difficult to distinguish from one another. Allowing for more complex neighborhood structures allows us to assume instead that it is much more difficult to tell some pairs of states apart than others. Note that under the general formalism (19), this is true not only because some pairs of states share a neighborhood while others do not — and more generally, that the length of the chain of neighborhoods required to link two states differs for different pairs of states

³³Stipulation (ii) is added in order to ensure that Condition 6 is satisfied. For this it suffices that D_i be bounded below by a Bregman divergence for each i . But as explained in section appendix C.1, any invariant divergence is locally equivalent (for p near q) to a positive multiple of D_{KL} . Hence in order for D_i to be bounded below by a Bregman divergence, it must be bounded below by a positive multiple of D_{KL} .

³⁴The definition is here stated only for the case $\alpha \neq 1$. When $\alpha = 1$, the α -divergence is simply the KL divergence, and Condition 6 is weakly satisfied. If $\alpha > 1$, the α -divergence satisfies $D_\alpha(p||q) > D_{KL}(p||q)$ for all $p \neq q$, so that the strong form of Condition 6 is satisfied, implying a strict preference for gradual learning.

— but also because the divergences D_i can be different for different neighborhoods.

By the results of Chentsov (1982), the fact that D_i is an invariant divergence implies that its Hessian matrix is proportional to the Fisher information matrix. As a result, the approximation described in equation (18) applies, but only within each neighborhood. That is,

$$C_N(p, q; S) \approx \frac{1}{2} \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in S} (e_s^T p q) (q_{i,s} - q_i)^T g(q_i) (q_{i,s} - q_i), \quad (20)$$

where the $c_i > 0$ are positive constants. This implies the following structure for the information-cost matrix:

Lemma 4. *The information-cost matrix function $k_N(q)$ associated with the neighborhood-based cost function is*

$$k_N(q) = \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i E_i^T g^+(q_i) E_i,$$

where g^+ is the inverse Fisher information matrix and the constant $c_i > 0$ for each i .

Proof. See the appendix, section D.8. □

We can use the information-cost matrix function in our continuous-time problem (the problem defined in section 2).³⁵ It satisfies the equation necessary for the results of Theorem 1 to apply (equation (11)). As a result, there is a generalized entropy function, $H_N(q)$, associated Bregman divergence, $D_N(p||q)$, and posterior-separable static information-cost function, $C_N^{static}(p, q; S)$, that can be used to define the static rational-inattention problem the choice probabilities of which coincide with the solution to the dynamic model. The following lemma describes these functions:

³⁵Our derivation of the continuous-time model from the discrete-time model applies only to cost functions satisfying Conditions 1-6. We have established these conditions only for neighborhood structures that include a neighborhood containing all states. However, the constant c_i associated with this neighborhood can be arbitrarily small, and in what follows we will ignore this requirement.

Lemma 5. Let $H^{\text{Shannon}}(q)$ be Shannon's entropy (8). The generalized entropy function $H_N(q)$ and Bregman divergence $D_N(q_s||q)$ associated with the neighborhood-based information-cost matrix function $k_N(q)$ are

$$H_N(q) = - \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i H^{\text{Shannon}}(q_i),$$

$$D_N(q_s||q) = \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_{i,s} D_{KL}(q_{i,s}||q_i).$$

The posterior-separable static information cost function derived from the neighborhood-based generalized entropy can then be written as

$$\begin{aligned} C_N^{\text{static}}(p, q; \mathcal{S}) &= \sum_{s \in \mathcal{S}} (e_s^T p q) D_N(q_s||q), \\ &= \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in \mathcal{S}} \bar{p}_{i,s} D_{KL}(q_{i,s}||q_i) \\ &= \sum_{i \in \mathcal{I}(q)} c_i \sum_{x \in X_i} (e_x^T q) D_{KL}(p e_x || p E_i^T q_i). \end{aligned}$$

Proof. See the appendix, section D.9. □

The fact that $\bar{k}_N(q)$ is the Hessian of a convex function $H_N(q)$ means that we can apply the sufficient condition (17) in order to verify that Condition 6 is satisfied by any cost function in this family. For a flow cost function of the form (19), the marginal cost of increasing the probability of a jump to an arbitrary posterior q' is given by the divergence $D_N^*(q'||q) = \sum_{i \in \mathcal{I}(q)} \bar{q}'_i D_i(q'_i||q_i)$. Under our assumption that D_i is bounded below by $c_i D_{KL}$ for each i , it follows that $D_N^*(q'||q) \geq D_N(q'||q)$, and condition (17) is verified. If for each i , D_i is a positive multiple of an α -divergence with $\alpha_i > 1$, then (17) holds with a strict inequality for any $q' \neq q$. In this case, the cost function satisfies the strong form of Condition 6, so that there is a strict preference for gradual learning.

Lemma 5 allows us to write the static rational inattention problem (Theorem 1) directly in terms of an optimization over choice probabilities $\{\pi_x\}$ so as to maximize

$$\sum_{x \in X} e_x^T q_0 \sum_{a \in A} e_a^T \pi_x u_{x,a} - \theta C(\{\pi_x\}_{x \in X}, q_0; A). \quad (21)$$

As discussed previously, in the special case in which there is only a single neighborhood, this is the standard rational inattention problem. The relevance of alternative assumptions about the neighborhood structure is illustrated by the following result.

Lemma 6. *Consider a rational inattention problem (21) with a neighborhood-based information-cost function, and let x, x' be two states with the property that (i) $u_{a,x} = u_{a,x'}$ for all actions $a \in A$, and (ii) the set of neighborhoods $\{X_i\}$ such that $x \in X_i$ is the same as the set such that $x' \in X_i$. Then under the optimal policy, $\pi_x^* = \pi_{x'}^*$.*

Proof. The result follows directly from the problem in (21) and the alternative expression for the cost function in Lemma 5. □

The significance of Lemma 6 can be seen if we consider the predictions of rational inattention for a standard form of perceptual discrimination experiment, an application we describe in appendix section A.1. In these experiments, payments are based on correct and incorrect responses. As a result, two states in which the correct response is identical will (for a single-neighborhood cost function) have the same likelihood of a correct response. Experimental evidence (intuitively) shows that in some states it is more difficult to determine the correct response than in other states.

In appendix section A.2, we consider an alternative neighborhood structure, in which states can be totally ordered and it is costly to discriminate between neighboring states. We derive an infinite state limit, and show that in this limit the cost function is equivalent to one based on Fisher information, an alternative measure of information costs that penalizes

sharply discriminating between nearby states. We connect our result to the work of Morris and Yang (2016), who show that this property is related to uniqueness in global games. We also note that, like mutual information, Fisher information is a cost function that can be applied in many contexts and has only a single free parameter.

We believe that these results show the usefulness of our continuous-time model of evidence accumulation as a micro-foundation for interesting classes of static rational-inattention problems, with properties that are relevant for economic applications. It remains for us to explain the justification for our proposed formulation of the continuous-time model of evidence accumulation itself.

6 Derivation of the Continuous-Time Model

We now show how the continuous-time model of section 2 can be obtained as the limit of a discrete-time model of sequential evidence accumulation, with a sequence of endogenous signals as in dynamic rational inattention models like that of Steiner et al. (2017), and an information cost function for each of the individual signals that satisfies the properties proposed for flow cost functions in section 4. In particular, we will justify the link proposed above between a second-order approximation to the information cost function for an individual signal and the information-cost matrix function defined in section 2.

We study a dynamic problem in which the DM has repeated opportunities to gather information before making a decision. The state of the world, $x \in X$, remains constant over time. At each time t , the DM can either stop and take an action $a \in A$, or continue and receive a signal drawn from the information structure $\{p_{t,x} \in \mathcal{P}(S)\}_{x \in X}$, for some signal alphabet S . We assume that the number of potential actions is weakly less than the number of states, $|A| \leq |X|$.

We also assume that the signal alphabet S is finite and fixed over time, with $|S| \geq$

$2|X| + 1$. However, the information structure $\{p_{t,x}\}_{x \in X}$ is a choice variable that can be state- and time-dependent. Fixing the signal alphabet S has no economic meaning, because the information content of receiving a particular signal $s \in S$ can change between periods. The assumption allows us to assume a finite information structure and invoke the results from section 4.³⁶ As a technical device, we assume that S contains one signal, \bar{s} , that is required to be uninformative. This assumption is a technical device to ensure that the DM can choose to mix any arbitrary signal structure with an uninformative one, even if she has already used up her “useful” signals.

The DM’s prior beliefs at time t , before receiving the signal, are denoted q_t . Each time period has a length Δ . Let τ denote the time at which the DM stops and makes a decision, with $\tau = 0$ corresponding to making a decision without acquiring any information. At this time, the DM receives utility $u(x, a) - \kappa\tau$ if she takes action a at time τ and the true state of the world is x . As in the previous sections, let $\hat{u}(q_\tau)$ be the utility (not including the penalty for delay) associated with taking an optimal action under beliefs q_τ . The parameter κ governs the size of the penalty the DM faces from delaying his decision. The reason the DM does not make a decision immediately is that she is able to gather information, and make a more-informed decision. The setup thus far is essentially identical to the continuous-time model described previously.

The DM can choose an information structure that depends on the current time and past history of the signals received. As we will see, the problem has a Markov structure, and the current time’s “prior,” q_t , summarizes all of the relevant information that the DM needs

³⁶As mentioned previously, the work of Ay et al. (2014) discusses how to extend the Chentsov (1982) theorems to infinite-dimensional structures. We conjecture that their results would allow us to extend our theorems to infinite signal spaces, but do not attempt such an extension here.

to design the information structure. The DM is constrained to satisfy

$$E_0\left[\frac{\Delta}{\rho} \sum_{j=0}^{\tau\Delta^{-1}-1} C(\{p_{\Delta j}, q_{\Delta j}; S\})^\rho\right]^{\frac{1}{\rho}} \leq \Delta c E_0[\tau], \quad (22)$$

if the DM choose to acquire any information at all ($\tau > 0$ always in this case). In words, the L^ρ -norm of the flow information cost function $C(\cdot)$ over time and possible histories must be less than the constant c per unit time.

In the limit as $\rho \rightarrow \infty$, this would approach a per-period constraint on the amount of information the DM can obtain. For finite values of ρ , the DM can allocate more information gathering to states and times in which it is more advantageous to gather more information. We assume, however, that $\rho > 1$, to ensure that it is optimal for the DM to gather information gradually, rather than all at once.³⁷ We also assume that the flow cost function $C(\cdot)$ satisfies Conditions 1-6 stated in section 4.

Let $V(q_0; \Delta)$ denote the value obtained in the sequence problem for a DM with prior beliefs q_0 , and let q_τ denote the DM's beliefs when stopping to act. The DM's problem is

$$V(q_0; \Delta) = \max_{\{p_{\Delta j}\}, \tau} E_0[\hat{u}(q_\tau) - \kappa\tau],$$

subject to the information-cost constraint (22). The dual version of this problem is

$$W(q_0, \lambda; \Delta) = \max_{\{p_{\Delta j}\}, \tau} E_0[\hat{u}(q_\tau) - \kappa\tau] - \lambda E_0\left[\Delta^{1-\rho} \sum_{j=0}^{\tau\Delta^{-1}-1} \left\{\frac{1}{\rho} C(p_{\Delta j}, q_{\Delta j}; S)^\rho - \Delta^\rho c^\rho\right\}\right]. \quad (23)$$

Here, the function $W(q_0, \lambda; \Delta)$ can be thought of as the value function of a different prob-

³⁷Our assumption of $\rho > 1$ is similar to the convex cost of the rate of experimentation assumed by Moscarini and Smith (2001).

lem, in which there is a cost of gathering information proportional to $\lambda \frac{1}{\rho} C(\cdot)^\rho$. We guess and verify that $\lambda \in (0, \kappa c^{-\rho})$. In our proofs, we demonstrate that there is no duality gap in the continuous time limit of this problem, and that our assumption about λ is without loss of generality. We describe our result below, and outline the proof in appendix section C.2.

Theorem 3. *Let $n \in \mathbb{N}$ index a sub-sequence of policies described in Lemma 10. There exists a $\lambda^* \in (0, \kappa c^{-\rho})$ such that*

$$\lim_{n \rightarrow \infty} W(q_t, \lambda^*; \Delta_n) = \lim_{n \rightarrow \infty} V(q_0; \Delta_n) = V(q_0),$$

where $V(q_0)$ is the solution to the continuous-time problem described in section 2, with $\chi = \rho^{\rho-1} c$ and $\mu = \kappa$. There exists a sequence of policies in the discrete-time models that achieve, in the limit, the value function $V(q_0)$ and for which the associated belief process, $q_{t,n}$, and stopping time τ_n converges in law to a belief process q_t^* and stopping time τ^* that are induced by an optimal policy in the continuous-time model (and hence q_t^* is a diffusion). If the cost function exhibits a strict preference for gradual learning, every convergent sub-sequence of belief processes $q_{t,n}^*$ associated with optimal policies in the discrete-time model converges in law to a diffusion.

Proof. See the appendix, section C.2 and section D.16. □

We have shown that the DM's behavior in the continuous-time problem can be thought of as an approximation of her behavior in discrete-time problems with flow cost functions drawn from a very general class. These convergence results can be viewed as offering a sort of micro-foundation for the continuous-time model, and in particular for our assumptions in section 2 about the information-cost matrix function.

7 Conclusion

We have derived a continuous-time rational-inattention model as the limit of a discrete-time sequential evidence accumulation problem. In the limit of a very large number of successive signals, each of which is only minimally informative, only the local properties of the flow cost function matter. These properties can be summarized by a matrix-valued function defined on the space of possible posterior beliefs, which we call the information-cost matrix function. This function summarizes the degree to which it is costly to further distinguish between different pairs of possible states, given the decision maker's current beliefs.

For a broad class of information-cost functions, we are able to solve the resulting continuous time problem, and show that the solution is equivalent to the solution of a static rational-inattention problem, with a particular posterior-separable information-cost function. The use of posterior-separable cost functions in static rational-inattention problems can thus be justified as summarizing the implications of a dynamic evidence accumulation process. Among the static cost functions that can be justified in this way is the mutual-information cost function proposed by Sims and the neighborhood-based cost function that we introduce. Unlike mutual information, the neighborhood-based cost function incorporates the idea that “nearby” states are more difficult to distinguish from one another. We argue that this property is appealing in the context of both perceptual experiments and economic applications.

References

- Shun-ichi Amari and Hiroshi Nagaoka. *Methods of information geometry*, volume 191. American Mathematical Soc., 2007.
- Nihat Ay, Jürgen Jost, Hông Vân Lê, and Lorenz Schwachhöfer. Information geometry and sufficient statistics. *Probability Theory and Related Fields*, pages 1–38, 2014.
- Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.
- Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Unpublished manuscript*, July 2017.
- Yeon-Koo Che and Konrad Mierendorff. Optimal sequential decision with limited attention. *Unpublished manuscript*, 2016.
- Nikolai Nikolaevich Chentsov. *Statistical decision rules and optimal inference*. Number 53. American Mathematical Soc., 1982.
- John A. Clithero. Improving out-of-sample predictions using response times and a model of the decision process. *Unpublished manuscript*, June 2016.
- Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
- Henrique De Oliveira, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek. Rationally inattentive preferences and hidden information costs. *Theoretical Economics*, 12:621–624, 2017.
- Ambuj Dewan and Nate Neligh. Estimating information cost functions in models of rational inattention. *Unpublished manuscript*, January 2017.
- Ernst Fehr and Antonio Rangel. Neuroeconomic foundations of economic choice — recent advances. *Journal of Economic Perspectives*, 25(4):3–30, 2011.
- Mogens Fosgerau, Emerson Melo, and Matthew Shum. Discrete choice and rational inattention: A general equivalence result. *Unpublished manuscript*, 2016.
- Drew Fudenberg, Philipp Strack, and Tomasz Strzalecki. Stochastic choice and optimal sequential sampling. *Unpublished manuscript*, 2015.
- Benjamin Hébert. Moral hazard and the optimality of debt. *Unpublished manuscript*, 2014.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- Ian Krajbich, Bastiaan Oud, and Ernst Fehr. Benefits of neuroeconomics modeling: New policy interventions and predictors of preference. *American Economic Review*, 104(5):501–506, 2014.
- Filip Matějka, Alisdair McKay, et al. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98, 2015.
- Stephen Morris and Philipp Strack. The Wald problem and the equivalence of sequential sampling and static information costs. *Unpublished manuscript*, June 2017.
- Stephen Morris and Ming Yang. Coordination and the relative cost of distinguishing nearby

- states. *Unpublished manuscript*, 2016.
- Giuseppe Moscarini and Lones Smith. The optimal level of experimentation. *Econometrica*, 69(6):1629–1644, 2001.
- Michael Shadlen and Daphna Shohamy. Decision making and sequential sampling from memory. *Neuron*, 90(5):927–939, 2016.
- Christopher A Sims. Rational inattention and monetary economics. *Handbook of Monetary Economics*, 3:155–181, 2010.
- Jakub Steiner, Colin Stewart, and Filip Matějka. Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553, 2017.
- Satohiro Tajima, Jan Drugowitsch, and Alexandre Pouget. Optimal policy for value-based decision-making. *Nature communications*, 7, 2016.
- Tim van Erven and Peter Harremoës. Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory*, 60:3797–3820, 2014.
- Michael Woodford. Inattentive valuation and reference-dependent choice. *Unpublished manuscript*, May 2012.
- Michael Woodford. An optimizing neuroeconomic model of discrete choice. Technical report, National Bureau of Economic Research, February 2014.
- Weijie Zhong. Optimal dynamic information acquisition. *Unpublished manuscript*, March 2017.

A Applications of Neighborhood-Based Cost Functions

A.1 Psychometric Functions

Suppose that the different states $X = \{1, 2, \dots, N\}$ represent different stimuli that may be presented to the subject, and that the subject is asked to classify the stimulus that is presented as one of two types (L or R); R is the correct answer if and only if $x > (N + 1)/2$. For example, the stimuli might be visual images with different orientations relative to the vertical, with increasing values of x corresponding to increasingly clockwise orientations; the subject is asked whether the image is tilted clockwise or counter-clockwise relative to the vertical. In such experiments, the subject's goal is often simply to give as many correct responses as possible; hence we suppose that $u_{x,a} = 1$ if $a = R$ and $x > (N + 1)/2$ or if $a = L$ and $x < (N + 1)/2$, while $u_{x,a} = 0$ in all other cases. We shall assume that each of the possible stimuli is presented with equal prior probability, and hence (assuming that N is odd) that both responses have an equal ex ante probability of being correct.

The standard theory of rational inattention, in which the static information cost is mutual information, corresponds to a special case of a neighborhood-based cost function, in which all states belong to the unique neighborhood. Hence condition (ii) of Lemma 6 holds for any pair of states. Lemma 6 thus implies that if any two states result in the same payoff regardless of the action chosen, the frequency with which different actions will be chosen under an optimal policy must be the same in the two states.

In the problem just posed, this implies that the probability of response R must be the same for all states $x < (N + 1)/2$, and also the same (but higher) for all states $x > (N + 1)/2$. Changing the severity of the information constraint changes the degree to which the probability of responding R is higher when $x > (N + 1)/2$, but it cannot change the prediction that the response probabilities should depend only on whether x is greater or less than $(N + 1)/2$. This is illustrated in Appendix Figure 1, which plots the optimal response

frequencies as a function of x , for alternative values of the cost parameter θ , in a numerical example in which C is given by mutual information and $N = 20$.

Alternatively, consider a posterior-separable neighborhood-based cost function in which the neighborhoods are given by

$$X_i = \{x_i, x_{i+1}\} \quad (24)$$

for $i = 1, 2, \dots, N - 1$. Thus two states belong to a common neighborhood if and only if they are either identical or one comes immediately after the other in the sequence. This captures the idea that the available measurement technologies all respond similarly in states that are “similar,” in the sense of being at nearby positions in the sequence, so that repeated measurements are necessary to reliably distinguish between two states if and only if they are near each other in the sequence. Suppose further that $c_i = 1$ for all i , implying that it is equally difficult to distinguish two neighboring states at all points in the sequence.³⁸ These assumptions suffice to completely determine a static information cost function (Lemma 5).

With this alternative neighborhood structure, Lemma 6 no longer requires that the response frequencies be identical for any two states. Moreover, because the cost function penalizes large differences in signal frequencies (and hence in response frequencies) in the case of neighboring states, in this case an optimal policy involves a gradual increase in the probability of response R as x increases, even though the payoffs associated with the different actions jump abruptly at a particular value of x . This is illustrated in Appendix Figure 2, which again shows the optimal response frequencies as a function of x , for alternative values of θ , in the case of the alternative neighborhood structure (24). The sigmoid functions predicted by rational inattention with this cost function — with the property that response frequencies differ only modestly from 50 percent when the stimuli are near the threshold of being correctly classified one way or the other, and yet approach zero or one in the case

³⁸If c_i is the same for all i , we can without loss of generality set it equal to one, as the multiplier θ can still be used to scale the overall magnitude of information costs.

of stimuli that are sufficiently extreme — are characteristic of measured “psychometric functions” in perceptual experiments of this kind.³⁹

The continuity of choice probabilities across points at which there are discrete changes in payoffs is also an important issue for the global games literature (Morris and Yang (2016)). However, this literature typically assumes a continuum of states, and many of the perceptual experiments that we have just referred to are naturally modeled with a continuum of states as well. In the next sub-section, we consider a continuous-state limit of the example just analyzed, that can be used as a model of imprecise perception in such examples.

A.2 Global Games and The Fisher-Information Cost Function

In this subsection, we continue our discussion of the neighborhood-based cost function proposed in the previous subsection, and consider the limit as the number of states of the world, $|X|$, becomes infinite. This example is motivated by the work of Yang (2015) and Morris and Yang (2016), who study global games (e.g. Morris and Shin (2001)) with endogenous information acquisition. However, we derive our limiting result for arbitrary action spaces and utility functions.

The result corresponds to a static rational inattention problem with a continuum of states, in which the information cost function is given by the average value of the Fisher information, a measure of the precision with which an information structure allows nearby

³⁹For the general concept of a psychometric function, see, for example, Gabbiani and Cox (2010), chap. 25, especially Figures 25.1 and 25.2, and discussion on p. 360; or Gold and Heekeren (2014), p. 356. For an example of an empirical psychometric function for the kind of task discussed in the text (classification of a field of moving dots as to which of two opposing directions is the dominant direction of motion), see Shadlen et al. (2007), Figure 10.1A. Note not only that the curve is monotonically increasing, with many data points corresponding to different response probabilities between zero and one, but also that in this experiment the subject’s reward function is clearly of the kind assumed in the text: only two possible reward levels (for correct vs. incorrect responses), with a discontinuous change in the reward where the sign of the “motion strength” changes from negative to positive.

states to be distinguished from each other (Cover and Thomas (2012)). Like Sims' proposal of a cost function proportional to Shannon's mutual information, the Fisher-information cost function is a single-parameter cost function, and it can also be applied in almost any context, as long as the state space is continuous. But unlike Shannon's mutual information, our measure of the informativeness of an information structure based on the Fisher information depends on the topological structure of the state space.

This is of considerable significance for the literature on global games. In the well-known analysis of Morris and Shin (2001), with exogenous private information, there is a unique equilibrium despite the incentives for coordination across DMs (subject to some caveats and details that are not relevant for our discussion). Instead Yang (2015) demonstrates that allowing for endogenous information acquisition, with mutual information as the information cost, restores a multiplicity of equilibria.

The key to Yang's result is that DMs can tailor the signals they receive to sharply discriminate between nearby states of the world, as discussed in our previous example. As a result, they can all coordinate their decision (say, to invest or not) on a particular threshold, and there are many such thresholds that can represent equilibria if coordinated upon. But this result depends on the fact that the mutual-information cost function does not make it costly to have abrupt changes in signal probabilities as the state of the world changes continuously. Morris and Yang (2016) develop the complementary result, showing that even in the case of an endogenous information structure, if signal probabilities must vary continuously with the state, there is again a unique equilibrium.

Here we show that a neighborhood-based cost function can provide a justification for the kind of continuity condition that the result of Morris and Yang (2016) requires. However, our results in the previous subsection cannot be applied directly to the model of Morris and Yang (2016), because the global games model in that paper assumes a continuum of states, whereas our analysis above supposes that $|X|$ is finite. To bridge this gap, we study

an example of the static model implied by Theorem 1 with a particular neighborhood-based cost function, and consider the limit as the number of states becomes unboundedly large. We show that the example model converges to a static rational inattention model with a particular cost function, similar in certain respects to the leading example of Morris and Yang (2016), that satisfies the continuous choice condition established by those authors.

