Rational Inattention
with Sequential Information Sampling*

Benjamin Hébert †
Stanford University

Michael Woodford ‡
Columbia University

May 17, 2017

Preliminary and Incomplete. Comments welcome.

Abstract

What should the cost of information be in rational inattention problems (Sims [2010])? We argue that the cost of information should summarize the results of a dynamic evidence accumulation process, and that the costs should reflect that fact that some states of the world are similar to (different from) other states, and therefore should be harder (easier) to distinguish. We introduce a continuous time model of sequential information sampling. This model is derived from an optimal information accumulation problem in which a large number of successive samples of information about a decision situation (each only minimally informative by itself) can be taken before a decision must be made. At each stage in the sampling process, many different experiments are possible, and the experimentation undertaken can also be contingent upon the evidence accumulated to that point. Having introduced the continuous time model, we show that, in a broad class of cases, we can establish that the choice frequencies resulting from optimal information accumulation are the same as those implied by a static rational inattention problem with a particular static information cost function. In other words, certain static cost functions in the rational inattention model summarize the results of a dynamic evidence accumulation process. The mutual information cost function employed by Sims (2010) has this property, but there are other cost functions that can also be justified in this way. We introduce a class of “neighborhood-based” cost functions, which also summarize the results of dynamic

*The authors would like to thank Mira Frick, Matthew Gentzkow, Emir Kamenica, Divya Kirti, Ran Shorrer, Stephen Morris, Ming Yang, Doron Ravid (the discussant), Mark Dean, Pietro Ortoleva, Sebastian Di Tella, Ilya Segal, Jacob Leshno, 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. 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.
evidence accumulation, and (unlike mutual information) incorporate a conception of the similarity of states to one another, and make it more costly to undertake experiments that can produce different results in similar but non-identical states. With this alternative form of cost function, optimal information accumulation results in choice frequencies that are similar in similar states; in a continuous-state extension of the model, optimality implies choice frequencies that vary continuously with the state, even when the payoffs for given responses jump discontinuously with variation in the state. This feature of our variant rational inattention model conforms with evidence from perceptual discrimination experiments.
1 Introduction

The theory of rational inattention, proposed by Christopher Sims and surveyed in Sims [2010], seeks to endogenize 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) can choose an arbitrary signal structure with which to gather information about her situation before making a decision; the cost of each possible signal structure is described by a cost function, which in Sims’ theory is an increasing function of the Shannon mutual information between the state of the world (that determines the DM’s reward from choosing different actions) and the DM’s signal.

Shannon’s mutual information, a measure of the degree to which the signal is informative about the state of the world, plays a central role in information theory (Cover and Thomas [2012]), as a consequence of powerful mathematical results that are considerable practical relevance in communications engineering. It not obvious, though, that the theorems that justify the use of mutual information in communications engineering provide any warrant for using it as a cost function in a theory of attention allocation, in the case of either economic decisions or perceptual judgments.\(^1\) Moreover, 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 learn about any two states that occur with the same probability. For example, in the case of an experimental task of the kind discussed in Caplin and Dean [2015]--- in which subjects are presented with an

\(^1\)As 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 decodable by the recipient. In the kind of situations with which we are concerned here, instead, there is a constraint on the informativeness of each individual transmission, that can be measured independently of the signals that are sent earlier and later.
array of 100 red and blue balls, and must determine whether there are more red balls or more blue --- Sims’ theory of rational inattention implies that given that 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. In fact, it is much easier to quickly and reliably determine that there are more red balls in the case of some arrays in this class (for example, one with 98 red balls and only two blue balls) than others (for example, one with 51 red balls and 49 blue balls, relatively evenly dispersed), and subjects make more payoff-maximizing responses in the former case.

One response to unappealing features of the mutual information cost function is to develop a theory of rational inattention that makes only much weaker assumptions about the cost function (consistent with mutual information, but not requiring it), as authors such as De Oliveira et al. [2014], Caplin and Dean [2015], and Huettner et al. [2016] have done. This approach results in a theory with correspondingly weaker predictions. We seek instead to motivate a more specific class of information-cost functions, so as to allow more definite conclusions, while still including cases that we regard as more realistic specifications than mutual information.

Our approach exploits 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 particularly emphasize the limiting case in which each individual experiment is only minimally informative, but a very large number of independent experiments can be performed within a given time interval. In the continuous-time limit of our sequential sampling process, 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 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 recently, some authors have proposed that data on stochastic choice and response time in economic contexts can be similarly modeled.\(^2\) In the case of preferential choice between goods, the sampling is presumably a process of drawing from memory associations (or recollections of past experiences) that bear upon the assignment of value to the presented options, rather than a sequence of repeated observations of the item presented.

Much of the empirical literature that models stochastic choice as the outcome of a sequential sampling process aims simply to provide a process model of observed behavior, but a number of recent papers endogenize at least some aspects of the information-sampling process. Fudenberg et al. [2015] consider the optimal stopping problem for an information sampling process described by the sample path of a Brownian motion with a drift that depends on the unknown state of the world.\(^3\) This can be thought of as a problem in which a given experiment (with the probability of positive or negative outcomes dependent on the state of the world) can be repeated an indefinite number of times, with a fixed cost per repetition of the experiment. The sequence of outcomes of the successive experiments becomes a Brownian motion in the limiting case in which individual experiments require only an infinitesimal amount of time (and hence involve only an infinitesimal cost, as a fixed cost of sampling per unit time is assumed), and are correspondingly minimal in the information that they reveal about the state (because the difference in the mean outcome of an individual experiment across different states of the world is tiny relative to the standard

---

\(^2\)In addition to the references in Fehr and Rangel [2011], recent examples include Clithero and Rangel [2013] and Krajbich et al. [2014].

\(^3\)See also Tajima et al. [2016] for analysis of a related class of models.
deviation of the outcome). In the case assumed by Fudenberg et al. [2015], there is no choice about the type of experiment that can be repeatedly performed, but the decision when to stop collecting further information is optimized.

Woodford [2014] instead takes as given a stopping rule (motivated by the empirical psychology and neuroscience literatures), but rather than assuming a given Brownian motion for the evidence process as long as information sampling continues, this process is endogenized as in theories of rational inattention. The assumed stopping rule makes the decision whether to continue sampling (and the event action chosen) a function only of a single number, the cumulative excess of positive over negative outcomes from the sequence of experiments; but at each point in the sequential process, it is assumed to be possible to vary the probability of a positive response conditional upon the true state, subject to an information-cost function that makes more informative experiments more costly. Under the assumed cost function, Woodford [2014] finds that optimal information sampling results in an evidence process that evolves as a continuous stochastic diffusion process, and that it is optimal for the drift of this process to be an increasing function of the relative value of the two choice options, as assumed by Fudenberg et al. [2015]; but rather than this process being a Brownian motion with a constant drift, as in Fudenberg et al. [2015], it is generally optimal for the drift also to depend on the current belief state.

Our approach differs from these earlier efforts in seeking to endogenize both the nature of the information that is accumulated at each stage of the information-sampling process and the stopping rule that determines how much information 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 both Fudenberg et al. [2015] and Woodford [2014] do.) 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 Fudenberg et al.
[2015]. Among other results, we describe a class of information-cost 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].

Our most general results characterize the solution to our dynamic model of information
sampling, assuming only that the flow information-cost function (specifying the cost of in-
dividual experiments) satisfies a relatively general set of conditions. We obtain tighter pre-
dictions (despite relatively flexible assumptions about the cost of individual experiments)
when we pass to a limiting case in which individual experiments are each only barely in-
formative but also each have only an infinitesimal cost. In this continuous limit of our
dynamic model, the optimal information-sampling problem can be modeled as a problem
of optimal control of a diffusion process on the probability simplex (the set of possible pos-
terior beliefs), with sampling stopping when certain (endogenous determined) boundaries
are reached. We show quite generally that optimal information sampling can be charac-
terized using a value function defined on the set of possible posterior beliefs, which value
function is a (weak) solution to a Hamilton-Jacobi-Bellman equation.

In addition to characterizing optimal sequential information sampling in our dynamic
model, we show that in the case of a relatively flexible family of possible cost functions for
individual experiments, the dynamic model’s predictions with regard to choice frequencies
conditional on the state of the world are the same as those of a static rational-inattention
model with an appropriately chosen information-cost function for the choice of a single
signal. (The finite set of possible signal values in the equivalent static model correspond
to the different possible terminal information states in the dynamic model, each of which
corresponds to one of the possible actions.) Under assumptions about the flow information-
cost function that we state below, 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], that do not rely on any analogy with rate-distortion theory in communications engineering.

Moreover, 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]) under certain assumptions about the flow information-cost function, we show that different predictions can be obtained under other, very plausible specifications of the flow cost function. In particular, we discuss the implications of an attractive family of flow information-cost functions, that 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 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 cost functions of this kind from the mutual information cost function assumed by Sims. 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).

As a still more specific special case, we consider a neighborhood-based cost function that can be defined in the case of a continuum of states that can be ordered on a line (this is a case of considerable interest, both for economic applications and for applications to perceptual psychology). We consider the infinite state limit of static rational inattention problems with an example of our neighborhood-based cost functions, and show that it is equivalent to assuming a continuous-state static rational inattention problem with the Fisher
information in the place of Shannon’s entropy. As emphasized in the information theory and physics literatures (Cover and Thomas [2012], Frieden [2004]), the Fisher information captures “local” variations, as opposed to the “global” variation summarized by Shannon’s entropy.

Our paper builds upon the rational inattention literature, surveyed in Sims [2010]. In its use of axioms to characterize the assumed form of the flow information-cost function, it is particularly close to Caplin and Dean [2013], Caplin and Dean [2015], and De Oliveira et al. [2014]. The Chentsov [1982] theorems used to characterize the properties of general rational inattention cost functions were also used by Hébert [2014], in a different context. We also use techniques developed by Kamenica and Gentzkow [2011] and Matějka et al. [2015] in characterizing the solution to our problem.

Section 2 begins by describing our continuous time model. We postpone the derivation of this model as the limit of a discrete time dynamic evidence accumulation problem until section §6. After introducing the continuous time model, in theorem 1 we show 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 a static rational inattention model. The information costs that appear in the static rational inattention model are “posterior-separable,” in the terminology of Caplin and Dean [2013] (and related to the “GERI” costs described by Fosgerau et al. [2016]). In section §4, we discuss the general conditions on information costs the we assume. In section §5, we introduce a new class of cost functions that satisfy these general conditions, are posterior-separable (and therefore can be justified by our dynamic model), and capture notions of “distance” between different states of the world. In subsection 5.1 of this section, we derive a particular cost function, based on the Fisher information, that has these properties and is applicable to a setting with a continuum of states. Section 6 shows that any of the cost functions satisfying the conditions we impose leads, in the discrete time dynamic evidence accumulation problem, to the continuous time
2 A Continuous Time Model of Sequential Evidence Accumulation

We begin our paper, in some sense, in the “middle,” by introducing our continuous time model of sequential evidence accumulation. After we introduce the model, we will discuss its derivation and its solutions. The model builds on the rational inattention framework of Sims [2010]. Let $x \in X$ be the underlying state of the nature, and $a \in A$ be the action taken by the agent. 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 agent’s utility from taking action $a$ in state $x$ at time $t$ is $u_{a,x} - \mu t$. The parameter $\mu > 0$ governs the penalty for delaying making a decision; the agent does not discount the future. We use this penalty, rather than time discounting, for tractability reasons.

The agent does not perfectly observe the state $x \in X$. At each time $t$, the agents holds beliefs $q_t \in \mathcal{P}(X)$, where $\mathcal{P}(X)$ denotes the probability simplex over $X$. Let $q_{x,t}$ be the probability, under the agent’s beliefs at time $t$, of state $x$. Time begins at $t = 0$, when the agent holds prior beliefs $q_0$. At each moment in time, the agent 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 $q_\tau$ at time $\tau$, the agent 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 $\hat{u}(q_\tau) - \mu \tau$.

When the agent gathers information, she chooses the variance-covariance matrix of her beliefs, subject to a constraint on the “size” of that matrix. That is, the agent’s beliefs
evolve as
\[ dq_{x,t} = q_{x,t} \sigma_{x,t} \cdot dB_t, \]  
(1)

where \( dB_t \) is an \(|X|\)-dimensional Brownian motion, \( \sigma_t \) is a matrix chosen by the agent, and \( \sigma_{x,t} \) is a particular row of that matrix.

The agent’s choice of \( \sigma_t \) 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 agent can acquire. The trivial restriction is that
\[ t^T \cdot dq_t = 0 \]
always, where \( t \) 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 agent can acquire at each moment, is
\[ \frac{1}{2} \text{tr}[\sigma_t \sigma_t^T k(q_t)] \leq \chi, \]  
(2)

where \( k(q_t) \) is an \(|X| \times |X|\) dimensional matrix-valued function we will refer to as the “information cost matrix”, \( \text{tr}[\cdot] \) is the trace, and \( \chi \) is a positive constant. After we describe the agent’s objective function, we will discuss this constraint, and the information cost matrix, in more detail. For now, it is sufficient to know that the information cost matrix function satisfies certain properties: it is positive semi-definite, and a vector of ones \( (t) \) is its null space.
Using her control of the volatility of her beliefs, and subject to the constraints imposed by the information cost matrix, our agent attempts to maximize her payoff. Her sequence problem can be written, given beliefs $q_t$ at time $t$,

$$V(q_t) = \sup_{\{\sigma_t \in M(q_t)\}, \tau \geq t} E_t[\tilde{\mu}(q_\tau) - \mu \tau],$$

where $\tau$ is the agent’s endogenous stopping time, subject to the constraints listed previously. The value function $V(q_t)$ is defined on the probability simplex, $\mathcal{P}(X)$, but it will simplify notation to extend its domain to $\mathbb{R}_{+}^{|X|} \setminus \{0\}$ by assuming that it is homothetic of degree one. That is, for all $\alpha > 0$ and $q_t \in \mathcal{P}(X)$,

$$V(\alpha q_t) = \alpha V(q_t).$$

Wherever this value function is twice-differentiable and the agent does not choose to stop, the problem has a simple recursive representation:

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

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

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

The following lemma describes the HJB equation associated with the sequence problem. It is derived by showing that the information constraint binds. The maximum eigenvalue appears in place of a maximization over $\sigma_t$, but this is just a compact way of expressing the idea that the agent is choosing in which direction(s) to update her beliefs.
Lemma 1. Anywhere the value function $V(q_t)$ is twice-differentiable, it satisfies

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

where $\theta = \chi^{-1}\mu$ and $\lambda_1(\cdot)$ denotes the largest eigenvalue of a matrix.

Proof. See the appendix, section A.1. \qed

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 $\theta$ 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 $\theta$. 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—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 agent’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 agent 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 agent is just indifferent between two actions with distinct state-contingent payoffs, the stopping payoff is non-differentiable.\footnote{At this point, we have also not shown that $V(q)$ is differentiable everywhere, but this is proven in the} However, it will never be optimal for
the agent 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 agent can acquire, because it limits the volatility of her beliefs. Our agent is a Bayesian, meaning that she can never expect to revise her beliefs in a particular direction—her beliefs must be a martingale, explaining why there is 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. This embeds the idea that, as one looks at smaller and smaller time intervals, the informativeness of the signals the agent is observing scales down. The continuity of the agents’ beliefs in our model is something between an assumption and a result—we prove continuity is a feature of the optimal information-gathering strategy we use to derive the continuous time model (see section §6), but do so by making assumptions that lead the agent to want to smooth the quantity of information gathered across time.

We derive this cost from a model in which the agent can choose any signal structure she desires at each time period, as in standard rational inattention models. One result of our derivation is the observation that the agent can choose any volatility matrix $\sigma_t$. 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 agent 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

---

5Che and Mierendorff [2016] explore a related model with jumps in beliefs.
smooth her information gathering over time, and the instantaneous volatility is sufficient to characterize the process for beliefs.

Relatedly, our derivation shows that the conditional dynamics of beliefs (conditional on some true state $x'$) can be written as a function of the agent’s optimal choice of volatility matrix, $\sigma^*_t$. Let $e_{x'}$ denote a basis vector, equal to 1 for that state $x'$ and zero otherwise. Conditional on the true state $x'$, assuming the agent believes the true state is possible ($q_{x',t} > 0$), the agent’s beliefs evolve, for all $x \in X$, as

$$
\frac{dq_{x,t}}{q_{x,t}} = \sigma^*_{x,t} \sigma^*_T e_{x'} dt + \sigma^*_{x,t} \cdot dB_t.
$$

(4)

Intuitively, states $x \in X$ for which $q_{x,t}$ positively covaries with $q_{x',t}$ under the optimal policy experience an upward drift in their likelihood, and states that negatively covary experience a negative drift. Drift-diffusion models, which have been studied extensively in the psychology literature and used more recently in economics (Krajbich et al. [2014], Fudenberg et al. [2015]) are typically expressed in terms of their conditional dynamics. Our continuous time model involves dynamics that, in certain respects, resemble these models. A more detailed exploration of the relationship between these models is beyond the scope of the present paper.

Next, we will describe in more detail the information cost matrix function, $k(q_t)$, which is central to our model. The information cost matrix function describes the relative difficulty of learning about different states of the world. In many economic applications, there is a natural ordering of the states of the world. As an example, suppose the states of the world represent possible returns of the stock market, relative to a bank account, and the decision problem is a portfolio choice problem about how much to invest in the stock market. In this case, it seems natural to assume that, when the agent learns about the likelihood of stock returns being 10%, she also learns some information about the likelihood of stock
returns being 9% or 11%, but does not learn much about the likelihood of stock returns being 50% or -50%. Put another way, there are complementarities with respect to learning about the 10% stock market return state and the 9% state. In the information cost matrix, these complementarities are represented by negative off-diagonal elements. Positive off-diagonal elements would represent substitutabilities. These notions of complementarity and substitutability are not present in the standard rational inattention cost function (mutual information).

Additionally, there may be some states that are simply easier or harder to learn about. Continuing with the stock market example, it may be easy to acquire information about the relative likelihood of a 10% stock return and a 5% stock return, but very difficult to acquire information about the relative likelihood of a 10% stock return and a -50% stock return. In the information cost matrix, the relative difficult of learning about a particular state $x \in X$ is represented by the value of diagonal element, $k_{xx}(q_t)$. Again, the idea that certain states can be easier or harder to learn about is not present in the standard rational inattention cost function.

Below, we show an example of an information cost matrix. This example matrix embeds some notion of “distance” between different states of the world:

$$k(q) = \begin{bmatrix}
\frac{1}{2} & -\frac{1}{2} & 0 & \ldots & 0 \\
-\frac{1}{2} & 1 & -\frac{1}{2} & \ddots & \vdots \\
0 & -\frac{1}{2} & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & 1 & -\frac{1}{2} \\
0 & \ldots & 0 & -\frac{1}{2} & \frac{1}{2}
\end{bmatrix}.$$  

Each adjacent state in the matrix is “close,” in there sense that there are complementarities (negative off-diagonal elements) when learning about whether this state occurs and whether
another close state has occurred. Except for the “end” states, this matrix does not exhibit smaller or larger costs of learning about particular states, although it is easy to construct examples that do exhibit this feature. This function also has a notable feature: it is “prior-invariant,” meaning that the matrix is not in fact of a function of the current beliefs \((q)\). Such a cost is interesting, conceptually, but the solutions we discuss in the next section will not apply in the prior-invariant case.

Another example of an information cost matrix is the inverse Fisher information matrix. This matrix, which is related to Shannon’s entropy, mutual information, and the Kullback-Leibler divergence (all of which appear in the standard model of rational inattention and are defined in section §4) can be written as

\[
k(q) = \begin{bmatrix}
q_1(1 - q_1) & -q_1 q_2 & \cdots & -q_1 q_{|X|} \\
-q_1 q_2 & q_2(1 - q_2) & \cdots & -q_2 q_{|X|} \\
\vdots & \vdots & \ddots & \vdots \\
-q_1 q_{|X|} & -q_2 q_{|X|} & \cdots & q_{|X|}(1 - q_{|X|})
\end{bmatrix}.
\]

This information cost matrix imposes the assumption that the difficulty of learning about each state, and the complementaries or substitutabilities between states, are entirely functions of the current beliefs. There is no inherent difficulty of learning about a particular state of the world, or inherent complementarity or substitutability between learning about two states of the world. We view this as problematic in many economic applications. However, this information cost function falls into the class of problems we can solve, and in fact will be closely related to the standard static rational inattention problem with the mutual information cost function.
A third example is the following information cost matrix function:

\[
k(q) = \begin{bmatrix}
\frac{q_1 q_2}{q_1 + q_2} & \frac{q_1 q_2}{q_1 + q_2} & 0 & \ldots & 0 \\
\frac{-q_1 q_2}{q_1 + q_2} + \frac{q_2 q_3}{q_2 + q_3} & \frac{-q_1 q_2}{q_1 + q_2} + \frac{q_2 q_3}{q_2 + q_3} & \ddots & \ddots & \vdots \\
0 & \frac{-q_2 q_3}{q_2 + q_3} & \ddots & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & \frac{-q_{|X|} q_{|X|-1}}{q_{|X|} + q_{|X|-1}} & \frac{-q_{|X|} q_{|X|-1}}{q_{|X|} + q_{|X|-1}} \\
\end{bmatrix}
\]

This information cost function combines two appealing features of the previous two examples. There is complementarity between “nearby” states, but not “far away” states, as in the first example we described. Moreover, this cost function can be analyzed using the solution methods described in the next section, like the inverse Fisher matrix. This information cost function is an example of a more general class, which we will describe in section §5, which all feature notions of states that are “close” or “far,” and can be analyzed with the solution methods we introduce.

Our derivation of the continuous time problem imposes some additional restrictions on the information cost matrix function. We assume that it is differentiable with respect to \(q_t\), and that the second smallest eigenvalue is strongly positive. We describe these assumptions in more detail in section §6. For a large class information cost matrix functions \(k(q_t)\), we will solve the sequence problem described in this section, and show that the solution is a variant of the standard static rational inattention problem. We present these results in the next section.
3 The Equivalence of Static and Dynamic Rational Inattention Problems

In this section, 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 that, for all $q_t$ in the interior of the simplex,

$$D(q_t)^{-1}k(q_t)D(q_t)^{-1} = H_{qq}(q_t). \hspace{1cm} (5)$$

This class includes two information cost matrix functions of interest: when $k(q_t)$ is the inverse Fisher information matrix, which we will show corresponds to the standard rational inattention model, and when $k(q_t)$ is the “neighborhood-based” function we introduce in the next section. The assumption is restrictive, however– it does not cover the case of $k(q_t) = K$, for some constant matrix $K$, which we refer to as the “prior-invariant” case. We will refer to the function $H$ as the “generalized entropy function,” for reasons that will become clear below.