For each of a sequence of values for the finite integer N , we assume a neighborhood structure of the kind discussed in the previous subsection for a model with $N + 1$ states. The set of states is ordered, $X^N = \{0, 1, \dots, N\}$, and each pair of adjacent states forms a neighborhood, $X_j = \{i, i + 1\}$, for all $j \in \{0, 1, \dots, N - 1\}$. We will also assume that there is an $N + 1$ st neighborhood containing all of the states. Note that N indexes both the number of states and the number of neighborhoods, which is always equal to the number of states. We consider the limit as $N \rightarrow \infty$.

To study this limit, we need to define how the initial beliefs, q_N , and the magnitude of the information costs vary with N . For the initial beliefs, we shall assume that there is a differentiable probability density function $f : [0, 1] \rightarrow \mathbb{R}^+$, with full support on $[0, 1]$, with a derivative that is Lipschitz continuous. Using this function, we define, for any $i \in X^N$,

$$e_i^T q_N = \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} f(x) dx.$$

That is, for each value of N , the prior q_N is assumed to be a discrete approximation to the Lipschitz-continuous p.d.f. $f(x)$, which becomes increasingly accurate as $N \rightarrow \infty$.

For our neighborhood structures, we assume that the constants associated with the cost of each neighborhood, c_j , are equal to N^2 for all $j < N$, and N^{-1} for $j = N$. In this particular example, the scaling ensures that the DM is neither able to determine the state with certainty, nor prevented from gathering any useful information, even as N is made arbitrarily large; moreover, the scaling ensures that the neighborhood containing all states

plays no role in the limiting behavior, so that in the limit all information costs are local. We also scale the entire cost function by a constant, $\bar{\theta} > 0$.

We also need to define the set of actions, and the utility from those actions. We will assume the set of actions, A , remains fixed as N grows, and define the utility from a particular action, in a particular state, as

$$e_i^T u_{a,N} = \frac{\int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} f(x) u_a(x) dx}{e_i^T q_N}.$$

Here, the utility $u_a : [0, 1] \rightarrow \mathbb{R}$ is a bounded measurable function for each action $a \in A$.⁴⁰ In other words, as N grows large, the prior converges to $f(x)$ and the utilities converge to the functions $u_a(x)$.

Under these assumptions, the static model of Theorem 1 can be written as

$$V_N(q_N; \bar{\theta}) = \max_{\pi_N \in \mathcal{P}(A), \{q_{a,N} \in \mathcal{P}(X^N)\}_{a \in A}} \sum_{a \in A} \pi_N(a) (u_{a,N}^T \cdot q_{a,N}) - \bar{\theta} \sum_{a \in A} \pi_N(a) D_N(q_{a,N} || q_N), \quad (25)$$

subject to the constraint that

$$\sum_{a \in A} \pi_N(a) q_{a,N} = q_N.$$

Here D_N denotes the divergence associated with the neighborhood-based cost function introduced above, specialized to the particular neighborhood structure of this section:

$$D_N(q_{a,N} || q_N) = N^2 \sum_{j \in \mathbf{X}^N \setminus \{N\}} \bar{q}_{j,a,N} D_{KL}(q_{j,a,N} || q_{j,N}) + N^{-1} D_{KL}(q_{a,N} || q_N).$$

The following theorem shows that the solution to this problem, both in terms of the

⁴⁰Note that we do not require the payoff resulting from a action to be a continuous function of x at all points, though it will be continuous almost everywhere. This allows for the possibility that a DM's payoffs change discontinuously when the state x crosses some threshold, as in the kind of equilibria discussed by Yang (2015).

value function and the optimal policies, converges to the solution of a static rational inattention problem with a continuous state space.

Theorem 4. *Consider the sequence of finite-state-space static rational inattention problems (25), with progressively larger state spaces indexed by the natural numbers N . Then there exists a sub-sequence of integers $n \in \mathbb{N}$ for which the solutions to the sub-sequence of problems converge, in the sense that*

$$i) \lim_{n \rightarrow \infty} V_n(q_n; \bar{\theta}) = V(q; \bar{\theta});$$

$$ii) \lim_{n \rightarrow \infty} \pi_n^* = \pi^*; \text{ and}$$

$$iii) \text{ for all } a \in A \text{ and all } x \in [0, 1], \lim_{n \rightarrow \infty} \sum_{i=0}^{\lfloor xn \rfloor} e_i^T q_{a,n}^* = \int_0^x f_a^*(y) dy.$$

Moreover, the limiting value function $V(q; \bar{\theta})$ is the value function for the following continuous-state-space static rational inattention problem:

$$\begin{aligned} V(f; \bar{\theta}) = & \sup_{\pi \in \mathcal{P}(A), \{f_a \in \mathcal{P}_{LipG}([0,1])\}_{a \in A}} \sum_{a \in A} \pi(a) \int_0^1 u_a(x) f_a(x) dx \\ & - \frac{\bar{\theta}}{4} \sum_{a \in A} \left\{ \pi(a) \int_0^1 \frac{(f'_a(x))^2}{f_a(x)} dx \right\} + \frac{\bar{\theta}}{4} \int_0^1 \frac{(f'(x))^2}{f(x)} dx, \end{aligned}$$

subject to the constraint that, for all $x \in [0, 1]$,

$$\sum_{a \in A} \pi(a) f_a(x) = f(x), \tag{26}$$

and where $\mathcal{P}_{LipG}([0, 1])$ denotes the set of differentiable probability density functions with full support on $[0, 1]$, whose derivatives are Lipschitz-continuous. Furthermore, the limiting action probabilities $\pi^*(a)$ and posteriors f_a^* are the optimal policies for the continuous-state-space problem.

Proof. See the appendix, section D.11. □

This theorem demonstrates that the value function, choice probabilities, and posterior beliefs of the discrete state problem converge to the value function, choice probabilities, and posterior beliefs associated with a continuous state problem. The continuous state problem uses a particular cost function, the expected value of the Fisher information $I^{Fisher}(x; p)$, defined locally for each element of the continuum of possible states x , with the expectation taken with respect to the prior over possible states.⁴¹ The posterior beliefs in the continuous state problem, $f_a(x)$, are required to be differentiable and have full support on $[0, 1]$, with a Lipschitz-continuous derivative. This is a result; the limiting posterior beliefs of the discrete state problem will have these properties. This restriction also ensures that the Fisher information is finite, so that the optimization associated with the continuous state problem is well-behaved.

This cost function, unlike mutual information, depends only on the degree to which the information structure allows states to be distinguished from ones extremely close to them (under the topology of the real line); and unlike the rational inattention problem based on mutual information, this static mutual information problem will generate the smoothness of responses across discrete changes in payoffs shown in Figure 2. For these reasons, we believe that the Fisher-information cost function is likely to be more appropriate than mutual information in a wide range of settings. It should also be noted that, as in the case of Sims' theory of rational inattention, the Fisher-information cost function has only a single degree of freedom. We thus obtain a rational inattention theory for problems with a continuous space that yields highly specific predictions, albeit different ones from Sims' theory.

The static rational inattention problem for the limiting case of a continuous state space can be given an alternative, equivalent formulation, in which the objects of choice are the

⁴¹This aggregate Fisher information has also proven useful in a variety of physics applications (Frieden (2004)).

conditional probabilities of taking different actions in the different possible states, rather than the posteriors associated with different actions.

Lemma 7. *Consider the alternative continuous-state-space static rational inattention problem:*

$$\bar{V}(f; \bar{\theta}) = \sup_{p \in \mathcal{P}_{LipG}(A)} \int_0^1 f(x) \sum_{a \in A} p_a(x) u_a(x) dx - \frac{\bar{\theta}}{4} \int_0^1 f(x) I^{Fisher}(x; p) dx,$$

where $\mathcal{P}_{LipG}(A)$ is the set of mappings $p : [0, 1] \rightarrow \mathcal{P}(A)$ such that for each action a , the function $p_a(x)$ ⁴² is a differentiable function of x with a Lipschitz-continuous derivative, and for any information structure $p \in \mathcal{P}_{LipG}(A)$, the Fisher information at state $x \in X$ is defined as

$$I^{Fisher}(x; p) \equiv \sum_{a \in A} \frac{(p'_a(x))^2}{p_a(x)}.$$

This problem is equivalent to the one defined in Theorem 4, in the sense that the information structure p^* that solves this problem defines action probabilities and posteriors

$$\pi^*(a) = \int_0^1 f(x) p_a^*(x), \quad f_a^*(x) = \frac{f(x) p_a^*(x)}{\pi^*(a)} \quad (27)$$

that solve the problem in Theorem 4, and conversely, the action probabilities and posteriors $\{\pi^*(a), f_a^*\}$ that solve the problem stated in the theorem define state-contingent action probabilities

$$p_a^*(x) = \frac{\pi^*(a) f_a^*(x)}{f(x)} \quad (28)$$

that solve the problem stated here. Moreover, the maximum achievable value is the same for both problems: $\bar{V}(f; \bar{\theta}) = V(f; \bar{\theta})$.

Proof. See the appendix, section D.12. □

⁴²Here for any $x \in [0, 1]$, we use the notation $p_a(x)$ to indicate the probability of action a implied by the probability distribution $p(x) \in \mathcal{P}(A)$.

We can apply this result to the problem considered in Morris and Yang (2016). Those authors study a global game with two possible actions, “invest” and “not-invest,” with equilibrium behavior characterized by a probability $s(x)$ of investing when the state is x . Their equilibrium uniqueness result depends on an assumption of continuous choice, meaning that for all $\bar{\theta} > 0$ and all parameterizations of the relevant utility function, $s(x)$ is absolutely continuous. Our Theorem 4 provides an example of more primitive assumptions that would guarantee continuous choice in this sense.

B Figures

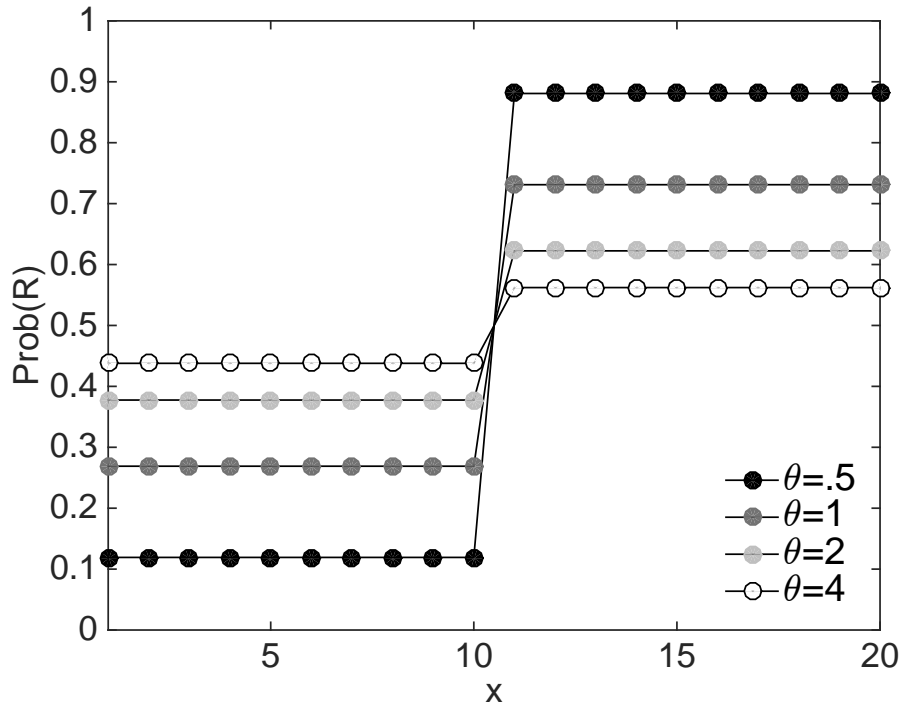


Figure 1: Predicted response probabilities with a mutual-information cost function, for alternative values of the cost parameter θ .

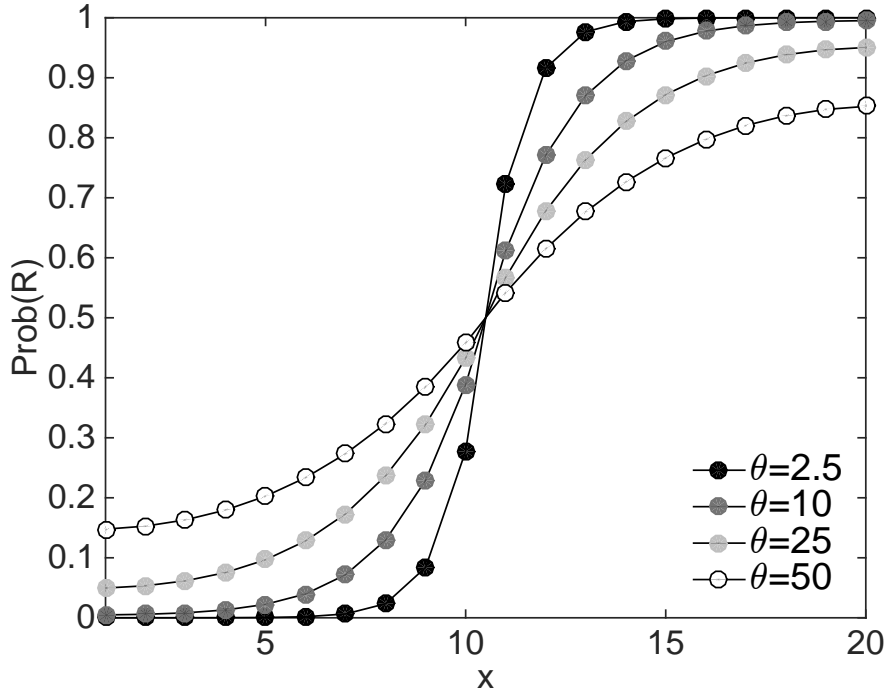


Figure 2: Predicted response probabilities with a neighborhood-based cost function, in which each neighborhood consists only of two adjacent states.

C Proof Outlines

C.1 Approximations of the Cost Function

We describe the local (second-order) properties of any information cost function satisfying our conditions. The condition requiring that Blackwell-dominant information structures cost weakly more (Condition 3) is of particular importance. To understand why, it is first useful to recall Blackwell's theorem.

Theorem. (Blackwell (1953)) *The information structure $\{p_x\}_{x \in X}$, with signal alphabet S , is more informative, in the Blackwell sense, than $\{p'_x\}_{x \in X}$, with signal alphabet S' , if and*

only if there exists a Markov transition matrix $\Pi : S \rightarrow S'$ such that, for all $s' \in S'$ and $x \in X$,

$$p'_x = \Pi p_x. \quad (29)$$

This Markov transition matrix is known as the “garbling” matrix. Another way of interpreting Condition 3 is that garbled signals are (weakly) less costly than the original signal.

There are certain kinds of garbling matrices that don’t really garble the signals. These garbling matrices have left inverses that are also Markov transition matrices. If we define an information structure $\{p_x\}_{x \in X}$, with signal alphabet S , and another information structure $\{p'_x\}_{x \in X}$, with signal alphabet S' , using one of these left-invertible matrices, via equation (29), then $\{p_x\}_{x \in X}$ is more informative than $\{p'_x\}_{x \in X}$, but $\{p'_x\}_{x \in X}$ is also more informative than $\{p_x\}_{x \in X}$. These two information structures are called “Blackwell-equivalent,” and it follows that the cost of these two information structures must be equal, by Condition 3. The left-invertible Markov transition matrices associated with Blackwell-equivalent information structures are called Markov congruent embeddings by Chentsov (1982). Chentsov (1982) studied tensors and divergences that are invariant to Markov congruent embeddings (we will say “invariant” for brevity).

An invariant divergence is a divergence that is invariant to these embeddings. Let Π be a Markov congruent embedding from $\mathcal{P}(S)$ to $\mathcal{P}(S')$. The KL divergence and the f -divergences more generally are invariant, meaning that

$$D_f(\Pi p || \Pi r) = D_f(p || r)$$

for all $p, r \in \mathcal{P}(S)$. There are also other, non-additively-separable invariant divergences.

Chentsov's theorem (Chentsov (1982)) states that, for any invariant divergence D_I ,

$$\frac{\partial^2 D_I(p||r)}{\partial p^i \partial p^j} \Big|_{p=r} = c \cdot g_{ij}(r), \quad (30)$$

where $c > 0$ is a positive constant and $g_{ij}(r)$ is the (i, j) -element of the Fisher information matrix evaluated at r .

However, the focus of this paper is not invariant divergences, but rather invariant information cost functions. By Condition 3, all information cost functions satisfying our conditions are invariant to Markov congruent embeddings. It necessarily follows that, for any Markov congruent embedding Π , that

$$C(\{p_x\}_{x \in X}, q; S) = C(\{\Pi p_x\}_{x \in X}, q; S').$$

Using this invariance, and results from Chentsov (1982), we will describe the local structure of all information cost functions satisfying our conditions.

Chentsov establishes the following results:⁴³

- i) Any continuous function that is invariant over the probability simplex is equal to a constant.
- ii) Any continuous, invariant 1-form tensor field over the probability simplex is equal to zero.
- iii) Any continuous, invariant quadratic form tensor field over the probability simplex is proportional to the Fisher information matrix.⁴⁴

⁴³See Lemma 11.1, Lemma 11.2, and Theorem 11.1 in Chentsov (1982). See also Proposition 3.19 of Ay et al. (2014), who demonstrate how to extend the Chentsov results to infinite sets X and S .

⁴⁴A 1-form tensor field on a probability simplex \mathcal{P} is a function $T : V \times \mathcal{P} \rightarrow \mathbb{R}$, where V is the tangent space of the simplex. Let $\Pi : \mathcal{P} \rightarrow \mathcal{P}'$ be a mapping from the simplex \mathcal{P} to the simplex \mathcal{P}' , let V' be the tangent space of the simplex \mathcal{P}' , and let $d\Pi : V \rightarrow V'$ be the pushforward of the mapping Π . The tensor field is invariant under Π if $T(d\Pi v, \Pi p) = T(v, p)$ for all $p \in \mathcal{P}$ and v in the tangent space at p , and a

These results allow us to characterize the local properties of rational inattention cost functions, via a Taylor expansion. Hold fixed the signal alphabet S , and consider an information structure $p_x(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = r + \boldsymbol{\varepsilon}\boldsymbol{\tau}_x + \boldsymbol{\nu}\boldsymbol{\omega}_x$, where $r \in \mathcal{P}(S)$. Here, $\boldsymbol{\tau}_x$ satisfies $\mathbf{1}^T \boldsymbol{\tau}_x = 0$ for all x , and, for all $s \in S$, $e_s^T \boldsymbol{\tau}_x \neq 0$ only if $e_s^T r > 0$. That is, $\boldsymbol{\tau}_x$ is an element of the tangent space of the probability simplex at r , and the same holds true for $\boldsymbol{\omega}_x$. As a result, for values of the perturbation parameters $\boldsymbol{\varepsilon}$ and $\boldsymbol{\nu}$ sufficiently close to zero, $p_x \in \mathcal{P}(S)$ for all $x \in X$. In other words, the parameters $\boldsymbol{\varepsilon}$ and $\boldsymbol{\nu}$ index a two-parameter family of perturbations of an uninformative information structure (corresponding to $\boldsymbol{\varepsilon} = \boldsymbol{\nu} = 0$), in which the perturbed information structures will generally be informative; the $\boldsymbol{\tau}_x$ and $\boldsymbol{\omega}_x$ specify two directions of perturbation. Each of the perturbed information structures has the property that p_x is absolutely continuous with respect to r .

By Condition 1, $C(\{p_x(0, 0)\}_{x \in X}; q; S) = 0$. The first order term is

$$\frac{\partial}{\partial \boldsymbol{\varepsilon}} C(\{p_x(\boldsymbol{\varepsilon}, \boldsymbol{\nu})\}_{x \in X}, q; S)|_{\boldsymbol{\varepsilon}=\boldsymbol{\nu}=0} = \sum_{x \in X} C_x(\{r\}_{x \in X}, q; S) \cdot \boldsymbol{\tau}_x,$$

where C_x denotes the derivative with respect to p_x . This derivative, $C_x(\{r\}; q; S)$, forms a continuous 1-form tensor field over the probability simplex $\mathcal{P}(S)$. By the invariance of $C(\cdot)$, it also follows that C_x is invariant, and therefore, by Chentsov's results, it is equal to zero.

We repeat the argument for the second derivative terms. Those terms can be written as

$$\frac{\partial}{\partial \boldsymbol{\nu}} \frac{\partial}{\partial \boldsymbol{\varepsilon}} C(\{p_x(\boldsymbol{\varepsilon}, \boldsymbol{\nu})\}_{x \in X}, q; S)|_{\boldsymbol{\varepsilon}=\boldsymbol{\nu}=0} = \sum_{x' \in X} \sum_{x \in X} \boldsymbol{\omega}_{x'}^T \cdot C_{xx'}(\{r\}_{x \in X}, q; S) \cdot \boldsymbol{\tau}_x.$$

By the invariance of $C(\cdot)$, the quadratic form $C_{xx'}(\cdot)$ is invariant for all $x, x' \in X$, and therefore is proportional to the Fisher information matrix for all $x, x' \in X$. We can define a matrix

similar definition holds for quadratic form tensor fields. This means that if the quadratic form is extended to a quadratic form over $\mathbb{R}^{|X|}$ using the method defined in footnote 11, its matrix representation is proportional to the matrix defined in (15).

$k(q)$ consisting of the constants of proportionality associated with each $x, x' \in X$. That is,

$$\frac{\partial}{\partial \mathbf{v}} \frac{\partial}{\partial \boldsymbol{\varepsilon}} C(\{p(\cdot|\cdot; \boldsymbol{\varepsilon}, \mathbf{v})\}, q)|_{\boldsymbol{\varepsilon}=\mathbf{v}=0} = \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \boldsymbol{\omega}_{x'}^T g(r) \boldsymbol{\tau}_x,$$

where $g(r)$ is the Fisher information matrix evaluated at the unconditional distribution of signals $r \in \mathcal{P}(S)$. We note that the matrix $k(q)$ can depend on the prior q , but cannot depend on the unconditional distribution of signals, r ; otherwise, invariance would not hold.

This matrix $k(q)$ is the matrix referenced in Theorem 2, and is (by the results of section §6) the information cost matrix function described in the continuous time model. The corollary below expresses the results of Theorem 2 in terms of posterior beliefs.

Corollary 3. *Under the assumptions of Theorem 2, the posterior beliefs can be written, for any $s \in S$ such that $e_s^T r > 0$, as*

$$q_{s,n,x} = q_x + \Delta^{\frac{1}{2}} q_x \frac{e_s^T \boldsymbol{\tau}_{n,x}}{e_s^T r} + o(\Delta^{\frac{1}{2}}).$$

The cost function can be written as

$$C(\{p_x\}_{x \in X}, q; S) = \frac{1}{2} \sum_{s \in S: e_s^T r > 0} (e_s^T r) (q_s - q)^T \bar{k}(q) (q_s - q) + o(\Delta).$$

Proof. See the appendix, section D.5. □

C.2 Convergence of Discrete to Continuous Time

We begin by describing the recursive representation for the value function $W(q_t, \lambda; \Delta)$, and discussing certain technical lemmas that are necessary to establish our main results. The

value function has a recursive representation:

$$W(q_t, \lambda; \Delta) = \max_{p_t} \left\{ \max_{p_t} -\kappa\Delta + \lambda\Delta^{1-\rho} (\Delta^\rho c^\rho - \frac{1}{\rho} C(p_t, q_t; S)^\rho) + \sum_{s \in S} (e_s^T p_t q_t) W(q_{t+\Delta, s}, \lambda; \Delta), \hat{u}(q_t) \right\},$$

where $q_{t+\Delta, s}$ is pinned down by Bayes' rule. In standard rational inattention problems, it is without loss of generality to equate signals and actions. In this problem, when the DM does not stop and make a decision, the “action” is updating one's beliefs. Rather than consider a probability distribution over signals, and then an updating of beliefs by Bayes' rule, one can consider the DM to be choosing a probability distribution over posteriors, subject to the constraint that the expectation of the posterior is equal to the prior.⁴⁵

To begin our analysis, we note that the value function $W(q_t, \lambda; \Delta)$ is well-behaved:

Lemma 8. *The value function $W(q_t, \lambda; \Delta)$ is bounded on $q_t \in \mathcal{P}(X)$, and convex in q . The optimal stopping time τ_Δ is bounded in expectation by a constant, $\bar{\tau}$, for all Δ :*

$$E_0[\tau_\Delta] \leq \bar{\tau}.$$

Proof. See the appendix, section D.13. □

The boundedness of the value function follows from the setup of the problem: ultimately, the DM will make a decision, and the utility from making the best possible decision in the best possible state of the world is finite. The convexity of the value function is what motivates the DM to acquire information. By updating her beliefs from q to either q' or q'' , with $q = \alpha q'' + (1 - \alpha)q'$ for some $\alpha \in (0, 1)$, the DM improves her welfare by enabling better decision making. That the optimal stopping time is bounded in expectation follows

⁴⁵The notion of choosing a probability distribution over posteriors appears in Kamenica and Gentzkow (2011), Caplin and Dean (2015), and Caplin et al. (2017), among other papers.

from an obvious point: waiting too long to make a decision will eventually become worse, even if the DM eventually makes the best possible decision, than making the worst possible decision immediately.