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:

$$V(q_0) = E_0[V(q_\tau) - \theta H(q_\tau) + \theta H(q_0)]$$

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

Using this idea, and the notion that, in an optimal stopping problem, the agent “chooses the boundaries,” we conjecture and verify the following result:
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, \(\pi^*\) and \(q^*_a\), such that \(\pi^*\) 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 agent will hold when stopping and choosing that action.

The divergence \(D_H\) is the Bregman divergence defined by the convex function \(H\),

\[
D_H(p || q) = H(p) - H(q) - (p^T - q^T) \cdot H_q(q).
\]

Proof. See the appendix, section A.2. \(\square\)

The sequential evidence accumulation problem can be thought of as a static rational inattention problem, with a posterior-separable information cost function (in the terminology of Caplin and Dean [2013]) constructed from a particular divergence. The expected value of the divergence embodies the expected time cost required for the agent to reach a decision. The probability distribution \(q^*_a \in \mathcal{P}(X)\) is the agent’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 agent can reach before choosing to stop and take a particular action. The further this belief is from the agent’s prior, \(q_0\), the more time it will take (in expectation) for the agent to arrive at this belief before acting. The divergence encodes this notion of the belief \(q^*_a\) being “far” or “close to” \(q_0\).
The term
\[ \sum_{a \in A} \pi(a) D_H(q_a || q_0) \]
in the statement of theorem 1 can be thought of as the “information cost” in the static problem described in that theorem. In a standard rational inattention problem, the “information cost” is mutual information, the Bregman divergence is the Kullback-Leibler divergence, the generalized entropy function is the negative of Shannon’s entropy, and the corresponding information cost matrix function is the inverse Fisher information matrix.

This illustrates the “two-way” aspect of our results. Given an information cost matrix function \( k(q_t) \) satisfying equation (5), there is a posterior-separable information cost function that, when used in a static rational inattention problem, summarizes the results of the dynamic model. Conversely, given a static rational inattention model with a posterior-separable information cost function, there is a dynamic model with the corresponding information cost matrix function that delivers the same choice probabilities and beliefs conditional on choosing a particular action.

This also suggests an interpretation of the continuous time model. Consider the information cost function in the statement of theorem 1, and suppose that each \( q_a \) was very close to \( q_0 \). We would have, ignoring higher order terms,
\[ \sum_{a \in A} \pi(a) D_H(q_a || q_0) \approx \frac{1}{2} \sum_{a \in A} \pi(a)(q_a - q_0)^T D(q_0)^{-1} k(q_0) D(q_0)^{-1} (q_a - q_0). \]

This cost is reminiscent of the constraint in the continuous time model (equation (2)). Speaking loosely, the continuous time model resembles a sequence of static models, in which the agent gathers a small amount of information each period. We will show this formally in section §6.

This argument depends only on the local (second-order) properties of the information
cost, because it presumes that the amount of information the agent gathers each period is small. We have shown that there is a one-to-one relationship between posterior-separable information costs and information cost matrix functions, but there are in fact many information costs that generate the same information cost matrix function, because all of these information costs resemble, to second-order, some posterior-separable information cost.

Put another way, a sequence of static rational inattention models involving an information cost converges, as time grows short and the agent gathers little information in each time period, to a continuous time model with an information cost matrix function. Solving that continuous time model (assuming the information cost matrix function satisfies equation (5)) leads back to a static rational inattention model, with a posterior-separable information cost. Our results provide a sort of micro-foundation which maps information costs to posterior-separable information costs. Posterior-separable information costs are the fixed points of this mapping: use one to generate an information cost matrix, and solve the model, and you will recover the same posterior-separable information cost.

Thus far, we have been vague about what, exactly, is an information cost. In other words, what is the domain of the mapping described above? In the next section, we will define what we mean by information costs, providing both posterior-separable and non-posterior-separable information costs.

4 Information Costs

In this section, we provide a definition of information costs. These information costs are characterized by a set of conditions that are relevant or useful in economic applications. We show that all of the information costs in this class are approximately the same, in a sense, for signals that convey very little information.

We begin by defining information cost functions in general. As in the previous sections,
let $x \in X$ be the underlying state of the nature, and let $s \in S$ be a signal the agent 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 agent’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 a “signal structure.” After receiving signal $s$, the agent will hold a posterior, $q_s \in \mathcal{P}(X)$, which is a function of $p$ and $q$, defined by Bayes’ rule.

In the static rational inattention problem studied in most of the literature (e.g. Sims [2010]) and derived in theorem 1, the agent chooses an action $a \in A$ whose payoff is determined by the state of nature. As in previous sections, let $\hat{u}(q_s)$ be the agents’ payoff from choosing the optimal action, conditional on having posterior beliefs $q_s$. The standard static rational inattention problem, given the signal alphabet $S$,\(^6\) is

$$
\max_{\{p_x \in \mathcal{P}(S)\}_{x \in X}} \sum_{s \in S} \hat{u}(q_s) - \theta I(\{p_x\}_{x \in X}, q; S),
$$

where $I(\cdot)$ denotes the mutual information between the signal and the state. In this problem, mutual information is the information cost.

There are several equivalent definitions of mutual information. For our purposes, the most convenient formulation uses Shannon’s entropy,

$$
H_S(q) = - \sum_{x \in X} (e_x^T q) \ln(e_x^T q).
$$

Shannon’s entropy is a concave function, and can be used to define a Bregman divergence

\(^6\) The full problem includes a choice over the signal alphabet $S$. A standard result is that $|S| = |A|$ is sufficient.
(the Kullback-Leibler divergence),

\[ D_{KL}(q_s || q) = H_S(q) - H_S(q_s) + (q_s - q)^T H_{S,q}(q). \]

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

\[ I(p, q; S) = \sum_{s \in S} (e_s^T pq) D_{KL}(q_s || q). \]

The information cost functions we consider are generalizations of mutual information. Given a signal alphabet \( S \), an information cost function is a mapping from a signal structure \( \{ p_x \}_{x \in X} \) and prior \( q \) to a cost,

\[ C(\cdot; S) : \mathcal{P}(S)^{|X|} \times \mathcal{P}(X) \rightarrow \mathbb{R}. \]

The cost functions we study, like mutual information, are defined for all finite signal alphabets \( S \). However, mutual information is also defined over alternative sets of states of nature, \( X \). We will not impose this requirement on our more general cost functions— all of our analysis will take the states of nature as given.

We assume four conditions that characterize the family of information cost functions we consider. All of these conditions are satisfied by mutual information, but also by many other cost functions.

**Condition 1.** Signal structures that convey no information \( (p_x = p_{x'} \text{ for all } x, x' \in X) \) have zero cost. All other signal structures have a cost weakly greater than zero.

This condition ensures that the least costly strategy for the agent 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 utility cost is a normalization.
The next condition is called mixture feasibility by Caplin and Dean [2015]. Consider two signal 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 signal 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 \(S_M\), if \(s_1 \in S_1\),

\[
e_s^T p_M = \lambda e_{s_1}^T p_1,
\]
and otherwise \(e_s^T p_M = 0\). Likewise, for each \(s = (s_2, 2)\), if \(s_2 \in S_2\),

\[
e_s^T p_M = (1 - \lambda)e_{s_2}^T p_2,
\]
and otherwise \(e_s^T p_M = 0\). That is, this signal structure results, with probability \(\lambda\), in a posterior associated with signal structure \(p_1\), and with probability \(1 - \lambda\) in a posterior associated with signal structure \(p_2\). In other words, the distribution of posteriors under the mixed signal structure is a convex combination of the distributions of posteriors under the two signal structures. It is as if the agent flipped a coin, observed the result, and then randomly chose one of the two signal structures. The mixture feasibility condition requires that choosing a mixed signal structure costs no more than the cost of randomizing over signal structures (using a mixed strategy in the rational inattention problem).

**Condition 2.** Given two signal 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 signal structure is weakly less than the weighted average of the cost of the separate signal 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 signal structure Blackwell dominates the second signal structure if, for all utility functions $u(a, x)$ and all priors $q \in \mathcal{P}(X)$,

$$\sup_{a(s), x \in X, s \in S} \sum_{x \in X} q_{x} p_{s,x} u(a(s), x) \geq \sup_{a(s'), x \in X, s' \in S'} \sum_{x \in X} q_{x} p'_{s',x} u(a(s'), x).$$

If one signal 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 signal structures do not dominate each other in this sense. However, when a signal structure does Blackwell dominate another signal structure, we assume the dominant signal structure is more costly.

**Condition 3.** If the signal 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 an agent 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 agent pursuing mixed strategies over signal 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. [2014].

The mixture feasibility condition (2) and Blackwell monotonicity condition (3) are equivalent to requiring that the cost function be convex over signal structures and Blackwell monotone. We summarize this equivalent in the following lemma.
Lemma 2. Let $p$ and $p'$ be signal structures with signal alphabet $S$. A cost function is convex in signal 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 (2 and 3) if and only if it is convex in signal structures and satisfies Blackwell monotonicity.

Proof. See the appendix, section A.8. \qed

The fourth condition we assume, which is not imposed by Caplin and Dean [2015] and De Oliveira et al. [2014], is a differentiability condition that will allow us to characterize the local properties of our cost functions.

Condition 4. For all signal alphabets $S$ and $q \in \mathcal{P}(X)$, the information cost function is continuously twice-differentiable in signal structures $\{p_x\}_{x \in X}$, in a neighborhood around any uninformative signal structure, in all directions that do not change the support of the signal distribution.

The last condition we assume, which is also not imposed by Caplin and Dean [2015] and De Oliveira et al. [2014], is a sort of local strong convexity. We will assume that the cost function exhibits strong convexity, in the neighborhood of an uninformative signal structure, with respect to signal 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 signal structures that are sufficiently close to uninformative ($C(p,q;S) < B$),

$$C(p,q;S) \geq \frac{m}{2} \sum_{S \in S} (e^T_S p) ||q_s - q||_X^2.$$
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 1; it essentially “local strong convexity” instead of local convexity. It implies that all informative signal structures have a strictly positive cost, and that (regardless of the agents’ current beliefs) there are no informative signal structures that are “almost free.”

From these five conditions, we derive a result about the second-order properties of the cost function. Condition 4 is not completely general; for example, it rules out the case in which the agent is constrained to use only signals in a parametric family of probability distributions, and the cost of other signal distributions is infinite. Condition 4 also rules out other proposed alternatives, such as the channel capacity constraint suggested by Woodford [2012].\(^7\) Mutual information satisfies each of these five conditions. However, it is not the only cost function to do so.

Consider the following example, constructed from the “f-divergences.” This class of divergences, which includes the Kullback-Leibler divergence, can be defined in our context as

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

for some strictly convex, twice-differentiable function \( f \) with \( f(1) = f'(1) = 0 \) and \( f''(1) = 1 \). The KL divergence corresponds to \( f(u) = u \ln u - u + 1 \). Using this family of divergences, we can define a family of information costs,

\[
I_f(p, q; S) = \sum_{s \in S} (e_s^T pq) D_f(q_s||q),
\]

which nest mutual information as a special case. It is relatively easy to observe that this

\(^7\)We speculate that it may be possible to apply our methods to generalized versions of the channel capacity.
family of information costs satisfies all five of the conditions described above (this is proven as a special case of lemma 3 in the next section).

Next, we discuss the local (second-order) properties of any information cost satisfying the conditions described above. The condition requiring that Blackwell-dominant signal structures cost weakly more (3) is of particular importance. To understand why, it is first useful to recall Blackwell’s theorem.

**Theorem 2. (Blackwell [1953])** If, and only if, the signal 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 there exists a Markov transition matrix \( \Pi : S \to S' \) such that, for all \( s' \in S' \) and \( x \in X \),

\[
p'_x = \Pi p_x. \tag{8}
\]

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 a signal structure \( \{ p_x \}_{x \in X} \), with signal alphabet \( S \), and another signal structure \( \{ p'_x \}_{x \in X} \), with signal alphabet \( S' \), using one of these left-invertible matrices, via equation (8), 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 signal structures are called “Blackwell-equivalent,” and it follows that the cost of these two signal structures must be equal, by condition 3. The left-invertible Markov transition matrices associated with Blackwell-equivalent signal 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
An invariant divergence is a divergence that is invariant to these embeddings. Let $\Pi$ be a Markov congruent embedding from $S$ to $S'$. 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 $D_I$,

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

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

$$g(r) = D(r)^+ - t t^T.$$ 

Here, $D^+(r)$ denotes the pseudo-inverse of the $|S| \times |S|$ diagonal matrix whose diagonal elements are $r$ and $t$ is a vector of ones of length $|S|$.

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 $\Pi$, 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 proved the following
1. Any continuous function that is invariant over the probability simplex is equal to a constant.

2. Any continuous, invariant 1-form tensor field over the probability simplex is equal to zero.

3. Any continuous, invariant quadratic form tensor field over the probability simplex is proportional to the Fisher information matrix.

These results will allow us to characterize the local properties of rational inattention cost functions, via a Taylor expansion. Hold fixed the signal alphabet $S$, and consider a signal structure $p_x(\varepsilon, \nu) = r + \varepsilon \tau_x + \nu \omega_x$, where $r \in \mathcal{P}(S)$. Here, $\tau_x$ and $\omega_x$ represent informative signal structures such that $e^T_\varepsilon \tau_x \neq 0$ only if $e^T_\varepsilon r > 0$, and likewise for $\omega$. The scalars $\varepsilon$ and $\nu$ are perturbations in those directions, away from an uninformative signal structure.

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

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

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, meaning that, for any Markov congruent embedding $\Pi : S \rightarrow S'$,

$$C_x(\{r\}_{x \in X}, q; S) \cdot \tau_x = C_x(\{\Pi r\}_{x \in X}, q; S'); \Pi \tau_x.$$

Therefore, by Chentsov’s results, it is equal to zero.

---

8Lemma 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 $S$. 

29
We repeat the argument for the second derivative terms. Those terms can be written as

\[
\frac{\partial}{\partial \nu} \frac{\partial}{\partial \varepsilon} C(\{p_x(\varepsilon, \nu)\}_{x \in X}, q; S) |_{\varepsilon=\nu=0} = \sum_{x' \in X} \sum_{x \in X} \omega_x^T \cdot C_{xx'}(\{r\}_{x \in X}, q; S) \cdot \tau_x.
\]

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

\[
\frac{\partial}{\partial \nu} \frac{\partial}{\partial \varepsilon} C(\{p(\cdot|\varepsilon, \nu)\}, q) |_{\varepsilon=\nu=0} = \sum_{x' \in X} \sum_{x \in X} (e_x^T k(q) e_{x'}) \omega_x^T g(r) \tau_x,
\]

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. In the case of the mutual information cost, the matrix \(k(q)\) is itself the inverse Fisher information matrix,

\[
k(q) = g^+(q) = D(q) - qq^T.
\]

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. Our choice of notation hints at what we will prove in section §6: that this matrix-valued function is the information cost matrix function described in section §2.

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 \(\Delta\) (we use \(\Delta\) because in the next sub-section, we will be looking at small time intervals).

**Theorem 3.** Suppose that a signal structure \(\{p_x\}_{x \in X}\), with signal alphabet \(S\), is described by the equation

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

30
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 a rational inattention cost function that satisfies conditions 1-4. Then, for $\Delta$ sufficiently small,

1. The cost of this signal structure is, for some matrix $k(q)$,

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

2. The gradient of this cost function with respect to $\tau_x$, for a given $x \in X$,

$$C_x(\cdot) = \Delta^{0.5} \sum_{x' \in X} (e_s^T k(q) e_{x'}) \tau_{x'}^T g(r) + o(\Delta^{1/2}).$$

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

Suppose that the cost function also satisfies condition 5. Then, for some constant $m_g > 0$, 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 appendix, section A.9.

The results of theorem 3 characterize the cost of a small amount of information, for any rational inattention cost function satisfying our conditions. The theorem substantially restricts the local structure of the cost function, 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 conditions 1, 2, and 4 (but not the Blackwell’s ordering condition, condition 3), the cost function would locally resemble an inner product with respect to a positive semi-definite, $N \times N$ matrix. By imposing condition 3, the results of theorem 3 show that we can restrict this matrix to the $k(q)$ matrix, an $|X| \times |X|$ matrix. If the agent were only allowed binary
signals ($|S| = 2$), this restriction would be trivial. When the agent is allowed to contemplate more general signal 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 redefine the cost function using the posteriors and unconditional signal probabilities, rather than the prior. The results are described in the corollary below.

**Corollary 1.** Under the assumptions of theorem 3, the posterior beliefs can be written, for any $s \in S$ such that $e_s^T r > 0$, as

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

Define the matrix

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

where $D^+(\cdot)$ is the pseudo-inverse of the diagonal matrix. Adopting the convention that $q_s = q$ for any $s \in S$ such that $e_s^T r = 0$, the cost function can be written as

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

**Proof.** See the appendix, section A.10. \qed

Consider the family of information costs built from f-divergences defined above (equation (7)). All of the cost functions in this family resemble mutual information, to second order, in the sense defined by corollary 1. Assuming that the posteriors induced by the signal 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,

$$I_f(p, q; S) \approx \frac{1}{2} \sum_{s \in S} (e_s^T pq)(q_s - q)^T D(q)^+ g^+(q) D(q)^+(q_s - q).$$  \hspace{1cm} (10)$$

In other words, in a sense that we will show formally in section §6, all of these information costs will induce the same information cost matrix function in the continuous time problem. As a result, by theorem 1, all of 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 information cost originally is the KL divergence or not, the KL divergence will appear in the solution to the continuous time problem. In fact, this result applies to the larger class of invariant divergences, which includes the f-divergences, and follows from Chentsov’s theorem (equation (9)).

Of course, we argued in section §2 that the inverse Fisher information matrix, when used as the information cost matrix function, lacks certain desirable properties related to the distance between different states of the world. In the next section, we will introduce a new family of cost functions, all of which induce information cost matrix functions that do capture these notions. Moreover, these information cost matrix functions satisfy equation (5), and therefore theorem 1 applies. We will 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 matters in economic applications.

### 5 The Neighborhood-Based Cost Function

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 \( \{x_0, \ldots, x_n\} \) 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 \( I(q) \) 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 \( q \in \mathcal{P}(X_i) \) 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 signal structure \( p \) and prior \( q \):
the form

\[ C_N(p, q; S) = \sum_{i \in \mathcal{I}(q)} \tilde{q}_i \sum_{s \in S} e_i^T \tilde{p}_i D_i(q_i, s || q_i), \]  

(11)

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.\(^9\) 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 Kullback-Leibler divergence, so that

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

The information costs discussed in the previous sections, based on f-divergences, are also single-neighborhood versions of neighborhood-based cost functions. The following lemma shows that all cost functions with a neighborhood structure satisfy the conditions defined in the previous section.

**Lemma 3.** All cost functions with a neighborhood structure (11) satisfy the conditions 1-4 described 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 A.3. \( \square \)

An implication of this lemma is that all posterior-separable cost functions (defined by Caplin and Dean [2013] and discussed in section §3) with strongly convex generalized entropy functions satisfy our conditions.

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 requirement that the divergences \( D_i(\cdot || \cdot) \) be invariant. This family can have complex neighborhood

---

\(^9\)The f-divergences defined previously satisfy these condition (Amari and Nagaoka [2007]).
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 all states belong to a single neighborhood (the case that includes the mutual information cost function), 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 (11), 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 — but also because the divergences $D_i$ can be different for different neighborhoods.

As discussed above, 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 (10) applies, but only within each neighborhood. That is,

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

(12)

where $c_i > 0$ are positive constants (this approximation is proven in the proof of lemma 4, below). 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 $c_i > 0$ are positive constants.

**Proof.** See the appendix, section A.4.
We can use the information cost matrix function in our continuous time problem (the problem defined in section §2). It satisfies the equation necessary for the results of theorem 1 to apply (equation (5)). As a result, there is a generalized entropy function, $H_N(q)$, associated Bregman divergence, $D_N(p||q)$, and posterior-separable information cost, $C_{NPS}(p,q;S)$, that can be used to define the static rational inattention problem whose choice probabilities coincide with the solution to the dynamic model. The following lemma describes these functions:

**Lemma 5.** Let $H_S(q)$ denote Shannon’s entropy. The generalized entropy function $H_N(q)$, associated with the neighborhood-based information cost matrix function $k_N(q)$, is

$$H_N(q) = - \sum_{i \in I(q)} c_i \tilde{q}_i H_S(q_i).$$

The posterior-separable information cost developed from the neighborhood-based generalized entropy can be written as

$$C_N(p,q;S) = \sum_{i \in I(q)} c_i \tilde{q}_i \sum_{x \in X} (e^T q)_x D_{KL}(p_{x|x \in X_i} || p_{E_i^T q_i}),$$

or alternatively as

$$C_N(p,q;S) = \sum_{i \in I(q)} c_i \sum_{x \in X} (e^T q)_x D_{KL}(p_{e_x} || p_{E_i^T q_i}).$$

**Proof.** See the appendix, section A.5.

This result allows us to write the static rational inattention problem (theorem 1) directly

---

10 Our derivation of the continuous time model from the discrete time model applies only to cost functions satisfying conditions 1-5. 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.
in terms of an optimization over choice probabilities \( \{\pi_x\} \) so as to maximize

\[
\sum_{x \in X} e_x^T q_0 \sum_{a \in A} e_a^T \pi_x u_{x.a} - \theta C(\{\pi_x\}_{x \in X}, q_0; A).
\] (13)

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 (13) 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 direction from the problem in 13 and the alternative expression of the cost function in lemma 5. \qed

The significance of lemma 6 can be seen if we consider the predictions of rational inattention for a standard form of perceptual discrimination experiment. Suppose that the different states \( X = \{1, 2, \ldots, N\} \) 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 \((L \text{ or } R)\); \( R \) 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 figure 1, which plots the optimal response frequencies as a function of \( x \), for alternative values of the cost parameter \( \theta \), 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

\[
X_i = \{x_i, x_{i+1}\}
\]

for \( i = 1, 2, \ldots, N - 1 \). 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.\(^{11}\) 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 figure 2, which again shows the optimal response frequencies as a function of \( x \), for alternative values of \( \theta \), in the case of the alternative neighborhood structure \((14)\). The sigmoid functions predicted by rational inattention with this cost function — with the property that response frequen-

\(^{11}\)If \( c_i \) is the same for all \( i \), we can without loss of generality set it equal to one, as the multiplier \( \theta \) can still be used to scale the overall magnitude of information costs.
Figure 2: Predicted response probabilities with a neighborhood-based cost function, in which each neighborhood consists only of two adjacent states.
cies 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 of stimuli that are sufficiently extreme — are characteristic of measured “psychometric functions” in perceptual experiments of this kind. [Get references.]

The continuity of choice probabilities across 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 referenced frequently are naturally modeled with a continuum of states. In the next sub-section, we consider the infinite-state limit of example just described, in which each neighborhood consists of a pair of adjacent states.

5.1 The Infinite-State Limit

In this subsection, we provide an example of our neighborhood-based cost function, and consider the limit of this example 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 will derive our limit for arbitrary action spaces and utility functions. The result, which will be a static rational inattention problem with a continuum of states, will use an information cost based on the expected change in the Fisher information measure (which we will define below; see also Cover and Thomas [2012]). The expected change in Fisher information, like the expected change in Shannon’s entropy (mutual information) is a single-parameter information cost that can be applied in almost any context. Unlike Shannon’s entropy, the Fisher information measure captures notions of distance or localness, a property that makes it useful in a variety of physics applications (Frieden [2004]).

In standard global games models, with exogenous private information, there is a unique
equilibrium despite the incentives for coordination across agents (subject to some caveats and details that are not relevant for our discussion). Yang [2015] demonstrates that allowing for endogenous information acquisition in global games, with mutual information as the information cost, restores a multiplicity of equilibria. The intuition is that agents 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, invest or not invest) on a particular threshold, and there are many such thresholds that are feasible. Morris and Yang [2016] develop the complementary result, showing that when the probability of investment must vary smoothly across states, there is a unique equilibrium.

The requirement that the agents’ investment choice vary smoothly across states is closely related to the notion of distance between states that the neighborhood-based cost function seeks to capture. However, our results cannot be applied directly to the model of Morris and Yang [2016], because the global games model in that paper presumes a continuum of states, whereas our analysis 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 infinite. We will 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.

We begin by discussing our example neighborhood structure. The set of states is ordered, $X^N = \{0, 1, \ldots, N\}$, and each pair of adjacent states forms a neighborhood, $X_j = \{i, i+1\}$, for all $j \in \{0, 1, \ldots, N-1\}$. We will also assume that there is an $N^{th}$ 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 study 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 will assume there is a differentiable PDF \( f : [0, 1] \to \mathbb{R}^+ \), with full support on \([0, 1]\), and whose derivative 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, each initial belief \( q_N \) is a discrete approximation of a Lipschitz continuous PDF \( f(x) \), which becomes increasingly accurate as \( N \to \infty \).

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, this scaling ensures that the agent is neither able to determine the state with certainty or prevented from gathering any useful information, and that the neighborhood containing all states plays no role in the limiting behavior, implying that all costs are local. We will scale the entire cost function by a constant, \( \bar{\theta} \).

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 \( N \to \infty \), 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, \( u_a : [0, 1] \to \mathbb{R} \) is a bounded and continuous utility function defined for each action \( a \in A \). 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

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

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

The following theorem shows that the solution to this problem, both in terms of the value function and the optimal policies, converges to the solution of a static rational inattention problem over a continuum of states.

**Theorem 4.** Consider the following static rational inattention problem:

\[
V(f; \bar{\theta}) = \sup_{\pi \in \mathcal{P}(A), \{f_a \in \mathcal{P}_{LipG}([0,1])\}_{a \in A}} \sum_{a \in A} \pi_N(a) \int_0^1 u_a(x)f_a(x)dx \\
- \frac{\bar{\theta}}{2} \sum_{a \in A} \{\pi(a) \int_0^1 (f_a'(x))^2dx \} + \frac{\bar{\theta}}{2} \int_0^1 \frac{(f'(x))^2}{f(x)}dx,
\]

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

\[ \sum_{a \in A} \pi(a)f_a(x) = f(x), \]

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. Let \( \pi^*(a) \) and \( f^*_a \) denote optimal policies. Then there exists a sequence of increasing \( N \), which we denote by \( n \in \mathbb{N} \), such that

1. \( \lim_{n \to \infty} V_n(q_n; \bar{\theta}) = V(q; \bar{\theta}) \)

2. \( \lim_{n \to \infty} \pi^*_n = \pi^* \)
3. For all $a \in A$ and all $x \in [0,1]$, \( \lim_{n \to \infty} \sum_{i=0}^{\lfloor xn \rfloor} e_i^T q_{a,n} = \int_x^0 f_a^*(y)dy \).

**Proof.** See the appendix, section A.7.

This theorem shows that the limit of our neighborhoods problem converges to a static rational inattention problem with a particular cost function. That cost function is the expected value of something that (unfortunately, for the purpose of clarity in this paper) is called the Fisher information measure (Cover and Thomas [2012], Frieden [2004]). This cost function, unlike mutual information, will generate the smoothness of responses across discrete changes in payoffs shown in figure 2. For this reason, we believe it is more appropriate than mutual information in a wide range of settings and, like mutual information, has only a single degree of freedom.

We now turn to the question motivating this example: does this static rational inattention problem satisfy the continuous choice condition of Morris and Yang [2016]? Those authors study a global game with two actions, “invest” and “not-invest.” They express their condition in terms of the likelihood of investing in particular state, which can be expressed (by Bayes’ rule) as

\[
s(x) = \frac{\pi^*(\text{invest}) f^*_{\text{invest}}(x)}{f(x)},
\]

where \( \pi^* \) and \( f^*_{\text{invest}}(x) \) are optimal policies for the problem defined in theorem 4. Morris and Yang [2016] require continuous choice, meaning that for all \( \bar{\theta} > 0 \) and all parameterizations of the relevant utility function, \( s(x) \) is absolutely continuous. That continuous choice is satisfied in our example follows immediately from the statement of theorem 4.

Thus far, we have described and solved our continuous time model, and introduced neighborhood-based cost functions as an alternative to mutual information. We now circle back to the beginning, and show that the continuous time model described in section §2 is the limit of a discrete time model with sequential evidence accumulation. We will justify the link made previously between the second-order approximation of any information cost
(section §4) and the information cost matrix function defined in section §2.

6 Derivation of the Continuous Time Model

In this section, we will derive our continuous time model from a discrete time model that resembles discrete-time rational inattention models (e.g. Steiner et al. [2016]). We study a dynamic problem in which the agent 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 agent can either stop and take an action \( a \in A \), or continue and receive a signal 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 \( |S| > |X| \). However, signal structure \( \{p_{t,x}\}_{x \in X} \) is a choice variable that can be state- and time-dependent. Fixing the signal structure \( S \) 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 signal structure and invoke the results from the previous sub-section.\(^{12}\)

The agent’s prior beliefs at time \( t \), before receiving the signal, are denoted \( q_t \). Each time period has a length \( \Delta \). Let \( \tau \) denote the time at which the agent stops and makes a decision, with \( \tau = 0 \) corresponding to making a decision without acquiring any information. At this time, the agent receives utility \( u(x, a) - \kappa \tau \) if she takes action \( a \) at time \( \tau \) and the true state of the world is \( x \). As in the previous sections, let \( \tilde{u}(q_\tau) \) be the utility (not including the penalty for delay) associated with taking an optimal action under beliefs \( q_\tau \). The parameter \( \kappa \) governs the size of the penalty the agent faces from delaying his decision. The reason the agent does not make a decision immediately is that she is able to gather information,\(^{12}\)

\(^{12}\)As mentioned previously, the work of Ay et al. [2014] discusses how to extend the Chentsov [1982] theorems to infinite dimensional structures. We speculate that their results would allow us to extend our theorems to infinite dimensional signal spaces.
and make a more-informed decision. Note that this setup closely resembles the continuous
time model described previously, except that we use $\kappa$ instead of $\mu$ to describe the penalty
for delay. This choice is intentional– we will derive a value for the parameter $\mu$ in the
continuous time model that depends in part on $\kappa$, but also on other parameters.

The agent can choose a signal 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 agent needs to design the
signal structure. The agent is constrained to satisfy

$$E_0[\frac{\Delta^{\tau \Delta^{-1}-1}}{\rho} \sum_{j=0} C(\{p_{\Delta jx}x' \in X, q_{\Delta j}; S\})^\rho]^{\frac{1}{\rho}} \leq \Delta c E_0[\tau], \quad (15)$$

if the agent choose to acquire any information at all ($\tau > 0$ always in this case). In words,
the $L^p$-norm of the flow information cost function $C(\cdot)$ over time and possible histories
must be less that the constant $c$ per unit time. In the limit as $\rho \to \infty$, this would approach a
per-period constraint on the amount of information the agent can obtain. For finite values
of $\rho$, the agent can allocate more information gathering to states and times in which it is
more advantageous to gather more information. We will assume, however, that $\rho > 1$, for
reasons that we will discuss later. Because we are holding the signal alphabet fixed, we
will generally omit the argument $S$ when we write the flow information cost function $C(\cdot)$.

Let $V(q_0; \Delta)$ denote the value of the solution to the sequence problem for an agent with
prior beliefs $q_0$, and let $q_\tau$ denote the agent’s beliefs when stopping to make a decision.

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

subject to the information cost constraint, equation (15). The dual version of this problem
can be written, assuming the agent acquires some information, as

\[
W(q_0, \lambda; \Delta) = \max_{\{p_{\Delta}, \tau\}} E_0[\hat{u}(q_\tau) - \kappa \tau] - \\
\lambda E_0[\Delta^{-\rho} \sum_{j=0}^{\Delta^{-1}-1} \left\{ \frac{1}{\rho} C\left( \left\{ p_{\Delta_j, x}\right\}_{x \in X}, q_{\Delta_j}(\cdot) \right) \rho - \Delta^{\rho} c^\rho \right\}].
\]  

(16)

Here, the function \( W(q_0, \lambda; \Delta) \) can be thought of as the value function of a different problem, in which there is a cost of gathering information proportional to \( \lambda \frac{1}{\rho} C(\cdot)^\rho \). In what follows, we will refer to the function \( W \) as the value function, bearing in mind that \( \lambda \) is not actually exogenous to the problem. We will proceed under the assumption that \( \lambda \in (0, \kappa c^{-\rho}) \). Below, we demonstrate that there is no duality gap in the continuous time limit of this problem, and that our assumption about \( \lambda \) is without loss of generality.

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:

\[
W(q_t, \lambda; \Delta) = \max_{\{p_{x,s}\}_{x \in X}} -\kappa\Delta + \lambda \Delta^{-\rho} (\Delta^{\rho} c^\rho - \frac{1}{\rho} C(\cdot)^\rho) + \\
\sum_{x \in X} \sum_{s \in S} p_{t,s,x} q_{t,s} W(q_{t+\Delta,s}, \lambda; \Delta, x), \hat{u}(q_t) \right\},
\]

where \( q_{t+\Delta,x} \) is pinned down by Bayes’ rule and \( W(q_t, \lambda; \Delta, x) \) is the “state-specific” value function (the value function conditional on the true state being \( x \)). The state-specific value function also has a recursive representation, in the region in which the agent continues to gather information:
\[ W(q_t, \lambda; \Delta, x) = -\kappa \Delta + \lambda \Delta^{1-\rho} (\Delta^\rho c^\rho - \frac{1}{\rho} C(\cdot)^\rho) + \sum_{s \in S} \rho_{t,s,x}^* W(q_{t+\Delta,s}^*, \lambda; \Delta, x). \]

In this equation, we take the optimal signal structure as given. Note that, by construction, wherever the agent does not choose to stop, the expected value of the state-specific value functions is equal to the value function.

\[ \sum_{x \in X} q_{t,x} W(q_t, \lambda; \Delta, x) = W(q_t, \lambda; \Delta). \]

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

**Lemma 7.** The value function \( W(q_t, \lambda; \Delta) \) is bounded on \( q_t \in \mathcal{P}(X) \), and convex in \( q \).

**Proof.** See the appendix, section A.11.

The boundedness of the value function follows from the setup of the problem: ultimately, the agent 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 agent 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 agent improves her welfare by enabling better decision making.

The boundedness and convexity of the value function are sufficient to establish that a second-order Taylor expansion of the value function exists almost everywhere. Using this result, we can characterize the optimal information gathering policy (or policies, as there is not necessarily a unique optimum) as the time period shrinks. In particular, we establish that an optimal signal structure \( p_{t,\Delta}^*(\cdot|\cdot) \) converges to an uninformative signal structure as
Lemma 8. Let $\Delta_m, m \in \mathbb{N}$, denote a sequence such that $\lim_{m\to\infty} \Delta_m = 0$. At each time $t$ and prior $q_t \in \mathcal{P}(X)$, the optimal policy $p^*_{t,\Delta_m}$ has a convergent sub-sequence, denoted $p^*_{t,n}$, that can be expressed as

$$p^*_{t,n,x} = r^*_t + \phi^*_t + \Delta_n^\frac{1}{2} \tau^*_t + o(\Delta_n^\frac{1}{2}),$$

where $r^*_t \in \mathcal{P}(S), \lim_{n\to\infty} \phi^*_{t,n,s} = 0$ for all $s \in S$, $t^T \phi^*_t = t^T \tau^*_t = 0$ for all $x \in X$, $e^T \tau^*_t q_t = 0$ for all $s \in S$, and $\|\tau^*_s D(q_t)\|$ is bounded.

For all $s$ such that $e^T s r^*_t > 0$, the posterior associated with receiving signal $s$ satisfies

$$q^*_{t,n,s} = q_t + \Delta_n^\frac{1}{2} e^T s \tau^*_t D(q_t) e^T s r^*_t + o(\Delta_n^\frac{1}{2}).$$

Proof. See appendix, section A.13.

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 signal structure to an uninformative one, as the time period shrinks, allows us to use the approximation described in the previous sub-section of this paper 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 agent 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. [2016] and Che and Mierendorff [2016]).

In standard rational inattention problems, it is without loss of generality to equate signals and actions. In this problem, when the agent 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 agent to be choosing a probability distribution over posteriors, subject to the constraint that the expectation of the posterior is equal to the prior. For any convergent sub-sequence of optimal policies, we can define the revision to the posterior (up to order $\Delta^2$) as

$$z^*_t, s = \frac{e^T_s \tau^*_t D(q_t)}{e^T_s r^*_t}.$$

The almost-everywhere differentiability of the value function invites us to consider approximations to the Bellman equations described previously.

**Lemma 9.** For any $q_t \in \mathcal{P}(X)$ and an associated convergent sub-sequence of optimal policies (as defined in lemma 8), if the value function is twice-differentiable with respect to $q$ at $q_t$, then

$$\theta \sum_{s \in S} (e^T_s r^*_t) z^*_{t,s} k(q_t) z^*_t, s = \sum_{s \in S} (e^T_s r^*_t) z^*_{t,s} \cdot W_{qq}(q_t, \lambda; \Delta) \cdot z^*_t, s + o(1),$$

where $\theta = \lambda \left( \frac{\kappa - \lambda c^p}{\lambda(p - 1)} \right)^{\rho - 1}$. The scale of the vector $z^*_t, s$ is such that, up to order $\Delta$, the flow information cost is constant everywhere:

$$\frac{1}{\rho} \Delta^\rho C(\cdot)^\rho = \frac{\kappa - \lambda c^p}{\lambda(p - 1)} + o(1).$$

**Proof.** See appendix, section A.14. 

The results of this proposition describe an approximate difference equation that the value function must satisfy, for any update to the posterior belief, $z^*_t, s$, that would be observed under the optimal policy. It is tempting to neglect the higher-order terms in that

---

13 The notion of choosing a probability distribution over posteriors appears in Kamenica and Gentzkow [2011] and Caplin and Dean [2015], among other papers.
proposition and attempt to solve the resulting continuous time problem. We next establish that the value function associated with the discrete time problem does in fact converge to a function characterized by a version of this equation that neglects higher-order terms. We will call this function “the” continuous time value function, although we have not yet proved it is unique. We also have yet to show that this function is the solution to the continuous-time rational inattention problem; thus far, it is a function which is very close to the solution of the discrete time rational inattention problem, as the time interval becomes small.

We begin by defining the covariance matrix of the posterior beliefs under the optimal policy, given some time interval $\Delta_n$ (as in lemma 8):

$$\Omega^*_{t,n} = \sum_{s \in S} (e^T_s p^*_t q_t - q_t)(q^*_t - q_t)^T,$$

As mentioned previously, given some beliefs $q_t$, there may be multiple optimal policies, and therefore multiple matrices $\Omega_{t,n}$ that characterize the covariance matrix of updates to beliefs under an optimal policy. We begin by showing that there exists a convergent subsequence of time intervals such that a limiting value function and stochastic process for beliefs exist.

**Lemma 10.** Let $\Delta_m$, $m \in \mathbb{N}$, denote a sequence such that $\lim_{m \to \infty} \Delta_m = 0$. Let $q_{t,m}$ denote the stochastic process for the agent’s beliefs at time $t$, under the optimal policy, given $\Delta = \Delta_m$. There exists a sub-sequence $n \in \mathbb{N}$ such that

1. The value function $W(q, \lambda; \Delta_n)$ converges uniformly on $q \in \mathcal{P}(X)$ to a bounded and convex function, $W(q, \lambda)$. 

53
2. The law of the stochastic process \( q_{t,n} \) converges, in the sense of weak convergence, to

\[
q_t = q_0 + \int_0^t D(q_s)\sigma_s^* dB_s,
\]

where \( B_s \) is an \(|X|\)-dimensional Brownian motion, each element of \( D(q_s)\sigma_s^*\sigma_s^{*T}D(q_s) \) is a uniformly integrable stochastic process adapted to the filtration generated by \( q_t \), and

\[
\lim_{n \to \infty} \Delta_n^{-1}\Omega_{s,n} = D(q_s)\sigma_s^*\sigma_s^{*T}D(q_s).
\]

Proof. See the appendix, section A.15.

A key implication of lemma 10 is the existence of a limit for \( \Delta_n^{-1}\Omega_{t,n} \). Thus far, we have avoided claiming that there is a unique optimal policy; the results in the previous section consider only convergent sub-sequences of policies at a particular point. The results in lemma 10 prove something stronger— that under a sub-sequence of optimal policies, the stochastic process for beliefs converges to a particular martingale. As a result, we can speak of a convergent sub-sequence of policies (function of beliefs), as opposed to a point-wise convergent sub-sequence.

To finish the proof, we resolve several issues. We show that the limiting value function \( W \) is unique, that duality holds (\( V = W \) for a suitable choice of lambda), and that the the limit of \( V \) is the solution to the continuous time problem described in section §2.

**Theorem 5.** Let \( n \in \mathbb{N} \) index a sub-sequence of policies described in lemma 10. There exists a \( \lambda^* \in (0, \kappa c^{-p}) \) such that

\[
\lim_{n \to \infty} W(q_t, \lambda^*; \Delta_n) = \lim_{n \to \infty} V(q_0; \Delta_n) = V(q_0),
\]
where \( V(q_0) \) is the solution to the continuous time problem described in section §2, with \( \chi = \rho^{p^{-1}} c \) and \( \mu = \kappa \). The optimal stopping times \( \tau_n \) and stochastic process \( q_{t,n} \) converge (in the sense of weak convergence) to optimal policies in the continuous time model.

**Proof.** See the appendix, section A.16. The proof relies on the results of Amin and Khanna [1994].

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

Lemma 5 demonstrates the stochastic process for \( q_{t,n} \) converges to the martingale described in section §2. This is the “unconditional” stochastic process, rather than the stochastic process conditional on any particular state. The following corollary demonstrates that, conditional on some particular true state \( x' \in X \), the stochastic process for \( q_{t,n} \) converges to the conditional stochastic process described in section §2.

**Corollary 2.** Let \( n \in \mathbb{N} \) index a sub-sequence of policies described in lemma 10. Assume that the true state is \( x' \in X \) and that \( e_{x'}^T q_0 > 0 \). Let \( \tau \) be the lesser of some time \( t \) and the first hitting time of \( e_{x'}^T q_r = 0 \). The law of the stochastic process \( q_{\tau,n} \) converges, in the sense of weak convergence, to

\[
q_\tau = q_0 + \int_0^\tau D(q_r) \sigma_r^* e_r^* d\tau + \int_0^\tau D(q_r) \sigma_r^* d\tilde{B}_r,
\]

where \( \tilde{B}_r \) is an \( |X| \)-dimensional Brownian motion.

**Proof.** See the appendix, section A.17. \( \square \)
The statement of corollary 2 is written to avoid the assumption that the agent will never put zero probability on the true state of the world. In fact, we believe that the agent’s beliefs will always remain on the interior of the simplex (assuming they start in the interior) and that these caveats are unnecessary. However, proving this was not necessary for any of our other results. We leave a detailed study of the belief dynamics in our model to future work.

7 Conclusion

We have derived a continuous time model of rational inattention as the limit of a discrete time sequential evidence accumulation problem. In this discrete time problem, we assumed a cost function derived from a small set of conditions. Using this conditions, and taking the continuous time limit, we demonstrated that the continuous time problem is characterized by an information cost matrix function. This function captures assumptions about the similarity or differences between states, and how the cost of acquiring information varies with the agents’ prior beliefs. For a large class of information cost functions, we have shown that the continuous time model is equivalent to a static rational inattention problem, with a posterior-separable cost function. That is, posterior-separable cost functions, when employed in static rational inattention problems, can be justified as summarizing a dynamic evidence accumulation process. Mutual information, the standard cost function assumed in the literature, falls into this class. However, mutual information lacks notions of “distance” between different states of the world. We introduce the neighborhood-based cost functions, which capture this notion, and also fall into the posterior-separable class. We argue that this more accurately capture behavior in perceptual experiments. We show a particular example, involving a continuous state limit, in which the Fisher information emerges as the limit of a neighborhood-based cost function.
References


Mark Caplin and Andrew Dean. The behavioral implications of rational inattention with shannon entropy. 2013.


Frank Huettner, Tamer Boyaci, and Yalcin Akcay. Consumer choice under limited attention when options have different information costs. 2016.


A Proofs

A.1 Proof of lemma 1

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

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

subject to

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

First, suppose that the constraint does not bind:

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

where $\sigma_t^*$ 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

$$\frac{1}{2} tr[\sigma_T^T D(q_t) V_{qq}(q_t) D(q_t) \sigma_t] = c^2 \mu > \frac{1}{2} tr[\sigma_T^* T D(q_t) V_{qq}(q_t) D(q_t) \sigma_t^*] = \mu.$$
a contradiction. Therefore, the constraint binds under the optimal policy.