Next, we show that, because of the curvature (ρ) that we impose, the DM will choose, under any optimal policy, to gather only a small amount of information in each time period, as the length of each time period shrinks.

Lemma 9. *Let $n \in \mathbb{N}$ denote a sequence such that $\lim_{n \rightarrow \infty} \Delta_n = 0$. Any associated sequence of optimal policies $p_{t,n}^*$ satisfies, for all elements of the sequence,*

$$C(p_{t,n}^*, q_{t,n}; \mathcal{S}) \leq \left(\frac{\theta}{\lambda}\right)^{\frac{1}{\rho-1}} \Delta_n,$$

where $\theta = \lambda \left(\rho \frac{\kappa - \lambda c \rho}{\lambda(\rho-1)}\right)^{\frac{\rho-1}{\rho}}$.

Proof. See appendix, section D.14. □

The key step in proving this lemma is demonstrating that, as the time period shrinks, the optimal quantity of information acquired vanishes at a sufficiently fast rate. The convergence of the information structure to an uninformative one, as the time period shrinks, allows us to use the approximation described in Theorem 2 to study the continuous-time limit of the sequential evidence accumulation model. The assumption that $\rho > 1$ is critical to generating this result. When $\rho = 1$, the DM has no particular desire to smooth the quantity of information gathered over time, and might choose to gather a large quantity of information in a single period (as in Steiner et al. (2017)).

We next discuss the convergence of an arbitrary sequence of stochastic processes for beliefs (denoted $q_{t,m}$) and of stopping times (denoted τ_m) to their continuous-time limits, under the assumption that the policies generating them satisfy the bound in Lemma 9 and the bound on expected stopping times. This lemma applies to a sequence of optimal poli-

cies, but also to sequences of sub-optimal policies. The lemma describes the convergence of the beliefs process to a martingale, which is not necessarily a diffusion (it may have jumps, or even be a semi-martingale that is not a jump-diffusion).

Lemma 10. *Let Δ_m , $m \in \mathbb{N}$, denote a sequence such that $\lim_{m \rightarrow \infty} \Delta_m = 0$. Let $p_m(q)$ denote a sequence of Markov policies satisfying the bound in Lemma 9. Let $q_{t,m}$ denote the stochastic process for the DM's beliefs at time t , under such a policy, and let τ_m be a sequence of stopping policies such that $E_0[\tau_m] \leq \bar{\tau}$.*

There exists a sub-sequence $n \in \mathbb{N}$ and a probability space such that:

- i) The beliefs $q_{t,n}$ and the stopping time τ_n converge almost surely to a martingale q_t and a stopping time τ .*
- ii) The martingale q_t can be represented in terms of its semi-martingale characteristics,*

$$B_t = - \int_0^t \left(\int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) x dx \right) dA_s$$

$$C_t = \int_0^t D(q_{s^-}) \sigma_s \sigma_s^T D(q_{s^-}) dA_s$$

$$v_t(x) = dA_t \psi_t(x),$$

where σ_s is an $|X| \times |X|$ matrix-valued predictable stochastic process, satisfying $q_{s^-}^T \sigma_s = \vec{0}$, ψ_s is a measure on $\mathbb{R}^{|X|} \setminus \{0\}$ such that $q_{s^-} + x \in \mathcal{P}(X)$ and $q_{s^-} + x \ll q_{s^-}$ for all x in the support of ψ_s , and dA_s is the increment of a weakly increasing process.

- iii) For all stopping times T ,*

$$E_t \left[\int_t^T \left\{ \frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\} dA_s \right] \leq$$

$$\left(\frac{\theta}{\lambda} \right)^{\frac{1}{\rho-1}} E_t[T - t].$$

iv) *The limit of the cumulative information cost is bounded below,*

$$E_t \left[\int_0^\tau \left\{ \frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\}^\rho \left(\frac{dA_s}{ds} \right)^\rho ds \right] \geq \lim_{n \rightarrow \infty} E_0 \left[\int_0^{\tau_n} \Delta_n^{1-\rho} C(p_n(q_{t,n}), q_{t,n}; S)^\rho dt \right]$$

Proof. See the appendix, section D.15. □

In essence, the stochastic process $q_{t,n}$ converges to a jump-diffusion process. The semi-martingale characteristics, B_t, C_t, ν_t , summarize the DM's policy function. They have a representation as a function of σ_t, ψ_t, A_t because of the need for beliefs to remain the simplex, and the property that, once a state $x \in X$ has been assigned zero probability, it will be assigned zero probability forever after.

To finish the proof, we resolve several issues. We show that the constraint given in Lemma 9 binds. We show that the limiting value function W is unique, that duality holds ($V = W$ for a suitable choice of λ), and that the limit of V is the solution to the continuous-time problem described in section 2. We also show that there is a sequence of (possibly sub-optimal) policies in the discrete-time model that achieve, in the limit, the optimal utility and converge to a diffusion. Moreover, if the cost function $C(p, q; S)$ exhibits a strict preference for gradual learning (it satisfies Condition 6 strictly for $q' \neq q$), then all sequences of optimal policies converge to diffusions that are optimal policies of the continuous-time model.

D Proofs

D.1 Proof of Lemma 1

The problem in the continuation region is (everywhere the value function is twice differentiable)

$$\sup_{\sigma_t \in M(q_t)} \frac{1}{2} \text{tr}[\sigma_t^T D(q_t) V_{qq}(q_t) D(q_t) \sigma_t] = \kappa,$$

subject to

$$\frac{1}{2} \text{tr}[\sigma_t^T k(q_t) \sigma_t] \leq \chi.$$

First, suppose that the constraint does not bind and a maximizing optimal policy exists:

$$\frac{1}{2} \text{tr}[\sigma_t^{*T} k(q_t) \sigma_t^*] = a\chi,$$

where σ_t^* is a maximizer, for some $a \in [0, 1)$ ($a \geq 0$ by the positive semi-definiteness of $k(q_t)$). For any $c \in (1, a^{-1})$, with $a^{-1} = \infty$ for $a = 0$, if we used $\sigma_t = c\sigma_t^*$ instead, the policy would be feasible and we would have

$$\frac{1}{2} \text{tr}[\sigma_t^T D(q_t) V_{qq}(q_t) D(q_t) \sigma_t] = c^2 \kappa > \frac{1}{2} \text{tr}[\sigma_t^{*T} D(q_t) V_{qq}(q_t) D(q_t) \sigma_t^*] = \kappa,$$

a contradiction by the fact that $\kappa > 0$. Therefore, either the constraint binds under the optimal policy or an optimal policy does not exist. The latter would require that, for some vector $z \in \mathbb{R}^{|X|}$ with $zz^T \in M(q_t)$,

$$z^T D(q_t) V_{qq}(q_t) D(q_t) z > 0$$

and $z^T k(q_t) z = 0$, but the null space of $k(q_t)$ consists only of vectors whose elements are

constant over the support of q_t , and therefore satisfy $q^T z \neq 0$, implying that $zz^T \notin M(q_t)$. Therefore, the constraint binds.

Using θ as defined in the lemma, it must be the case (anywhere the DM chooses not to stop and the value function is twice differentiable) that

$$\sup_{\sigma_t \in M(q_t)} \frac{1}{2} \text{tr}[\sigma_t \sigma_t^T (D(q_t) V_{qq}(q_t) D(q_t) - \theta k(q_t))] = 0.$$

Because of the homogeneity assumption on V ,

$$q_t^T V_q(q_t) = V(q_t).$$

Differentiating again,

$$q_t^T V_{qq}(q_t) = 0.$$

It follows that, for any $\alpha \in \mathbb{R}$,

$$\begin{aligned} \frac{1}{2} \text{tr}[(\sigma_t \sigma_t^T + \alpha \mathbf{1} \mathbf{1}^T)(D(q_t) V_{qq}(q_t) D(q_t) - \theta k(q_t))] = \\ \frac{1}{2} \text{tr}[(\sigma_t \sigma_t^T)(D(q_t) V_{qq}(q_t) D(q_t) - \theta k(q_t))]. \end{aligned}$$

Suppose that we relax the requirement that $q_t^T \sigma_t = \vec{0}$ and simply require that σ_t be a square matrix. Let $\tilde{\sigma}_t$ be any square matrix. Setting

$$\alpha = -q_t^T \tilde{\sigma}_t \tilde{\sigma}_t^T q_t,$$

and performing an eigendecomposition,

$$VDV^T = \tilde{\sigma}_t \tilde{\sigma}_t^T + \alpha \mathbf{1} \mathbf{1}^T,$$

we construct a matrix

$$\sigma_t = VD^{\frac{1}{2}}$$

that achieves the same utility and satisfies $\sigma_t \in M(q_t)$. Therefore, ignoring this restriction is without loss of generality.

It immediately follows that, in the continuation region, the maximum eigenvalue of

$$D(q_t)V_{qq}(q_t)D(q_t) - \theta k(q_t)$$

must be equal to zero. If it were less than zero, we would always have

$$\frac{1}{2}tr[(\sigma_t \sigma_t^T)(D(q_t)V_{qq}(q_t)D(q_t) - \theta k(q_t))] < 0,$$

and if it were greater than zero, we could achieve

$$\frac{1}{2}tr[(\sigma_t \sigma_t^T)(D(q_t)V_{qq}(q_t)D(q_t) - \theta k(q_t))] > 0$$

by setting $\sigma_t = v_1 e_1^T$, where v_1 is an associated eigenvector of the maximal eigenvalue.

Finally, note that the DM would always choose to stop if $V(q_t) < \hat{u}(q_t)$, and therefore we must have $V(q_t) \geq \hat{u}(q_t)$. If $V(q_t) > \hat{u}(q_t)$, the DM must choose to continue, and (assuming twice-differentiability) the HJB equation must hold.

D.2 Proof of Theorem 1

Define $\phi(q_t)$ as the function described in the statement of the theorem (we will prove that it is indeed equal to $V(q_t)$, the value function of the dynamic problem). We will first show that $\phi(q_t)$ satisfies the HJB equation, can be implemented by a particular strategy for the DM, and that any other strategy for the DM achieves weakly less utility.

We begin by observing that

$$\iota^T k(q_t) D(q_t)^{-1} = 0 = \iota^T D(q_t) H_{qq}(q_t) = q_t^T H_{qq}(q_t).$$

We claim that, without loss of generality, we can assume that $H(q_t)$ is homogeneous of degree one,

$$H(\alpha q_t) = \alpha H(q_t)$$

for all $\alpha \in \mathbb{R}^+$ and $q_t \in \mathcal{P}(X)$. Differentiating with respect to α and then with respect to q_t , and evaluating at $\alpha = 1$, implies that

$$q_t^T H_{qq}(q_t) = 0,$$

consistent with the claim above.

We start by showing that the function $\phi(q_t)$ is twice-differentiable in certain directions.

The function is

$$\phi(q_0) = \max_{\pi \in \mathcal{P}(A), \{q_a \in \mathcal{P}(X)\}_{a \in A}} \sum_{a \in A} \pi(a) u_a^T \cdot q_a - \theta \sum_{a \in A} \pi(a) D_H(q_a || q_0),$$

subject to the constraint that

$$\sum_{a \in A} \pi(a) q_a = q_0.$$

Substituting the definition of the divergence, we can rewrite the problem as

$$\phi(q_0) = \max_{\pi \in \mathcal{P}(A), \{q_a \in \mathcal{P}(X)\}_{a \in A}} \sum_{a \in A} \pi(a) u_a^T \cdot q_a + \theta H(q_0) - \theta \sum_{a \in A} \pi(a) H(q_a),$$

subject to the same constraint. Define a new choice variable,

$$\hat{q}_a = \pi(a)q_a.$$

By definition, $\hat{q}_a \in \mathbb{R}_+^{|X|}$, and the constraint is $\sum_{a \in A} \hat{q}_a = q_0$. By the homogeneity of H , the objective is

$$\phi(q_0) = \max_{\pi \in \mathcal{P}(A), \{q_a \in \mathcal{P}(X)\}_{a \in A}, \{\hat{q}_a \in \mathcal{P}(X)\}_{a \in A}} \sum_{a \in A} u_a^T \cdot \hat{q}_a + \theta H(q_0) - \theta \sum_{a \in A} H(\hat{q}_a).$$

Any choice of \hat{q}_a satisfying the constraint can be implemented by some choice of π and q_a in the following way: set

$$\pi(a) = \iota^T \hat{q}_a,$$

and (if $\pi(a) > 0$) set

$$q_a = \frac{\hat{q}_a}{\pi(a)}.$$

If $\pi(a) = 0$, set $q_a = q_0$. By construction, the constraint will require that $\pi(a) \leq 1$, $\sum_{a \in A} \pi(a) = 1$, and the fact that the elements of q_a are weakly positive will ensure $\pi(a) \geq 0$. Similarly, $\iota^T q_a = 1$ for all $a \in A$, and the elements of q_a are weakly greater than zero. Therefore, we can implement any set of \hat{q}_a satisfying the constraints.

Rewriting the problem in Lagrangian form,

$$\begin{aligned} \phi(q_0) = & \max_{\{\hat{q}_a \in \mathbb{R}^{|X|}\}_{a \in A}} \min_{\kappa \in \mathbb{R}^{|X|}, \{v_a \in \mathbb{R}_+^{|X|}\}_{a \in A}} \sum_{a \in A} u_a^T \cdot \hat{q}_a + \theta H(q_0) \\ & - \theta \sum_{a \in A} H(\hat{q}_a) + \kappa^T (q_0 - \sum_{a \in A} \hat{q}_a) + \sum_{a \in A} v_a^T \hat{q}_a. \end{aligned}$$

We begin by observing that $\phi(q_0)$ is convex in q_0 . Suppose not: for some $q = \lambda q_0 + (1 - \lambda)q_1$, with $\lambda \in (0, 1)$, $\phi(q) < \lambda \phi(q_0) + (1 - \lambda)\phi(q_1)$. Consider a relaxed version of the

problem in which the DM is allowed to choose two different \hat{q}_a for each a . Observe that, because of the convexity of H , even with this option, the DM will set both of the \hat{q}_a to the same value, and therefore the relaxed problem reaches the same value as the original problem. However, in the relaxed problem, choosing the optimal policies for q_0 and q_1 in the original problem, scaled by λ and $(1 - \lambda)$ respectively, is feasible. It follows that $\phi(q) \geq \lambda \phi(q_0) + (1 - \lambda)\phi(q_1)$. Note also that $\phi(q_0)$ is bounded on the interior of the simplex. It follows by Alexandrov's theorem that ϕ is twice-differentiable almost everywhere on the interior of the simplex.

By the convexity of H , the objective function is concave, and the constraints are affine and a feasible point exists. Therefore, the KKT conditions are necessary. Moreover, the objective function is continuously differentiable in the choice variables and in q_0 , and therefore the envelope theorem applies. We have, by the envelope theorem,

$$\phi_q(q_0) = \theta H_q(q_0) + \kappa,$$

and the first-order conditions (for all $a \in A$),

$$u_a - \theta H_q(\hat{q}_a) - \kappa + v_a = 0.$$

Define $\hat{q}_a(q_0)$, $\kappa(q_0)$, and $v_a(q_0)$ as functions that are solutions to the first-order conditions and constraints.

Consider an alternative prior, $\tilde{q}_0 \in \mathcal{P}(X)$, such that

$$\tilde{q}_0 = \sum_{a \in A} \alpha(a) \hat{q}_a(q_0)$$

for some $\alpha(a) \geq 0$. Conjecture that $\hat{q}_a(\tilde{q}_0) = \alpha(a) \hat{q}_a(q_0)$, $\kappa(\tilde{q}_0) = \kappa(q_0)$, and $v_a(\tilde{q}_0) =$

$v_a(q_0)$. By the homogeneity property,

$$H_q(\alpha(a)\hat{q}_a(q_0)) = H_q(\hat{q}_a(q_0)),$$

and therefore the first-order conditions are satisfied. By construction, the constraint is satisfied, the complementary slackness conditions are satisfied, and \hat{q}_a and v_a are weakly positive. Therefore, all necessary conditions are satisfied, and by the concavity of the problem, this is sufficient. It follows that the conjecture is verified.

Consider a perturbation

$$q_0(\varepsilon; z) = q_0 + \varepsilon z,$$

with $z \in \mathbb{R}^{|X|}$, such that $q_0(\varepsilon; z)$ remains in $\mathcal{P}(X)$ for some $\varepsilon > 0$. If z is in the span of $\hat{q}_a(q_0)$, then there exists a sufficiently small $\varepsilon > 0$ such that the above conjecture applies. It follows in this case that κ is constant, and therefore $\phi_q(q_0(\varepsilon; z))$ is directionally differentiable with respect to ε . If $q_0(-\varepsilon; z) \in \mathcal{P}(X)$ for some $\varepsilon > 0$, then ϕ_q is differentiable, with

$$\phi_{qq}(q_0) \cdot z = \theta H_{qq}(q_0) \cdot z,$$

proving twice-differentiability in this direction. This perturbation exists anywhere the span of $\hat{q}_a(q_0)$ is strictly larger than the line segment connecting zero and q_0 (in other words, all $\hat{q}_a(q_0)$ are not proportional to q_0). Define this region as the continuation region, Ω . Outside of this region, all $\hat{q}_a(q_0)$ are proportional to q_0 , implying that

$$\phi(q_0) = \max_{a \in A} u_a^T \cdot q_0,$$

as required for the stopping region. Within the continuation region, the strict convexity of

$H(q_0)$ in all directions orthogonal to q_0 implies that

$$\phi(q_0) > \max_{a \in A} u_a^T \cdot q_0,$$

as required.

Now consider an arbitrary perturbation z such that $q_0(\varepsilon; z) \in \mathbb{R}_+^{|\mathcal{X}|}$ and $q_0(-\varepsilon; z) \in \mathbb{R}_+^{|\mathcal{X}|}$ for some $\varepsilon > 0$. Observe that, by the constraint,

$$\varepsilon z = \sum_{a \in A} (\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)).$$

It follows that

$$(\kappa^T(q_0(\varepsilon; z)) - \kappa^T(q_0))\varepsilon z = \sum_{a \in A} (\kappa^T(q_0(\varepsilon; z)) - \kappa^T(q_0))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)).$$

By the first-order condition,

$$\begin{aligned} & (\kappa^T(q_0(\varepsilon; z)) - \kappa^T(q_0))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)) = \\ & [\theta H_q(\hat{q}_a(q_0)) - \theta H_q(\hat{q}_a(\varepsilon; z)) + \mathbf{v}_a^T(q_0(\varepsilon; z)) - \mathbf{v}_a^T(q_0)](\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)). \end{aligned}$$

Consider the term

$$(\mathbf{v}_a^T(q_0(\varepsilon; z)) - \mathbf{v}_a^T(q_0))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)) = \sum_{x \in \mathcal{X}} (\mathbf{v}_a^T(q_0(\varepsilon; z)) - \mathbf{v}_a^T(q_0)) e_x e_x^T (\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)).$$

By the complementary slackness condition,

$$(\mathbf{v}_a^T(q_0(\varepsilon; z)) - \mathbf{v}_a^T(q_0))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)) = -\mathbf{v}_a^T(q_0(\varepsilon; z))\hat{q}_a(q_0) - \mathbf{v}_a^T(q_0)\hat{q}_a(\varepsilon; z) \leq 0.$$

By the convexity of H ,

$$\theta(H_q(\hat{q}_a(q_0)) - \theta H_q(\hat{q}_a(\varepsilon; z)))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)) \leq 0.$$

Therefore,

$$(\kappa^T(q_0(\varepsilon; z)) - \kappa^T(q_0))\varepsilon z \leq 0.$$

It follows that anywhere ϕ is twice differentiable (almost everywhere on the interior of the simplex),

$$\phi_{qq}(q) \preceq \theta H_{qq}(q),$$

with equality in certain directions. Therefore, it satisfies the HJB equation almost everywhere in the continuation region. Moreover, by the convexity of ϕ ,

$$(H_q(q_0(\varepsilon; z)) - H_q(q_0))^T \varepsilon z \geq (\phi_q(q_0(\varepsilon; z)) - \phi_q(q_0))^T \varepsilon z \geq 0,$$

implying that the ‘‘Hessian measure’’ (see Villani (2003)) associated with ϕ_{qq} has no pure point component. This implies that ϕ is continuously differentiable.

Next, we show that there is a strategy for the DM in the dynamic problem which can implement this value function. Suppose the DM starts with beliefs q_0 , and generates some $\hat{q}_a(q_0)$ as described above. As shown previously, this can be mapped into a policy $\pi(a, q_0)$ and $q_a(q_0)$, with the property that

$$\sum_{a \in A} \pi(a, q_0) q_a(q_0) = q_0.$$

We will construct a policy such that, for all times t ,

$$q_t = \sum_{a \in A} \pi_t(a) q_a(q_0)$$

for some $\pi_t(a) \in \mathcal{P}(A)$. Let Ω (the continuation region) be the set of q_t such that a $\pi_t \in \mathcal{P}(A)$ satisfying the above property exists and $\pi_t(a) < 1$ for all $a \in A$. The associated stopping rule will be the stop whenever $\pi_t(a) = 1$ for some $a \in A$.

For all $q_t \in \Omega$, there is a linear map from $\mathcal{P}(A)$ to Ω , which we will denote $Q(q_0)$:

$$Q(q_0)\pi_t = q_t.$$

Therefore, we must have

$$Q(q_0)d\pi_t = D(q_t)\sigma_t dB_t.$$

By the assumption that $|X| \geq |A|$, there exists a $|A| \times |X|$ matrix $\sigma_{\pi,t}$ such that

$$Q(q_0)\sigma_{\pi,t} = D(q_t)\sigma_t$$

and

$$d\pi_t = \sigma_{\pi,t} dB_t.$$

Define

$$\tilde{\phi}(\pi_t) = \phi(q_t).$$

As shown above,

$$Q^T(q_0)\phi_{qq}(q_t)Q(q_0)$$

exists everywhere in Ω , and therefore

$$\tilde{\phi}(\pi_t) - \theta H(Q(q_0)\pi_t)$$

is a martingale.

We also have to scale $\sigma_{\pi,t}$ to respect the constraint,

$$\frac{1}{2}tr[\sigma_t \sigma_t^T k(q_t)] = \chi > 0.$$

This can be rewritten as

$$\frac{1}{2}tr[\sigma_{\pi,t} \sigma_{\pi,t}^T Q^T(q_0) D^+(Q(q_0)\pi_t) k(Q(q_0)\pi_t) D^+(Q(q_0)\pi_t) Q(q_0)] = \chi,$$

where D^+ denotes the pseudo-inverse.

By the positive-definiteness of k in all directions orthogonal to ι , we will always have $\frac{1}{2}tr[\sigma_{\pi,t} \sigma_{\pi,t}^T] > 0$. Under the stopping rule described previously, the boundary will be hit a.s. as the horizon goes to infinity. As a result, by the martingale property described above, initializing $\pi_0(a) = \pi(a, q_0)$,

$$\tilde{\phi}(\pi_0) = E_0[\tilde{\phi}(\pi_\tau) - \theta H(Q(q_0)\pi_\tau) + \theta H(Q(q_0)\pi_0)].$$

By Ito's lemma,

$$\theta H(Q(q_0)\pi_\tau) - \theta H(Q(q_0)\pi_0) = \int_0^\tau \chi \theta dt = \mu \tau.$$

By the value-matching property of ϕ , $\tilde{\phi}(\pi_\tau) = \hat{u}(Q(q_0)\pi_\tau)$. It follows that

$$\phi(q_0) = \tilde{\phi}(\pi_0) = E_0[\hat{u}(q_\tau) - \mu \tau],$$

as required.

Finally, we verify that alternative policies are sub-optimal. Consider an arbitrary control process σ_t and stopping rule described by the stopping time τ . We have, by the convexity

of ϕ and the generalized Ito formula for convex functions (noting that we have shown that the Hessian measure associated with ϕ_{qq} has no pure point component), interpreting ϕ_{qq} in a distributional sense,

$$E_0[\phi(q_\tau)] - \phi(q_0) = \frac{1}{2}E_0\left[\int_0^\tau \text{tr}[\sigma_t^T D(q_t)\phi_{qq}(q_t)D(q_t)\sigma_t]dt\right].$$

By the feasibility of the policies, anywhere in the continuation region of the optimal policy,

$$\frac{1}{2}\text{tr}[\sigma_t^T D(q_t)\phi_{qq}(q_t)D(q_t)\sigma_t] \leq \frac{1}{2}\theta\text{tr}[\sigma_t^T k(q_t)\sigma_t] \leq \theta\chi.$$

In the stopping region of the optimal policy,

$$\frac{1}{2}\text{tr}[\sigma_t^T D(q_t)\phi_{qq}(q_t)D(q_t)\sigma_t] = 0 < \theta\chi.$$

Therefore,

$$\phi(q_0) \geq E_0[\phi(q_\tau)] - \int_0^\tau \theta\chi dt.$$

By the inequality

$$\phi(q_\tau) \geq \hat{u}(q_\tau),$$

we have

$$\phi(q_0) \geq E_0[\hat{u}(q_\tau) - \mu\tau]$$

for all policies, verifying optimality.

D.3 Proof of Lemma 2

Let p and p' be information structures with signal alphabet S . First, we will show that mixture feasibility and Blackwell monotonicity imply convexity. By mixture equivalence,

letting p_M denote the mixture information structure and S_M the signal alphabet,

$$C(p_M, q; S_M) \leq \lambda C(p, q; S) + (1 - \lambda) C(p', q; S).$$

Consider the garbling $\Pi : S \times \{1, 2\} \rightarrow S$, which maps each $(s, i) \in S_M$ to $s \in S$. By Blackwell monotonicity,

$$C(p_M, q; S_M) \geq C(\Pi p_M, q; S).$$

By construction,

$$e_s^T \Pi p_M = \lambda e_s^T p + (1 - \lambda) e_s^T p',$$

and the result follows.

Now we show the other direction: that convexity and Blackwell monotonicity imply mixture feasibility. Let p_1 and p_2 be information structures with signal alphabets S_1 and S_2 . Because the cost function satisfies Blackwell monotonicity, it is invariant to Markov congruent embeddings. Define $S_M = (S_1 \cup S_2) \times \{1, 2\}$. There exists an embedding $\Pi_1 : S_1 \rightarrow S_M$ such that, for some $s_M = (s, i) \in S_M$,

$$e_{s_M}^T \Pi_1 p_1 = \begin{cases} 0 & i = 2 \\ 0 & s \notin S_1 \\ e_s^T p_1 & \text{otherwise} \end{cases} .$$

Define an embedding Π_2 along similar lines, and note that these embeddings are left-invertible. It follows by invariance that

$$C(\Pi_1 p_1, q; S_M) = C(p_1, q; S_1),$$

and likewise that

$$C(\Pi_2 p_2, q; S_M) = C(p_2, q; S_2).$$

By convexity,

$$C(\lambda \Pi_1 p_1 + (1 - \lambda) \Pi_2 p_2; q; S_M) \leq \lambda C(\Pi_1 p_1, q; S_M) + (1 - \lambda) C(\Pi_2 p_2, q; S_M).$$

Observing that

$$\lambda \Pi_1 p_1 + (1 - \lambda) \Pi_2 p_2 = p_M$$

proves the result.

D.4 Proof of Theorem 2

Parts 1 and 2 of the theorem follow from a Taylor expansion of the cost function. Using the lemmas and theorem of Chentsov (1982), cited in the text, we know that for any invariant cost function with continuous second derivatives,

$$C(p, q; S) = \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \tau_{x'}^T g(r) \tau_x + o(\Delta).$$

The second claim follows by a similar argument.

We next demonstrate the claimed properties of $k(q)$. First, $k(q)$ is symmetric and continuous in q , by the symmetry of partial derivatives and the assumption of continuous second derivatives (Condition 4). Recall the assumption that

$$p_x = r + \Delta^{\frac{1}{2}} \tau_x + o(\Delta^{\frac{1}{2}}),$$

which implies that $\sum_{s \in S} e_s^T r = 1$ and $\sum_{s \in S} e_s^T \tau_x = 0$ for all $x \in X$. Consider an information structure for which $\tau_x = \phi e_x^T v$, where $v \in \mathbb{R}^{|X|}$ and $\phi \in \mathbb{R}^{|S|}$, with $\sum_{s \in S} e_s^T \phi = 0$. Suppose

that both v and ϕ are not zero. For this information structure,

$$C(p, q; S) = \frac{1}{2} \Delta \bar{g} v^T k(q) v + o(\Delta),$$

where $\phi^T g(r) \phi = \bar{g} > 0$. Suppose the information structure is uninformative for all Δ . This would be the case if v is proportional to ι , and therefore

$$\iota^T k(q) \iota = 0$$

by Condition 1. Regardless of whether the information structure is informative, by Condition 1, we must have

$$v^T k(q) v \geq 0,$$

implying that $k(q)$ is positive semi-definite. If z and $-z$ are in the tangent space of the simplex at q , there exists an x, x' $e_x^T z \neq e_{x'}^T z$ with x, x' in the support of q . Using z in the place of v above, by Condition 1, we must have

$$z^T k(q) z > 0.$$

Suppose now that the cost function satisfies Condition 5. Let v be as above, non-zero, and not proportional to ι . We have

$$C(p, q; S) = \frac{1}{2} \Delta \bar{g} v^T k(q) v + o(\Delta),$$

and therefore for the B defined in Condition 5 there exists a Δ_B such that, for all $\Delta < \Delta_B$, $C(p, q; S) < B$. Therefore, we must have

$$C(\{p_x\}_{x \in X}, q) \geq \frac{m}{2} \sum_{s \in S} (e_s^T p q) \|q_s - q\|_X^2.$$

By Bayes' rule, for any signal that is received with positive probability,

$$q_s - q = \frac{(D(q) - qq^T)p^T e_s}{q^T p^T e_s}.$$

By convention, $q_s = q$ for any s such that $e_s^T pq = 0$.

The support of q_s is always a subset of the support of q , and therefore (by the equivalence of norms),

$$C(\{p_x\}_{x \in X}, q) \geq \frac{m_g}{2} \sum_{s \in S} (e_s^T pq)(q_s - q)^T D^+(q)(q_s - q)$$

for some constant $m_g > 0$.

For sufficiently large Δ , $e_s^T pq > 0$ if $e_s^T r_s > 0$, and therefore

$$C(\{p_x\}_{x \in X}, q) \geq \frac{m}{2} \sum_{s \in S: e_s^T r_s > 0} \frac{(e_s^T p(D(q) - qq^T)D^+(q)(D(q) - qq^T)p^T e_s)}{(e_s^T pq)},$$

or,

$$C(\{p_x\}_{x \in X}, q) \geq \frac{m}{2} \Delta \sum_{s \in S: e_s^T r_s > 0} (e_s^T \phi)^2 \frac{v^T (D(q) - qq^T)D^+(q)(D(q) - qq^T)v}{(e_s^T r)} + o(\Delta).$$

Noting that

$$\sum_{s \in S: e_s^T pq > 0} \frac{(e_s^T \phi)^2}{(e_s^T pq)} = \phi^T g(r)\phi = \bar{g},$$

and that

$$(D(q) - qq^T)D^+(q)(D(q) - qq^T) = g^+(q),$$

we have

$$C(\{p_x\}_{x \in X}, q) \geq \frac{m_g}{2} \Delta \bar{g} v^T g^+(q)v + o(\Delta).$$

It follows that we must have

$$\frac{1}{2}v^T k(q)v \geq \frac{m_g}{2}v^T g^+(q)v$$

for all v .

D.5 Proof of Corollary 3

Under the stated assumptions,

$$p_x = r + \Delta^{\frac{1}{2}}\tau_x + o(\Delta^{\frac{1}{2}}).$$

By Bayes' rule, for any $s \in S$ such that $e_s^T p q > 0$,

$$q_s = \frac{D(q)p^T e_s}{q^T p^T e_s}.$$

It follows immediately that

$$\lim_{\Delta \rightarrow 0^+} q_s = D(q) \frac{r^T e_s}{r_s^T} = q.$$

Next,

$$\begin{aligned} \Delta^{-\frac{1}{2}}(q_s - q) &= \Delta^{-\frac{1}{2}} \frac{(D(q) - qq^T)p^T e_s}{q^T p^T e_s} \\ &= D(q) \frac{\tau^T e_s - \mathbf{1}q^T \tau^T e_s + o(1)}{q^T p^T e_s}. \end{aligned}$$

For any s such that $q^T p^T e_s > 0$,

$$\lim_{\Delta \rightarrow 0^+} \Delta^{-\frac{1}{2}}(q_s - q) = D(q) \frac{\tau^T e_s - \mathbf{1}q^T \tau^T e_s}{r^T e_s}.$$

By Theorem 2,

$$C(p, q; S) = \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \tau_{x'}^T g(r) \tau_x + o(\Delta).$$

By the result that $\iota^T k(q) = 0$, we have

$$\begin{aligned} C(p, q; S) &= \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} e_x^T k(q) e_{x'} \cdot (\tau_{x'} - q\tau)^T g(r) (\tau_x - q\tau) \\ &\quad + o(\Delta). \end{aligned}$$

By the definition of the Fisher matrix, and the observation that $\iota^T \tau_x = 0$ for all $x \in X$,

$$(\tau_{x'} - q\tau)^T g(r) (\tau_x - q\tau) = \sum_{s \in S: e_s^T r > 0} (e_s^T r) \frac{(\tau_{x'} - q\tau)^T}{(e_s^T r)} e_s e_s^T \frac{(\tau_x - q\tau)}{(e_s^T r)}.$$

Substituting in the result regarding the posterior,

$$C(p, q; S) = \frac{1}{2} \sum_{s \in S: e_s^T r > 0} (e_s^T r) (q_s - q)^T D^+(q) k(q) D^+(q) (q_s - q) + o(\Delta),$$

which is the result.

D.6 Proof of Corollary 2

The cost function is directionally differentiable with respect to signals that add to the support of the signal distribution.

By directional differentiability and the continuity of the directional derivatives, there

exists a function

$$f(\omega, r, q; S) = \lim_{\Delta \rightarrow 0^+} \frac{C(\bar{p}_\Delta + \Delta \omega, q; S) - C(\bar{p}_\Delta, q; S)}{\Delta}.$$

Observe that, if ωe_x is in the support of r for all x in the support of q , we must have $f(\omega, \bar{p}, q; S) = 0$, by the results of Theorem 2. Relatedly, if ω and ω' differ only with respect to the frequency of signals in the support of r for all x in the support of q , we must have

$$f(\omega, r, q; S) = f(\omega', r, q; S).$$

Assuming there are signals not in the support of \bar{p} , we can write $\omega = \omega_1 + \omega_2 + \dots$, where each ω_i is a perturbation that contains only one signal not the support of $\bar{p}q$. Let $N \leq |S|$ denote the number of these perturbations. We can define

$$f_i(\omega_i, r, q; S) = \lim_{\Delta \rightarrow 0^+} \frac{C(p_{i-1} + \Delta \omega_i, q; S) - C(p_{i-1}, q; S)}{\Delta},$$

where $p_{i-1} = \bar{p}_\Delta + \Delta \sum_{j=1}^{i-1} \omega_j$. By the assumption of the continuity of the directional derivatives,

$$f_i(\omega_i, r, q; S) = f(\omega_i, r, q; S).$$

It follows that

$$f(\omega, r, q; S) = \sum_{i=1}^N f(\omega_i, r, q; S).$$

By invariance, the function $f(\omega_i, r, q; S)$ does not depend on r or S . By the argument above, it is only a function of $e_{s_i} \omega_i$, where $s_i \in S$ is the unique signal in ω_i with $e_{s_i}^T r = 0$. By Bayes' rule,

$$e_{s_i} \omega_i = (e_{s_i} \omega_i q) D(q)^+ q_{s_i},$$

where q_{s_i} is the posterior associated with signal s_i . By the homogeneity of the directional derivative, we can rewrite this as

$$f(\omega_i, r, q; S) = (e_{s_i} \omega_i q) F(q_{s_i}, q).$$

By the requirement that the cost of an uninformative signal structure is zero, and everything else is costly, we must have

$$F(q, q) = 0,$$

$$F(q', q) > 0$$

for all $q' \neq q$. Therefore, F is a divergence, which we write $D^*(q' || q)$. The finiteness of $D^*(q' || q)$ is implied by the existence of the directional derivative. The approximation of the cost function follows from this result and Corollary 3.

By invariance, there exists a Markov congruent embedding that splits each signal in S into $M > 1$ distinct signals in S' . As M becomes arbitrarily large, the probability of each signal becomes small — and in particular, can be of order Δ . It follows for all $s \in S'$ such that $\|q_s - q\| = O(\Delta^{\frac{1}{2}})$ (e.g. the signals described in Corollary 3), we must have

$$D^*(q_s || q) = \frac{1}{2} \Delta (q_s^T - q) \bar{k}(q) (q_s - q) + O(\Delta).$$

Moreover, by this argument, $D^*(q' || q)$ must be twice differentiable for q' in the neighborhood of q .

D.7 Proof of Lemma 3

We will show that Conditions 1-5 are satisfied. Recall the definition:

$$C_N(p, q; S) = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S} e_s^T \bar{p}_i D_i(q_{i,s} || q_i)$$

D.7.1 Condition 1

Condition 1 requires that if the information structure is uninformative, the cost is zero, and if it is not, the cost is weakly positive. If the signal is uninformative, for any signal received with positive probability,

$$q_{i,s} = q_i,$$

and by our convention that $q_{i,s} = q_i$ if $\bar{q}_{i,s} = 0$, this also holds for zero-probability signals. By the definition of a divergence, $D_i(q_i || q_i) = 0$ for all q_i , and therefore the cost of an uninformative information structure is zero.

The cost is weakly positive by the definition of a divergence (being weakly positive) and the fact that probabilities are weakly positive.

D.7.2 Condition 2

Mixture feasibility requires that

$$C(p_M, q; S_M) \leq \lambda C(p_1, q; S_1) + (1 - \lambda) C(p_2, q; S_2).$$

By definition,

$$\bar{p}_{i,M} = \frac{\sum_{x \in X_i} p_M e_x e_x^T q}{\bar{q}_i}$$

and

$$q_{i,s,M} = \frac{E_i q_{s,M}}{\sum_{x \in X_i} e_x^T q_{s,M}}$$

for any s such that $\bar{q}_{i,s,M} > 0$. For any $(s, 1) \in S_M$, if $\bar{q}_{i,s,M} > 0$, we must have $\bar{q}_{i,s} > 0$, and therefore $q_{i,s,M} = q_{i,s,1}$ (denoting the posterior under p_1). The same argument holds for the second information structure.

It follows that

$$\begin{aligned} C(p_M, q; S_M) &= \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S_M} e_s^T \bar{p}_{i,M} D_i(q_{i,s,M} \| q_i) \\ &= \sum_{i \in \mathcal{I}(q)} \bar{q}_i (\lambda \sum_{s \in S_1} e_s^T \bar{p}_{i,1} D_i(q_{i,s,1} \| q_i) + (1 - \lambda) \sum_{s \in S_2} e_s^T \bar{p}_{i,2} D_i(q_{i,s,2} \| q_i)) \\ &= \lambda C(p_1, q; S_1) + (1 - \lambda) C(p_2, q; S_2), \end{aligned}$$

verifying that the condition holds.

D.7.3 Condition 3

By Blackwell's theorem, for any Markov mapping $\Pi : S \rightarrow S'$, we require that

$$C(\Pi p, q; S') \leq C(p, q; S).$$

Consider a neighborhood $i \in \mathcal{I}(q)$. By definition,

$$\bar{p}'_i = \frac{\sum_{x \in X_i} \Pi p e_x e_x^T q}{\bar{q}_i} = \Pi \bar{p}_i$$

and

$$\begin{aligned}
q_{i,s'} &= \frac{E_i q_{s'}}{\sum_{x \in X_i} e_x^T q_{s'}} \\
&= \frac{E_i D(q) p^T \Pi^T e_{s'}}{\sum_{x \in X_i} e_x^T D(q) p^T \Pi^T e_{s'}} \\
&= \frac{D(q_i) E_i p^T \Pi^T e_{s'}}{\bar{p}_i^T \Pi^T e_{s'}}
\end{aligned}$$

where the second step follows by Bayes' rule,

$$D(q) p^T \Pi^T e_{s'} = (e_{s'}^T \Pi p q) q_{s'}.$$

Also by Bayes' rule,

$$\begin{aligned}
D(q_i) E_i p^T e_s &= (e_s^T p E_i^T q_i) q_{i,s} \\
&= (e_s^T \bar{p}_i) q_{i,s}.
\end{aligned}$$

and therefore

$$q_{i,s'} = \frac{\sum_{s \in S} q_{i,s} \bar{p}_i^T \Pi^T e_{s'}}{\bar{p}_i^T \Pi^T e_{s'}}.$$

It follows by the convexity of D_i in its first argument that

$$(\bar{p}_i^T \Pi^T e_{s'}) D_i(q_{i,s'} || q_i) \leq \sum_{s \in S} \bar{p}_i^T \Pi^T e_{s'} D_i(q_{i,s} || q_i).$$

Therefore,

$$\begin{aligned} C(\Pi p, q; S') &= \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s' \in S'} e_{s'}^T \Pi \bar{p}_i D_i(q_{i,s'} || q_i) \\ &\leq \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s' \in S'} \sum_{s \in S} \bar{p}_i^T \Pi^T e_{s'} D_i(q_{i,s} || q_i). \end{aligned}$$

By definition,

$$\sum_{s' \in S'} \Pi^T e_{s'} = 1$$

and therefore

$$C(\Pi p, q; S') \leq C(p, q; S).$$

D.7.4 Condition 4

By the definition of the neighborhood structure,

$$C_N(p, q; S) = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S} e_s^T \bar{p}_i D_i(q_{i,s} || q_i),$$

and the twice-differentiability of D_i in its first argument, it is sufficient to show that \bar{p}_i and $q_{i,s}$ are both twice-differentiable with respect to perturbations to p , in the neighborhood of an uninformative information structure.

Suppose that

$$p(\varepsilon) = r t^T + \varepsilon \tau + \nu \omega,$$

where $r \in \mathcal{P}(S)$ and the support of τe_x is in the support of r , and likewise for ωe_x , for all $x \in X$.

By Bayes' rule, for all $s \in S$ such that $e_s^T r > 0$,

$$q_s(\varepsilon, \nu) = \frac{D(q)p(\varepsilon, \nu)^T e_s}{q^T p(\varepsilon, \nu)^T e_s}.$$

Simplifying,

$$q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = q \frac{r^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} + \frac{\boldsymbol{\varepsilon} D(q) \boldsymbol{\tau}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} + \frac{\boldsymbol{\nu} D(q) \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s}.$$

In the neighborhood around $\boldsymbol{\varepsilon} = \boldsymbol{\nu} = 0$, the denominator is strictly positive, and therefore

$$\frac{\partial}{\partial \boldsymbol{\nu}} q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = -q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) \frac{q^T \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} + \frac{D(q) \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s}$$

and

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\varepsilon}} \frac{\partial}{\partial \boldsymbol{\nu}} q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) &= q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) \frac{q^T \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \frac{q^T \boldsymbol{\tau}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \\ &\quad - \frac{q^T \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \frac{D(q) \boldsymbol{\tau}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \\ &\quad - q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) \frac{q^T \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \frac{q^T \boldsymbol{\tau}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \\ &\quad - \frac{D(q) \boldsymbol{\omega}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s} \frac{q^T \boldsymbol{\tau}^T e_s}{r^T e_s + \boldsymbol{\varepsilon} q^T \boldsymbol{\tau}^T e_s + \boldsymbol{\nu} q^T \boldsymbol{\omega}^T e_s}. \end{aligned}$$

For $s \in S$ such that $e_s^T r = 0$, $q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = q$, and therefore $\frac{\partial}{\partial \boldsymbol{\varepsilon}} \frac{\partial}{\partial \boldsymbol{\nu}} q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = 0$. Therefore, $\frac{\partial}{\partial \boldsymbol{\nu}} q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu})$ can be written as a quadratic form in $\text{vec}(\boldsymbol{\tau})$ and $\text{vec}(\boldsymbol{\omega})$. It follows that $q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu})$, in the neighborhood of an uninformative information structure, is twice-differentiable in the directions that do not change the support of the distribution of signals.

For all $i \in \mathcal{I}(q)$, define $\tilde{q}_i \in \mathcal{P}(X)$ as

$$e_x^T \tilde{q}_i = \begin{cases} \frac{e_x^T q}{\tilde{q}_i} & x \in X_i \\ 0 & \text{otherwise.} \end{cases}$$

By definition,

$$\bar{p}_i(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = p\tilde{q}_i = r + \boldsymbol{\varepsilon}\boldsymbol{\tau}\tilde{q}_i + \boldsymbol{\nu}\boldsymbol{\omega}\tilde{q}_i.$$

and therefore is twice-differentiable in the required directions. Moreover, by construction, if $e_s^T r = 0$, then $e_s^T \bar{p}_i(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = 0$, and if $e_s^T r > 0$, then $e_s^T \bar{p}_i(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) > 0$ in the neighborhood around $\boldsymbol{\varepsilon} = \boldsymbol{\nu} = 0$.

By definition,

$$q_{i,s}(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) = \frac{E_i q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu})}{\sum_{x \in X_i} e_x^T q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu})}.$$

For all $i \in \mathcal{I}(q)$, in the neighborhood of an uninformative information structure, $\sum_{x \in X_i} e_x^T q_s(\boldsymbol{\varepsilon}, \boldsymbol{\nu}) \approx \bar{q}_i > 0$, and therefore the twice-differentiability of q_s in the required directions implies the twice-differentiability of $q_{i,s}$ in those directions.

D.7.5 Condition 5

This condition requires that, for some $m > 0$ and $B > 0$, for all $C(p, q; S) < B$,

$$C(p, q; S) \geq \frac{m}{2} \sum_{s \in S} (e_s^T p q) \|q_s - q\|_X^2,$$

where $\|\cdot\|_X$ is an arbitrary norm on the tangent space of $\mathcal{P}(X)$. It follows immediately by the strong convexity of the divergence for the neighborhood that contains all states.

D.8 Proof of Lemma 4

Consider Corollary 3. Under the stated assumptions,

$$p_x = r + \Delta^{\frac{1}{2}} \boldsymbol{\tau}_x + o(\Delta^{\frac{1}{2}})$$

$$q_{s,x} = q_x + \Delta^{\frac{1}{2}} q_x \frac{e_s^T (\tau_x - \sum_{x' \in X} \tau_{x'} q_{x'})}{e_s^T r} + o(\Delta^{\frac{1}{2}}).$$

By definition,

$$\bar{k}(q) = D^+(q)k(q)D^+(q),$$

and the cost function can be written as

$$C(\{p_x\}_{x \in X}, q; S) = \frac{1}{2} \sum_{s \in S} (e_s^T r)(q_s - q)^T \bar{k}(q)(q_s - q) + o(\Delta).$$

Now consider the definition of neighborhood cost function (19):

$$C_N(\{p_x\}_{x \in X}, q; S) = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S} e_s^T \bar{p}_i D_i(q_{i,s} || q_i).$$

By definition,

$$\begin{aligned} \bar{q}_i \bar{p}_i &= \sum_{x \in X_i} p_x e_x^T q \\ &= r \bar{q}_i + o(1). \end{aligned}$$

Note that

$$pq = r + o(1)$$

as well.

By Chentsov's theorem (Chentsov (1982)) and the approximation above,

$$D_i(q_{i,s} || q_i) = c_i(q_{i,s} - q_i)^T g(q_i)(q_{i,s} - q_i) + o(\Delta).$$

The approximation described in equation (20) follows.

Define the $|X| \times |X_i|$ matrix E_i that selects the elements of X that are contained in X_i .