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

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

Differentiating the homotheticity 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 $\alpha \in \mathbb{R}$,

$$\frac{1}{2} tr[(\sigma_t \sigma_T^T + \alpha t t^T) (D(q_t) V_{qq}(q_t) D(q_t) - \theta k(q_t))] = \frac{1}{2} tr[(\sigma_t \sigma_T^T) (D(q_t) V_{qq}(q_t) D(q_t) - \theta k(q_t))].$$

Suppose that we relax the requirement that $q_t^T \sigma_t = \vec{0}$ and simply require that $\sigma_t$ by 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,

$$VDV^T = \tilde{\sigma}_t \tilde{\sigma}_t^T + \alpha t t^T.$$
with $q_t^TV = \tilde{0}$, we construct a matrix

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

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 $\sigma_t = v_1e_1^T$, where $v_1$ is the associated eigenvector.

Finally, note that the agent 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 agent must choose to continue, and (assuming twice-differentiability) the HJB condition must hold.

### A.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
agent, and that any other strategy for the agent achieves weakly less utility.

We begin by observing that

\[ i^T k(q_t)D(q_t)^{-1} = 0 = i^T D(q_t)H_{qq}(q_t) = q_t^T H_{qq}(q_t). \]

We claim that, without loss of generality, we can assume that \( H(q_t) \) is homothetic 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 \( \alpha \) 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, \( \hat{q}_a \in \mathbb{R}^{\vert X \vert} \), and the constraint is \( \sum_{a \in A} \hat{q}_a = q_0 \). By the homotheticity of \( H \), the objective is

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

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

\[ \pi(a) = t^T \hat{q}_a, \]

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 \( \pi(a) \geq 0 \).

Therefore, we can implement any set of \( \hat{q}_a \) satisfying the constraints.

Rewriting the problem in Lagrangian form,

\[
\phi(q_0) = \max_{\{q_a \in \mathbb{R}^{\vert X \vert}\}_{a \in A}} \min_{\kappa \in \mathbb{R}^{\vert X \vert}, \{V_a \in \mathbb{R}^{\vert X \vert}\}_{a \in A}} \sum_{a \in A} u_a^T \cdot \hat{q}_a + \theta H(q_0) - \theta \sum_{a \in A} H(\hat{q}_a) + \kappa^T (q_0 - \sum_{a \in A} \hat{q}_a) + \sum_{a \in A} V_a^T \hat{q}_a.
\]

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 agent 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 agent 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 $\lambda$ 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 is 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 $a \in A$),

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

Define $\hat{q}_a(q_0)$, $\kappa(q_0)$, and $\nu_a(q_0)$ 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 $\nu_a(\tilde{q}_0) =$
νₐ(q₀). By the homotheticity property,

\[ H_q(\alpha(a)\hat{q}_a(q₀)) = H_q(\hat{q}_a(q₀)), \]

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 \( ν_a \) 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

\[ q₀(ε;z) = q₀ + εz, \]

with \( z \in \mathbb{R}^{|X|} \), such that \( q₀(ε;z) \) remains in \( \mathcal{P}(X) \) for some \( ε > 0 \). If \( z \) is in the span of \( \hat{q}_a(q₀) \), then there exists a sufficiently small \( ε > 0 \) such that the above conjecture applies. It follows in this case that \( κ \) is constant, and therefore \( \phi_q(q₀(ε;z)) \) is directionally differentiable with respect to \( ε \). If \( q₀(−ε;z) \in \mathcal{P}(X) \) for some \( ε > 0 \), then \( \phi_q \) is differentiable, with

\[ \phi_{qq}(q₀) \cdot z = θH_{qq}(q₀) \cdot z, \]

proving twice-differentiability in this direction. This perturbation exists anywhere the span of \( \hat{q}_a(q₀) \) is strictly larger than the line segment connecting zero and \( q₀ \) (in other words, all \( \hat{q}_a(q₀) \) are not proportional to \( q₀ \)). Define this region as the continuation region, \( Ω \). Outside of this region, all \( \hat{q}_a(q₀) \) are proportional to \( q₀ \), implying that

\[ \phi(q₀) = \max_{a ∈ A} u_a^T \cdot q₀, \]

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

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

as required.

Now consider an arbitrary perturbation $z$ such that $q_0(\varepsilon; z) \in \mathbb{R}_+^{|X|}$ and $q_0(-\varepsilon; z) \in \mathbb{R}_+^{|X|}$ 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,

$$(\kappa^T(q_0(\varepsilon; z)) - \kappa^T(q_0))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)) = [\theta H_q(\hat{q}_a(q_0)) - \theta H_q(\hat{q}_a(\varepsilon; z)) + v_a^T(q_0(\varepsilon; z)) - v_a^T(q_0)](\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)).$$

Consider the term

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

By the complementary slackness condition,

$$(v_a^T(q_0(\varepsilon; z)) - v_a^T(q_0))(\hat{q}_a(\varepsilon; z) - \hat{q}_a(q_0)) = -v_a^T(q_0(\varepsilon; z))\hat{q}_a(q_0) - v_a^T(q_0)\hat{q}_a(\varepsilon; z) \leq 0.$$
By the convexity of $H$,

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

Therefore,

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

It follows that anywhere $\phi$ 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 $\phi$,

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

implying that the “Hessian measure” (see Villani [2003]) associated with $\phi_{qq}$ is absolutely continuous (there is no singular component). In other words, $\phi$ is continuously differentiable.

Next, we show that there is a strategy for the agent in the dynamic problem which can implement this value function. Suppose the agent 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 $\Omega$ (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 $\Omega$, 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 $\Omega$, 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,
\[
\frac{1}{2} tr[\sigma_{\pi,t}\sigma_{\pi,t}^T k(q_t)] = \chi > 0.
\]
This can be rewritten as
\[
\frac{1}{2} tr[\sigma_{\pi,t}\sigma_{\pi,t}^T \left( Q^+(Q(q_0)\pi_t)k(Q(q_0)\pi_t) \right) + Q^+(Q(q_0)\pi_t)Q(q_0)] = \chi,
\]
where $D^+$ denotes the pseudo-inverse.

By the positive-definiteness of $k$ in all directions orthogonal to $t$, 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 $\phi$, $\tilde{\phi}(\pi_\tau) = \hat{u}(Q(q_0)\pi_\tau)$. 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 $\sigma_t$ and stopping rule described by the stopping time $\tau$. We have, by the convexity of $\phi$ and the generalized Ito formula (nothing that we have shown that the Hessian measure $\phi_{qq}$ has no singular component),

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

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

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

In the stopping region of the optimal policy,

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

Therefore,

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

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.

A.3 Proof of lemma 3

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

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

A.3.1 Condition 1

Condition 1 requires that if the signal 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 signal 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.

A.3.2 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,

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

It follows that

\[
C(p_M, q; S_M) = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S_M} e_s^T \bar{p}_{i,M} D_i(q_{i,s,M} || q_i)
\]

\[
= \sum_{i \in \mathcal{I}(q)} \bar{q}_i \left( \lambda \sum_{s \in S_1} e_s^T \bar{p}_{i,1} D_i(q_{i,s,1} || q_i) + (1 - \lambda) \sum_{s \in S_2} e_s^T \bar{p}_{i,2} D_i(q_{i,s,2} || q_i) \right)
\]

\[
= \lambda C(p_1, q; S_1) + (1 - \lambda) C(p_2, q; S_2),
\]

verifying that the condition holds.

**A.3.3 Condition 3**

By Blackwell’s theorem, for any Markov mapping \( \Pi : S \to S' \), we require that

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

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

\[
\tilde{p}_i' = \frac{\sum_{x \in X} \Pi p e_x e_x^T q}{\bar{q}_i} = \Pi \tilde{p}_i
\]
and

\[ q_{i,s'} = \frac{E_i q_{s'}}{\sum_{x \in X_i} e_x^T q_{s'}} = \frac{E_i D(q) p^T \Pi^T e_{s'}}{\sum_{x \in X_i} e_x^T D(q) p^T \Pi^T e_{s'}} = \frac{D(q_i) E_i p^T \Pi^T e_{s'}}{\tilde{p}_{i1}^T \Pi^T e_{s'}} \]

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,

\[ D(q_i) E_i p^T e_s = (e_s^T p E_i^T q_{i,s}) q_{i,s} = (e_s^T \tilde{p}_i) q_{i,s}. \]

and therefore

\[ q_{i,s'} = \frac{\sum_{s \in S} q_{i,s} \tilde{p}_{i1}^T \Pi^T e_{s'}}{\tilde{p}_{i1}^T \Pi^T e_{s'}}. \]

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

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

\[
C(\Pi p, q; S') = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s' \in S'} e^T_s \Pi \bar{p}_i D_i(q, i, s'|q_i) \\
\leq \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s' \in S'} \sum_{s \in S} \bar{p}^T_i \Pi T e_s D_i(q, i, s'|q_i).
\]

By definition,

\[
\sum_{s' \in S} \Pi T e_{s'} = 1
\]

and therefore

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

**A.3.4 Condition 4**

By the definition of the neighborhood structure,

\[
C_N(p, q; S) = \sum_{i \in \mathcal{I}(q)} \bar{q}_i \sum_{s \in S} e^T_s \bar{p}_i D_i(q, i, s'|q_i),
\]

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 signal structure.

Suppose that

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

where \( r \in \mathcal{P}(S) \) and the support of \( \tau e_x \) is in the support of \( r \), and likewise for \( \omega e_x \), for all \( x \in X \).

By Bayes’ rule, for all \( s \in S \) such that \( e^T_s r > 0 \),

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

\[ q_s(\epsilon, \nu) = q^T \epsilon e_s + \frac{\epsilon D(q) \tau^T e_s}{r^T e_s + \epsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} + \frac{\nu D(q) \omega^T e_s}{r^T e_s + \epsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} \]

In the neighborhood around \( \epsilon = \nu = 0 \), the denominator is strictly positive, and therefore

\[
\frac{\partial}{\partial \nu} q_s(\epsilon, \nu) = -q_s(\epsilon, \nu) \frac{q^T \omega^T e_s}{r^T e_s + \epsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} + \frac{D(q) \omega^T e_s}{r^T e_s + \epsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}
\]

and

\[
\frac{\partial}{\partial \epsilon} \frac{\partial}{\partial \nu} q_s(\epsilon, \nu) = q_s(\epsilon, \nu) \frac{q^T \omega^T e_s}{r^T e_s + \epsilon q^T \tau^T e_s + \nu q^T \omega^T e_s} - \frac{q^T \tau^T e_s}{r^T e_s + \epsilon q^T \tau^T e_s + \nu q^T \omega^T e_s}
\]

For \( s \in S \) such that \( e_s^T r = 0 \), \( q_s(\epsilon, \nu) = q \), and therefore \( \frac{\partial}{\partial \epsilon} \frac{\partial}{\partial \nu} q_s(\epsilon, \nu) = 0 \). Therefore, \( \frac{\partial}{\partial \tau} q_s(\epsilon, \nu) \) can be written as a quadratic form in \( vec(\tau) \) and \( vec(\omega) \). It follows that \( q_s(\epsilon, \nu) \), in the neighborhood of an uninformative signal 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

\[
e_x^T \tilde{q}_i = \begin{cases} 
\frac{e_i^T q}{\bar{q}_i} & x \in X_i \\
0 & \text{otherwise}
\end{cases}
\]
By definition,

\[ \tilde{p}_i(\epsilon, \nu) = p\tilde{q}_i = r + \epsilon \tau \tilde{q}_i + \nu \omega \tilde{q}_i. \]

and therefore is twice-differentiable in the required directions. Moreover, by construction, if \( e^T_s r = 0 \), then \( e^T_s \tilde{p}_i(\epsilon, \nu) = 0 \), and if \( e^T_s r > 0 \), then \( e^T_s \tilde{p}_i(\epsilon, \nu) > 0 \) in the neighborhood around \( \epsilon = \nu = 0 \).

By definition,

\[ q_{i,s}(\epsilon, \nu) = \frac{E_i q_s(\epsilon, \nu)}{\sum_{x \in X} e^T x q_s(\epsilon, \nu)}. \]

For all \( i \in \mathcal{J}(q) \), in the neighborhood of an uninformative signal structure, \( \sum_{x \in X} e^T x q_s(\epsilon, \nu) \approx \tilde{q}_i > 0 \), and therefore the twice-differentiability of \( q_s \) in the required directions implies the twice-differentiability of \( q_{i,s} \) in those directions.

**A.3.5 Condition 5**

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

\[ C(p, q; S) \geq \frac{m}{2} \sum_{s \in S} (e^T_s pq) ||q_s - q||^2_X, \]

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.

**A.4 Proof of lemma 4**

Consider corollary 1. Under the stated assumptions,

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

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 (11):

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

By definition,

$$\bar{q}_i \bar{p}_i = \sum_{x \in X_i} p_x e_s^T q$$

$$= r \bar{q}_i + o(1).$$

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 (12) follows.

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

78
We have

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

That is,

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

Using this,

\[
(q_{i,s} - q_i)^T g(q_i)(q_{i,s} - q_i) = (q_{i,s} - q_i)^T D^+(q_i)(q_{i,s} - q_i)
\]

\[
= \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) E_i(q_s - q) - \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) q_i D^+(q_i) E_i(q_s - q)
\]

\[
- \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) q_i D^+(q_i) E_i(q_s - q)
\]

\[
+ \frac{1}{(\bar{q}_i)^2} (q_s - q)^T E_i^T D^+(q_i) q_i D^+(q_i) E_i(q_s - q) + o(\Delta).
\]

Therefore,

\[
C_N(\{p_x\}_{x \in X}, q; S) = \sum_{i \in \mathcal{F}(q)} c_i \bar{q}_i \sum_{s \in S} (e_s^T r)(q_{i,s} - q_i)^T g(q_i)(q_{i,s} - q_i) + o(\Delta)
\]

\[
= \Delta \sum_{i \in \mathcal{F}(q)} c_i \bar{q}_i \sum_{s \in S} (e_s^T r)(q_s - q)^T \tilde{k}_i(q)(q_s - q) + o(\Delta),
\]

where

\[
\tilde{k}_i(q) = \frac{1}{(\bar{q}_i)^2} E_i^T (D^+(q_i) - D^+(q_i) q_i q_i^T D^+(q_i)) E_i.
\]
The $\bar{k}(q)$ matrix is

$$\bar{k}_N(q) = \sum_{i \in \mathcal{I}(q)} c_i\bar{q}_i\bar{k}_i(q)$$

$$= \sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i}E_i^T(D^+(q_i) - D^+(q_i)q_iq_i^TD^+(q_i))E_i. \quad (17)$$

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

$$k_N(q) = D(q)\bar{k}(q)D(q)$$

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

$$\quad = \sum_{i \in \mathcal{I}(q)} \{c_iE_i^TD(q)E_i - c_i\bar{q}_iE_i^Tq_iq_i^TE_i\}$$

$$\quad = \sum_{i \in \mathcal{I}(q)} c_i\bar{q}_iE_i^TS^+(q_i)E_i.$$  

### A.5 Proof of Lemma 5

Using equation (17) from the proof of Lemma 4, we have

$$\bar{k}_N(q) = \sum_{i \in \mathcal{I}(q)} \frac{c_i}{\bar{q}_i}E_i^T(D^+(q_i) - D^+(q_i)q_iq_i^TD^+(q_i))E_i. \quad \text{(17)}$$

Consider the function

$$H_N(q) = \sum_{i \in \mathcal{I}(q)} c_i \left\{ \sum_{x \in \mathcal{X}_i} (e_x^Tq) \ln(e_x^Tq) - \left( \sum_{x \in \mathcal{X}_i} (e_x^Tq) \right) \ln\left( \sum_{x \in \mathcal{X}_i} (e_x^Tq) \right) \right\}$$

$$\quad = \sum_{i \in \mathcal{I}(q)} c_i \sum_{x \in \mathcal{X}_i} (e_x^Tq) \ln(q_i,x)$$

$$\quad = -\sum_{i \in \mathcal{I}(q)} c_i\bar{q}_iH_S(q_i).$$
Differentiating, 
\[
\frac{\partial H_N(q)}{\partial q_{x'}} = (\ln(q_{x'}) + 1) \sum_{i \in \mathcal{I}(q):x' \in X_i} c_i - \sum_{i \in \mathcal{I}(q):x' \in X_i} c_i (1 + \ln(\sum_{x \in X_i} (e^T_x q))).
\]

Differentiating again, 
\[
\frac{\partial^2 H_N(q)}{\partial q_{x'} \partial q_{x''}} = \frac{\delta_{x',x''}}{q} \sum_{i \in \mathcal{I}(q):x' \in X_i} c_i - \sum_{i \in \mathcal{I}(q):x',x'' \in X_i} \frac{c_i}{\sum_{x \in X_i} (e^T_x q)},
\]

where $\delta_{x',x''}$ is the Kronecker delta. By definition, 
\[
\sum_{i \in \mathcal{I}(q)} \frac{c_i}{q_i} e^T_{i} E^T_i D^+(q_i) q_i D^+(q_i) E_i e_{x''} = \sum_{i \in \mathcal{I}(q):x',x'' \in X_i} \frac{c_i}{\sum_{x \in X_i} (e^T_x q)}
\]

and 
\[
\sum_{i \in \mathcal{I}(q)} \frac{c_i}{q_i} e^T_{i} E^T_i D^+(q_i) E_i e_{x''} = \delta_{x',x''} \sum_{i \in \mathcal{I}(q):x',x'' \in X_i} \frac{c_i}{(e^T_x q)},
\]

proving that $\tilde{k}_N(q)$ is the Hessian of $H_N(q)$.

The posterior-separable cost function is defined as 
\[
C_N(p, q; S) = \sum_{s \in S} (e^T_s pq)(H_N(q_s) - H_N(q)).
\]

Using the definitions above, 
\[
C_N(p, q; S) = -\sum_{s \in S} (e^T_s pq) \sum_{i \in \mathcal{I}(q_s)} c_i \bar{q}_{i,s} H_S(q_{i,s}) + \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i H_S(q_i).
\]

81
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

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

By Bayes’ rule,

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

and by definition,

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

and therefore

$$C_N(p,q;S) = - \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in S} \bar{p}_{i,s} (H_S(q_{i,s}) - H_s(q_i))$$

$$= \sum_{i \in \mathcal{I}(q)} c_i \bar{q}_i \sum_{s \in S} \bar{p}_{i,s} D_{KL}(q_{i,s} || q).$$

The claim that

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

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

### A.6 Additional Definition and Lemmas

**Definition 1.** Let $X^N$ be a sequence of state spaces, as described in section 5.1. A sequence of policies $\{p_N \in \mathcal{P}(X^N)\}_{N \in \mathbb{N}}$ satisfies the “convergence condition” if:

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

$$\frac{c_H}{N+1} \geq e_i^T p_N \geq \frac{c_H}{N+1}.$$
2. The sequence satisfies, for some constant $K_1 > 0$, all $N$, and all $i \in X^N \setminus \{0, N\}$,

$$N^3 \left| \frac{1}{2}(e^T_{i+1} + e^T_{i-1} - 2e^T_i) p_N \right| \leq K_1,$$

and

$$N^2 \left| \frac{1}{2}(e^T_N - e^T_{N-1}) p_N \right| \leq K_1$$

and

$$N^2 \left| \frac{1}{2}(e^T_1 - e^T_0) p_N \right| \leq K_1.$$

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

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

where $X^N$ is the state space described in section 5.1. 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 $i \in X^N \setminus \{0, N\}$,

$$N^2 \left| \ln\left(\frac{1}{2}(e^T_{i+1} + e^T_i) q_N\right) + \ln\left(\frac{1}{2}(e^T_{i-1} + e^T_i) q_N\right) - 2 \ln(e^T_i q_N)\right| \leq K$$

and

$$N \left| \ln\left(\frac{1}{2}(e^T_1 + e^T_0) q_N\right) - \ln(e^T_0 q_N)\right| < K$$

and

$$N \left| \ln\left(\frac{1}{2}(e^T_N + e^T_{N-1}) q_N\right) - \ln(e^T_N q_N)\right| < K.$$

**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 \( i \in X^N \setminus \{N\} \),
\[
(e_{i+1}^T - e_i^T) p_N = \int_{x = \frac{i+1}{N+1}}^{x = \frac{i+1}{N+1}} (p(x + \frac{1}{N+1}) - p(x)) dx \\
= \int_{x = \frac{i+1}{N+1}}^{x = \frac{i+1}{N+1}} \int_{y = 0}^{y = \frac{1}{N+1}} p'(x+y) dy dx
\]
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 \( i \in X^N \setminus \{0, N\} \),
\[
(e_{i+1}^T + e_{i-1}^T - 2e_i^T) p_N = \int_{x = \frac{i+1}{N+1}}^{x = \frac{i+1}{N+1}} (p(x + \frac{1}{N+1}) + p(x - \frac{1}{N+1}) - 2p(x)) dx \\
= \int_{x = \frac{i+1}{N+1}}^{x = \frac{i+1}{N+1}} \int_{y = 0}^{y = \frac{1}{N+1}} (p'(x+y) - p'(x-y)) dy dx.
\]
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

\[ (e_{i+1}^T + e_i^T - 2e_i^T)p_N = \int_{\frac{i}{N+1}}^{\frac{i+1}{N+1}} \int_0^{\frac{1}{N+1}} \int_{-y}^y (p''(x+z))dzdydx. \]

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

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

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 \( \ln(x) \leq x - 1 \),

\[ \ln\left( \frac{1}{2} (e_{i+1}^T + e_i^T) / e_i^T \right) \leq \frac{1}{2} \ln \left( \frac{1}{2} (e_{i+1}^T + e_i^T) / e_i^T \right) \leq 2 \ln \left( \frac{1}{2} (e_{i+1}^T + e_i^T - 2e_i^T) / e_i^T \right) \leq \frac{1}{2} \ln \left( \frac{1}{2} (e_{i+1}^T + e_i^T - 2e_i^T) / e_i^T \right) \]

Therefore, by the bounds above,

\[ \ln\left( \frac{1}{2} (e_{i+1}^T + e_i^T) / e_i^T \right) + \ln\left( \frac{1}{2} (e_{i-1}^T + e_i^T) / e_i^T \right) \leq \frac{(N+1)K}{N^3cL} \leq \frac{2K}{N^2cL}. \]

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

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

We can rewrite this as

\[ \ln\left( \frac{1}{2} (e_{i+1}^T + e_i^T) / e_i^T \right) + \ln\left( \frac{1}{2} (e_{i-1}^T + e_i^T) / e_i^T \right) \geq \frac{1}{2} (e_{i+1}^T + e_i^T - 2e_i^T) / e_i^T + \frac{1}{2} (e_{i-1}^T - e_i^T) / e_i^T - \frac{1}{2} \ln\left( \frac{1}{2} (e_{i+1}^T + e_i^T) / e_i^T \right). \]
By the bounds above,
\[
\frac{1}{2}(e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N \geq -\frac{2K}{N^2 c_L}
\]
and
\[
\frac{1}{2}(e_i^T - e_{i-1}^T)p_N \frac{1}{2}(e_{i+1}^T + e_i^T)p_N - \frac{1}{2}(e_i^T - e_{i-1}^T)p_N \frac{1}{2}(e_{i+1}^T + e_i^T)p_N 
\leq -\frac{N^2}{c_L^2} \frac{1}{(N+1)^4}(K_2)^2
\leq -(\frac{K_2}{2Nc_L})^2.
\]

Therefore,
\[
N^2 |\ln(\frac{1}{2}(e_{i+1}^T + e_i^T)p_N/e_i^T p_N) + \ln(\frac{1}{2}(e_{i-1}^T + e_i^T)p_N/e_i^T p_N)| \leq \frac{2K}{c_L} + (\frac{K_2}{2c_L})^2.
\]

For the end-points,
\[
\frac{1}{2}(e_1^T - e_0^T)q_N \leq \ln(\frac{1}{2}(e_1^T + e_0^T)q_N) \leq \frac{1}{2}(e_1^T - e_0^T)q_N
\]
and therefore
\[
|\ln(\frac{1}{2}(e_1^T + e_0^T)q_N)| \leq \frac{K_2}{(N+1)c_L} \leq \frac{K_2}{Nc_L}.
\]

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

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 \( x \in \)
\[
\hat{p}_N(x) = (N + 1)((N + 1)x + \frac{1}{2} - \lfloor (N + 1)x + \frac{1}{2} \rfloor)p_N + \frac{1}{2} - (N + 1)x + \lfloor (N + 1)x + \frac{1}{2} \rfloor) e^T_{\lfloor (N + 1)x + \frac{1}{2} \rfloor - 1} p_N;
\]

And, for \( x \in [0, \frac{1}{2(N + 1)}) \),
\[
\hat{p}_N(x) = (N + 1)e^T_0 q_N,
\]

And, for \( x \in [1 - \frac{1}{2(N + 1)}, 1] \),
\[
\hat{p}_N(x) = (N + 1)e^T_N q_N.
\]

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

1. \( p_n(x) \) 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 \( x \in [0, 1] \),

2. the following sum converges:
\[
\lim_{n \to \infty} \sum_{i \in X^n \setminus \{n\}} \{g(e^T_i p_N) + g(e^T_{i+1} p_N) - 2g(\frac{1}{2}(e^T_i + e^T_{i+1}) p_N)\} = \frac{1}{4} \int_0^1 \left( \frac{p'(x)}{p(x)} \right)^2 dx;
\]

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

3. for all \( a \in A \), \( \lim_{n \to \infty} u_{a,n}^T p_n = \int_0^1 u_a(x) p(x) dx \), and

4. if the sequence \( \{p_N \in \mathcal{P}(X^N)\} \) 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 $[\frac{1}{2(N+1)}, 1-\frac{1}{2(N+1)}]$, 

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

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

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

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

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

$$TV(f'_N) = \sum_{i=1}^{N-1} (N+1)^2|e^T_{i+1} + e^T_{i-1} - 2e^T_i) p_N| +$$

$$+ (N+1)^2|e^T_N - e^T_{N-1}) p_N| + (N+1)^2|e^T_1 - e^T_0) p_N|.$$ 

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 subsequence, which we denote by $n$, such that $f'_n(x)$ converges point-wise to some $\rho'(x)$. 