We have

$$\begin{aligned}
q_{i,s,x} &= \frac{q_{s,x}(\Delta)}{\sum_{x' \in X_i} q_{s,x'}(\Delta)} \\
&= \frac{q_x}{\sum_{x' \in X_i} q_{x'}} + \Delta^{\frac{1}{2}} \frac{q_x}{\sum_{x' \in X_i} q_{x'}} \frac{e_s^T (\tau_x - \sum_{x' \in X} \tau_{x'} q_{x'})}{e_s^T r} \\
&\quad - \Delta^{\frac{1}{2}} \frac{q_x}{(\sum_{x' \in X_i} q_{x'})^2} \sum_{x' \in X_i} q_{x'} \frac{e_s^T (\tau_{x'} - \sum_{x'' \in X} \tau_{x''} q_{x''})}{e_s^T r} + o(\Delta^{\frac{1}{2}}).
\end{aligned}$$

That is,

$$q_{i,s} = q_i + \frac{1}{\bar{q}_i} E_i(q_s - q) - \frac{1}{\bar{q}_i} q_i q_i^T D^+(q_i) E_i(q_s - q) + o(\Delta^{\frac{1}{2}}),$$

Using this,

$$\begin{aligned}
(q_{i,s} - q_i)^T g(q_i) (q_{i,s} - q_i) &= (q_{i,s} - q_i)^T D^+(q_i) (q_{i,s} - q_i) \\
&= \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) E_i (q_s - q) - \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) q_i q_i^T D^+(q_i) E_i (q_s - q) \\
&\quad - \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) q_i q_i^T D^+(q_i) E_i (q_s - q) \\
&\quad + \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) q_i q_i^T D^+(q_i) E_i (q_s - q) + o(\Delta).
\end{aligned}$$

Therefore,

$$\begin{aligned}
C_N(\{p_x\}_{x \in X}, q; S) &= \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in S} (e_s^T r) (q_{i,s} - q_i)^T g(q_i) (q_{i,s} - q_i) + o(\Delta) \\
&= \Delta \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in S} (e_s^T r) (q_s - q)^T \bar{k}_i(q) (q_s - q) + o(\Delta),
\end{aligned}$$

where

$$\bar{k}_i(q) = \frac{1}{(\bar{q}_i)^2} E_i^T (D^+(q_i) - D^+(q_i) q_i q_i^T D^+(q_i)) E_i.$$

The $\bar{k}(q)$ matrix is

$$\begin{aligned}\bar{k}_N(q) &= \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \bar{k}_i(q) \\ &= \sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i} E_i^T (D^+(q_i) - D^+(q_i) q_i q_i^T D^+(q_i)) E_i.\end{aligned}\quad (31)$$

Thus, the associated $k(q)$ matrix is

$$\begin{aligned}k_N(q) &= D(q) \bar{k}(q) D(q) \\ &= \sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i} D(q) E_i^T (D^+(q_i) - D^+(q_i) q_i q_i^T D^+(q_i)) E_i D(q) \\ &= \sum_{i \in \mathcal{I}(q)} \{c_i E_i^T D(q) E_i - c_i \bar{q}_i E_i^T q_i q_i^T E_i D\} \\ &= \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i E_i^T g^+(q_i) E_i.\end{aligned}$$

D.9 Proof of Lemma 5

Using equation (31) from the proof of Lemma 4, we have

$$\bar{k}_N(q) = \sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i} E_i^T (D^+(q_i) - D^+(q_i) q_i q_i^T D^+(q_i)) E_i.$$

Consider the function

$$\begin{aligned}H_N(q) &= \sum_{i \in \mathcal{I}(q)} c_i \left[\sum_{x \in X_i} (e_x^T q) \ln(e_x^T q) - \left(\sum_{x \in X_i} (e_x^T q) \right) \ln \left(\sum_{x \in X_i} (e_x^T q) \right) \right] \\ &= \sum_{i \in \mathcal{I}(q)} c_i \sum_{x \in X_i} (e_x^T q) \ln(q_{i,x}) \\ &= - \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i H^{\text{Shannon}}(q_i).\end{aligned}$$

Differentiating,

$$\frac{\partial H_N(q)}{\partial q_{x'}} = (\ln(q_{x'}) + 1) \sum_{i \in \mathcal{I}(q): x' \in X_i} c_i - \sum_{i \in \mathcal{I}(q): x' \in X_i} c_i (1 + \ln(\sum_{x \in X_i} (e_x^T q))).$$

Differentiating again,

$$\frac{\partial^2 H_N(q)}{\partial q_{x'} \partial q_{x''}} = \frac{\delta_{x', x''}}{q_{x'}} \sum_{i \in \mathcal{I}(q): x' \in X_i} c_i - \sum_{i \in \mathcal{I}(q): x', x'' \in X_i} \frac{c_i}{\sum_{x \in X_i} (e_x^T q)},$$

where $\delta_{x', x''}$ is the Kronecker delta. By definition,

$$\sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i} e_{x'}^T E_i^T D^+(q_i) q_i q_i^T D^+(q_i) E_i e_{x''} = \sum_{i \in \mathcal{I}(q): x', x'' \in X_i} \frac{c_i}{\sum_{x \in X_i} (e_x^T q)}$$

and

$$\sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i} e_{x'}^T E_i^T D^+(q_i) E_i e_{x''} = \delta_{x', x''} \sum_{i \in \mathcal{I}(q): x', x'' \in X_i} \frac{c_i}{(e_{x'}^T q)},$$

proving that $\bar{k}_N(q)$ is the Hessian of $H_N(q)$. Differentiation of $H_N(q)$ then yields the form given in the lemma for the associated Bregman divergence.

The posterior-separable static information-cost function is defined as

$$C_N^{\text{static}}(p, q; \mathcal{S}) = \sum_{s \in \mathcal{S}} (e_s^T p q) (H_N(q_s) - H_N(q)).$$

Using the definitions above,

$$\begin{aligned} C_N^{\text{static}}(p, q; \mathcal{S}) &= - \sum_{s \in \mathcal{S}} (e_s^T p q) \sum_{i \in \mathcal{I}(q_s)} c_i \bar{q}_{i,s} H^{\text{Shannon}}(q_{i,s}) \\ &\quad + \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i H^{\text{Shannon}}(q_i). \end{aligned}$$

Note that $\bar{q}_{i,s} = 0$ for $i \in \mathcal{I}(q) \setminus \mathcal{I}(q_s)$, and $\mathcal{I}(q_s) \subseteq \mathcal{I}(q)$, and therefore

$$C_N^{static}(p, q; \mathcal{S}) = - \sum_{s \in \mathcal{S}} (e_s^T p q) \sum_{i \in \mathcal{I}(q)} c_i (\bar{q}_{i,s} H^{Shannon}(q_{i,s}) - \bar{q}_i H^{Shannon}(q_i)).$$

By Bayes' rule,

$$(e_s^T p q) \bar{q}_{i,s} = \bar{q}_i \bar{p}_{i,s}$$

and by definition,

$$\sum_{s \in \mathcal{S}} \bar{p}_{i,s} = 1,$$

and therefore

$$\begin{aligned} C_N^{static}(p, q; \mathcal{S}) &= - \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in \mathcal{S}} \bar{p}_{i,s} (H^{Shannon}(q_{i,s}) - H^{Shannon}(q_i)) \\ &= \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in \mathcal{S}} \bar{p}_{i,s} D_{KL}(q_{i,s} || q_i). \end{aligned}$$

The claim that

$$C_N^{static}(p, q; \mathcal{S}) = \sum_{i \in \mathcal{I}(q)} c_i \sum_{x \in X: x \in X_i} (e_x^T q) D_{KL}(p e_x || p E_i^T q_i)$$

follows from the usual alternative ways of expressing mutual information and definitions.

D.10 Additional Definition and Lemmas

Definition 1. Let X^N be a sequence of state spaces, as described in section A.2. A sequence of policies $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$ satisfies the ‘‘convergence condition’’ if:

- i) The sequence satisfies, for some constants $c_H > c_L > 0$, all N , and all $i \in X^N$,

$$\frac{c_H}{N+1} \geq e_i^T p_N \geq \frac{c_L}{N+1}.$$

ii) The sequence satisfies, for some constant $K_1 > 0$, all N , and all $i \in X^N \setminus \{0, N\}$,

$$N^3 \left| \frac{1}{2} (e_{i+1}^T + e_{i-1}^T - 2e_i^T) p_N \right| \leq K_1,$$

and

$$N^2 \left| \frac{1}{2} (e_N^T - e_{N-1}^T) p_N \right| \leq K_1$$

and

$$N^2 \left| \frac{1}{2} (e_1^T - e_0^T) p_N \right| \leq K_1.$$

Lemma 11. Given a function $p \in \mathcal{P}([0, 1])$, define the sequence $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$,

$$e_i^T p_N = \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} p(x) dx,$$

where X^N is the state space described in section A.2. If the function p is strictly greater than zero for all $x \in [0, 1]$, differentiable, and its derivative is Lipschitz continuous, then the sequence $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$ satisfies the convergence condition, and satisfies, for some constant $K > 0$, all N , and all $i \in X^N \setminus \{0, N\}$,

$$N^2 \left| \ln\left(\frac{1}{2}(e_{i+1}^T + e_i^T)q_N\right) + \ln\left(\frac{1}{2}(e_{i-1}^T + e_i^T)q_N\right) - 2\ln(e_i^T q_N) \right| \leq K$$

and

$$N \left| \ln\left(\frac{1}{2}(e_1^T + e_0^T)q_N\right) - \ln(e_0^T q_N) \right| < K$$

and

$$N \left| \ln\left(\frac{1}{2}(e_N^T + e_{N-1}^T)q_N\right) - \ln(e_N^T q_N) \right| < K.$$

Proof. The function p is strictly greater than zero, and continuous, and therefore attains a maximum and minimum on $[0, 1]$, which we denote with c_H and c_L , respectively. By

construction,

$$e_i^T p_N \geq \frac{c_L}{N+1}$$

and likewise for c_H , satisfying the bounds.

For all $i \in X^N \setminus \{N\}$,

$$\begin{aligned} (e_{i+1}^T - e_i^T)p_N &= \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \left(p\left(x + \frac{1}{N+1}\right) - p(x) \right) dx \\ &= \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \int_0^{\frac{1}{N+1}} p'(x+y) dy dx \end{aligned}$$

and therefore, letting K_2 be the maximum of the absolute value of p' on $[0, 1]$ (which exists by the continuity of p'), we have

$$|(e_{i+1}^T - e_i^T)p_N| \leq \frac{1}{(N+1)^2} K_2,$$

satisfying the convergence condition for the endpoints.

For all $i \in X^N \setminus \{0, N\}$,

$$\begin{aligned} (e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N &= \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \left(p\left(x + \frac{1}{N+1}\right) + p\left(x - \frac{1}{N+1}\right) - 2p(x) \right) dx \\ &= \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \int_0^{\frac{1}{N+1}} (p'(x+y) - p'(x-y)) dy dx. \end{aligned}$$

By the Lipschitz continuity of p' , it is absolutely continuous, and therefore

$$p'(x+y) = p'(x) + \int_0^y p''(x+z) dz.$$

It follows that

$$(e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N = \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \int_0^{\frac{1}{N+1}} \int_{-y}^y (p''(x+z)) dz dy dx.$$

Let K_3 denote the Lipschitz constant associated with p' . It follows that

$$|(e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N| \leq \frac{2K_3}{(N+1)^3}.$$

Therefore, the convergence condition is satisfied for $K = \max(\frac{1}{2}K_2, K_3)$.

By the concavity of the log function, and the inequality $\ln(x) \leq x - 1$,

$$\begin{aligned} \ln\left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{e_i^T p_N}\right) + \ln\left(\frac{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N}{e_i^T p_N}\right) &\leq 2\ln\left(\frac{\frac{1}{4}(e_{i+1}^T + e_{i-1}^T + 2e_i^T)p_N}{e_i^T p_N}\right) \\ &\leq \frac{\frac{1}{2}(e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N}{e_i^T p_N}. \end{aligned}$$

Therefore, by the bounds above,

$$\ln\left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{e_i^T p_N}\right) + \ln\left(\frac{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N}{e_i^T p_N}\right) \leq \frac{(N+1)K}{N^3 c_L} \leq \frac{2K}{N^2 c_L}.$$

By the inequality $-\ln(\frac{1}{x}) \leq x - 1$,

$$\ln\left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{e_i^T p_N}\right) + \ln\left(\frac{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N}{e_i^T p_N}\right) \geq \frac{\frac{1}{2}(e_{i+1}^T - e_i^T)p_N}{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N} + \frac{\frac{1}{2}(e_{i-1}^T - e_i^T)p_N}{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N}.$$

We can rewrite this as

$$\begin{aligned} \ln\left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{e_i^T p_N}\right) + \ln\left(\frac{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N}{e_i^T p_N}\right) &\geq \\ &\left(\frac{\frac{1}{2}(e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N}{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N} + \frac{\frac{1}{2}(e_{i-1}^T - e_i^T)p_N}{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N} \left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N} - 1\right)\right). \end{aligned}$$

By the bounds above,

$$\frac{\frac{1}{2}(e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N}{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N} \geq -\frac{2K}{N^2 c_L}$$

and

$$\begin{aligned} \frac{\frac{1}{2}(e_{i-1}^T - e_i^T)p_N}{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N} \left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N} - 1 \right) &= \frac{\frac{1}{2}(e_{i-1}^T - e_i^T)p_N}{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N} \left(\frac{\frac{1}{2}(e_{i+1}^T - e_{i-1}^T)p_N}{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N} \right) \\ &\geq -\frac{N^2}{c_L^2} \frac{1}{(N+1)^4} (K_2)^2 \\ &\geq -\left(\frac{K_2}{2Nc_L}\right)^2. \end{aligned}$$

Therefore,

$$N^2 \left| \ln\left(\frac{\frac{1}{2}(e_{i+1}^T + e_i^T)p_N}{e_i^T p_N}\right) + \ln\left(\frac{\frac{1}{2}(e_{i-1}^T + e_i^T)p_N}{e_i^T p_N}\right) \right| \leq \frac{2K}{c_L} + \left(\frac{K_2}{2c_L}\right)^2.$$

For the end-points,

$$\frac{\frac{1}{2}(e_1^T - e_0^T)q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N} \leq \ln\left(\frac{\frac{1}{2}(e_1^T + e_0^T)q_N}{e_0^T q_N}\right) \leq \frac{\frac{1}{2}(e_1^T - e_0^T)q_N}{e_0^T q_N}$$

and therefore

$$\left| \ln\left(\frac{\frac{1}{2}(e_1^T + e_0^T)q_N}{e_0^T q_N}\right) \right| \leq \frac{K_2}{(N+1)c_L} \leq \frac{K_2}{Nc_L}.$$

A similar property holds for the other endpoint, and therefore the claim holds for $K_1 = \max\left(\frac{K_2}{c_L}, \frac{2K}{c_L} + \left(\frac{K_2}{2c_L}\right)^2\right)$. \square

Lemma 12. Let $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$ be a sequence of probability distributions over the state spaces associated with Theorem 4. Define the functions $\hat{p}_N \in \mathcal{P}([0, 1])$ as, for $x \in$

$$\left[\frac{1}{2(N+1)}, 1 - \frac{1}{2(N+1)}\right),$$

$$\begin{aligned}\hat{p}_N(x) &= (N+1)\left((N+1)x + \frac{1}{2} - \lfloor (N+1)x + \frac{1}{2} \rfloor\right) e^T_{\lfloor (N+1)x + \frac{1}{2} \rfloor} p_N + \\ &\quad + (N+1)\left(\frac{1}{2} - (N+1)x + \lfloor (N+1)x + \frac{1}{2} \rfloor\right) e^T_{\lfloor (N+1)x + \frac{1}{2} \rfloor - 1} p_N,\end{aligned}$$

and, for $x \in [0, \frac{1}{2(N+1)})$,

$$\hat{p}_N(x) = (N+1)e_0^T q_N,$$

and, for $x \in [1 - \frac{1}{2(N+1)}, 1]$,

$$\hat{p}_N(x) = (N+1)e_N^T q_N.$$

If the sequence $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$ satisfies the convergence condition (Definition 1), then there exists a sub-sequence, whose elements we denote by n , such that:

- i) $p_n(x)$ converges point-wise to a differentiable function $p(x) \in \mathcal{P}([0, 1])$, whose derivative is Lipschitz-continuous, with $p(x) > 0$ for all $x \in [0, 1]$,
- ii) the following sum converges:

$$\lim_{n \rightarrow \infty} n^2 \sum_{i \in X^n \setminus \{n\}} \{g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T) p_N)\} = \frac{1}{4} \int_0^1 \frac{(p'(x))^2}{p(x)} dx,$$

where $g(x) = x \ln(x)$,

- iii) for all $a \in A$, $\lim_{n \rightarrow \infty} u_{a,n}^T p_n = \int_0^1 u_a(x) p(x) dx$, and

- iv) if the sequence $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$ is constructed from some function $\tilde{p}(x)$, as in Lemma 11, then $p(x) = \tilde{p}(x)$ for all $x \in [0, 1]$.

Proof. We begin by noting that the functions $\hat{p}_N(x)$ are absolutely continuous. Almost

everywhere in $[\frac{1}{2(N+1)}, 1 - \frac{1}{2(N+1)}]$,

$$\hat{p}'_N(x) = (N+1)^2 (e^T_{\lfloor (N+1)x + \frac{1}{2} \rfloor} - e^T_{\lfloor (N+1)x + \frac{1}{2} \rfloor - 1}) p_N,$$

and outside this region, $\hat{p}'_N(x) = 0$. Let $f'_N(x)$ denote the right-continuous Lebesgue-integrable function on $[0, 1]$ such that

$$\hat{p}_N(x) = \hat{p}_N(0) + \int_0^x f'_N(y) dy,$$

which is equal to $\hat{p}'_N(x)$ anywhere the latter exists.

The total variation of $f'_N(x)$ is equal to

$$\begin{aligned} TV(f'_N) &= \sum_{i=1}^{N-1} (N+1)^2 |(e^T_{i+1} + e^T_{i-1} - 2e^T_i) p_N| + \\ &+ (N+1)^2 |(e^T_N - e^T_{N-1}) p_N| + (N+1)^2 |(e^T_1 - e^T_0) p_N|. \end{aligned}$$

By the convergence condition,

$$TV(f'_N) \leq \frac{(N+1)^3}{N^3} 2K_1,$$

and therefore the sequence of functions $f'_N(x)$ has uniformly bounded variation. The function is also uniformly bounded at the end points, and therefore Helly's selection theorem applies. That is, there exists a sub-sequence, which we denote by n , such that $f'_n(x)$ converges point-wise to some $p'(x)$.

For any $1 - \frac{1}{2(N+1)} > x > y \geq \frac{1}{2(N+1)}$, the quantity

$$\begin{aligned} |f'_N(x) - f'_N(y)| &= (N+1)^2 \left| \sum_{i=\lfloor (N+1)y+\frac{1}{2} \rfloor}^{\lfloor (N+1)x+\frac{1}{2} \rfloor} (e_{i+1}^T + e_{i-1}^T - 2e_i^T) p_N \right| \\ &\leq \frac{(N+1)^2((N+1)(x-y)+2)}{N^3} 2K_1. \end{aligned}$$

At the end points, for all $x \in [0, \frac{1}{2(N+1)})$,

$$|f'_N(\frac{1}{2(N+1)}) - f'_N(x)| \leq \frac{2K_1}{N+1},$$

and for all $x \in [1 - \frac{1}{2(N+1)}, 1]$,

$$|f'_N(x) - \lim_{y \uparrow 1 - \frac{1}{2(N+1)}} f'_N(y)| \leq \frac{2K_1}{N+1}.$$

Therefore, by the point-wise convergence of f'_n to f'_n , for all $x > y$,

$$|f'(x) - f'(y)| \leq 2K_1(x-y),$$

meaning that f' is Lipschitz-continuous. By the fact that $f'(0) = 0$, this implies that $|f'(x)| \leq 2K_1$ for all $x \in [0, 1]$.

By the convergence condition, $c_L \leq \hat{p}_N(0) \leq c_H$. Therefore, there exists a convergent sub-sequence. We now use n to denote the sub-sequence for which $\lim_{n \rightarrow \infty} \hat{p}_n(0) = p(0)$ and for which $f'_n(x)$ converges point-wise to $p'(x)$. By the dominated convergence theorem, for all $x \in [0, 1]$,

$$\lim_{n \rightarrow \infty} \hat{p}_n(x) = \lim_{n \rightarrow \infty} \left\{ \hat{p}_n(0) + \int_0^x f'_n(y) dy \right\} = p(0) + \int_0^x p'(y) dy.$$

Define the function $p(x) = p(0) + \int_0^x p'(y)dy$ for all $x \in [0, 1]$. By the convergence conditions, this function is bounded, $0 < c_L \leq p(x) \leq c_H$, by construction it is differentiable, and its derivative is Lipschitz continuous. Moreover,

$$\int_0^1 p(x)dx = 1,$$

and therefore $p \in \mathcal{P}([0, 1])$.

Next, consider the limiting cost function. We have, Taylor-expanding,

$$g(y) = g(x) + g'(x)(y-x) + \frac{1}{2}g''(cy + (1-c)x)(y-x)^2$$

for some $c \in (0, 1)$. Therefore,

$$\begin{aligned} g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g\left(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N\right) = \\ \frac{1}{8}g''(c_1 e_i^T p_N + (1-c_1)\frac{1}{2}(e_i^T + e_{i+1}^T)p_N)((e_{i+1}^T - e_i^T)p_N)^2 \\ + \frac{1}{8}g''(c_2 e_i^T p_N + (1-c_2)\frac{1}{2}(e_i^T + e_{i+1}^T)p_N)((e_{i+1}^T - e_i^T)p_N)^2 \end{aligned}$$

for constants $c_1, c_2 \in (0, 1)$. Note that, by the boundedness $\hat{p}_N(x)$ from below, $e_i^T p_N \geq (N+1)^{-1}c_L$ for all $i \in X^N$. It follows that

$$g''(c_1 e_i^T p_N + (1-c_1)\frac{1}{2}(e_i^T + e_{i+1}^T)p_N) = \frac{1}{c_1 e_i^T p_N + (1-c_1)\frac{1}{2}(e_i^T + e_{i+1}^T)p_N} \leq (N+1)c_L.$$

Therefore,

$$0 \leq g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g\left(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N\right) \leq \frac{(N+1)c_L}{4}((e_{i+1}^T - e_i^T)p_N)^2.$$

By construction,

$$e_i^T p_N = \frac{1}{(N+1)} \hat{p}_N\left(\frac{2i+1}{2(N+1)}\right).$$

Therefore,

$$(N+1)(g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N)) = \\ g(\hat{p}_N(\frac{2i+1}{2(N+1)})) + g(\hat{p}_N(\frac{2i+3}{2(N+1)})) - 2g(\hat{p}_N(\frac{2i+2}{2(N+1)})).$$

and

$$g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N) \leq \frac{c_L}{4(N+1)} (\hat{p}(\frac{2i+3}{2(N+1)}) - \hat{p}(\frac{2i+1}{2(N+1)}))^2.$$

By the boundedness of $f'_N(x)$,

$$g(\hat{p}(\frac{2i+1}{2(N+1)})) + g(\hat{p}(\frac{2i+3}{2(N+1)})) - 2g(\hat{p}(\frac{2i+2}{2(N+1)})) \leq \frac{K_1^2 c_L}{(N+1)^2}.$$

Writing the limiting cost as an integral, and switching to the sub-sequence n defined above,

$$n^2 \sum_{i \in X^n \setminus \{n\}} \{g(e_i^T p_n) + g(e_{i+1}^T p_n) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_n)\} = \\ \frac{n^3}{n+1} \int_0^1 \{g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)})) + g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}))\} dx.$$

By the bound above,

$$\frac{n^3}{n+1} \int_0^1 \{g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)})) + g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}))\} dx \leq \\ \frac{n^3}{(n+1)^3} \int_0^1 K_1^2 c_L dx.$$

Applying the dominated convergence theorem,

$$\lim_{n \rightarrow \infty} n^2 \sum_{i \in X^n \setminus \{n\}} \{g(e_i^T p_n) + g(e_{i+1}^T p_n) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_n)\} =$$

$$\int_0^1 \lim_{n \rightarrow \infty} \frac{n^3}{n+1} \{g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)})) + g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}))\} dx.$$

By the Taylor expansion above,

$$\lim_{n \rightarrow \infty} \frac{n^3}{n+1} \{g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)})) + g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}))\} =$$

$$\lim_{n \rightarrow \infty} \frac{1}{8} \frac{n^3}{n+1} \{g''(\cdot) + g''(\cdot)\} (\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)}) - \hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)}))^2.$$

By definition,

$$(n+1)(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)}) - \hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)})) = f'_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)})$$

and

$$\lim_{n \rightarrow \infty} g''(\hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}) + c_n(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)}) - \hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}))) = \frac{1}{p(x)},$$

with $c_n \in (0, 1)$ for all n , and therefore

$$\lim_{n \rightarrow \infty} \frac{n^3}{n+1} \{g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 1}{2(n+1)})) + g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2\lfloor nx \rfloor + 2}{2(n+1)}))\} =$$

$$\lim_{n \rightarrow \infty} \frac{1}{4} \frac{(p'(x))^2}{p(x)},$$

proving the second claim.