88
For any $1 - \frac{1}{2(N+1)} > x > y \geq \frac{1}{2(N+1)}$, the quantity 

$$|f'_N(x) - f'_N(y)| = (N+1)^2 \sum_{i=\lfloor(N+1)y + \frac{1}{2}\rfloor}^{\lfloor(N+1)x + \frac{1}{2}\rfloor} (e_{i+1}^T + e_{i-1}^T - 2e_i^T)p_N|$$

$$\leq \frac{(N+1)^2((N+1)(x-y) + 2)}{N^3}2K_1.$$

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 subsequence. We now use $n$ to denote the subsequence for which $\lim_{n \to \infty} \hat{p}_n(0) = p(0)$ and for which $f'_n(x)$ converges point-wise to $p'(x)$. By the dominated convergence theorem, for all $x \in [0, 1],$

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

\[
g(y) = g(x) + g'(x)(y-x) + \frac{1}{2}g''(c_y + (1-c)x)(y-x)^2
\]

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

\[
g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N) = \frac{1}{8}g''(c_1 e_i^T p_N + (1-c_1)\frac{1}{2}(e_i^T + e_{i+1}^T)p_N)((e_{i+1}^T - e_i^T)p_N)^2 + \frac{1}{8}g''(c_2 e_i^T p_N + (1-c_2)\frac{1}{2}(e_i^T + e_{i+1}^T)p_N)((e_{i+1}^T - e_i^T)p_N)^2
\]

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,

\[
0 \leq g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N) \leq \frac{(N+1)c_L}{4}((e_{i+1}^T - e_i^T)p_N)^2.
\]
By construction,
\[ e_i^T p_N = \frac{1}{(N+1)} \hat{p}_N \left( \frac{2i+1}{2(N+1)} \right). \]

Therefore,
\[
(N + 1)(g(e_i^T p_N) + g(e_{i+1}^T p_N) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_N)) =
\]
\[
g(\hat{p}_N \left( \frac{2i+1}{2(N+1)} \right)) + g(\hat{p}_N \left( \frac{2i+3}{2(N+1)} \right)) - 2g(\hat{p}_N \left( \frac{2i+2}{2(N+1)} \right)).
\]
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 subsequence \( n \) defined above,
\[
\sum_{i \in \mathbb{N} \setminus \{n\}} \left\{ g(e_i^T p_n) + g(e_{i+1}^T p_n) - 2g(\frac{1}{2}(e_i^T + e_{i+1}^T)p_n) \right\} =
\]
\[
\frac{n^3}{n+1} \int_0^1 \left\{ g(\hat{p}_n(\frac{2|nx|+1)}{2(n+1)}) + g(\hat{p}_n(\frac{2|nx|+3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2|nx|+2}{2(n+1)})) \right\} dx.
\]

By the bound above,
\[
\frac{n^3}{n+1} \int_0^1 \left\{ g(\hat{p}_n(\frac{2|nx|+1}{2(n+1)})) + g(\hat{p}_n(\frac{2|nx|+3}{2(n+1)})) - 2g(\hat{p}_n(\frac{2|nx|+2}{2(n+1)})) \right\} dx \leq
\]
\[
\frac{n^3}{(n+1)^3} \int_0^1 K_1^2 c_L dx.
\]
Applying the dominated convergence theorem,

\[
\lim_{n \to \infty} n^2 \sum_{i \in X^n \setminus \{n\}} \{g(e_i^T p_n) + g(e_{i+1}^T p_n) - 2g\left(\frac{1}{2}(e_i^T + e_{i+1}^T)p_n\right)\} =
\]

\[
\int_0^1 \lim_{n \to \infty} \frac{n^3}{n+1} \left\{g\left(\hat{p}_n\left(\frac{2|nx| + 1}{2(n+1)}\right)\right) + g\left(\hat{p}_n\left(\frac{2|nx| + 3}{2(n+1)}\right)\right) - 2g\left(\hat{p}_n\left(\frac{2|nx| + 2}{2(n+1)}\right)\right)\right\} dx.
\]

By the Taylor expansion above,

\[
\lim_{n \to \infty} \frac{n^3}{n+1} \left\{g\left(\hat{p}_n\left(\frac{2|nx| + 1}{2(n+1)}\right)\right) + g\left(\hat{p}_n\left(\frac{2|nx| + 3}{2(n+1)}\right)\right) - 2g\left(\hat{p}_n\left(\frac{2|nx| + 2}{2(n+1)}\right)\right)\right\} =
\]

\[
\lim_{n \to \infty} \frac{1}{8} \frac{n^3}{n+1} \left\{g''(\cdot) + g''(\cdot)\right\} \left(\hat{p}_n\left(\frac{2|nx| + 3}{2(n+1)}\right) - \hat{p}_n\left(\frac{2|nx| + 1}{2(n+1)}\right)\right)^2.
\]

By definition,

\[
(n+1)\left(\hat{p}_n\left(\frac{2|nx| + 3}{2(n+1)}\right) - \hat{p}_n\left(\frac{2|nx| + 1}{2(n+1)}\right)\right) = f''\left(\frac{2|nx| + 2}{2(n+1)}\right)
\]

and

\[
\lim_{n \to \infty} g''\left(\hat{p}_n\left(\frac{2|nx| + 2}{2(n+1)}\right) + c_n\left(\hat{p}_n\left(\frac{2|nx| + 3}{2(n+1)}\right) - \hat{p}_n\left(\frac{2|nx| + 2}{2(n+1)}\right)\right)\right) = \frac{1}{p(x)},
\]

with \(c_n \in (0, 1)\) for all \(n\), and therefore

\[
\lim_{n \to \infty} \frac{n^3}{n+1} \left\{g\left(\hat{p}_n\left(\frac{2|nx| + 1}{2(n+1)}\right)\right) + g\left(\hat{p}_n\left(\frac{2|nx| + 3}{2(n+1)}\right)\right) - 2g\left(\hat{p}_n\left(\frac{2|nx| + 2}{2(n+1)}\right)\right)\right\} =
\]

\[
\lim_{n \to \infty} \frac{1}{4} \frac{(p'(x))^2}{p(x)},
\]

proving the second claim.
Turning to the third claim, recall that, by definition,

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

We define the function, for \( x \in [0, 1) \), as

\[ u_{a,N}(x) = e_{[N+1]x]}^T u_{a,N}, \]

and let \( u_{a,N}(1) = e_i^T u_{a,N} \). We also define the function

\[ \tilde{x}(x) = \frac{2[(N+1)x] + 1}{2(N+1)}. \]

By construction, \( \hat{p}_N(\tilde{x}(x)) = (N+1)e_i^T e_{[N+1]x]} p_{a,N} \) for all \( x \in [0, 1) \), and equals \( e_i^T p_{a,N} \) for \( x = 1 \). Therefore,

\[
\begin{align*}
 u_{a,N}^T p_N &= \sum_{i \in X^N} (e_i^T u_{a,N})(e_i^T p_N) \\
 &= \int_{0}^{1} \hat{p}_N(\tilde{x}(x)) u_{a,N}(x) \, dx.
\end{align*}
\]

By the boundedness of utilities and the dominated convergence theorem,

\[
\lim_{n \to \infty} u_{a,n}^T p_n = \int_{0}^{1} p(x) u_{a}(x) \, dx.
\]

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 \( \lim_{n \to \infty} \hat{p}_{a,N}(x) = \hat{p}(x) \) for all \( x \in X \), and therefore \( \hat{p}(x) = p(x) \). \( \square \)

**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 5.1. For each \( a \in A \), the sequence \( \{ q_{a,N} \} \) satisfies the convergence condition (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,

\[
\frac{c_H}{N + 1} \geq e^T q_N \geq \frac{c_L}{N + 1},
\]

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

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

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

\[ V_N(q_N; \tilde{\theta}) = \max_{\{p,N \in \mathcal{P}(A)\}} \sum_{i=0}^{N-1} e_i^T p D(q_i) u_N e_a \]

\[-\tilde{\theta} N^2 \sum_{i=0}^{N-1} (e_i^T q_N) D_{KL}(p_{i:N} || p_{i+1:N}(e_i^T q_N) + p_{i+1:N}(e_i^T q_N)) \]

\[-\tilde{\theta} N^2 \sum_{i=1}^{N} (e_i^T q_N) D_{KL}(p_{i:N} || p_{i-1:N}(e_i^T q_N)) \]

\[-\tilde{\theta} N^{-1} \sum_{i=0}^{N-1} (e_i^T q_N) D_{KL}(p_{i:N} || p_{N+1:N}). \]

The FOC for this problem is, for all \( i \in [1, N - 1] \) and \( a \in A \) such that \( e_a^T p_{i:N} > 0 \),

\[ e_i^T u_{a,N} - \tilde{\theta} N^2 \ln \left( \frac{e_i^T p_{i:N}(e_i^T q_N + e_i^T q_N)}{e_a^T (p_{i:N}(e_i^T q_N + p_{i+1:N}(e_i^T q_N)))} \right) \]

\[-\tilde{\theta} N^2 \ln \left( \frac{e_i^T p_{i:N}(e_i^T q_N + e_{i-1:N}(e_i^T q_N))}{e_a^T (p_{i:N}(e_i^T q_N + p_{i-1:N}(e_i^T q_N))))} \right) - \tilde{\theta} \ln \left( \frac{e_i^T p_{i:N}}{e_a^T p_{N+1:N}} \right) - e_i^T \kappa_N = 0, \]

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 \( \pi_N(a) > 0 \),

\[ e_i^T (u_{a,N} - \kappa_N) = -\tilde{\theta} N^2 \ln \left( \frac{e_i^T q_{a,N}}{e_i^T q_{a,N} + e_i^T q_N} \right) - \tilde{\theta} N^2 \ln \left( \frac{e_i^T q_{a,N}}{e_i^T q_{a,N} + e_{i+1} q_N} \right) - \tilde{\theta} \ln N^{-1} \left( \frac{e_i^T p_{i:N}}{e_a^T p_{N+1:N}} \right) \]

\[ = \tilde{\theta} N^2 \ln \left( 1 + \frac{e_i^T q_{a,N}}{e_i^T q_{a,N}} \right) - \tilde{\theta} N^2 \ln \left( 1 + \frac{e_i^T q_{a,N}}{e_i^T q_N} \right) - \tilde{\theta} \ln N^{-1} \left( \frac{e_i^T q_{a,N}}{e_i^T q_N} \right) \]

\[ - \tilde{\theta} N^2 \ln \left( 1 + \frac{e_{i-1} q_N}{e_i^T q_N} \right) - \tilde{\theta} \ln N^{-1} \left( \frac{e_i^T q_{a,N}}{e_i^T q_N} \right) \]

\[ = \tilde{\theta} N^2 \left( \ln \left( \frac{1}{2} (e_i^T q_{a,N}) + \ln \left( \frac{1}{2} (e_i^T q_{a,N}) - (2 + N^{-3}) \ln (e_i^T q_{a,N}) + 2 \ln 2 \right) \right) \right) \]

\[ - \tilde{\theta} N^2 \left( \ln \left( \frac{1}{2} (e_i^T q_{a,N}) + \ln \left( \frac{1}{2} (e_i^T q_{a,N}) - (2 + N^{-3}) \ln (e_i^T q_{a,N}) + 2 \ln 2 \right) \right) \right). \]
Using lemma 11, for all $i \in X^N \setminus \{0, N\}$,

$$N^2|\ln\left(\frac{1}{2}(e^T_{i+1} + e^T_i)q_{a,N}\right) + \ln\left(\frac{1}{2}(e^T_{i-1} + e^T_i)q_{a,N}\right) - 2\ln(e^T_{i} q_{N})| \leq K.$$  

By the boundedness of the utility function,

$$e^T_i \kappa_N \geq -\bar{u} - \bar{\theta}K + \bar{\theta}N^2(\ln\left(\frac{e^T_i q_{a,N}}{\frac{1}{2}(e^T_{i+1} + e^T_i)q_{a,N}}\right) + \ln\left(\frac{e^T_i q_{a,N}}{\frac{1}{2}(e^T_{i-1} + e^T_i)q_{a,N}}\right)) + \bar{\theta}N^{-1}\ln\left(\frac{e^T_{i} q_{a,N}}{e^T_{i} q_{N}}\right).$$

By the concavity of the log function,

$$\ln\left(\frac{1}{2}(e^T_{i+1} + e^T_i)q_{a,N}\right) + \ln\left(\frac{1}{2}(e^T_{i-1} + e^T_i)q_{a,N}\right) + N^{-3}\ln(e^T_i q_{N}) \leq (2 + N^{-3})\ln\left(\frac{1}{2(2 + N^{-3})}(e^T_{i+1} + e^T_{i-1} + 2e^T_i)q_{a,N} + \frac{N^{-3}}{2 + N^{-3}}e^T_i q_{N}\right).$$

and therefore

$$\ln\left(\frac{1}{2}(e^T_{i+1} + e^T_i)q_{a,N}\right) + \ln\left(\frac{1}{2}(e^T_{i-1} + e^T_i)q_{a,N}\right) + N^{-3}\ln(e^T_i q_{N}) - (2 + N^{-3})\ln(e^T_i q_{a,N}) \leq (2 + N^{-3})\ln\left(\frac{1}{2(2 + N^{-3})}(e^T_{i+1} + e^T_{i-1} + 2e^T_i)q_{a,N} + \frac{N^{-3}}{2 + N^{-3}}e^T_i q_{N}\right).$$

It follows that

$$e^T_i \kappa_N \geq -\bar{u} - \bar{\theta}K - (2 + N^{-3})\bar{\theta}N^2\ln\left(\frac{1}{2(2 + N^{-3})}(e^T_{i+1} + e^T_{i-1} + 2e^T_i)q_{a,N} + \frac{N^{-3}}{2 + N^{-3}}e^T_i q_{N}\right).$$
Exponentiating,

\[(e_i^T q_{a,N}) \exp\left(-\frac{1}{2+N^{-3}} \overline{\theta}^{-1} N^{-2} (\bar{u} + \overline{\theta}K + e_i^T \kappa_N)\right) \leq \frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_{a,N} + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N. \] (18)

Summing over \(a\), weighted by \(\pi_N(a)\),

\[(e_i^T q_N) \exp\left(-\frac{1}{2+N^{-3}} \overline{\theta}^{-1} N^{-2} (\bar{u} + \overline{\theta}K + e_i^T \kappa_N)\right) \leq \frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_N + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N. \]

Taking logs,

\[\frac{1}{2+N^{-3}} \overline{\theta}^{-1} N^{-2} (\bar{u} + \overline{\theta}K + e_i^T \kappa_N) \leq \ln\left(\frac{1}{2(2+N^{-3})} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_N + \frac{N^{-3}}{2+N^{-3}} e_i^T q_N \right) \leq \ln(1 + \frac{N^{-3}}{2+N^{-3}} + \frac{1}{2+N^{-3}} \frac{K_1 N^{-3}}{c_L N^{-1}}), \]

where the last step follows by lemma 11, recalling that \(c_L\) is the lower bound on \(f(x)\). We have

\[e_i^T \kappa_N \geq -2 \overline{\theta} N^2 \ln(1 + \frac{N^{-3}}{2+N^{-3}} + \frac{1}{2+N^{-3}} \frac{K_1}{c_L} N^{-2}) - \bar{u} - \overline{\theta}K \]

\[\geq -\bar{u} - \overline{\theta}K - \frac{N^{-1}}{2+N^{-3}} - \frac{1}{2+N^{-3}} \frac{K_1}{c_L} \]

\[\geq -\bar{u} - \overline{\theta}K - \frac{1}{2} \frac{K_1}{c_L}. \]

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

\[ e_0^T (u_{a,N} - \kappa_N) = \tilde{\theta} N^2 \left( \ln \left( \frac{1}{2} (e_1^T + e_0^T) q_{a,N} \right) - \ln(e_0^T q_{a,N}) \right) \]

\[ - \tilde{\theta} N^2 \left( \ln \left( \frac{1}{2} (e_1^T + e_0^T) q_{N} \right) - \ln(e_0^T q_{N}) \right) - \tilde{\theta} N^{-1} \ln\left( \frac{e_0^T q_{a,N}}{e_0^T q_{N}} \right). \]

By the concavity of the log function,

\[ \ln \left( \frac{1}{2} (e_1^T + e_0^T) q_{a,N} \right) + N^{-3} \ln(e_0^T q_{N}) - (1 + N^{-3}) \ln(e_0^T q_{a,N}) \leq (1 + N^{-3}) \ln \left( \frac{\frac{1}{2} (e_1^T + e_0^T) q_{a,N} + N^{-3} e_0^T q_{N}}{e_0^T q_{a,N}} \right). \] (19)

Therefore,

\[ \tilde{\theta} n^2 \ln \left( \frac{1}{2} (e_1^T + e_0^T) q_{a,N} \right) + \tilde{\theta} n^{-1} \ln \left( \frac{e_0^T q_{N}}{e_0^T q_{a,N}} \right) - \tilde{\theta} K \]

\[ \leq e_0^T (u_{a,N} - \kappa_N) + \tilde{\theta} N^2 \left( \ln \left( \frac{1}{2} (e_1^T + e_0^T) q_{N} \right) - \ln(e_0^T q_{N}) \right) \]

\[ \leq \tilde{\theta} (1 + N^{-3}) \ln \left( \frac{\frac{1}{2} (e_1^T + e_0^T) q_{a,N} + N^{-3} e_0^T q_{N}}{e_0^T q_{a,N}} \right) + \tilde{\theta} K. \]

By the boundedness of the utility function,

\[ -\tilde{\theta} (1 + N^{-3}) \ln \left( \frac{\frac{1}{2} (e_1^T + e_0^T) q_{a,N} + N^{-3} e_0^T q_{N}}{e_0^T q_{a,N}} \right) - \bar{u} \]

\[ \leq e_0^T \kappa_N + \tilde{\theta} N^2 \ln \left( \frac{e_0^T q_{N}}{\frac{1}{2} (e_1^T + e_0^T) q_{N}} \right) \]

\[ \leq -\tilde{\theta} N^2 \ln \left( \frac{\frac{1}{2} (e_1^T + e_0^T) q_{a,N}}{e_0^T q_{a,N}} \right) + \tilde{\theta} N^{-1} \ln \left( \frac{e_0^T q_{a,N}}{e_0^T q_{N}} \right) + \tilde{\theta}. \]
By the inequality \( \ln(x) \leq x - 1 \),

\[
\tilde{\theta} N^{-1} \ln \left( \frac{e_0^T q_{a,N}}{e_0^T q_N} \right) \leq \tilde{\theta} N^{-1} \left( \frac{e_0^T q_{a,N}}{e_0^T q_N} - 1 \right) \leq \tilde{\theta} c_L^{-1},
\]

where the latter follows from \( e_0^T q_N \geq c_L N^{-1} \). Exponentiating,

\[
(e_0^T q_{a,N}) \exp \left( -\tilde{\theta}^{-1} (1 + N^{-3})^{-1} N^{-2} \tilde{\mu} \right) \leq \left( \frac{1}{1 + N^{-3}} \right)^{\frac{1}{2}} (e_1^T + e_0^T) q_{a,N} \frac{N^{-3}}{1 + N^{-3} e_0^T q_N} \exp \left( \tilde{\theta}^{-1} (1 + N^{-3})^{-1} N^{-2} e_0^T \kappa_N \right) \frac{e_0^T q_N}{2(e_1^T + e_0^T) q_N}
\]

and

\[
\left( \frac{1}{2} (e_1^T + e_0^T) q_{a,N} \right) \exp \left( \tilde{\theta}^{-1} N^{-2} e_0^T \kappa_N \right) \frac{e_0^T q_N}{2(e_1^T + e_0^T) q_N} \leq (e_0^T q_{a,N}) \exp \left( \tilde{\theta}^{-1} N^{-2} (\tilde{\mu} + \tilde{\theta} c_L^{-1}) \right).
\]