Turning to the third claim, recall that, by definition,

$$e_i^T u_{a,N} = \frac{\int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} u_a(x) f(x) dx}{\int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} f(x) dx.}$$

We define the function, for $x \in [0, 1)$, as

$$u_{a,N}(x) = e_{\lfloor (N+1)x \rfloor}^T u_{a,N},$$

and let $u_{a,N}(1) = e_N^T u_{a,N}$. We also define the function

$$\tilde{x}(x) = \frac{2\lfloor (N+1)x \rfloor + 1}{2(N+1)}.$$

By construction, $\hat{p}_N(\tilde{x}(x)) = (N+1)e_{\lfloor (N+1)x \rfloor}^T p_{a,N}$ for all $x \in [0, 1)$, and equals $e_N^T p_{a,N}$ for $x = 1$. Therefore,

$$\begin{aligned} u_{a,N}^T p_N &= \sum_{i \in X^N} (e_i^T u_{a,N})(e_i^T p_N) \\ &= \int_0^1 \hat{p}_N(\tilde{x}(x)) u_{a,N}(x) dx. \end{aligned}$$

By the measurability of $u_a(x)$,

$$\lim_{N \rightarrow \infty} u_{a,N}(x) = u_a(x).$$

Therefore, by the boundedness of utilities and the dominated convergence theorem,

$$\lim_{n \rightarrow \infty} u_{a,n}^T p_n = \int_0^1 p(x) u_a(x) dx.$$

Finally, suppose that, for all N

$$e_i^T p_{a,N} = \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \tilde{p}(x) dx.$$

It follows that $\lim_{n \rightarrow \infty} \hat{p}_{a,N}(x) = \tilde{p}(x)$ for all $x \in X$, and therefore $\tilde{p}(x) = p(x)$. \square

Lemma 13. *Let $\pi_N(a) \in \mathcal{P}(A)$ and $\{q_{a,N} \in \mathcal{P}(X^N)\}_{a \in A}$ denote optimal policies in the discrete state setting described in section A.2. For each $a \in A$, the sequence $\{q_{a,N}\}$ satisfies the convergence condition (Definition 1).*

Proof. We begin by noting that the conditions given for the function $f(x)$ satisfy the conditions of Lemma 11, and therefore the sequence q_N satisfies the convergence condition. We will use the constants c_H and c_L to refer to its bounds,

$$\frac{c_H}{N+1} \geq e_i^T q_N \geq \frac{c_L}{N+1},$$

and the constants K_1 and K to refer to the constants described by convergence condition and Lemma 11 for the sequence q_N . By the convention that $q_{a,N} = q_N$ if $\pi_N(a) = 0$, $q_{a,N}$ also satisfies the convergence condition whenever $\pi_N(a) = 0$.

The problem of size N is

$$V_N(q_N; \bar{\theta}) = \max_{\pi_N \in \mathcal{P}(A), \{q_{a,N} \in \mathcal{P}(X^N)\}_{a \in A}} \sum_{a \in A} \pi_N(a) (u_{a,N}^T \cdot q_{a,N}) - \bar{\theta} \sum_{a \in A} \pi_N(a) D_N(q_{a,N} || q_N)$$

subject to

$$\sum_{a \in A} \pi_N(a) q_{a,N} = q_N.$$

Let u_n denote that $|X^N| \times |A|$ matrix whose columns are $u_{a,N}$. Using Lemma 5, we can

rewrite the problem as

$$\begin{aligned}
V_N(q_N; \bar{\theta}) &= \max_{\{p_{x,N} \in \mathcal{P}(A)\}_{i \in X}} \sum_{a \in A} e_a^T p D(q) u_N e_a \\
&\quad - \bar{\theta} N^2 \sum_{i=0}^{N-1} (e_i^T q_N) D_{KL}(p_{i,N} \parallel \frac{p_{i,N}(e_i^T q_N) + p_{i+1,N}(e_{i+1}^T q_N)}{(e_i^T + e_{i+1}^T) q_N}) \\
&\quad - \bar{\theta} N^2 \sum_{i=1}^N (e_i^T q_N) D_{KL}(p_{i,N} \parallel \frac{p_{i,N}(e_i^T q_N) + p_{i-1,N}(e_{i-1}^T q_N)}{(e_i^T + e_{i-1}^T) q_N}) \\
&\quad - \bar{\theta} N^{-1} \sum_{i=0}^{N-1} (e_i^T q_N) D_{KL}(p_{i,N} \parallel p_N q_N).
\end{aligned}$$

The FOC for this problem is, for all $i \in [1, N-1]$ and $a \in A$ such that $e_a^T p_{i,N} > 0$,

$$\begin{aligned}
& e_i^T u_{a,N} - \bar{\theta} N^2 \ln\left(\frac{e_a^T p_{i,N}(e_i^T + e_{i+1}^T) q_N}{e_a^T (p_{i,N}(e_i^T q_N) + p_{i+1,N}(e_{i+1}^T q_N))}\right) \\
& - \bar{\theta} N^2 \ln\left(\frac{e_a^T p_{i,N}(e_i^T + e_{i-1}^T) q_N}{e_a^T (p_{i,N}(e_i^T q_N) + p_{i-1,N}(e_{i-1}^T q_N))}\right) - \bar{\theta} \ln\left(\frac{e_a^T p_{i,N}}{e_a^T p_N q_N}\right) - e_i^T \kappa_N = 0,
\end{aligned}$$

where $\kappa_N \in \mathbb{R}^{N+1}$ are the multipliers (scaled by $e_i^T q_N$) on the constraints that $\sum_{a \in A} e_a^T p_{i,N} = 1$ for all $i \in X$. Defining $q_{-1,N} = q_{N+1,N} = 0$, and defining $p_{-1,N}$ and $p_{N+1,N}$ in arbitrary fashion, we can recover this FOC for all $i \in X$.

Rewriting the FOC in terms of the posteriors, for any a such that $\pi_N(a) > 0$,

$$\begin{aligned}
e_i^T (u_{a,N} - \kappa_N) &= -\bar{\theta} N^2 \ln\left(\frac{(e_i^T q_{a,N})(1 + \frac{e_{i+1}^T q_N}{e_i^T q_N})}{(e_{i+1} + e_i)^T q_{a,N}}\right) - \bar{\theta} N^2 \ln\left(\frac{(e_i^T q_{a,N})(1 + \frac{e_{i-1}^T q_N}{e_i^T q_N})}{(e_{i-1} + e_i)^T q_{a,N}}\right) - \bar{\theta} \ln N^{-1} \left(\frac{e_a^T p_{i,N}}{e_a^T p_N q_N}\right) \\
&= \bar{\theta} N^2 \ln\left(1 + \frac{e_{i+1}^T q_{a,N}}{e_i^T q_{a,N}}\right) - \bar{\theta} N^2 \ln\left(1 + \frac{e_{i+1}^T q_N}{e_i^T q_N}\right) + \bar{\theta} N^2 \ln\left(1 + \frac{e_{i-1}^T q_{a,N}}{e_i^T q_{a,N}}\right) \\
&\quad - \bar{\theta} N^2 \ln\left(1 + \frac{e_{i-1}^T q_N}{e_i^T q_N}\right) - \bar{\theta} \ln N^{-1} \left(\frac{e_i^T q_{a,N}}{e_i^T q_N}\right) \\
&= \bar{\theta} N^2 \left(\ln\left(\frac{1}{2}(e_{i+1}^T + e_i^T) q_{a,N}\right) + \ln\left(\frac{1}{2}(e_{i-1}^T + e_i^T) q_{a,N}\right) - (2 + N^{-3}) \ln(e_i^T q_{a,N}) + 2 \ln 2\right) \\
&\quad - \bar{\theta} N^2 \left(\ln\left(\frac{1}{2}(e_{i+1}^T + e_i^T) q_N\right) + \ln\left(\frac{1}{2}(e_{i-1}^T + e_i^T) q_N\right) - (2 + N^{-3}) \ln(e_i^T q_N) + 2 \ln 2\right).
\end{aligned}$$

Using Lemma 11, for all $i \in X^N \setminus \{0, N\}$,

$$N^2 |\ln(\frac{1}{2}(e_{i+1}^T + e_i^T)q_N) + \ln(\frac{1}{2}(e_{i-1}^T + e_i^T)q_N) - 2\ln(e_i^T q_N)| \leq K.$$

By the boundedness of the utility function,

$$e_i^T \kappa_N \geq -\bar{u} - \bar{\theta}K + \bar{\theta}N^2 (\ln(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}) + \ln(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}})) + \bar{\theta}N^{-1} \ln(\frac{e_i^T q_{a,N}}{e_i^T q_N}).$$

By the concavity of the log function,

$$\begin{aligned} \ln(\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}) + \ln(\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}) + N^{-3} \ln(e_i^T q_N) \leq \\ (2 + N^{-3}) \ln(\frac{1}{2(2 + N^{-3})}(e_{i+1}^T + e_{i-1}^T + 2e_i^T)q_{a,N}) + \frac{N^{-3}}{2 + N^{-3}} \ln(e_i^T q_N). \end{aligned}$$

and therefore

$$\begin{aligned} \ln(\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}) + \ln(\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}) + N^{-3} \ln(e_i^T q_N) - (2 + N^{-3}) \ln(e_i^T q_{a,N}) \\ \leq (2 + N^{-3}) \ln(\frac{\frac{1}{2(2 + N^{-3})}(e_{i+1}^T + e_{i-1}^T + 2e_i^T)q_{a,N} + \frac{N^{-3}}{2 + N^{-3}} e_i^T q_N}{e_i^T q_{a,N}}). \end{aligned}$$

It follows that

$$e_i^T \kappa_N \geq -\bar{u} - \bar{\theta}K - (2 + N^{-3})\bar{\theta}N^2 \ln(\frac{\frac{1}{2(2 + N^{-3})}(e_{i+1}^T + e_{i-1}^T + 2e_i^T)q_{a,N} + \frac{N^{-3}}{2 + N^{-3}} e_i^T q_N}{e_i^T q_{a,N}}).$$

Exponentiating,

$$(e_i^T q_{a,N}) \exp\left(-\frac{1}{2+N^{-3}} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + e_i^T \kappa_N)\right) \leq \frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_{a,N} + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N. \quad (32)$$

Summing over a , weighted by $\pi_N(a)$,

$$(e_i^T q_N) \exp\left(-\frac{1}{2+N^{-3}} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + e_i^T \kappa_N)\right) \leq \frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_N + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N.$$

Taking logs,

$$\begin{aligned} -\frac{1}{2+N^{-3}} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + e_i^T \kappa_N) &\leq \ln\left(\frac{\frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_N + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N}{(e_i^T q_N)}\right) \\ &\leq \ln\left(1 + \frac{N^{-3}}{2+N^{-3}} + \frac{1}{2+N^{-3}} \frac{K_1 N^{-3}}{c_L N^{-1}}\right), \end{aligned}$$

where the last step follows by Lemma 11, recalling that c_L is the lower bound on $f(x)$. We have

$$\begin{aligned} e_i^T \kappa_N &\geq -2\bar{\theta} N^2 \ln\left(1 + \frac{N^{-3}}{2+N^{-3}} + \frac{1}{2+N^{-3}} \frac{K_1}{c_L} N^{-2}\right) - \bar{u} - \bar{\theta} K \\ &\geq -\bar{u} - \bar{\theta} K - \frac{N^{-1}}{2+N^{-3}} - \frac{1}{2+N^{-3}} \frac{K_1}{c_L} \\ &\geq -\bar{u} - \bar{\theta} K - \frac{1}{2} - \frac{1}{2} \frac{K_1}{c_L}. \end{aligned}$$

where the second step follows by the inequality $\ln(1+x) < x$ for $x > 0$.

Turning to the end points, the FOC can be simplified to

$$\begin{aligned} e_0^T(u_{a,N} - \kappa_N) &= \bar{\theta}N^2(\ln(\frac{1}{2}(e_1^T + e_0^T)q_{a,N}) - \ln(e_0^T q_{a,N})) \\ &\quad - \bar{\theta}N^2(\ln(\frac{1}{2}(e_1^T + e_0^T)q_N) - \ln(e_0^T q_N)) - \bar{\theta}N^{-1} \ln(\frac{e_0^T q_{a,N}}{e_0^T q_N}). \end{aligned}$$

By the concavity of the log function,

$$\begin{aligned} \ln(\frac{1}{2}(e_1^T + e_0^T)q_{a,N}) + N^{-3} \ln(e_0^T q_N) - (1 + N^{-3}) \ln(e_0^T q_{a,N}) \\ \leq (1 + N^{-3}) \ln(\frac{\frac{1}{(1+N^{-3})} \frac{1}{2}(e_1^T + e_0^T)q_{a,N} + \frac{N^{-3}}{1+N^{-3}} e_0^T q_N}{e_0^T q_{a,N}}). \quad (33) \end{aligned}$$

Therefore,

$$\begin{aligned} &\bar{\theta}n^2 \ln(\frac{\frac{1}{2}(e_1^T + e_0^T)q_{a,N}}{e_0^T q_{a,N}}) + \bar{\theta}n^{-1} \ln(\frac{e_0^T q_N}{e_0^T q_{a,N}}) - \bar{\theta}K \\ &\leq e_0^T(u_{a,N} - \kappa_N) + \bar{\theta}N^2(\ln(\frac{1}{2}(e_1^T + e_0^T)q_N) - \ln(e_0^T q_N)) \\ &\leq \bar{\theta}(1 + N^{-3}) \ln(\frac{\frac{1}{(1+N^{-3})} \frac{1}{2}(e_1^T + e_0^T)q_{a,N} + \frac{N^{-3}}{1+N^{-3}} e_0^T q_N}{e_0^T q_{a,N}}) + \bar{\theta}K. \end{aligned}$$

By the boundedness of the utility function,

$$\begin{aligned} &-\bar{\theta}(1 + N^{-3}) \ln(\frac{\frac{1}{(1+N^{-3})} \frac{1}{2}(e_1^T + e_0^T)q_{a,N} + \frac{N^{-3}}{1+N^{-3}} e_0^T q_N}{e_0^T q_{a,N}}) - \bar{u} \\ &\leq e_0^T \kappa_N + \bar{\theta}N^2 \ln(\frac{e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}) \\ &\leq -\bar{\theta}N^2 \ln(\frac{\frac{1}{2}(e_1^T + e_0^T)q_{a,N}}{e_0^T q_{a,N}}) + \bar{\theta}N^{-1} \ln(\frac{e_0^T q_{a,N}}{e_0^T q_N}) + \bar{u}. \end{aligned}$$

By the inequality $\ln(x) \leq x - 1$,

$$\begin{aligned}\bar{\theta}N^{-1}\ln\left(\frac{e_0^T q_{a,N}}{e_0^T q_N}\right) &\leq \bar{\theta}N^{-1}\left(\frac{e_0^T q_{a,N}}{e_0^T q_N} - 1\right) \\ &\leq \bar{\theta}c_L^{-1},\end{aligned}$$

where the latter follows from $e_0^T q_N \geq c_L N^{-1}$. Exponentiating,

$$\begin{aligned}(e_0^T q_{a,N}) \exp(-\bar{\theta}^{-1}(1+N^{-3})^{-1}N^{-2}\bar{\mu}) &\leq \\ \left(\frac{1}{(1+N^{-3})}\frac{1}{2}(e_1^T + e_0^T)q_{a,N} + \frac{N^{-3}}{1+N^{-3}}e_0^T q_N\right) \exp(\bar{\theta}^{-1}(1+N^{-3})^{-1}N^{-2}e_0^T \kappa_N) &\frac{e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}\end{aligned}$$

and

$$\left(\frac{1}{2}(e_1^T + e_0^T)q_{a,N}\right) \exp(\bar{\theta}^{-1}N^{-2}e_0^T \kappa_N) \frac{e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N} \leq (e_0^T q_{a,N}) \exp(\bar{\theta}^{-1}N^{-2}(\bar{\mu} + \bar{\theta}c_L^{-1})).$$

Summing over a , weighted by $\pi_N(a)$,

$$\begin{aligned}(e_0^T q_N) \exp(-\bar{\theta}^{-1}(1+N^{-3})^{-1}N^{-2}\bar{\mu}) &\leq \\ \left(\frac{1}{(1+N^{-3})}\frac{1}{2}(e_1^T + e_0^T)q_N + \frac{N^{-3}}{1+N^{-3}}e_0^T q_N\right) \exp(\bar{\theta}^{-1}(1+N^{-3})^{-1}N^{-2}e_0^T \kappa_N) &\frac{e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N},\end{aligned}$$

$$\left(\frac{1}{2}(e_1^T + e_0^T)q_N\right) \exp(\bar{\theta}^{-1}N^{-2}e_0^T \kappa_N) \frac{e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N} \leq (e_0^T q_N) \exp(\bar{\theta}^{-1}N^{-2}(\bar{\mu} + \bar{\theta}c_L^{-1})).$$

Taking logs,

$$-\bar{\theta}N^2(1+N^{-3})\left(\ln\left(\frac{\frac{1}{(1+N^{-3})}\frac{1}{2}(e_1^T + e_0^T)q_N + \frac{N^{-3}}{1+N^{-3}}e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}\right) - \bar{\mu}\right) \leq e_0^T \kappa_N \leq \bar{\mu} + \bar{\theta}c_L^{-1}.$$

We can write

$$\begin{aligned} \ln\left(\frac{\frac{1}{(1+N^{-3})} \frac{1}{2}(e_1^T + e_0^T)q_N + \frac{N^{-3}}{1+N^{-3}} e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}\right) &= \ln\left(\frac{1}{1+N^{-3}} + \frac{\frac{N^{-3}}{1+N^{-3}} e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}\right) \\ &\leq \frac{1}{1+N^{-3}} + \frac{2N^{-3}}{1+N^{-3}} - 1. \end{aligned}$$

Therefore,

$$-\bar{\theta}N^2(1+N^{-3})\left(\ln\left(\frac{\frac{1}{(1+N^{-3})} \frac{1}{2}(e_1^T + e_0^T)q_N + \frac{N^{-3}}{1+N^{-3}} e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}\right)\right) \geq -\bar{\theta}N^{-1} \geq -\bar{\theta}.$$

By Lemma 11,

$$-\bar{\theta} - \bar{u} \leq e_0^T \kappa_N \leq \bar{u} + \bar{\theta}c_L^{-1}.$$

A similar argument applies to the other end-point ($e_N^T \kappa_N$). Summarizing, $e_i^T \kappa_N \geq -B_L$ for some constant $B_L > 0$, and $e_i^T \kappa_N \leq B_H$ for some $B_H > 0$ if $i \in \{0, N\}$.

Returning to the FOC, for all $i \in X^N \setminus \{0, N\}$,

$$e_i^T \kappa_N \leq \bar{u} + \bar{\theta}K + \bar{\theta}N^2\left(\ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right) + \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}}\right)\right) + \bar{\theta}N^{-1} \ln\left(\frac{e_i^T q_{a,N}}{e_i^T q_N}\right),$$

and as argued above,

$$\bar{\theta}N^{-1} \ln\left(\frac{e_i^T q_{a,N}}{e_i^T q_N}\right) \leq \bar{\theta}c_L^{-1}.$$

Using this bound,

$$\bar{\theta}N^2\left(\ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right) + \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}}\right)\right) \geq -(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).$$

For the end-points, the FOC requires that

$$e_0^T \kappa_N \leq \bar{u} - \bar{\theta} N^2 \ln\left(\frac{e_0^T q_N}{\frac{1}{2}(e_1^T + e_0^T)q_N}\right) + \bar{\theta} N^2 \ln\left(\frac{e_0^T q_{a,N}}{\frac{1}{2}(e_1^T + e_0^T)q_{a,N}}\right) + \bar{\theta} N^{-1} \ln\left(\frac{e_0^T q_{a,N}}{e_0^T q_N}\right)$$

and

$$e_N^T \kappa_N \leq \bar{u} - \bar{\theta} N^2 \ln\left(\frac{e_N^T q_{a,N}}{\frac{1}{2}(e_N^T + e_{N-1}^T)q_{a,N}}\right) + \bar{\theta} N^2 \ln\left(\frac{e_N^T q_{a,N}}{\frac{1}{2}(e_N^T + e_{N-1}^T)q_{a,N}}\right) + \bar{\theta} N^{-1} \ln\left(\frac{e_N^T q_{a,N}}{e_N^T q_N}\right).$$

Using Lemma 11, we can rewrite these inequalities as

$$\begin{aligned} \bar{\theta} N \ln\left(\frac{e_N^T q_{a,N}}{\frac{1}{2}(e_N^T + e_{N-1}^T)q_{a,N}}\right) &\geq -N^{-1}(\bar{u} + B_L + \bar{\theta} c_L^{-1}) + \bar{\theta} N \ln\left(\frac{e_N^T q_{a,N}}{\frac{1}{2}(e_N^T + e_{N-1}^T)q_{a,N}}\right) \\ &\geq -N^{-1}(\bar{u} + B_L + \bar{\theta} c_L^{-1}) - \bar{\theta} K \\ &\geq -(\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}), \end{aligned}$$

and likewise

$$\bar{\theta} N \ln\left(\frac{e_0^T q_{a,N}}{\frac{1}{2}(e_1^T + e_0^T)q_{a,N}}\right) \geq -(\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}).$$

Define $\hat{q}_{a,N}(x)$ as in Lemma 12. Define the function

$$l_{a,N}(x) = (N+1) \left(\ln(\hat{q}_{a,N}(x)) - \ln\left(\hat{q}_{a,N}\left(x - \frac{1}{2(N+1)}\right)\right) \right)$$

for any $x \in [\frac{1}{2(N+1)}, 1]$. For any $i \in X^N \setminus \{0\}$,

$$l_{a,N}\left(\frac{2i+1}{2(N+1)}\right) = (N+1) \ln\left(\frac{(N+1)e_i^T q_{a,N}}{\frac{1}{2}(N+1)(e_i^T + e_{i-1}^T)q_{a,N}}\right),$$

and for any $i \in X^N \setminus \{N\}$,

$$l_{a,N}\left(\frac{2i+2}{2(N+1)}\right) = (N+1) \ln\left(\frac{\frac{1}{2}(N+1)(e_i^T + e_{i+1}^T)q_{a,N}}{(N+1)e_i^T q_{a,N}}\right).$$

Therefore, for any $i \in X^N \setminus \{0, N\}$, the lower bound can be written as

$$\bar{\theta} \frac{N^2}{N+1} (l_{a,N}\left(\frac{2i+2}{2(N+1)}\right) - l_{a,N}\left(\frac{2i+1}{2(N+1)}\right)) \leq (\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).$$

The lower endpoint bound is

$$\bar{\theta} \frac{N}{N+1} l_{a,N}\left(\frac{1}{(N+1)}\right) \leq (\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).$$

The upper endpoint bound is

$$\bar{\theta} \frac{N}{N+1} l_{a,N}(1) \geq -(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).$$

We also have, for all $i \in X^N \setminus \{N\}$

$$\begin{aligned} & \bar{\theta} \frac{N^2}{N+1} (l_{a,N}\left(\frac{2i+3}{2(N+1)}\right) - l_{a,N}\left(\frac{2i+2}{2(N+1)}\right)) \\ &= \bar{\theta} N^2 \left(\ln\left(\frac{(N+1)e_{i+1}^T q_{a,N}}{\frac{1}{2}(N+1)(e_{i+1}^T + e_i^T)q_{a,N}}\right) - \ln\left(\frac{\frac{1}{2}(N+1)(e_i^T + e_{i+1}^T)q_{a,N}}{(N+1)e_i^T q_{a,N}}\right) \right) \\ & \leq 0, \end{aligned}$$

by the concavity of the log function. Therefore, for all $j \in \{2, 3, \dots, 2(N+1)\}$

$$\bar{\theta} \frac{N^2}{N+1} (l_{a,N}\left(\frac{j+1}{2(N+1)}\right) - l_{a,N}\left(\frac{j}{2(N+1)}\right)) \leq (\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).$$

It follows that, for all $j \in \{2, 3, \dots, 2(N+1)\}$

$$\begin{aligned} l_{a,N}\left(\frac{j}{2(N+1)}\right) &= l_{a,N}\left(\frac{2}{2(N+1)}\right) + \sum_{k=2}^{j-1} \left(l_{a,N}\left(\frac{k+1}{2(N+1)}\right) - l_{a,N}\left(\frac{k}{2(N+1)}\right) \right) \\ &\leq \bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}) \frac{N+1}{N} \left(1 + \frac{j-2}{N}\right). \end{aligned}$$

Similarly, for all $j \in \{2, 3, \dots, 2(N+1)\}$,

$$l_{a,N}(1) = l_{a,N}\left(\frac{j}{2(N+1)}\right) + \sum_{k=j-1}^{2N} \left(l_{a,N}\left(\frac{k+1}{2(N+1)}\right) - l_{a,n}\left(\frac{k}{2(N+1)}\right) \right)$$

and therefore

$$\begin{aligned} -l_{a,N}\left(\frac{j}{2(N+1)}\right) &= -l_{a,n}(1) + \sum_{k=j-1}^{2N} \left(l_{a,N}\left(\frac{k+1}{2(N+1)}\right) - l_{a,n}\left(\frac{k}{2(N+1)}\right) \right) \\ &\leq \bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}) \frac{N+1}{N} \left(1 + \frac{2(N+1)-j}{N^2}\right). \end{aligned}$$

It follows that

$$\begin{aligned} |l_{a,N}\left(\frac{j}{2(N+1)}\right)| &\leq 2\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}) \frac{N+1}{N} \\ &\leq 4\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}). \end{aligned}$$

Note that there must exist some $\tilde{i}_{a,N} \in X^N$ such that $e_{\tilde{i}_{a,N}}^T q_{a,N} \geq \frac{1}{N+1}$, implying that

$$\ln((N+1)e_{\tilde{i}_{a,n}}^T q_{a,N}) \geq 0.$$

By the definition of $l_{a,N}$, for any $i \in X^N \setminus \{0\}$,

$$l_{a,N}\left(\frac{2i+1}{2(N+1)}\right) + l_{a,N}\left(\frac{2i}{2(N+1)}\right) = (N+1) \ln\left(\frac{(N+1)e_i^T q_{a,N}}{(N+1)e_{i-1}^T q_{a,N}}\right).$$

For any $i > \tilde{i}_{a,N}$,

$$\begin{aligned}
\ln((N+1)e_i^T q_{a,N}) &= \ln((N+1)e_{\tilde{i}_{a,n}}^T q_{a,N}) + \sum_{j=\tilde{i}_{a,n}+1}^i \ln\left(\frac{(N+1)e_j^T q_{a,N}}{(N+1)e_{j-1}^T q_{a,N}}\right) \\
&= \ln((N+1)e_{\tilde{i}_{a,n}}^T q_{a,N}) + \frac{1}{N+1} \sum_{j=\tilde{i}_{a,n}+1}^i l_{a,N}\left(\frac{2j+1}{2(N+1)}\right) + l_{a,N}\left(\frac{2j}{2(N+1)}\right) \\
&\geq -\frac{1}{N+1} \sum_{j=\tilde{i}_{a,n}+1}^i 8\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}) \\
&\geq -8\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).
\end{aligned}$$

Similarly, for any $i < \tilde{i}_{a,N}$,

$$\ln((N+1)e_{\tilde{i}_{a,n}}^T q_{a,N}) = \ln((N+1)e_i^T q_{a,N}) + \sum_{j=i+1}^{\tilde{i}_{a,n}} \ln\left(\frac{(N+1)e_{j+1}^T q_{a,N}}{(N+1)e_j^T q_{a,N}}\right).$$

Therefore,

$$\begin{aligned}
\ln((N+1)e_i^T q_{a,N}) &\geq -\sum_{j=i+1}^{\tilde{i}_{a,n}} \ln\left(\frac{(N+1)e_{j+1}^T q_{a,N}}{(N+1)e_j^T q_{a,N}}\right) \\
&\geq -8\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).
\end{aligned}$$

Repeating this argument, there must be some $\hat{i}_{a,N}$ such that $e_{\hat{i}_{a,N}}^T q_{a,N} \leq N^{-1}$, and using the bounds on $l_{a,N}$ in similar fashion yields

$$\ln((N+1)e_{\hat{i}_{a,N}}^T q_{a,N}) \leq 8\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}).$$

It follows that there exists a constant $c \in (0, 1)$ such that, for all $N, a \in A$ such that $\pi_N(a) > 0$, and $i \in X^N$,

$$\frac{c^{-1}}{(N+1)} \geq e_i^T q_{a,N} \geq \frac{c}{N+1},$$

demonstrating that $q_{a,N}$ satisfies the first part of the convergence condition.

Using the bound on $l_{a,N}$, and a Taylor expansion, for some $a \in (0, 1)$

$$\begin{aligned} |(N+1) \ln\left(\frac{\frac{1}{2}(N+1)(e_i^T + e_{i+1}^T)q_{a,N}}{(N+1)e_i^T q_{a,N}}\right)| &= \frac{(N+1) \left| \frac{1}{2}(e_{i+1}^T - e_i^T)q_{a,N} \right|}{e_i^T q_{a,N} + \frac{a}{2}(e_{i+1}^T - e_i^T)q_{a,N}} \\ &\leq 4\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}), \end{aligned}$$

and therefore, by the bound on $e_i^T q_{a,N}$,

$$(N+1)^2 \left| \frac{1}{2}(e_{i+1}^T - e_i^T)q_{a,N} \right| \leq B$$

for some $B > 0$. By a similar argument,

$$(N+1)^2 \left| \frac{1}{2}(e_{i+1}^T - e_{i-1}^T)q_{a,N} \right| \leq 4B.$$

Returning to the first-order condition, for $i \in X^N \setminus \{0, N\}$, and using some of the bounds employed previously,

$$e_i^T \kappa_N \leq \bar{u} + \bar{\theta}K + \bar{\theta}c_L + \bar{\theta}N^2 \left(\ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right) + \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}}\right) \right).$$

By the inequality $\ln(x) \leq x - 1$,

$$e_i^T \kappa_N \leq \bar{u} + \bar{\theta}K + \bar{\theta}c_L + \bar{\theta}N^2 \left(\frac{\frac{1}{2}(e_i^T - e_{i+1}^T)q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}} + \frac{\frac{1}{2}(e_i^T - e_{i-1}^T)q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}} \right)$$

Multiplying through,

$$\begin{aligned}
& \frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}(e_i^T \kappa_N - \bar{u} - \bar{\theta}K - \bar{\theta}c_L) \\
& \leq \bar{\theta}N^2 \left(\frac{1}{2}(e_i^T - e_{i+1}^T)q_{a,N} + \frac{1}{2}(e_i^T - e_{i-1}^T)q_{a,N} \frac{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}} \right). \\
& \leq \bar{\theta}N^2 \left(\frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} + \frac{1}{2}(e_i^T - e_{i-1}^T)q_{a,N} \left(\frac{\frac{1}{2}(e_{i+1}^T - e_{i-1}^T)q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}} \right) \right).
\end{aligned}$$

Using the bounds above,

$$\begin{aligned}
\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}(e_i^T \kappa_N - \bar{u} - \bar{\theta}K - \bar{\theta}c_L) & \leq \bar{\theta}N^2 \left(\frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} + \frac{B}{(N+1)^2} \left(\frac{4B}{\frac{c}{N+1}} \right) \right) \\
& \leq \bar{\theta}N^2 \left(\frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} \right) + \frac{4B^2N^2}{c(N+1)^3}.
\end{aligned}$$

Therefore,

$$c(e_i^T \kappa_N - \bar{u} - \bar{\theta}K - \bar{\theta}c_L) \leq \bar{\theta} \frac{N+1}{N} N^3 \left(\frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} \right) + \frac{4B^2}{c}.$$

Summing over a , weighted by $\pi_N(a)$, and applying Lemma 11,

$$c(e_i^T \kappa_N - \bar{u} - \bar{\theta}K - \bar{\theta}c_L) \leq 2\bar{\theta}K_1 + \frac{4B^2}{c}.$$

Therefore, $|e_i^T \kappa_N|$ is bounded below by some $B_\kappa > 0$ for all $i \in X^N$ (recalling that this was shown for $i \in \{0, N\}$ previously). It also follows the term

$$\begin{aligned}
(N+1)^3 \left(\frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} \right) & \geq \frac{(N+1)^2}{N^2} c(e_i^T \kappa_N - \bar{u} - \bar{\theta}K - \bar{\theta}c_L - \frac{4B^2}{c^2}) \\
& \geq -2c(B_\kappa + \bar{u} + \bar{\theta}K + \bar{\theta}c_L + \frac{4B^2}{c^2})
\end{aligned}$$

is bounded below.

Recalling equation (32), and employing the upper bound on $|e_i^T \kappa_N|$,

$$\begin{aligned} (e_i^T q_{a,N}) \exp\left(-\frac{1}{2+N^{-3}} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + B_\kappa)\right) \\ \leq \frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_{a,N} + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N. \end{aligned}$$

Rewriting this,

$$\begin{aligned} (e_i^T q_{a,N}) \left(\exp\left(-\frac{1}{2+N^{-3}} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + B_\kappa)\right) - 1 \right) \\ \leq \frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_{a,N} + \frac{N^{-3}}{2+N^{-3}} e_i^T (q_N - q_{a,N}) \end{aligned}$$

By the upper bound on $e_i^T q_N \leq \frac{c_H}{N+1}$ and $e_i^T q_{a,N} \geq \frac{c}{N+1}$,

$$\begin{aligned} \frac{(N+1)^3}{2} (e_{i+1}^T + e_{i-1}^T - 2e_i^T) q_{a,N} \geq \\ (2+N^{-3})(N+1)^2 \left(\exp\left(-\frac{1}{2+N^{-3}} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + B_\kappa)\right) - 1 \right) - \frac{c_H - c}{N^3} (N+1)^2. \end{aligned}$$

By the inequality $\exp(x) - 1 \geq x$,

$$\begin{aligned} \frac{(N+1)^3}{2} (e_{i+1}^T + e_{i-1}^T - 2e_i^T) q_{a,N} \geq -\frac{(N+1)^2}{N^2} \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_\kappa) - \frac{c_H - c}{N^3} (N+1)^2 \\ \geq -2\bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_\kappa) - 2c_H + c. \end{aligned}$$

Therefore, the first statement in the second part of the convergence condition (Definition 1) is satisfied.

Finally, we consider the endpoints. The first-order condition is

$$\begin{aligned} \bar{\theta} N^2 (\ln(\frac{1}{2}(e_1^T + e_0^T)q_{a,N}) - \ln(e_0^T q_{a,N})) = \\ e_0^T (u_{a,N} - \kappa_N) + \bar{\theta} N^2 (\ln(\frac{1}{2}(e_1^T + e_0^T)q_N) - \ln(e_0^T q_N)) + \bar{\theta} N^{-1} \ln(\frac{e_0^T q_{a,N}}{e_0^T q_N}). \end{aligned}$$

We can bound this as

$$\begin{aligned} & -N^{-1}(\bar{u} + B_\kappa) - \bar{\theta} K + \bar{\theta} N^{-2} \ln(\frac{c}{c_H}) \\ & \leq \bar{\theta} N (\ln(\frac{1}{2}(e_1^T + e_0^T)q_{a,N}) - \ln(e_0^T q_{a,N})) \\ & \leq N^{-1}(\bar{u} + B_\kappa + \bar{\theta} c_L^{-1}) + \bar{\theta} K, \end{aligned}$$

and note that because $\sum_{i \in X^N} e_i^T q_{a,N} = \sum_{i \in X^N} e_i^T q_N = 1$, we must have $c_H \geq c$. Therefore,

$$\bar{\theta} \ln(\frac{c}{c_H}) \leq \bar{\theta} N^{-2} \ln(\frac{c}{c_H}).$$

Using a Taylor expansion,

$$\ln(\frac{1}{2}(e_1^T + e_0^T)q_{a,N}) - \ln(e_0^T q_{a,N}) = \frac{\frac{1}{2}(e_1^T - e_0^T)q_{a,N}}{e_0^T q_{a,N} + \frac{a}{2}(e_1^T + e_0^T)q_{a,N}}$$

for some $a \in (0, 1)$. Therefore,

$$N^2 |\frac{1}{2}(e_1^T - e_0^T)q_{a,N}| \leq \frac{c}{\bar{\theta}} (\bar{u} + B_\kappa + \bar{\theta} K + \bar{\theta} \max(\ln(\frac{c_H}{c}), c_L^{-1})).$$

A similar logic holds for the other endpoint, and therefore the convergence condition is satisfied. \square

D.11 Proof of Theorem 4

By the boundedness of $\mathcal{P}(A)$, there exists a convergent sub-sequence of the optimal policy $\pi_N(a)$, which we denote by n . Define

$$\pi(a) = \lim_{n \rightarrow \infty} \pi_n(a).$$

By Lemma 13, for all $a \in A$, each sequence of optimal policies $\{q_{a,N}\}$ satisfies the convergence condition (Definition 1). Therefore, by Lemma 12, each sequence $\{\hat{q}_{a,N}(x)\}$ has a convergent sub-sequence that converges to a differentiable function $f_a^*(x)$, whose derivative is Lipschitz continuous, with full support on $[0, 1]$. We can construct a sub-sequence in which $\pi_n(a)$ and all $\{\hat{q}_{a,n}(x)\}$ converge by iteratively applying this argument. Denote this sequence by n .

We can write the discrete value function as, using Lemma 5, as

$$\begin{aligned} V_N(q_N; \bar{\theta}) &= \max_{\{p_{x,N} \in \mathcal{P}(A)\}_{i \in X}} \sum_{a \in A} e_a^T p D(q) u_N e_a \\ &\quad - \bar{\theta} N^2 \sum_{a \in A} (e_a^T p q) \sum_{i=0}^{N-1} [(e_i^T q_{a,N}) \ln(\frac{e_i^T q_{a,N}}{\bar{q}_{i,a,N}}) + (e_{i+1}^T q_{a,N}) \ln(\frac{e_{i+1}^T q_{a,N}}{\bar{q}_{i,a,N}})] \\ &\quad + \bar{\theta} N^2 \sum_{i=0}^{N-1} [(e_i^T q_N) \ln(\frac{e_i^T q_N}{\bar{q}_{i,a,N}}) + (e_{i+1}^T q_N) \ln(\frac{e_{i+1}^T q_N}{\bar{q}_{i,a,N}})] \\ &\quad - \bar{\theta} N^{-1} \sum_{i=0}^{N-1} (e_i^T q_N) D_{KL}(p_{i,N} || p_N q_N). \end{aligned}$$

We can re-arrange this to

$$\begin{aligned}
V_N(q_N; \bar{\theta}) &= \max_{\{p_{x,N} \in \mathcal{P}(A)\}_{i \in X}} \sum_{a \in A} e_a^T p D(q) u_N e_a \\
&\quad - \bar{\theta} N^2 \sum_{a \in A} (e_a^T p q) \sum_{i=0}^{N-1} [g(e_i^T q_{a,N}) + g(e_{i+1}^T q_{a,N}) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T) q_{a,N})] \\
&\quad + \bar{\theta} N^2 \sum_{i=0}^{N-1} [g(e_i^T q_N) + g(e_{i+1}^T q_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T) q_N)] \\
&\quad - \bar{\theta} N^{-1} \sum_{i=0}^{N-1} (e_i^T q_N) D_{KL}(p_{i,N} || p_N q_N).
\end{aligned}$$

By Lemma 12 and the boundedness of the KL divergence,

$$\begin{aligned}
\lim_{n \rightarrow \infty} V_n(q_n; \bar{\theta}) &= \sum_{a \in A} \pi(a) \int_0^1 u_a(x) f_a(x) dx \\
&\quad - \frac{\bar{\theta}}{4} \sum_{a \in A} \left\{ \pi(a) \int_0^1 \frac{(f'_a(x))^2}{f_a(x)} dx \right\} + \frac{\bar{\theta}}{4} \int_0^1 \frac{(f'(x))^2}{f(x)} dx.
\end{aligned}$$

Suppose that $\pi(a)$ and the $f_a(x)$ functions do not maximize this expression (subject to the constraints stated in Theorem 4). Let $\pi^*(a)$ and $f_a^*(x)$ be maximizers. Define, for all $N \in \mathbb{N}$,

$$\tilde{\pi}_N(a) = \pi^*(a),$$

$$e_i^T \tilde{q}_{a,N} = \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} f_a^*(x) dx.$$

Note that, by construction, $\tilde{q}_{a,N} \in \mathcal{P}(X^N)$ and $\sum_{a \in A} \tilde{\pi}_N(a) \tilde{q}_{a,N} = q_N$. That is, the constraints of the discrete-state problem are satisfied for all N . Denote the value function under these policies as $\tilde{V}_N(q_N; \bar{\theta})$.

Because of the constraints stated in Theorem 4, each f_a^* satisfies the conditions of Lemma 11, and therefore the sequence $\tilde{q}_{a,N}$ satisfies the convergence condition for all $a \in A$. It follows by Lemma 12 that this sequence of policies delivers, in the limit, the

value function $V(f; \bar{\theta})$. If this function is strictly larger than $\lim_{n \rightarrow \infty} V_n(q_n; \bar{\theta})$, there must exist some \bar{n} such that

$$\tilde{V}_{\bar{n}}(q_{\bar{n}}; \bar{\theta}) > V_{\bar{n}}(q_{\bar{n}}; \bar{\theta}),$$

contradicting optimality. Therefore, the functions $f_a(x)$ and $\pi(a)$ are maximizers.

It remains to show that

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{\lfloor xn \rfloor} e_i^T q_{a,n} = \int_0^x f_a(y) dy.$$

Note that

$$e_i^T q_{a,n} = (n+1) \int_{\frac{i}{n+1}}^{\frac{i+1}{n+1}} \hat{q}_{a,n} \left(\frac{2i+1}{2(n+1)} \right) dy,$$

where $\hat{q}_{a,n}$ is the function defined in Lemma 12. Therefore, the sum is equal to

$$\sum_{i=0}^{\lfloor xn \rfloor} e_i^T q_{a,n} = \int_0^{\frac{\lfloor xn \rfloor + 1}{n+1}} \hat{q}_{a,n} \left(\frac{\lfloor (n+1)y + \frac{1}{2} \rfloor + \frac{1}{2}}{(n+1)} \right) dy.$$

By the boundedness of $\hat{q}_{a,n}$ (which follows from the convergence condition) and the dominated convergence theorem,

$$\lim_{n \rightarrow \infty} \int_0^{\frac{\lfloor xn \rfloor + 1}{n+1}} \hat{q} \left(\frac{\lfloor (n+1)y + \frac{1}{2} \rfloor + \frac{1}{2}}{(n+1)} \right) dy = \int_0^x f_a(y) dy,$$

as required.

D.12 Proof of Lemma 7

We begin by observing that any information structure $p \in \mathcal{P}_{LipG}(A)$ defines unconditional action frequencies $\pi \in \mathcal{P}(A)$ and posteriors $f_a \in \mathcal{P}_{LipG}([0, 1])$ satisfying (26), using definitions (27). And conversely, any unconditional action frequencies and posteriors satisfy-

ing (26) define an information structure, using definitions (28). Hence the set of candidate structures is the same in both problems, and the problems are equivalent if the two objective functions are equivalent as well. It is also easily seen that in each problem, the first term of the objective function is the expected value of the DM's reward $u(x, a)$, integrating over the joint distribution for (x, a) . Hence it remains only to establish that the remaining terms of the objective function are equivalent as well.

Consider any information structure $p \in \mathcal{P}_{LipG}(A)$ and the corresponding unconditional action frequencies and posteriors, and let x be any point at which $f(x) > 0$, and at which $p_a(x)$ is twice differentiable for all a (and as a consequence, $f_a(x)$ is twice differentiable for all a as well). (We note that, given the Lipschitz continuity of the first derivatives, the set of x for which this is true must be of full measure.) Then the fact that $\sum_{a \in A} p_a(x) = 1$ for all x implies that

$$\sum_{a \in A} p_a''(x) = 0, \quad (34)$$

and similarly, constraint (26) implies that

$$\sum_{a \in A} \pi(a) f_a''(x) = f''(x). \quad (35)$$

At any such point, the definition of the Fisher information implies that

$$\begin{aligned}
I^{Fisher}(x) &\equiv \sum_{a \in A} \frac{(p'_a(x))^2}{p_a(x)} \\
&= \sum_a p''_a(x) - \sum_{a \in A} p_a(x) \frac{\partial^2 \log p_a(x)}{\partial x^2} \\
&= -\frac{\pi(a)f_a(x)}{f(x)} \frac{\partial^2}{\partial x^2} [\log \pi(a) + \log f_a(x) - \log f(x)] \\
&= \frac{1}{f(x)} \left[\sum_{a \in A} \pi(a) \frac{(f'_a(x))^2}{f_a(x)} - \sum_{a \in A} \pi(a) f''_a(x) - \frac{(f'(x))^2}{f(x)} + f''(x) \right] \\
&= \frac{1}{f(x)} \left[\sum_{a \in A} \pi(a) \frac{(f'_a(x))^2}{f_a(x)} - \frac{(f'(x))^2}{f(x)} \right].
\end{aligned}$$

Here the first line is the definition of the Fisher information (given in the lemma), and the second line follows from twice differentiating the function $\log p_a(x)$ with respect to x . In the third line, the first term from the second line vanishes because of (34); the remaining term from the second line is rewritten using (28). The fourth line follows from the third line by twice differentiating each of the terms inside the square brackets with respect to x . The fifth line then follows from (35).

Since this result holds for a set of x of full measure, we obtain expression

$$\int_0^1 f(x) I^{Fisher}(x) dx = \sum_{a \in A} \pi(a) \int_0^1 \frac{(f'_a(x))^2}{f_a(x)} dx - \int_0^1 \frac{(f'(x))^2}{f(x)} dx$$

for the mean Fisher information. This shows that the information-cost terms in both objective functions are equivalent, and hence the two problems are equivalent, and have equivalent solutions.

D.13 Proof of Lemma 8

Write the value function in sequence-problem form:

$$W(q_0, \lambda; \Delta) = \max_{\{p_{\Delta j}\}, \tau} E_0[\hat{u}(q_\tau) - \kappa\tau] - \lambda E_0[\Delta^{1-\rho} \sum_{j=0}^{\tau\Delta^{-1}} \{ \frac{1}{\rho} C(\{p_{\Delta j, x}\}_{x \in X}, q_{\Delta j}(\cdot))^\rho - \Delta^\rho c^\rho \}].$$

Define

$$\bar{u} = \max_{a \in A, x \in X} u(a, x).$$

By the weak positivity of the cost function $C(\cdot)$, it follows that

$$W(q_0, \lambda; \Delta) \leq \bar{u} + \max_{\tau} E_0[-\kappa\tau + \Delta \sum_{j=0}^{\tau\Delta^{-1}-1} \lambda c^\rho].$$

Because $\lambda \in (0, \kappa c^{-\rho})$, the expression

$$-\kappa\tau + \Delta \sum_{j=0}^{\tau\Delta^{-1}-1} \lambda c^\rho = (\lambda c^\rho - \kappa)\tau$$

is weakly negative, and therefore

$$W(q_0, \lambda; \Delta) \leq \bar{u}.$$

By a similar argument, there is a smallest possible decision utility \underline{u} , and because stopping now and deciding is always feasible,

$$W(q_0, \lambda; \Delta) \geq \underline{u}.$$

Therefore, $W(q_0, \lambda; \Delta)$ is bounded for all $\lambda \in (0, \kappa c^{-\rho})$ and all Δ . Note that this argument

also shows that

$$E_0[\tau](\kappa - \lambda c^\rho) \leq \bar{u} - W(q_0, \lambda; \Delta),$$

and hence that

$$E_0[\tau] \leq \frac{\bar{u} - \underline{u}}{(\kappa - \lambda c^\rho)}.$$

We can define the “state-specific” value function, $W(q_t, \lambda; \Delta, x)$, which is the value function conditional on the true state being x . The state-specific value function has a recursive representation, in the region in which the DM continues to gather information:

$$\begin{aligned} W(q_t, \lambda; \Delta, x) = & -\kappa\Delta + \lambda\Delta^{1-\rho}(\Delta^\rho c^\rho - \frac{1}{\rho}C(\cdot)^\rho) + \\ & \sum_{s \in S: e_s^T p_t e_x > 0} (e_s^T p_t^* e_x) W(q_{t+\Delta, s}^*, \lambda; \Delta, x). \end{aligned}$$

In this equation, we take the optimal information structure as given. Note that, by construction, wherever the DM does not choose to stop, the expected value of the state-specific value functions is equal to the value function.

$$\sum_{x \in X} q_{t,x} W(q_t, \lambda; \Delta, x) = W(q_t, \lambda; \Delta).$$

By the optimality of the policies, we have

$$W(q_t, \lambda; \Delta) \geq \sum_{x \in X} q_{t,x} W(q', \lambda; \Delta, x),$$

for any q' in $\mathcal{P}(X)$. Suppose not; then the DM could simply adopt the information structure associated with beliefs q' and achieve higher utility, contradicting the optimality of the policy.