Summing over \( a \), weighted by \( \pi_N(a) \),

\[
(e_0^T q_N) \exp \left( -\tilde{\theta}^{-1} (1 + N^{-3})^{-1} N^{-2} \tilde{\mu} \right) \leq \left( \frac{1}{1 + N^{-3}} \right)^{\frac{1}{2}} (e_1^T + e_0^T) q_N \frac{N^{-3}}{1 + N^{-3} e_0^T q_N} \exp \left( \tilde{\theta}^{-1} (1 + N^{-3})^{-1} N^{-2} e_0^T \kappa_N \right) \frac{e_0^T q_N}{2(e_1^T + e_0^T) q_N},
\]

\[
\left( \frac{1}{2} (e_1^T + e_0^T) q_N \right) \exp \left( \tilde{\theta}^{-1} N^{-2} e_0^T \kappa_N \right) \frac{e_0^T q_N}{2(e_1^T + e_0^T) q_N} \leq (e_0^T q_N) \exp \left( \tilde{\theta}^{-1} N^{-2} (\tilde{\mu} + \tilde{\theta} c_L^{-1}) \right).
\]

Taking logs,

\[
-\tilde{\theta} N^2 (1 + N^{-3}) \ln \left( \frac{1}{1 + N^{-3}} \right)^{\frac{1}{2}} (e_1^T + e_0^T) q_N + \left( \frac{1}{2} (e_1^T + e_0^T) q_N \right) \frac{N^{-3}}{1 + N^{-3} e_0^T q_N} \leq e_0^T \kappa_N \leq \tilde{\mu} + \tilde{\theta} c_L^{-1}.
\]
We can write
\[ \ln\left(\frac{1}{1+N^{-3}} \frac{1}{2}(e_1^T + e_0^T)q_N + \frac{N^{-3}}{1+N^{-3}}e_0^Tq_N\right) = \ln\left(\frac{1}{1+N^{-3}} + \frac{N^{-3}}{1+N^{-3}}e_0^Tq_N\right) \leq \frac{1}{1+N^{-3}} + \frac{2N^{-3}}{1+N^{-3}} - 1. \]

Therefore,
\[ -\tilde{\theta}N^2(1 + N^{-3})\ln\left(\frac{1}{1+N^{-3}} \frac{1}{2}(e_1^T + e_0^T)q_N + \frac{N^{-3}}{1+N^{-3}}e_0^Tq_N\right) \geq -\tilde{\theta}N^{-1} \geq -\tilde{\theta}. \]

By lemma 11,
\[ -\tilde{\theta} - \bar{u} = e_0^T \kappa_N \leq \bar{u} + \tilde{\theta}c_L^{-1}. \]

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 \(i \in X^N \setminus \{0, N\}\),
\[ e_i^T \kappa_N \leq \bar{u} + \tilde{\theta}K + \tilde{\theta}N^2 \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right) + \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}}\right) + \tilde{\theta}N^{-1} \ln\left(\frac{e_i^T q_{a,N}}{e_i^T q_N}\right), \]

and as argued above,
\[ \tilde{\theta}N^{-1} \ln\left(\frac{e_i^T q_{a,N}}{e_i^T q_N}\right) \leq \tilde{\theta}c_L^{-1}. \]

Using this bound,
\[ \tilde{\theta}N^2 \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right) + \ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}}\right) \geq -(\bar{u} + \tilde{\theta}K + B_L + \tilde{\theta}c_L^{-1}). \]
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 11, we can rewrite this inequalities as

\[ \bar{\theta} N \ln \left( \frac{e_N^T q_{a,N}}{\frac{1}{2}(e_N^T + e_{N-1}^T)q_{a,N}} \right) \geq -N^{-1} (\bar{u} + B_L + \bar{\theta} c_L^{-1}) + \bar{\theta} N \ln \left( \frac{e_N^T q_{a,N}}{\frac{1}{2}(e_N^T + e_{N-1}^T)q_{a,N}} \right) \]

\[ \geq -N^{-1} (\bar{u} + B_L + \bar{\theta} c_L^{-1}) - \bar{\theta} K \]

\[ \geq -(\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}) \]

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

\[ I_{a,N}(x) = (N + 1)(\ln(\hat{q}_{a,N}(x)) - \ln(\hat{q}_{a,N}(x - \frac{1}{2(N+1)}))) \]

for any \( x \in \left[\frac{1}{2(N+1)}, 1\right] \). For any \( i \in X^N \setminus \{0\} \),

\[ I_{a,N}(\frac{2i+1}{2(N+1)}) = (N + 1) \ln \left( \frac{(N + 1)e_i^T q_{a,N}}{\frac{1}{2}(N + 1)(e_i^T + e_{i-1}^T)q_{a,N}} \right) \]
and for any $i \in X^N \setminus \{N\}$,
\[
l_{a,N}(\frac{2i+2}{2(N+1)}) = (N+1) \ln(\frac{1}{(N+1)e_i^T q_{a,N}}).
\]

Therefore, for any $i \in X^N \setminus \{0, N\}$, the lower bound can be written as
\[
\tilde{\theta} \frac{N^2}{N+1} (l_{a,N}(\frac{2i+2}{2(N+1)}) - l_{a,N}(\frac{2i+1}{2(N+1)})) \leq (\bar{u} + \tilde{\theta} K + B_L + \tilde{\theta} c_L^{-1}).
\]

The lower endpoint bound is
\[
\tilde{\theta} \frac{N}{N+1} l_{a,N}(\frac{1}{(N+1)}) \leq (\bar{u} + \tilde{\theta} K + B_L + \tilde{\theta} c_L^{-1}).
\]

The upper endpoint bound is
\[
\tilde{\theta} \frac{N}{N+1} l_{a,N}(1) \geq -(\bar{u} + \tilde{\theta} K + B_L + \tilde{\theta} c_L^{-1}).
\]

We also have, for all $i \in X^N \setminus \{N\}$
\[
\tilde{\theta} \frac{N^2}{N+1} (l_{a,N}(\frac{2i+3}{2(N+1)}) - l_{a,N}(\frac{2i+2}{2(N+1)})) = \tilde{\theta} N^2 (\ln(\frac{(N+1)e_i^T q_{a,N}}{\frac{1}{2}(N+1)(e_i^T e_{i+1} + e_i^T) q_{a,N}}) - \ln(\frac{1}{(N+1)e_i^T q_{a,N}})) \leq 0,
\]

by the concavity of the log function. Therefore, for all $j \in \{2, 3, \ldots, 2(N+1)\}$
\[
\tilde{\theta} \frac{N^2}{N+1} (l_{a,N}(\frac{j+1}{2(N+1)}) - l_{a,N}(\frac{j}{2(N+1)})) \leq (\bar{u} + \tilde{\theta} K + B_L + \tilde{\theta} c_L^{-1}).
\]
It follows that, for all \( j \in \{2, 3, \ldots, 2(N+1)\} \)

\[
\frac{j}{2(N+1)} = \frac{2}{2(N+1)} + \sum_{k=2}^{j-1} \left( \frac{k+1}{2(N+1)} - \frac{k}{2(N+1)} \right)
\]

\[
\leq \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}) \frac{N+1}{N} \left( 1 + \frac{j-2}{N} \right).
\]

Similarly, for all \( j \in \{2, 3, \ldots, 2(N+1)\} \),

\[
l_{a.N}(1) = l_{a.N}(\frac{j}{2(N+1)}) + \sum_{k=j-1}^{2N} \left( l_{a.N}(\frac{k+1}{2(N+1)}) - l_{a.N}(\frac{k}{2(N+1)}) \right)
\]
and therefore

\[
-l_{a.N}(\frac{j}{2(N+1)}) = -l_{a.n}(1) + \sum_{k=j-1}^{2N} \left( l_{a.N}(\frac{k+1}{2(N+1)}) - l_{a.n}(\frac{k}{2(N+1)}) \right)
\]

\[
\leq \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}) \frac{N+1}{N} \left( 1 + \frac{2(N+1) - j}{N^2} \right).
\]

It follows that

\[
|l_{a.N}(\frac{j}{2(N+1)})| \leq 2 \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}) \frac{N+1}{N}
\]

\[
\leq 4 \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_L + \bar{\theta} c_L^{-1}).
\]

Note that there must exist some \( \tilde{l}_{a,N} \in X^N \) such that \( e_{l_{a,N}}^T q_{a,N} \geq \frac{1}{N+1} \), implying that

\[
\ln((N+1)e_{l_{a,N}}^T q_{a,N}) \geq 0.
\]

By the definition of \( l_{a,N} \), for any \( i \in X^N \setminus \{0\} \),

\[
l_{a.N}(\frac{2i+1}{2(N+1)}) + l_{a.N}(\frac{2i}{2(N+1)}) = (N+1) \ln \left( \frac{(N+1)e_i^T q_{a,N}}{(N+1)e_{i-1}^T q_{a,N}} \right).
\]
For any $i > \tilde{i}_{a,N}$,

\[
\ln((N+1)e_i^T q_{a,N}) = \ln((N+1)e_i^T q_{a,N}) + \sum_{j=i_{a,n}+1}^{i} \ln\left(\frac{(N+1)e_j^T q_{a,N}}{(N+1)e_{j-1}^T q_{a,N}}\right)
\]

\[
= \ln((N+1)e_i^T q_{a,N}) + \frac{1}{N+1} \sum_{j=i_{a,n}+1}^{i} l_{a,N}\left(\frac{2j+1}{2(N+1)}\right) + l_{a,N}\left(\frac{2j}{2(N+1)}\right)
\]

\[
\geq -\frac{1}{N+1} \sum_{j=i_{a,n}+1}^{i} 8\tilde{\theta}^{-1}(\bar{u} + \tilde{\theta}K + B_L + \tilde{\theta}c_{L}^{-1})
\]

\[
\geq -8\tilde{\theta}^{-1}(\bar{u} + \tilde{\theta}K + B_L + \tilde{\theta}c_{L}^{-1}).
\]

Similarly, for any $i < \tilde{i}_{a,N}$,

\[
\ln((N+1)e_i^T q_{a,N}) = \ln((N+1)e_i^T q_{a,N}) + \sum_{j=i+1}^{\tilde{i}_{a,n}} \ln\left(\frac{(N+1)e_j^T q_{a,N}}{(N+1)e_{j-1}^T q_{a,N}}\right).
\]

Therefore,

\[
\ln((N+1)e_i^T q_{a,N}) \geq -\sum_{j=i+1}^{\tilde{i}_{a,n}} \ln\left(\frac{(N+1)e_j^T q_{a,N}}{(N+1)e_j^T q_{a,N}}\right)
\]

\[
\geq -8\tilde{\theta}^{-1}(\bar{u} + \tilde{\theta}K + B_L + \tilde{\theta}c_{L}^{-1}).
\]

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

\[
\ln((N+1)e_i^T q_{a,N}) \leq 8\tilde{\theta}^{-1}(\bar{u} + \tilde{\theta}K + B_L + \tilde{\theta}c_{L}^{-1}).
\]

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 $i \in X^N$,

\[
\frac{e_i^{-1}}{(N+1)} \geq e_i^T q_{a,N} \geq \frac{c}{N+1}.
\]

104
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 \( a \in (0,1) \)

\[
|(N+1)\ln\left(\frac{\frac{1}{2}(N+1)(e_i^T + e_{i+1}^T)q_{a,N}}{(N+1)e_i^T q_{a,N}}\right)| = \frac{(N+1)|\frac{1}{2}(e_{i+1}^T - e_i^T)q_{a,N}|}{e_i^T q_{a,N} + \frac{a}{2}(e_{i+1}^T - e_i^T)q_{a,N}}
\]

\[
\leq 4\bar{\theta}^{-1}(\bar{u} + \bar{\theta}K + B_L + \bar{\theta}c_L^{-1}),
\]

and therefore, by the bound on \( e_i^T q_{a,N} \),

\[
(N+1)^2|\frac{1}{2}(e_{i+1}^T - e_i^T)q_{a,N}| \leq B
\]

for some \( B > 0 \). By a similar argument,

\[
(N+1)^2|\frac{1}{2}(e_{i+1}^T - e_{i-1}^T)q_{a,N}| \leq 4B.
\]

Returning to the first-order condition, for \( i \in X_N \setminus \{0,N\} \), and using some of the bounds employed previously,

\[
e_i^T \kappa_N \leq \bar{u} + \bar{\theta}K + \bar{\theta}c_L + \bar{\theta}N^2(\ln\left(\frac{e_i^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right) + \ln\left(\frac{e_{i+1}^T q_{a,N}}{\frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N}}\right))
\]

By the inequality \( \ln(x) \leq x - 1 \),

\[
e_i^T \kappa_N \leq \bar{u} + \bar{\theta}K + \bar{\theta}c_L + \bar{\theta}N^2\left(\frac{1}{2}(e_i^T - e_{i+1}^T)q_{a,N} + \frac{1}{2}(e_{i+1}^T - e_i^T)q_{a,N} + \frac{1}{2}(e_{i+1}^T - e_{i-1}^T)q_{a,N} + \frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}\right)
\]

105
Multiplying through,

\[
\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}(e_i^T \kappa_N - \bar{u} - \bar{\theta} K - \bar{\theta} c_L) \\
\leq \bar{\theta} N^2 \left( \frac{1}{2}(e_i^T - e_{i+1}^T)q_{a,N} + \frac{1}{2}(e_i^T - e_{i-1}^T)q_{a,N} \right) \left( \frac{1}{2}(e_{i+1}^T + e_i^T)q_{a,N} \right) \\
\leq \bar{\theta} N^2 \left( \frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} + \frac{1}{2}(e_i^T - e_{i-1}^T)q_{a,N} \left( \frac{1}{2}(e_{i+1}^T - e_{i-1}^T)q_{a,N} \right) \right).
\]

Using the bounds above,

\[
\frac{1}{2}(e_{i-1}^T + e_i^T)q_{a,N}(e_i^T \kappa_N - \bar{u} - \bar{\theta} K - \bar{\theta} c_L) \leq \bar{\theta} N^2 \left( \frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} + \frac{B}{(N+1)^2} \left( \frac{4B}{c(N+1)^2} \right) \right) \\
\leq \bar{\theta} N^2 \left( \frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} \right) + \frac{4B^2 N^2}{c(N+1)^3}.
\]

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 \( B_K > 0 \) for all \( i \in X^N \) (recalling that this was shown for \( i \in \{0,N\} \) previously). It also follows the term

\[
(N+1)^3 \left( \frac{1}{2}(2e_i^T - e_{i+1}^T - e_{i-1}^T)q_{a,N} \right) \geq \frac{(N+1)^2}{N^2} c(e_i^T \kappa_N - \bar{u} - \bar{\theta} K - \bar{\theta} c_L - \frac{4B^2}{c^2}) \\
\geq -2c(B_K + \bar{u} + \bar{\theta} K + \bar{\theta} c_L + \frac{4B^2}{c^2})
\]

is bounded below.
Recalling equation (18), and employing the upper bound on $|e_T^T \kappa_N|$, 

$$(e_i^T q_{a,N}) \exp(-\frac{1}{2+N-3} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + B_{K}))$$

$$\leq \frac{1}{2(2+N-3)} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_{a,N} + \frac{N^{-3}}{2+N-3} e_i^T q_N.$$

Rewriting this,

$$(e_i^T q_{a,N})(\exp(-\frac{1}{2+N-3} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + B_{K})) - 1)$$

$$\leq \frac{1}{2(2+N-3)} (e_{i+1}^T + e_{i-1}^T + 2e_i^T) q_{a,N} + \frac{N^{-3}}{2+N-3} e_i^T (q_N - q_{a,N}).$$

By the upper bound on $e_i^T q_N \leq c_{N+1}$ and $e_i^T q_{a,N} \geq \frac{c}{N+1}$,

$$(N+1)^3 (e_{i+1}^T + e_{i-1}^T - 2e_i^T) q_{a,N} \geq (2+N^{-3}) (N+1)^2 \left( \exp\left(-\frac{1}{2+N-3} \bar{\theta}^{-1} N^{-2} (\bar{u} + \bar{\theta} K + B_{K}) \right) - 1 \right) - \frac{c H - c}{N^3} (N+1)^2.$$

By the inequality $\exp(x) - 1 \geq x$,

$$\frac{(N+1)^3}{2} (e_{i+1}^T + e_{i-1}^T - 2e_i^T) q_{a,N} \geq -\frac{(N+1)^2}{N^2} \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_{K}) - \frac{c H - c}{N^3} (N+1)^2$$

$$\geq -2 \bar{\theta}^{-1} (\bar{u} + \bar{\theta} K + B_{K}) - 2c H + c.$$

Therefore, the first statement in the second part of the convergence condition (1) is satisfied.

Finally, we consider the endpoints. The first-order condition is

$$\bar{\theta} N^2 (\ln(\frac{1}{2} (e_1^T + e_0^T) q_{a,N}) - \ln(e_0^T q_{a,N})) =$$

$$e_0^T (u_{a,N} - \kappa_N) + \bar{\theta} N^2 (\ln(\frac{1}{2} (e_1^T + e_0^T) q_N) - \ln(e_0^T q_N)) + \bar{\theta} N^{-1} \ln(e_0^T q_{a,N} e_0^T q_N).$$
We can bound this as

\[-N^{-1}(\bar{u} + B_K) - \bar{\theta} K + \bar{\theta} N^{-2} \ln \left( \frac{C}{c_H} \right)\]

\[\leq \bar{\theta} N \left( \ln \frac{1}{2} (e_1^T + e_0^T)q_{a,N} - \ln(e_0^T q_{a,N}) \right)\]

\[\leq N^{-1}(\bar{u} + B_K + \bar{\theta} c^{-1}_L) + \bar{\theta} K,\]

and note that because \(\sum_{i \in \mathcal{X}N} e_i^T q_{a,N} = \sum_{i \in \mathcal{X}N} e_i^T q_N = 1\), we must have \(c_H \geq c\). Therefore,

\[\bar{\theta} \ln \left( \frac{C}{c_H} \right) \leq \bar{\theta} N^{-2} \ln \left( \frac{C}{c_H} \right).\]

Using a Taylor expansion,

\[\ln \left( \frac{1}{2} (e_1^T + e_0^T)q_{a,N} - \ln(e_0^T q_{a,N}) = \frac{1}{2} (e_1^T - e_0^T)q_{a,N}}{e_0^T q_{a,N} + \frac{a}{2} (e_1^T + e_0^T)q_{a,N}}\]

for some \(a \in (0, 1)\). Therefore,

\[N^2 \frac{1}{2} (e_1^T - e_0^T)q_{a,N} \leq \frac{c}{\bar{\theta}} (\bar{u} + B_K + \bar{\theta} K + \bar{\theta} \max(\ln \left( \frac{C_H}{c} \right), c^{-1}_L)).\]

A similar logic holds for the other endpoint, and therefore the convergence condition is satisfied.

\[\square\]

### A.7 Proof of theorem 4

By the boundedness of \(\mathcal{P}(A)\), there exists a convergent subsequence of the optimal policy \(\pi_N(a)\), which we denote by \(n\). Define

\[\pi(a) = \lim_{n \to \infty} \pi_n(a).\]
By lemma 13, for all \( a \in A \), each sequence of optimal policies \( \{q_{a,N}\} \) satisfies the convergence condition (1). Therefore, by lemma 12, each sequence \( \{\hat{q}_{a,N}(x)\} \) has a convergent subsequence that converges to a differentiable function \( f^*_a(x) \), whose derivative is Lipschitz continuous, with full support on \([0, 1]\). We can construct a subsequence in which \( \pi_n(a) \) and all \( \{\hat{q}_{a,n}(x)\} \) converge by iteratively applying this argument. Denote this sequence by \( n \).

Moreover, by lemma 12,

\[
\lim_{n \to \infty} V_n(q_n; \bar{\theta}) = \sum_{a \in A} \pi(a) \int_0^1 u_a(x)f_a(x)dx - \frac{\bar{\theta}}{4} \sum_{a \in A} \{\pi(a) \int_0^1 (f'_a(x))^2 f_a(x)dx + \frac{\bar{\theta}}{4} \int_0^1 (f'(x))^2 f(x)dx\}.
\]

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 \( \lim_{n \to \infty} V_n(q_n; \bar{\theta}) \), 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

$$\lim_{n \to \infty} \sum_{i=0}^{\lfloor x_n \rfloor} e_i^T q_{a,n} = \int_0^x f_a(y)dy.$$ 

Note that

$$e_i^T q_{a,n} = (n + 1) \int_{\frac{i}{n+1}}^{\frac{i+1}{n+1}} \hat{q}_{a,n}(\frac{2i + 1}{2(n+1)})dy,$$

where $\hat{q}_{a,n}$ is the function defined in lemma 12. Therefore, the sum is equal to

$$\sum_{i=0}^{\lfloor x_n \rfloor} e_i^T q_{a,n} = \int_0^{\lfloor x_n \rfloor + 1} \hat{q}_{a,n}(\frac{(n + 1)y + \frac{1}{2} + \frac{1}{2}(n+1)}{(n+1)})dy.$$ 

By the boundedness of $\hat{q}_{a,n}$ (which follows from the convergence condition) and the dominated convergence theorem,

$$\lim_{n \to \infty} \int_0^{\lfloor x_n \rfloor + 1} \hat{q}(\frac{(n + 1)y + \frac{1}{2} + \frac{1}{2}(n+1)}{(n+1)})dy = \int_0^x f_a(y)dy,$$

as required.

### A.8 Proof of lemma 2

Let $p$ and $p'$ be signal 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 signal 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 $\Pi : S \times \{1, 2\} \to S$, 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 signal 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 $\Pi_1 : S_1 \to S_M$ 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 $\Pi_2$ 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.

A.9 Proof of theorem 3

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,

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

The second claim follows by a similar argument.

We next demonstrate the claimed properties of \( k(q) \). First, \( k(q) \) is symmetric, 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 a signal structure, for which \( \tau_x = \phi e_s^T v \), 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 \( \phi \) are not zero. For this signal 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 signal structure is uninformative for all \( \Delta \). Then we must have \( v \) proportional to \( t \), and therefore

\[ t^T k(q) t = 0 \]

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

\[ v^T k(q) v \geq 0, \]

implying that \( k(q) \) is positive semi-definite.

Suppose now that the cost function satisfies 5. Let \( v \) be as above, non-zero, and not proportional to \( t \). 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 5 there exists a \( \Delta_B \) such that, for all \( \Delta < \Delta_B, C(p, q; S) < B \).

Therefore, we must have

\[ C(\{p_x\}_{x \in X}, q) \geq \frac{m}{2} \sum_{s \in S} (e^T s pq) ||q_s - q||_X^2. \]

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 \( e^T s pq = 0 \).
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 $\Delta$, $e_s^T pq > 0$ 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,

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

Noting that

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

and that

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

we have

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

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 $v$. 

114
A.10 Proof of corollary 1

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 \( e_s^T pq > 0 \),

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

It follows immediately that

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

Next,

\[ \Delta^{-\frac{1}{2}}(q_s - q) = \Delta^{-\frac{1}{2}} \frac{(D(q) - qq^T)p^T e_s}{q^T p^T e_s} \]
\[ = D(q) \frac{\tau^T e_s - tq^T \tau^T e_s + o(1)}{q^T p^T e_s}. \]

For any \( s \) such that \( q^T p^T e_s > 0 \),

\[ \lim_{\Delta \to 0^+} \Delta^{-\frac{1}{2}}(q_s - q) = D(q) \frac{\tau^T e_s - tq^T \tau^T e_s}{r^T e_s}. \]

By theorem 3,

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

By the result that \( t^T k(q) = 0 \), we have
\[ C(p, q; S) = \frac{1}{2} \Delta \sum_{x' \in X} \sum_{x \in X} e_{x'}^T k(q) e_{x} \cdot (\tau_{x'} - q \tau)^T g(r)(\tau_x - q \tau) \]
\[ + o(\Delta). \]

By the definition of the Fisher matrix, and the observation that \( i^T \tau_x = 0 \) for all \( x \in X \),
\[ (\tau_{x'} - q \tau)^T g(r)(\tau_x - q \tau) = \sum_{s \in S; e_s^T r > 0} (e_s^T r)(q_s - q) = \sum_{s \in S; e_s^T r > 0} (e_s^T r)(q_s - q) + o(\Delta), \]
which is the result.

**A.11 Proof of lemma 7**

Write the value function in sequence-problem form:

\[ W(q_0, \lambda; \Delta) = \max_{(p_{\Delta j}), \tau} E_0[\hat{u}(q_{\tau}) - \kappa \tau] - \]
\[ \lambda E_0[\Delta^{1-\rho} \sum_{j=0}^{\tau_{\Delta}^{-1}} \{ \frac{1}{\rho} C(\{p_{\Delta j, x} \}_{x \in X}, \cdot \})^\rho - \Delta^\rho c^\rho \}]. \]

Define
\[ \bar{u} = \max_{a \in A, x \in X} u(a, x). \]

116
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 $u$, and because stopping now and deciding is always feasible,

$$W(q_0, \lambda; \Delta) \geq u.$$

Therefore, $W(q_0, \lambda; \Delta)$ is bounded for all $\lambda \in (0, \kappa c^{-\rho})$ and all $\Delta$.

Next, recall that

$$\sum_{x \in X} q_t, x W(q_t, \lambda; \Delta, x) = W(q_t, \lambda; \Delta).$$

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 agent could simply adopt the signal structures associated with beliefs $q'$ and achieve higher utility, contradicting the optimality of the policy.
The convexity of the value function follows from this observation:

\[
W(\alpha q + (1 - \alpha)q', \lambda; \Delta) = \alpha \sum_{x \in X} q_x W(\alpha q + (1 - \alpha)q', \lambda; \Delta, x) + (1 - \alpha) \sum_{x \in X} q'_x W(\alpha q + (1 - \alpha)q', \lambda; \Delta, x),
\]

\[
W(\alpha q + (1 - \alpha)q', \lambda; \Delta) \leq \alpha \sum_{x \in X} q_x W(q, \lambda; \Delta, x) + (1 - \alpha) \sum_{x \in X} q'_x W(q', \lambda; \Delta, x),
\]

\[
W(\alpha q + (1 - \alpha)q', \lambda; \Delta) \leq \alpha W(q, \lambda; \Delta) + (1 - \alpha) W(q', \lambda; \Delta).
\]

A.12 Additional Lemma

Lemma 14. In the sequential evidence accumulation problem, for any norm ||·|| on the tangent space of signal structures, there exist constants \(B\) and \(\bar{\Delta}\) such that, for all \(\Delta < \bar{\Delta}\) and \(q_t \in \mathcal{P}(X)\),

\[
||p^*_t,\Delta - r^*_t,\Delta^T D(q_t)|| \leq B\Delta^{\frac{\rho - 1}{2\rho - 1}},
\]

where \(p^*_t,\Delta\) denotes an optimal signal structure given \(q_t\) and the time interval \(\Delta\), and \(r^*_t,\Delta = p^*_t,\Delta q_t\).

Proof. The agent’s problem, conditional on not stopping at the current time, is

\[
W(q_t, \lambda; \Delta) = \max_{\{p(s)\}} \sum_{x \in X} \sum_{s \in S} p_t, s, x q_t, s W(q_t, s, \lambda; \Delta, x),
\]

where \(q_t, s\) is the posterior associated with receiving signal \(s\), and is determined using Bayes’ rule, the prior \(q_t\), and \(p_t, s, x\). Let \(q^*_t, s\) denote the posteriors associated with the optimal policy. The optimal signal structure must achieve weakly higher utility than any other signal structure. Consider, in particular, an uninformative signal structure. We must have
\[
\sum_{s \in S} e_s^T r_{t,\Delta}^* [W(q_{t,s}^*, \lambda; \Delta) - W(q_t, \lambda; \Delta)] \geq \lambda \Delta^{1-p} \frac{1}{\rho} C(\cdot)^\rho.
\]

By the boundedness and convexity of \( W \) (lemma 7), it is Lipschitz-continuous, and therefore

\[
K \sum_{s \in S} e_s^T r_{t,\Delta}^* \|q_{t,s}^* - q_t\|_{X,2} \geq \sum_{s \in S} e_s^T r_{t,\Delta}^* [W(q_{t,s}^*, \lambda; \Delta) - W(q_t, \lambda; \Delta)],
\]

where \( K \) is the associated Lipschitz constant and \( \| \cdot \|_{X,2} : \mathbb{R}^{|X|} \to \mathbb{R}^+ \) is the Euclidean norm, defined on the tangent space of posteriors. By the concavity of the square root function,

\[
\sum_{s \in S} e_s^T r_{t,\Delta}^* \|q_{t,s}^* - q_t\|_{X,2} \leq \| (p_{t,\Delta}^* - r_{t,\Delta}^* t^T) D(q_t) \|_2,
\]

where \( \| \cdot \|_2 : \mathbb{R}^{|X| \times |A|} \to \mathbb{R}^+ \) denotes the Euclidean norm on the tangent space of joint distributions over signals and states. From this argument, we observe that

\[
K^{\rho-1} \Delta^{1-p} \| (p_{t,\Delta}^* - r_{t,\Delta}^* t^T) D(q_t) \|_2^\rho \geq (\frac{\lambda}{\rho})^{\rho-1} C(\cdot).
\]

By the finiteness of the simplex and the assumption that \( \rho > 1 \), it follows that for \( \Delta \) sufficiently small, \( C(\cdot) \) is bounded above by any positive constant. Therefore, by 5, for all \( \Delta < \bar{\Delta} \) such that 5 applies uniformly,

\[
C(\cdot) \geq m \sum_{s \in S} e_s^T r_{t,\Delta}^* \|q_{t,s}^* - q_t\|_{X,2}^2,
\]

for some positive constant \( m \) that does not depend on \( p_{t,\Delta}^* \). By the fact that \( r_{t,\Delta}^* \in \mathcal{P}(S) \),

\[
\sum_{s \in S} e_s^T r_{t,\Delta}^* \|q_{t,s}^* - q_t\|_{X,2}^2 \geq \sum_{s \in S} (e_s^T r_{t,\Delta}^*)^2 \|q_{t,s}^* - q_t\|_{X,2}^2,
\]

119
and therefore
\[
\sum_{s \in S} e_s^T r^*_t ||q^*_t - q_t||_2^2 \geq ||(p^*_t - r^*_t t^T)D(q_t)||_2^2.
\] (22)

Putting this together,
\[
\left(\frac{K\rho}{\lambda}\right)^{\rho-1} \Delta^{1-\rho} \geq m \left(\frac{1}{2}\right) \left(\frac{1}{2}\right)
\]

which proves the lemma for \(B = m^{-\rho} \left(\frac{K\rho}{\lambda}\right)^{\frac{1}{2}}\) and the Euclidean norm. The result holds for all norms by the equivalence of norms. 

\[\Box\]

### A.13 Proof of lemma 8

By lemma 14, any convergent sub-sequence of optimal policies converges uniformly to an uninformative signal structure. However, the rate of convergence implied by this lemma is not sufficiently fast for our purposes, so we proceed with an additional argument.

Consider a sequence of \(\Delta_n\) such that \(p^*_t,\Delta_n = p^*_t,\Delta\) converges (such a sequence exists by the boundedness of \(p^*_t,\Delta\)). Define \(r^*_t(s) = \lim_{n \to \infty} p^*_t,\Delta_n q_t\). Define the function

\[
f(p; qt, \Delta, \lambda) = -\sum_{s \in S} e_s^T p q_t [W(q_s(p, \lambda; \Delta)) - W(q_t, \lambda; \Delta)],
\]

with \(q_s(p, \lambda; \Delta)\) defined by Bayes’ rule, and the function

\[
g(p; qt, \Delta, \lambda) = -\kappa \Delta + \lambda \Delta^{1-\rho} \left(\Delta^\rho c^\rho - \frac{1}{\rho} C(p; qt)^\rho\right).
\]
Note that both functions are concave. The Bellman equation and optimality conditions require that, for an optimal signal structure,

\[ g((p^*_t, q_t; \Delta_n, \lambda)) - f((p^*_t, q_t; \Delta_n, \lambda)) = 0 \]

and for all signal structures,

\[ g(p; q_t, \Delta_n, \lambda) - f(p; q_t, \Delta_n, \lambda) \leq 0. \]

Define \( S_n \subseteq S \) as the support of \( p^*_t q_t \). By the convergence of \( ||(p^*_t - r^*_t (s) ^T)D(q_t)|| \) to zero (lemma 14), for each state \( x \in X \), \( p^*_{t,n,x} \) is in the interior of \( S_n \). By 4, the function \( g(\cdot) \) is twice-differentiable in the neighborhood of an uninformative structure, which is to say as \( n \) becomes sufficiently large. By Lemma 1 of Benveniste and Scheinkman [1979], the function \( f \) is differentiable, at \( \{p^*_t, q_t, \Delta_n\} \in x \in X \), with respect to signal structures for which the support of \( pq_t \) remains in \( S_n \).

By theorem 23.8 of Rockafellar [1970], it follows that at each posterior \( q^*_{t,n,s} \) that arises from an optimal signal structure, the function \( W(q, \lambda; \Delta) \) is differentiable with respect to \( q \). Suppose not: then for some signal realization, the sub-gradient of \( W(q, \lambda; \Delta) \) would contain multiple vectors, and it would follow by theorem 23.8 that the sub-gradient of the function \( f \) also contained multiple vectors, contradicting the differentiability result derived previously.

Denote the derivative as \( W_q(\cdot) \). We can write the optimality (first-order) condition as

\[ \sum_{s \in S} e_s ^T \omega D(q_t) W_q ^T (q^*_{t,n,s}, \lambda; \Delta_n) = \lambda \Delta_n ^{1-p} C(\cdot) ^{p-1} < \nabla C(\cdot), \omega >, \]

for any \( |S| \times |X| \) matrix \( \omega \) such that, for all \( x \in X \), \( \text{supp}(\omega(s|x)) \subseteq S_n \). Here, \( < \nabla C(\cdot), \omega > \) is the inner product over the tangent space of signal structures.
Specializing this result to $\omega = p_{t,n}^* - r_{t,n}^* t^T$, 

$$\sum_{s \in S} e_s^T (p_{t,n}^* - r_{t,n}^* t^T) D(q_t) W_q(q_{t,n,s}^*, \lambda; \Delta_n) = \lambda \Delta_n^{1-\rho} C(\cdot)^{\rho-1} < \nabla C(\cdot), p_{t,n}^* - r_{t,n}^* t^T > . \quad (23)$$

Next, observe that, for all $\lambda \in [0, \kappa c^{-\rho})$,

$$\sum_{s \in S} e_s^T r_{t,n}^* ||q_{t,n,s}^* - q_t||^2_\mathbb{X}_2 > 0.$$ 

Otherwise, $W(q_{t,n,s}^*, \lambda; \Delta_n) = W(q_t, \lambda; \Delta_n)$ for all $s \in S, C(\cdot) = 0$, and the Bellman equation (equation (20)) could not be satisfied. It follows that the set

$$\hat{Q}_{t,n} = \{ q \in \mathcal{P}(X) : \sum_{s \in S} e_s^T r_{t,n}^* ||q_{t,n,s}^* - q_t||^2_\mathbb{X}_2 < 2 \sum_{s \in S} e_s^T r_{t,n}^* ||q_{t,n,s}^*(\cdot, s) - q_t||^2_\mathbb{X}_2 \& \Vert q - q_t \Vert^2_\mathbb{X}_2 < \Delta_n \}$$

forms a non-empty open set containing the neighborhood around $q_t$.

By the boundedness and convexity of the value function, Alexandrov’s theorem holds, and, almost everywhere, the value function is first and second-order differentiable. Therefore, almost everywhere, and in particular for some $q_{t,n} \in \hat{Q}_{t,n},$
\[ W(q_{t,n,s}, \lambda; \Delta_n) = W(\hat{q}_{t,n}, \lambda; \Delta_n) + W'_q(\hat{q}_{t,n}) (q_{t,n,s} - \hat{q}_{t,n}) + \frac{1}{2} (q_{t,n,s} - \hat{q}_{t,n})^T A_n (q_{t,n,s} - \hat{q}_{t,n}) + o(||q_{t,n,s} - \hat{q}_{t,n}||^2_X) , \]

for some symmetric Hessian matrix \( A_n \). By the results of theorem 2.3 in Rockafellar [1999], we can write

\[ W_q(q_{t,n,s}, \lambda; \Delta) = W_q(\hat{q}_{t,n}, \lambda; \Delta_n) + < (q_{t,n,s} - \hat{q}_{t,n}) A_n + o(||q_{t,n,s} - \hat{q}_{t,n}||_{X,2}) . \]

That is, the points of second-differentiability of the value function are also the points at which the gradient has a Taylor expansion. Plugging this into equation (23),

\[ \sum_{s \in S} e_s^T (p_{t,n}^* - r_{t,n}^T) D(q_t) W_q(\hat{q}_{t,n}, \lambda; \Delta_n) + \]

\[ \sum_{s \in S} e_s^T (p_{t,n}^* - r_{t,n}^T) D(q_t) A_n (q_{t,n,s} - \hat{q}_{t,n}) + \]

\[ \sum_{s \in S} e_s^T (p_{t,n}^* - r_{t,n}^T) D(q_t) o(||q_{t+\Delta_n,n}(\cdot, s) - \hat{q}_{t,n}||_{X,2}) = \]

\[ \lambda \Delta^{1-n} C(\cdot)^{\rho-1} < \nabla C(\cdot), p_{t,n}^* - r_{t,n}^T > \]

Using Bayes’ rule, this simplifies to

123
\[ \sum_{s \in S} e_s^T r_{t,n}^* (q_{t,n,s}^* - \hat{q}_{t,n})^T A_n (q_{t,n,s}^* - \hat{q}_{t,n}) + \sum_{s \in S} e_s^T r_{t,n}^* o(||q_{t,n,s}^* - \hat{q}_{t,n}||_2^2) = \]

\[ \lambda \Delta_n^{1-\rho} C(\cdot)^{\rho-1} < \nabla C(\cdot), p_{t,n}^* - r_{t,n}^* t^T >. \quad (24) \]

Using the Alexandrov formula and the Bellman equation,

\[ W(\hat{q}_{t,n}, \lambda; \Delta_n) - W(q_t, \lambda; \Delta_n) + \frac{1}{2} \sum_{s \in S} e_s^T r_{t,n}^* (q_{t,n,s}^* - \hat{q}_{t,n})^T A_n (q_{t,n,s}^* - \hat{q}_{t,n}) + \]

\[ \sum_{s \in S} e_s^T r_{t,n}^* o(||q_{t,n,s}^* - \hat{q}_{t,n}||_2^2) = \frac{\lambda}{\rho} \Delta_n^{1-\rho} C(\cdot)^{\rho} + (\kappa - \lambda c^{\rho}) \Delta_n. \]

By the Lipschitz-Continuity of \( W(\cdot) \),

\[ |W(\hat{q}_{t,n}, \lambda; \Delta_n) - W(q_t, \lambda; \Delta_n)| \leq K ||\hat{q}_{t,n} - q_t||_{\mathcal{X}, 2}, \]

and therefore \( W(\hat{q}_{t,n}, \lambda; \Delta_n) - W(q_t, \lambda; \Delta_n) = o(\Delta_n) \). Moreover, note that by the definition of \( \hat{q}_{t,n} \),

\[ \sum_{s \in S} e_s^T r_{t,n}^* o(||q_{t,n,s}^* - \hat{q}_{t,n}||_{\mathcal{X}, 2}^2) = \sum_{s \in S} e_s^T r_{t,n}^* o(||q_{t,n,s}^* - q_t||_{\mathcal{X}, 2}^2). \]

Substituting out the Hessian term using equation (24),

\[ \lambda \Delta_n^{1-\rho} C(\cdot)^{\rho-1} (\frac{1}{2} < \nabla C(\cdot), p_{t,n}^* - r_{t,n}^* t^T > - \frac{1}{\rho} C(\cdot)) + \]

\[ o\left(\sum_{s \in S} e_s^T r_{t,n} ||q_{t,n,s}^* - \hat{q}_{t,n}||_{\mathcal{X}, 2}^2\right) = (\kappa - \lambda c^{\rho}) \Delta_n + o(\Delta_n). \]

By Taylor’s theorem,
\[
\frac{1}{2}(1 - \rho^{-1}) \lambda \Delta_n^{1 - \rho} C(\cdot)^{\rho - 1} \left( \sum_{s \in S} e_s^T r_{t,n}^* \| q_{t,n,s}^* - q_t \|^2_{X,k} \right) + o\left( \sum_{s \in S} r_{t,n}^*(s) \| q_{t,n,s}^* - q_t \|^2_{X,k} \right) = (\kappa - \lambda c^0) \Delta_n + o(\Delta_n). \tag{25}
\]
By the definition of little-o notation, this can be rewritten, given an arbitrary \( \varepsilon \in (0, 1) \), as

\[
\frac{1}{2}(1 - \rho^{-1})(1 - \varepsilon)\lambda \Delta_n^{1-\rho} C(\cdot)^{\rho-1} \sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2 + \\
+ o(\sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2) \leq (\kappa - \lambda c^\rho + \varepsilon)\Delta_n,
\]

for all \( n > n_\varepsilon \). Rescaling,

\[
\frac{1}{2}(1 - \rho^{-1})(1 - \varepsilon)\lambda \Delta_n^{1-\rho} C(\cdot)^{\rho-1} \sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2 + \\
+ \Delta_n^{\rho-1} o(\sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2) \leq (\kappa - \lambda c^\rho + \varepsilon)\Delta_n^\rho,
\]

Using the Taylor expansion of the cost function again, for some \( \hat{n}_\varepsilon \geq n_\varepsilon \), for all \( n > \hat{n}_\varepsilon \),

\[
C(\cdot) \geq (1 - \varepsilon)\frac{1}{2} \sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2.
\]

Therefore,

\[
(\frac{1}{2}(1 - \varepsilon))^\rho(1 - \rho^{-1})\lambda \left( \sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2 \right)^\rho + \\
+ \Delta_n^{\rho-1} o(\sum_{s \in S} e_s^t r_{t,n}^* ||q_{t,n,s} - q_t||_{\tilde{X},k}^2) \leq (\kappa - \lambda c^\rho + \varepsilon)\Delta_n^\rho.
\]

It follows that, for any \( \xi > 0 \), for \( n \) sufficiently large,
\[(\frac{1}{2}(1 - \varepsilon))^\rho (1 - \rho^{-1})\lambda (\sum_{s \in S} e^T_{s} r^*_{t,n} ||q^*_{t,n,s} - q_t||^2_{||X,k||})^\rho \leq (\kappa - \lambda c^\rho + \varepsilon)(\Delta^\rho_n + \xi \Delta^\rho_n - 1 \sum_{s \in S} e^T_{s} r^*_{t,n} ||q^*_{t,n,s} - q_t||^2_{||X,k||}).\]

Therefore, either

\[(\frac{1}{2}(1 - \varepsilon))^\rho (1 - \rho^{-1})\lambda (\sum_{s \in S} e^T_{s} r^*_{t,n} ||q^*_{t,n,s} - q_t||^2_{||X,k||})^\rho \leq 2(\kappa - \lambda c^\rho + \varepsilon)\Delta^\rho_n\]

or

\[(\frac{1}{2}(1 - \varepsilon))^\rho (1 - \rho^{-1})\lambda (\sum_{s \in S} e^T_{s} r^*_{t,n} ||q^*_{t,n,s} - q_t||^2_{||X,k||})^\rho \leq 2\xi \Delta^\rho_n - 1 \sum_{s \in S} e^T_{s} r^*_{t,n} ||q^*_{t,n,s} - q_t||^2_{||X,k||}.\]

Therefore,

\[\sum_{s \in S} e^T_{s} r^*_{t,n} ||q^*_{t,n,s} - q_t||^2_{||X,k||} \leq B\Delta_n,\]

where

\[B = \max\left(\frac{2(\kappa - \lambda c^\rho + \varepsilon)}{((\frac{1}{2}(1 - \varepsilon))^\rho (1 - \rho^{-1})\lambda)^\rho - 1}, \frac{2\xi}{((\frac{1}{2}(1 - \varepsilon))^\rho (1 - \rho^{-1})\lambda)^\rho - 1}\right) > 0.\]

By equation (22),

\[||P^*_{t,\Delta} - r^*_{t,\Delta} t^T||^2 \leq B_2 \Delta_n,\]

where \(B_2\) is a positive constant.
Armed with this result, we show that a convergent subsequence of optimal policies $p_{t,n}^*$ exists and has the claimed properties. Existence is guaranteed by the Bolzano-Weierstrauss theorem. Define

$$ r_t^* = \lim_{n \to \infty} p_{t,n}^* q_t. $$

It immediately follows that

$$ \lim_{n \to \infty} p_{t,n}^* = r_t^* t^T. $$

Define

$$ \tau_{t,n}^* = \Delta_n^{-\frac{1}{2}} (p_{t,n}^* - p_{t,n}^* q_t t^T). $$

By equation (26) above, $||\tau_{t,n}^* D(q_t)|| \leq B$, and therefore a convergent sub-sequence exists.