The convexity of the value function follows from the observation that

$$\begin{aligned} W(\alpha q + (1 - \alpha)q', \lambda; \Delta) &= \alpha \sum_{x \in X} q_x W(\alpha q + (1 - \alpha)q', \lambda; \Delta, x) + \\ &\quad (1 - \alpha) \sum_{x \in X} q'_x W(\alpha q + (1 - \alpha)q', \lambda; \Delta, x), \end{aligned}$$

and using the inequality above,

$$W(\alpha q + (1 - \alpha)q', \lambda; \Delta) \leq \alpha W(q, \lambda; \Delta) + (1 - \alpha)W(q', \lambda; \Delta).$$

D.14 Proof of Lemma 9

Consider an alternative policy that mixes (in the sense of Condition 2) the optimal signal structure and an uninformative signal, with probabilities $1 - a$ and a , respectively. We must have

$$-\sum_{s \in S} (e_s^T r_{t,n}^*) (W(q_{t,n,s}^*, \lambda; \Delta_n) - W(q_{t,n}, \lambda; \Delta_n)) - \lambda \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^{\rho-1} \frac{\partial C(p_{t,n}(a), q_{t,n})}{\partial a} \Big|_{a=0^+} \leq 0,$$

which is the first-order condition at the optimal policy in the direction of adding a little bit of the uninformative signal (decreasing a). By the convexity of $C(\cdot)$ and Condition 1,

$$C(p_{t,n}^*, q_{t,n}) + \frac{\partial C(p_{t,n}(a), q_{t,n})}{\partial a} \Big|_{a=0^+} \leq 0,$$

and therefore we must have

$$\sum_{s \in S} (e_s^T r_{t,n}^*) (W(q_{t,n,s}^*, \lambda; \Delta_n) - W(q_{t,n}, \lambda; \Delta_n)) \geq \lambda \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^\rho.$$

Applying the Bellman equation in the continuation region,

$$(\kappa - \lambda c^\rho) \Delta_n + \frac{\lambda}{\rho} \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^\rho \geq \lambda \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^\rho.$$

Therefore,

$$\lambda \left(1 - \frac{1}{\rho}\right) \Delta_n^{-\rho} C(p_{t,n}^*, q_{t,n})^\rho \leq (\kappa - \lambda c^\rho).$$

It follows by the assumption that $\lambda \in (0, \kappa c^{-\rho})$ and that $\rho > 1$ that

$$C(p_{t,n}^*, q_{t,n}) \leq \Delta_n \left(\frac{\theta}{\lambda}\right)^{\frac{1}{\rho-1}},$$

for the constant $\theta = \lambda \left(\rho \frac{\kappa - \lambda c^\rho}{\lambda(\rho-1)}\right)^{\frac{\rho-1}{\rho}} > 0$.

D.15 Proof of Lemma 10

We begin by discussing the convergence of stopping times. We have assumed that

$$E_0[\tau_n] \leq \bar{\tau},$$

for some strictly positive constant $\bar{\tau}$ and all n . It follows by the positivity of τ_n that the laws of τ_n are tight, and therefore there exists a sub-sequence that converges in measure. Pass to this sub-sequence (which we will also index by n), and let τ denote the limit of this sub-sequence.

The beliefs $q_{t,n}$ are a family of $\mathbb{R}^{|X|}$ -valued stochastic processes, with $q_{t,n} \in \mathcal{P}(X)$ for all $t \in [0, \infty)$ and $n \in \mathbb{N}$. Construct them as RCLL processes by assuming that $q_{\Delta_n j + \varepsilon, n} = q_{\Delta_n j, n}$ for all $m, \varepsilon \in [0, \Delta_n)$, and $j \in \mathbb{N}$. We next establish that the laws of $q_{t,n}$ are tight. By

Condition 5 and Lemma 9,

$$\frac{m}{2} \sum_{s \in S} (e_s^T p_n(q_{t,n}) q_{t,n}) \|q_{s,n}(q_{t,n}) - q_{t,n}\|_2^2 \leq C(p_n(q_{t,n}), q_{t,n}; S) \leq \Delta_n \left(\frac{\theta}{\lambda}\right)^{\frac{1}{\rho-1}},$$

where $q_{s,n}(q)$ is defined by $p_n(q)$ and Bayes' rule. It follows that, for any $\varepsilon > 0$, there exists an N_ε such that, for all $n > N_\varepsilon$,

$$P(\|q_{t+\Delta_n,n} - q_{t,n}\| > \varepsilon) \leq K_\varepsilon \Delta_n,$$

for the constant $K_\varepsilon = 2m^{-1} \varepsilon^{-2} \theta^{\frac{1}{\rho-1}}$. By Theorem 3.21, Condition 1 in chapter 6 of Jacod and Shiryaev (2013), and the boundedness of $q_{t,n}$, it follows that the laws of $q_{t,n}$ are tight. By Prokhorov's theorem (Theorem 3.9 in chapter 6 of Jacod and Shiryaev (2013)), it follows that there exists a convergent sub-sequence. Pass to this sub-sequence, and let q_t denote the limiting stochastic process. By Proposition 1.1 in chapter 9 of Jacod and Shiryaev (2013), q_t is a martingale with respect to the filtration it generates. By Skorohod's representation theorem, there exists a probability space and random variables (which we will also denote with $q_{t,n}$ and q_t) such the convergence is almost sure. We refer to this probability space and these random variables in what follows.

Note that, by Bayes' rule, if $e_x^T q_{t,n} = 0$ for some $x \in X$ and time t , then $e_x^T q_{s,n} = 0$ for all $s > t$. By Proposition 2.9 and Corollary 2.38 in chapter 2 of Jacod and Shiryaev (2013), we can write the "good" version of the martingale with characteristics

$$B = - \int_0^t \left(\int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) x dx \right) dA_s$$

$$C = \int_0^t \Sigma_s dA_s$$

$$v = dA_s \psi_s(x).$$

Because beliefs remain in the simplex, $\psi_s(x)$ has support only on x such that $q_s + x \in \mathcal{P}(X)$. Relatedly, $\iota^T \Sigma_s = 0$. By the property mentioned above, $q_s + x \ll q_s$, and Σ_s can be decomposed as $\Sigma_s = D(q_{s-}) \sigma_s \sigma_s^T D(q_{s-})$.

By the convexity of the cost function and Corollary 2,

$$C(p_n(q_{t,n}), q_{t,n}; S) \geq \sum_{s \in S} (e_s^T p_n(q_{t,n}) q_{t,n}) D^*(q_{s,n}(q_{t,n}) || q_{t,n}).$$

Defining the process, for arbitrary stopping time T ,

$$D_{s,n} = \lim_{\varepsilon \rightarrow 0^+} D^*(q_{s-+\varepsilon,n} || q_{s-,n}),$$

$$D_{t,T,n} = E_t \left[\int_t^T D_{s,n} ds \right] \leq \theta^{\frac{1}{\rho-1}} \Delta_n E_t \left[[\Delta_n^{-1} (T-t)] \right],$$

we have by Ito's lemma, almost sure convergence, and the dominated convergence theorem,

$$D_{t,T} = \lim_{n \rightarrow \infty} D_{t,T,n} = E_t \left[\int_t^T \left\{ \frac{1}{2} \text{tr} [\sigma_s \sigma_s^T k(q_{s-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s-} + x || q_{s-}) dx \right\} dA_s \right].$$

Hence, for all such stopping times T ,

$$E_t \left[\int_t^T \left\{ \frac{1}{2} \text{tr} [\sigma_s \sigma_s^T k(q_{s-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s-} + x || q_{s-}) dx \right\} dA_s \right] \leq \left(\frac{\theta}{\lambda} \right)^{\frac{1}{\rho-1}} E_t [T-t].$$

Note also by this argument that

$$\begin{aligned} & \lim_{n \rightarrow \infty} E_0 \left[\int_0^{\tau_n} \Delta_n^{1-\rho} C(p_n(q_{t,n}), q_{t,n}; S)^\rho dt \right] \\ & \geq E_t \left[\int_0^\tau \left\{ \frac{1}{2} \text{tr} [\sigma_s \sigma_s^T k(q_{s-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s-} + x || q_{s-}) dx \right\}^\rho \left(\frac{dA_s}{ds} \right)^\rho ds \right]. \end{aligned}$$

D.16 Proof of Theorem 3

Let m index a sequence of Markov optimal policies, $p_m^*(q)$, and of stopping times τ_m^* . Let $q_{t,n}^*$ denote the associated process for beliefs. By the uniform boundedness and convexity of the family of value functions $W(q, \lambda; \Delta_m)$, a uniformly convergent sub-sequence exists. Rockafellar (1970) Theorem 10.9 demonstrates that a uniformly convergent sub-sequence exists on the relative interior of the simplex, and Rockafellar (1970) Theorem 10.3 demonstrates that there is a unique extension to a convex and continuous function on the boundary of the simplex.

Pass to this sub-sequence, which (for simplicity) we also index by m , and let $W(q, \lambda)$ denote its limit. By Lemmas 8 and 9, the sequence of optimal policies and stopping time satisfies the conditions of Lemma 10. It follows by that lemma that

$$\begin{aligned} & \lim_{n \rightarrow \infty} E_0 \left[\int_0^{\tau_n^*} \Delta_n^{1-\rho} C(p_n^*(q_{t,n}^*), q_{t,n}^*; S)^\rho dt \right] \\ & \geq E_t \left[\int_0^\tau \left\{ \frac{1}{2} \text{tr}[\sigma_s^* \sigma_s^{*T} k(q_{s-}^*)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^*(x) D^*(q_{s-}^* + x | q_{s-}^*) dx \right\}^\rho \left(\frac{dA_s^*}{ds} \right)^\rho ds \right], \end{aligned}$$

where q_s^* is the limiting stochastic process and σ_s^* , ψ_s^* , dA_s^* are associated with the characteristics of the martingale q_s^* .

We also have, by weak convergence,

$$\lim_{n \rightarrow \infty} E_0 [\hat{u}(q_{\tau_n^*}^*) - (\kappa - \lambda c^\rho) \tau_n^*] = E_0 [\hat{u}(q_{\tau^*}) - (\kappa - \lambda c^\rho) \tau^*].$$

Recall also the bound, for any stopping time T measurable with respect filtration generated by q_s^* ,

$$E_t \left[\int_t^T \left\{ \frac{1}{2} \text{tr}[\sigma_s^* \sigma_s^{*T} k(q_{s-}^*)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^*(x) D^*(q_{s-}^* + x | q_{s-}^*) dx \right\} dA_s^* \right] \leq \left(\frac{\theta}{\lambda} \right)^{\frac{1}{\rho-1}} E_t [T - t].$$

It follows that

$$W(q, \lambda) \leq W^+(q, \lambda)$$

for all $q \in \mathcal{P}(X)$, where

$$W^+(q_t, \lambda) = \sup_{\{\sigma_s, \psi_s, dA_s, \tau\}} E_t[\hat{u}(q_\tau) - (\kappa - \lambda c^\rho)(\tau - t)] - \frac{\lambda}{\rho} E_t \left[\int_t^\tau \left\{ \frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\}^\rho \left(\frac{dA_s}{ds} \right)^\rho ds \right],$$

subject to the constraints, for all stopping times T measurable with respect filtration generated by q_s^* ,

$$E_t \left[\int_t^T \left\{ \frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_t)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\} dA_s \right] \leq \left(\frac{\theta}{\lambda} \right)^{\frac{1}{\rho-1}} E_t[T - t]$$

and

$$E_0[\tau] \leq \bar{\tau},$$

and the evolution of beliefs as implied by the characteristics derived from σ_s, ψ_s, dA_s . Observe, by the arguments in the proof of Lemma 8, that $W^+(q, \lambda)$ is convex in q .

Also note that, for W^+ , it is without loss of generality to set $dA_s = ds$. Scaling dA_s up and scaling $\sigma_s \sigma_s^T$ and ψ_s down, or vice versa, does not change the constraint, and setting $dA_s = 0$ is clearly sub-optimal by the assumption that $\kappa - \lambda c^\rho > 0$. Note also that there is a version of the optimal policies which satisfy the constraint everywhere:

$$\frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x | q_{s^-}) dx \leq \left(\frac{\theta}{\lambda} \right)^{\frac{1}{\rho-1}}.$$

The associated Bellman equation, in the continuation region, is

$$0 = \max_{\sigma_s, \psi_s} E[dW^+(q_s, \lambda)] - (\kappa - \lambda c^\rho) ds - \frac{\lambda}{\rho} \left\{ \frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\}^\rho.$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem (which we have yet to show are equal to σ_s^* and ψ_s^*). Suppose that the constraint does not bind, and consider a perturbation which scales $\sigma_s^+ \sigma_s^{+T}$ and ψ_s^+ by some constant $(1 + \varepsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1 + \varepsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $\kappa - \lambda c^\rho > 0$. The first order condition for this perturbation is

$$\begin{aligned} (\kappa - \lambda c^\rho) + \frac{\lambda}{\rho} \left\{ \frac{1}{2} \text{tr}[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\}^\rho = \\ \lambda \left\{ \frac{1}{2} \text{tr}[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(x) D^*(q_{s^-} + x | q_{s^-}) dx \right\}^\rho, \end{aligned}$$

which must hold at the optimal policies for this problem. It follows by the definition of θ (see the proof of Lemma 9) that the constraint binds.

Consider a sub-optimal policy which sets $\psi_s(x) = 0$ and satisfies the constraint. The above FOC applies, and therefore we must have

$$\text{tr}[\tilde{\sigma}_s \tilde{\sigma}_s^T (D(q_{s^-}) W_{qq}^+(q_{s^-}, \lambda) D(q_{s^-}) - \theta k(q_{s^-}))] \leq 0,$$

where W_{qq}^+ is understood in a distributional sense. It follows that, for all feasible x ,

$$W^+(q_{s^-} + x, \lambda) - W^+(q_{s^-}, \lambda) - x^T W_q^+(q_{s^-}, \lambda; -x) \leq \frac{1}{2} \int_0^1 x^T \bar{k}(q_{s^-} + lx) x dl.$$

By Condition 6, this implies that

$$W^+(q_{s^-} + x, \lambda) - W^+(q_{s^-}, \lambda) - x^T W_q^+(q_{s^-}, \lambda; -x) \leq \theta D^*(q_{s^-} + x | q_{s^-}).$$

Hence, it is without loss of generality to assume that $\psi_s^+(x) = 0$ for all x . Note that, if there is a strict preference for gradual learning, the above inequality is strict for all non-zero x . As a result, in this case $\psi_s^+(x) = 0$ for all x . Note also that our control problem involves direct control of the diffusion coefficients, and hence satisfies the standard requirements for the existence and uniqueness of a strong solution to the resulting SDE (Pham (2009) sections 1.3 and 3.2).

Noting that $W^+(q, \lambda) \geq W(q, \lambda)$, it follows that if there exists a sequence of policies that converge to the stochastic process q_t^+ , characterized by σ^+ , and whose costs $\Delta_n^{-1}C(\cdot)$ converge to $\theta^{\frac{1}{\rho-1}}$, then such a sequence of policies achieves, in the limit, at least as much utility as any other sequence of policies. It would then be the case that there must be sequence of optimal policies that converges a.s. (as in Lemma 10) to some optimal policy of W^+ (not necessarily σ^c and ψ^c , but this does not matter for our argument). Note, however, that if there is a strict preference for gradual learning, and W^+ is achievable, all optimal policies of W^+ generate diffusions, and hence all convergent sub-sequences of beliefs induced by optimal policies in the discrete-time model must converge to diffusions.

Define the function

$$\Sigma^+(q) = D(q)\sigma^+(q)\sigma^+(q)^T D(q).$$

We will construct a sequence that converges to this diffusion process.

Consider the eigenvector decomposition of the matrix

$$L(q)\Upsilon(q)L(q)^T = \alpha_n(q)\Sigma^+(q),$$

where $\alpha_n(q) > 0$ is a scalar function of q . For each pair $(s_i, s_{i+1}) \in S$, where $i \in \{1, 2, \dots, |X|\}$

is an even integer, set $e_{s_i}^T r_n = e_{s_{i+1}}^T r_n = \frac{1}{2|X|}$, and set

$$\begin{aligned} q_{s_i,n}(q) - q &= \\ q - q_{s_{i+1},n}(q) &= \\ L(q) \Upsilon^{\frac{1}{2}}(q) e_i. \end{aligned}$$

Set all other $e_s^T r_n = 0$. By construction,

$$\sum_{s \in S} (e_s^T r_{s,n})(q_{s,n}(q) - q) = 0,$$

and

$$\sum_{s \in S} (e_s^T r_n)(q_{s,n}(q) - q)(q_{s,n}(q) - q)^T = \alpha_n(q) \Sigma^+(q)$$

and

$$\sum_{s \in S} (e_s^T r_n) = 1.$$

We would like to have, for this policy, $C(p_n(q), q; S) = \Delta_n \theta^{\frac{1}{\rho-1}}$ always. Note that under this policy, $C(\cdot)$ is a function of α_n and q . By the convexity of $C(\cdot)$ and the definition of its derivatives,

$$C(\cdot) \geq \alpha_n(q) \frac{\partial C}{\partial \alpha} \Big|_{\alpha=0} = \alpha_n(q) \left(\frac{1}{2} \text{tr}[k(q) \sigma^+(q) (\sigma^+(q))^T] \right),$$

and hence

$$C(\cdot) \geq \alpha_n(q) \theta^{\frac{1}{\rho-1}}.$$

It follows that $\alpha_n(q) \leq \Delta_n$, it is feasible to have $C(p_n(q), q; S) = \Delta_n \theta^{\frac{1}{\rho-1}}$.

Note, by the finiteness of $\Sigma^+(q)$ (due the positive definiteness of $\bar{k}(q)$), that $q_{s,n}(q) - q =$

$O(\Delta_n^{\frac{1}{2}})$. It follows from lemmas 11.2.1 and 11.2.2 in Stroock and Varadhan (2007) that the law of q_n under this process converges to a solution to the martingale problem associated with the coefficients $\sigma^+(q)$. By the uniqueness of this solution established earlier, this law is the law of q_t^+ , a diffusion. Let τ_n be the first hitting time of q_n for the boundary defined by the stopping region of $W^+(q, \lambda)$ (that is, $W^+(q, \lambda) = \hat{u}(q)$). Therefore, by construction (and similar to the arguments of Amin and Khanna (1994)), the value functions of the discrete time problem converge to $W^+(q, \lambda)$. Therefore, this value function is achievable, and $W(q, \lambda) = W^+(q, \lambda)$. Note also that we have constructed a sequence of policies that converge to an optimal policy of $W(q, \lambda)$.

We next demonstrate equality of the primal and dual. We have shown that

$$W(q, \lambda) = E_0[\hat{u}(q_{\tau^*}) - (\kappa - \lambda c^\rho)\tau^*] - \frac{\lambda}{\rho} E_0\left[\int_0^{\tau^*} \left(\frac{\theta}{\lambda}\right)^{\frac{\rho}{\rho-1}} ds\right].$$

Recall the definition of θ ,

$$\theta = \lambda \left(\rho \frac{\kappa - \lambda c^\rho}{\lambda(\rho - 1)}\right)^{\frac{\rho-1}{\rho}}.$$

Define λ^* by

$$\frac{\kappa - \lambda^* c^\rho}{\lambda^*(\rho - 1)} = c^\rho,$$

which is

$$\lambda^* = \frac{\kappa}{\rho c^\rho}.$$

Note that $\lambda^* \in (0, \kappa c^{-\rho})$, as required. For this value of λ ,

$$W(q_0, \lambda^*) = E_0[\hat{u}(q_{\tau^*}) - \kappa \tau^*],$$

and the limit of the constraint is satisfied:

$$\frac{\lambda^*}{\rho} E_0 \left[\int_0^{\tau^*} \left(\frac{\theta}{\lambda^*} \right)^{\frac{\rho}{\rho-1}} ds \right] = \lambda^* E_0 \left[\int_0^{\tau^*} cds \right].$$

Consider a convergent sub-sequence of $V(q_0; \Delta_n)$ (which exists by the uniform boundedness and convexity of the problem), and denote its limit $V(q_0)$ (again, we will index this sequence by n). By the standard duality inequalities, for all λ ,

$$V(q_0; \Delta_n) \leq W(q_0, \lambda; \Delta_n),$$

for all n , and therefore

$$V(q_0) \leq W(q_0, \lambda).$$

Consider the value function $\tilde{V}(q_0)$, which is the value function under the feasible optimal policies for $W(q_0, \lambda^*)$. It follows that $\tilde{V}(q_0) = W(q_0, \lambda^*)$, and $\tilde{V}(q_0) \leq V(q_0)$, and therefore $V(q_0) = W(q_0, \lambda^*)$.

We can define

$$\begin{aligned} \theta^* &= \lambda^* \left(\rho \frac{\kappa - \lambda^* c^\rho}{\lambda^* (\rho - 1)} \right)^{\frac{\rho-1}{\rho}} \\ &= \lambda^* \rho^{\frac{\rho-1}{\rho}} c^{\rho-1} \\ &= \frac{\kappa}{c} \rho^{-\rho-1}. \end{aligned}$$

Note that every convergent sub-sequence of $V(q_0; \Delta_n)$ converges to the same function. By

the boundedness of value function, it follows that

$$\begin{aligned} V(q_0) &= \lim_{\Delta \rightarrow 0^+} V(q_0; \Delta). \\ &= E_0[\hat{u}(q_{\tau^*}) - \kappa \tau^*]. \end{aligned}$$

The constraint can be written as

$$\frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_{s^-})] \leq \left(\frac{\theta^*}{\lambda^*}\right)^{\frac{1}{\rho-1}},$$

with

$$\left(\frac{\theta^*}{\lambda^*}\right)^{\frac{1}{\rho-1}} = (\rho^{1-\rho^{-1}} c^{\rho-1})^{\frac{1}{\rho-1}} = c \rho^{\rho^{-1}} = \chi.$$

The optimal policy satisfies this constraint, and hence it follows that the value function is the maximized over all policies satisfying

$$\frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_s)] \leq \chi,$$

concluding the proof.

E Appendix References

- Kaushik Amin and Ajay Khanna. Convergence of american option values from discrete- to continuous-time financial models. *Mathematical Finance*, 4(4):289–304, 1994.
- Nihat Ay, Jürgen Jost, Hông Vân Lê, and Lorenz Schwachhöfer. Information geometry and sufficient statistics. *Probability Theory and Related Fields*, pages 1–38, 2014.
- David Blackwell. Equivalent comparisons of experiments. *The annals of mathematical statistics*, 24(2):265–272, 1953.
- Andrew Caplin and Mark Dean. Revealed preference, rational inattention, and costly information acquisition. *American Economic Review*, 105(7):2183–2203, 2015.
- Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing Shannon entropy. *Unpublished manuscript*, July 2017.

- Nikolai Nikolaevich Chentsov. *Statistical decision rules and optimal inference*. Number 53. American Mathematical Soc., 1982.
- Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.
- B Roy Frieden. *Science from Fisher information: a unification*. Cambridge University Press, 2004.
- Fabrizio Gabbiani and Steven J. Cox. *Mathematics for Neuroscientists*. Academic Press, 2010.
- Joshua I. Gold and Hauke R. Heekeren. Neural mechanisms for perceptual decision making. In Paul W. Glimcher and Ernst Fehr, editors, *Neuroeconomics: Decision Making and the Brain, 2d ed*. Academic Press, 2014.
- Jean Jacod and Albert Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.
- Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101(6):2590–2615, 2011.
- Stephen Morris and Hyun Song Shin. Global games: theory and applications. 2001.
- Stephen Morris and Ming Yang. Coordination and the relative cost of distinguishing nearby states. *Unpublished manuscript*, 2016.
- Huyên Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61. Springer Science & Business Media, 2009.
- R Tyrrell Rockafellar. *Convex analysis*. 1970.
- Michael N. Shadlen et al. The speed and accuracy of a perceptual decision: A mathematical primer. In K. Doya et al., editors, *Bayesian Brain: Probabilistic Approaches to Neural Coding*. M.I.T. Press, 2007.
- Jakub Steiner, Colin Stewart, and Filip Matějka. Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553, 2017.
- Daniel W Stroock and SR Srinivasa Varadhan. *Multidimensional diffusion processes*. Springer, 2007.
- Cédric Villani. *Topics in optimal transportation*. Number 58. American Mathematical Soc., 2003.
- Ming Yang. Coordination with flexible information acquisition. *Journal of Economic Theory*, 158:721–738, 2015.