Pass to this subsequence, which we will also denote with $n$. Define

$$ \tau_t^* = \lim_{n \to \infty} \tau_{t,n}^*. $$

It follows that $t^T \tau_t^* = 0$, that $e_s^T \tau_t^* q_t = 0$, and that $||\tau_t^* D(q_t)||$ is bounded.

Define $\phi_{t,n}^* = p_{t,n}^* q_t - r_t^*$. It immediately follows that $\lim_{n \to \infty} e_s^T \phi_{t,n}^* = 0$ for all $s \in S$ and that $t^T \phi_{t,n}^* = 0$. By construction, note that

$$ \lim_{n \to \infty} \frac{p_{t,n,x}^* - r_t^* - \phi_{t,n}^* - \Delta_n^{-\frac{1}{2}} \tau_{t,x}^*}{\Delta_n^{-\frac{1}{2}}} = 0. $$

Finally, we demonstrate the claim that, if $e_s^T r_t^* > 0$,

$$ q_{t,n,x}^* = q_t + \Delta_n^{-\frac{1}{2}} \frac{e_s^T \tau_t^* D(q_t)}{e_s^T r_t^*} + o(\Delta_n^{-\frac{1}{2}}). $$
By Bayes’ rule,

\[ q_{t,n,s}^* = \frac{e_s^T p_{t,n}^* D(q_t)}{e_s^T p_{t,n}^* q_t}. \]

It follows from the convergence of \( p_{t,n}^* \) to \( r_t^* i^T \) that, for all \( s \) such that \( e_s^T r_t^* > 0 \),

\[ \lim_{n \to \infty} q_{t,n,s}^* = q_t. \]

Next, note that

\[ \Delta_n^{-\frac{1}{2}} (q_{t,n,s}^* - q_t) = \Delta_n^{-\frac{1}{2}} \frac{e_s^T p_{t,n}^* (D(q_t) - q_t q_t^T)}{e_s^T p_{t,n}^* q_t} \]
\[ = \frac{e_s^T \tau_{t,n}^* D(q_t)}{e_s^T p_{t,n}^* q_t}. \]

It follows that, for all \( s \) such that \( e_s^T r_t^* > 0 \),

\[ \lim_{n \to \infty} \Delta_n^{-\frac{1}{2}} (q_{t,n,s}^* - q_t) = \frac{e_s^T \tau_{t,n}^* D(q_t)}{e_s^T r_t^*}, \]

proving the claim.

**A.14 Proof of proposition 9**

Let \( n \in \mathbb{N} \) index the convergent sequence of optimal policies. Using the Taylor-expansion for the first order condition (as inequation (24) in the proof of lemma 8), at any point \( q_t \) at
which $W$ is twice-differentiable,

$$\sum_{s \in S} e_s^T r_{t,n}^* (q_{t,n,s}^* - \hat{q}_{t,n})^T (\nabla^2 W(q_t, \lambda; \Delta_n))(q_{t,n,s}^* - \hat{q}_{t,n}) +$$

$$\sum_{s \in S} e_s^T r_{t,n}^* o(||q_{t,n,s}^* - \hat{q}_{t,n}||_2^2) =$$

$$\lambda \Delta_n^{-\rho} C(\cdot)^{\rho - 1} < \nabla C(\cdot), p_{t,n}^* - r_{t,n}^* t^T > .$$

By the convergence results lemma 8, for all $s$ such that $e_s^T r_{t,n}^* > 0$,

$$q_{t,n,s}^* - q_t = \Delta_n \frac{1}{2} z_{t,s}^* + o(\Delta_n^2),$$

and therefore

$$\sum_{s \in S} e_s^T r_{t,n}^* ||q_{t,n,s}^* - q_t||_k^2 = \Delta_n \sum_{s \in S} e_s^T r_{t,n}^* ||z_{t,s}^*||_k^2 + o(\Delta_n)$$

and

$$o(\sum_{s \in S} e_s^T r_{t,n}^* ||q_{t,n,s}^* - q_t||_k^2) \subseteq o(\Delta_n).$$

By the Taylor expansion of $C(\cdot)$,

$$C(p_{t,n}^*, q_t) = \frac{1}{2} \Delta_n \sum_{s \in S} e_s^T r_{t,n}^* ||z_{t,s}^*||_k^2 + o(\Delta_n).$$

Rewriting this first-order condition,

$$\Delta_n (\sum_{s \in S} e_s^T r_{t,n}^* e_s^T (\nabla^2 W(q_t, \lambda; \Delta_n)) z_{t,s}^* =$$

$$2\lambda \Delta_n (\frac{1}{2} \sum_{s \in S} e_s^T r_{t,n}^* ||z_{t,s}^*||_k^2)^{\rho - 1} + o(\Delta_n).$$

130
By equation (25) in the proof of lemma 8,

\[
\frac{1}{2} (1 - \rho^{-1}) \lambda \Delta_n \rho C(\cdot) \rho^{-1} \left( \sum_{s \in S} e_s^T r^*_t ||q_{t,n,s}^* - q_t||_{\hat{X},k}^2 + o(\Delta_n) \right) = (\kappa - \lambda c^\rho) \Delta_n + o(\Delta_n).
\]

It follows that

\[
(1 - \rho^{-1}) \lambda \Delta_n \left( \frac{1}{2} \sum_{s \in S} e_s^T r^*_t ||z_{t,s}^*||_{\hat{X},k}^2 \right) \rho = (\kappa - \lambda c^\rho) \Delta_n + o(\Delta_n),
\]

which can be written as

\[
\frac{\lambda}{\rho} \Delta_n \rho C(\cdot) \rho = \frac{\kappa - \lambda c^\rho}{\rho - 1} + o(1).
\]

Define

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

Putting this together with the first-order condition,

\[
\sum_{s \in S} e_s^T r^*_t z_{t,s}^T (\nabla^2 W(q_t, \lambda; \Delta_n)) z_{t,s}^* = 
\theta \sum_{s \in S} e_s^T r^*_t ||z_{t,s}^*||_{\hat{X},k}^2 + o(1).
\]

which proves the result for the given convergent sequence of policies.

**A.15 Proof of lemma 10**

We being by assuming, without loss of generality, that the optimal policies are Markovian with respect to the state variable \(q_t\). This implies that the covariance matrices \(\Omega_{t,\Delta}\) are measurable with respect to the filtration generated by the state variable \(q_t\).

By the uniform boundedness and convexity of the family of value functions \(W(q, \lambda; \Delta_m)\),
this family of value functions is uniformly Lipschitz-continuous. The existence of a convergent sub-sequence follows from the Arzela-Ascoli theorem (see also Rockafellar [1970] theorem 10.9).

Pass to this sub-sequence, which (for simplicity) we also index by $m$. The beliefs $q_{t,m}$ are a family of $\mathbb{R}^{|X|}$-valued stochastic processes, with $q_{t,m} \in \mathcal{P}(X)$ for all $t \in [0, \infty)$ and $m \in \mathbb{N}$. Construct them as RCLL processes by assuming that $q_{\Delta m j + \epsilon, m} = q_{\Delta m j, m}$ for all $m, \epsilon \in [0, \Delta m)$, and $j \in \mathbb{N}$. By chapter 6, theorem 3.21, condition 1 in Jacod and Shiryaev [2013], and the boundedness of $q_{t,m}$, it follows that the laws of $q_{t,n}$ are tight. By Prokhorov’s theorem (chapter 6, theorem 3.9 in Jacod and Shiryaev [2013]), it follows that there exists a convergent sub-sequence. Index this sub-sequence by $n \in \mathbb{N}$.

By lemma 14 and equation (21), for any $q_{t,n} \in \mathcal{P}(X)$, we must have

$$\sum_{s \in \mathcal{S}} e_s^T r_{t,n}^\ast ||q_{t,n,s}^\ast - q_{t,n}||_{X,2} \leq B \Delta_n^{p-1}. $$

Therefore, for any $\epsilon > 0$,

$$P^n(||q_{t+\Delta n,n} - q_{t,n}||_{X,2} > \epsilon) \leq \frac{B}{\epsilon \Delta_n^{p-1}},$$

where $P^n$ denotes the probability measure associated with the stochastic process $q_{t,n}$. It follows from chapter 6, theorem 3.26, condition 3 in Jacod and Shiryaev [2013] that the processes $q_{t,n}$ are “C-tight,” meaning that the laws of $q_{t,n}$ converge in measure to the law of some continuous stochastic process, which we denote $q_t$.

Under the optimal policy, the stochastic process $\Delta_n^{-1} C(p_{t,n}^\ast, q_{t,n})$ is $L^p$-integrable (otherwise, the agent would achieve negatively infinite utility). It follows from 5 that $\Delta_n^{-1} \Omega_{t,n}$ is $L^p$-integrable.

By Bayes’ rule, the processes $q_{t,n}$ are martingales, and therefore $q_t$ is also a martingale.
By chapter 7, theorem 3.4 in Jacod and Shiryaev [2013], convergence in law implies that

$$\lim_{n \to \infty} \Delta_n^{-1} \Omega_{s,n}^* = \Sigma_s^*,$$

where $\Delta_n^{-1} \Omega_{s,n}^*$ and $\Sigma_s^*$ are the second modified characteristics of the semi-martingales (in this case, martingales) $q_{s,n}$ and $q_s$. By the closedness of $L^p$, $\Sigma_s^*$ is $L^p$-integrable, and therefore uniformly integrable.

For any $x \in X$ such that $e^T_x q_t = 0$, we must have, for all $s \in S$, $e^T_x q_{t,n,s} = 0$ (otherwise, beliefs could not be martingales), and therefore $e^T_x \Omega_{s,n}^* e_x = e^T_x \Sigma_s^* e_x = 0$. By the fact that $\Sigma_s^*$ is positive semi-definite, it can be decomposed as

$$\Sigma_s^* = D(q_s) \sigma_s^* \sigma_s^{*T} D(q_s),$$

as required.

We construct a martingale with unit variance for its increments from the innovations in $q_{t,n}$, following Amin and Khanna [1994] (lemmas 3.1, 3.2, and 3.3, noting that the statement of lemma 3.3 contains a typo). The essence of the issue is that the conditional covariance matrix of $q_{t,n}$, $\Omega_{t,n}^*$, is not necessarily full rank. By the results in that paper, this martingale and $q_{t,n}$ jointly converge in measure to a Brownian motion $B_t$ and $q_t$, where the Brownian motion is measurable with respect to the filtration generated by $q_t$.

Fixing $q_0$, the converse also holds– $q_t$ is measurable with respect to the filtration generated by the Brownian motion (this follows from the Markov property of $\Omega_{s,n}$). By the martingale representation theorem,

$$q_t = q_0 + \int_0^t D(q_s) \sigma_s^* dB_s.$$
A.16 Proof of lemma 5

We begin by discussing the convergence of stopping times. Let \( \tau^*_n \) denote the sequence of stopping times associated with the optimal policies. Note that, by the boundedness of the value function, we must have

\[
E_0[\tau^*_n] \leq \bar{\tau},
\]

for some weakly positive constant \( \bar{\tau} \) and all \( n \). It follows by the positivity of \( \tau^*_n \) that the laws of \( \tau^*_n \) 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 \( \tau^* \) denote the limit of this sub-sequence.

Following the arguments of Amin and Khanna [1994], lemmas 3.1 through 3.5, we can construct the martingale described in the proof of lemma 10 so that the martingale, stopping times \( \tau^*_n \), and beliefs \( q_{t,n} \) converge in measure to \( \tau^* \), \( B_t \), and \( q_t \). By the continuous mapping theorem, it follows that

\[
\lim_{n \to \infty} E_0[\hat{u}(q_{\tau^*_n,n}) - \kappa \tau^*_n)] = E_0[\hat{u}(q_{\tau^*}) - \kappa \tau^*]].
\]

We now construct a possibly sub-optimal policy by rescaling the optimal Markov policy,

\[
\tilde{p}_{t,n} = \alpha_{t,n} p_{t,n}^* + (1 - \alpha_{t,n}) r_{t,n}^* T_t,
\]

to satisfy \( C(\tilde{p}_{t,n}, q) = \Delta_n (\rho \frac{\lambda - \lambda q}{\lambda (\rho - 1)}) \rho^{-1} \). Such a policy exists by the strong convexity condition, the continuity of the cost function, and the strict positivity of \( \Delta_n^{-1} C(\cdot) \) for the optimal policy. Note that \( \lim_{n \to \infty} \alpha_{t,n} = 1 \).

Define stopping boundaries using the value functions. That is, define the closed set

\[
Q_n = \{ q \in \mathcal{P}(X) : W(q, \lambda; \Delta_n) \leq \hat{u}(q) \}.
\]

134
Define the stopping times $\bar{\tau}_n$ as the first hitting time for the set $Q_n$. Consider the value function $\bar{W}(q_0;\lambda;\Delta_n)$ associated with this policy, and let $\bar{q}_{t,n}$ be the process for beliefs. By construction,

$$\lim_{n \to \infty} E_0[\hat{u}(\bar{q}_{\bar{\tau}_n,n}) - \kappa \bar{\tau}_n] = E_0[\hat{u}(q^{*}) - \kappa \tau^{*}].$$

Adopt the convention that after the stopping time, both the optimal signals $p_{t,n}^{*}$ and the possibly-suboptimal signals $\bar{p}_{t,n}$ are uninformative. It follows by optimality that we must have

$$\lim_{n \to \infty} \frac{\lambda}{\rho} E_0[\int_{0}^{\infty} \Delta_n^{-\rho} C(p_{s,n}^{*},q_{s,n})^{\rho} ds] \leq \lim_{n \to \infty} \frac{\lambda}{\rho} E_0[\int_{0}^{\infty} \Delta_n^{-\rho} C(\bar{p}_{s,n},\bar{q}_{s,n})^{\rho} ds].$$

By construction,

$$\frac{\lambda}{\rho} E_0[\int_{0}^{\infty} \Delta_n^{-\rho} C(\bar{p}_{s,n},\bar{q}_{s,n})^{\rho} ds] = \frac{\lambda}{\rho} E_0[\int_{0}^{\tau_n^{*}} \rho \frac{\kappa - \lambda c^{\rho}}{\lambda(\rho - 1)} ds].$$

By Fatou’s lemma and proposition 9,

$$\frac{\lambda}{\rho} E_0[\int_{0}^{\tau_n^{*}} \rho \frac{\kappa - \lambda c^{\rho}}{\lambda(\rho - 1)} ds] \leq \lim_{n \to \infty} \frac{\lambda}{\rho} E_0[\int_{0}^{\infty} \Delta_n^{-\rho} C(p_{s,n}^{*},q_{s,n})^{\rho} ds].$$

It follows by the convergence of $\tau_n^{*}$ to $\tau^{*}$ that

$$\frac{\lambda}{\rho} E_0[\int_{0}^{\tau_n^{*}} \rho \frac{\kappa - \lambda c^{\rho}}{\lambda(\rho - 1)} ds] = \lim_{n \to \infty} \frac{\lambda}{\rho} E_0[\int_{0}^{\infty} \Delta_n^{-\rho} C(p_{s,n}^{*},q_{s,n})^{\rho} ds].$$

Therefore,
\[
W(q_0, \lambda) = \lim_{n \to \infty} W(q_0, \lambda; \Delta_n) \\
= E_0[\hat{u}(q_{\tau^*}) - \kappa \tau^*] - \\
\lambda E_0[\int_0^{\tau^*} \left( \frac{\kappa - \lambda c^{\rho}}{\lambda (\rho - 1)} - c^{\rho} \right) ds].
\]

We next demonstrate equality of the primal and dual. Define \( \lambda^* \) by

\[
\frac{\kappa - \lambda^* c^\rho}{\lambda^*(\rho - 1)} = c^\rho,
\]

or equivalently

\[
\lambda^* = \frac{\kappa}{\rho c^\rho}.
\]

Note that \( \lambda^* \in (0, \kappa c^{-\rho}) \), as required. For this value of \( \lambda \),

\[
W(q_0, \lambda^*) = E_0[\hat{u}(q_{\tau^*}) - \kappa \tau^*],
\]

and the limit of the constraint is satisfied:

\[
\lim_{n \to \infty} \frac{1}{\rho} E_0[\int_{\Delta_n} p_s^{s,n} C(p_s^{s,n}, q_{s,n})^{\rho} ds] = E_0[\int_0^{\tau^*} c^{\rho} ds].
\]

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 by \( n \)). By the standard duality inequalities, for all \( \lambda \),

\[
V(q_0; \Delta_n) \leq W(q_0, \lambda; \Delta_n)
\]

136
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
\[
\theta^* = \lambda^* (\rho \frac{\kappa - \lambda^* c}{\lambda^*(\rho - 1)})^{\frac{\rho - 1}{\rho}}.
\]
\[
= \lambda^* \rho^{\frac{\rho - 1}{\rho}} c^{\rho - 1}
\]
\[
= \frac{\kappa}{c} \rho^{-1}.
\]

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
\[
V(q_0) = \lim_{\Delta \to 0^+} V(q_0; \Delta).
\]
\[
= E_0[\hat{u}(q_{\tau^*}) - \kappa \tau^*].
\]

By proposition 9 and lemma 10,
\[
\frac{1}{2} \text{tr}[\sigma^*_t \sigma^*_t^T k(q_t)] = (\rho \frac{\kappa - \lambda^* c}{\lambda^*(\rho - 1)})^{\rho - 1} = \rho^{\rho - 1} c.
\]
Define \( \chi = \rho^{\rho - 1} c \).

Consider now an arbitrary sequence of Markov policies in the discrete time model, \( p_{t,n} \).
such that, for all $q_t$ at which the agent does not stop,

$$\lim_{n \to \infty} \sum_{s \in S} (e_s^T p_{t,n} q_t) ||q_{t,n,s} - q_{t,n}||_2 = 0.$$ 

By the arguments of lemma 10, under these policies,

$$q_t = q_0 + \int_0^t D(q_s) \sigma_s dB_s$$

for some $\sigma_s$ such that

$$D(q_s) \sigma_s \sigma_s^T D(q_s) = \lim_{n \to \infty} \Delta_n^{-1} \sum_{s \in S} (e_s^T p_{t,n} q_t) (q_{t,n,s} - q_t) (q_{t,n,s} - q_t)^T.$$ 

Restrict attention to policies satisfying feasibility: for any stopping time $\tau$,

$$E_0[\int_0^\tau c ds] \geq E_0[\frac{\Delta}{\rho} \sum_{j=0}^{\Delta^{-1} - 1} \Delta_n^{-1} C(p_{s,n}, q_{s,n})^\rho].$$

By Fatou’s lemma,

$$\lim_{n \to \infty} \frac{1}{\rho} E_0[\int_0^\tau \Delta_n^{-1} C(p_{s,n}, q_{s,n})^\rho ds] \geq \frac{1}{\rho} E_0[\int_0^\tau (\frac{1}{2} \text{tr}[\sigma_s \sigma_s^T k(q_s)])^\rho ds].$$

It follows that any policies satisfying

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

for all $q_t$ are feasible in the limit. Note that the optimal policies fall into this class.

If there were a sequence of policies and stopping times satisfying this condition that achieved strictly higher utility in the limit, there would be a policy in the discrete time problem that achieved higher utility than the optimal policies, a contradiction. It follows
that the value function is the maximum over all policies satisfying

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

concluding the proof.

### A.17 Proof of corollary 2

We begin by defining the drift component. Conditional on the true state \(x'\), the probability of receiving signal \(s \in S\) at time \(r\) under the optimal policy is

\[ e_s^T p_{r,n}^* e_{x'} \cdot \]

Define the drift in beliefs as

\[ a_{r,n} = \sum_{s \in S} (e_s^T p_{r,n}^* e_{x'})(q_{r,n,s} - q_{r,n}). \]

Note that

\[ |e_s^T a_{r,n}| < 1. \]

Let \(S_{x'} \subseteq S\) be the set of signals such that \(e_{x'}^T p_{r,n}^* e_{x'} > 0\). By Bayes’ rule,

\[ (e_{x'}^T p_{r,n}^* e_{x'})(e_{x'}^T q_{r,n}) = (e_{x'}^T p_{r,n} q_{r,n})(e_{x'}^T q_{r,n,s}). \]
By assumption, $e_x^T q_{r,n} > 0$ (that is, that the agent places some support on the true state of the world). We have

$$
\Delta_n^{-1} a_{r,n} = \Delta_n^{-1} \sum_{s \in S} \frac{(e_s^T p_{r,n}^* q_{r,n})(e_s^T q_{r,n,s})}{(e_{x'}^T q_{r,n})} (q_{r,n,s} - q_{r,n})
$$

$$= \Delta_n^{-1} \frac{1}{e_{x'}^T q_{r,n}} \Omega_{r,n}^* e_{x'}.
$$

Therefore,

$$
\lim_{n \to \infty} \Delta_n^{-1} a_{r,n} = \frac{1}{e_{x'}^T q_r} D(q_r) \sigma_r^* \sigma_r^T D(q_r) e_{x'}
$$

$$= D(q_r) \sigma_r^* \sigma_r^T e_{x'}.
$$

For $t = \Delta j$ for some positive integer $j$, the process

$$
\tilde{q}_{\Delta j,n} = q_{\Delta j,n} - \sum_{i=0}^{j-1} \Delta_n a_{\Delta i,n}
$$

is a martingale with the same variance-covariance matrix as $q_{\Delta j,n}$.

The variance-covariance matrix of beliefs, conditional on the true state, is

$$
\Sigma_{r,n} = \sum_{s \in S} (e_s^T p_{r,n}^* e_{x'}) (q_{r,n,s} - q_{r,n}) (q_{r,n,s} - q_{r,n})^T.
$$

Observe, by the results of lemma 8, that

$$
\lim_{n \to \infty} \Delta_n^{-1} \Sigma_{r,n} = \lim_{n \to \infty} \Delta_n^{-1} \sum_{s \in S} (e_s^T r_{r,s}^*) z_{r,s}^* z_{r,s}^T + o(1)
$$

$$= D(q_r) \sigma_r^* \sigma_r^T D(q_r).
$$
Note that, because probabilities are bounded between zero and one,

\[ |e_x^T \Sigma_{ee} e_x'| \leq |S|. \]

It follows by the dominated convergence theorem and the arguments employed in lemma 10 that

\[ q_\tau = q_0 + \int_0^\tau D(q_r) \sigma_r^* \sigma_r^{*T} e_r dr + \int_0^\tau D(q_r) \sigma_r^* e_r d\tilde{B}_r, \]

as required.