Technical Appendix

B Proofs

Notation

In our proofs, we often use a matrix notation in which a prior or posterior q is a vector in $\mathbb{R}^{|X|}$ and a signal structure p is an $|S| \times |X|$ matrix. $p_x \in \mathbb{R}^{|S|}$ and $p_s \in \mathbb{R}^{|X|}$ refer to specific columns and rows of this matrix. We use the notation $e_x \in \mathbb{R}^{|X|}$ and $e_s \in \mathbb{R}^{|X|}$ to refer to basis vectors with a one in the element corresponding to $x \in X$ and $s \in S$ respectively, and zero otherwise. We use the notation ι to refer to a vector of ones (in both $\mathbb{R}^{|S|}$ and $\mathbb{R}^{|X|}$ contexts).

B.1 Proof of Lemma 1

Let p and p' be information structures with signal alphabet S. First, we will show that mixture feasibility and Blackwell monotonicity imply convexity. By mixture feasibility, letting p_M denote the mixture information structure and S_M the signal alphabet,

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

Consider the garbling $\Pi : S \times \{1,2\} \to S$, 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 information structures with signal alphabets S_1 and S_2. Because the cost function satisfies Blackwell monotonicity, it is invariant to Markov congruent embeddings. Define $S_M = (S_1 \cup S_2) \times \{1,2\}$. There exists an embedding $\Pi_1 : S_1 \to S_M$ such that, for some $s_M = (s,i) \in S_M$,

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

Define an embedding Π_2 along similar lines,

$$e_{s_M}^T \Pi_2 p_2 = \begin{cases} 0 & i = 1 \\ 0 & s \notin S_2 \\ e_s^T p_2 & otherwise, \end{cases}$$

and note that these embeddings are left-invertible. It follows by Blackwell monotonicity in both directions 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.

B.2 Proof of Theorem 1

To prove the theorem, we use Taylor's theorem to approximate the cost function and its gradient up to order Δ_m (a second-order approximation for the cost function, first-order for the gradient).

We start by describing the local (second-order) properties of any information cost function satisfying our conditions. The condition requiring that Blackwell-dominant information structures cost weakly more (Condition 3) is of particular importance. Recall Blackwell's theorem:

Theorem. *(Blackwell [1953]) The information structure $\{p_x\}_{x\in X}$, with signal alphabet S, is more informative, in the Blackwell sense, than $\{p'_x\}_{x\in X}$, with signal alphabet S', if and only if there exists a Markov transition matrix $\Pi : S \to S'$ such that, for all $s' \in S'$ and $x \in X$,*

$$p'_x = \Pi p_x. \tag{25}$$

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 actually garble the signals. These garbling matrices have left inverses that are also Markov transition matrices. If we define an information structure $\{p_x\}_{x\in X}$, with signal alphabet S, and another information structure $\{p'_x\}_{x\in X}$, with signal alphabet S', using one of these left-invertible matrices, via equation (25), then $\{p_x\}_{x\in X}$ is more informative than $\{p'_x\}_{x\in X}$, but $\{p'_x\}_{x\in X}$ is also more informative than $\{p_x\}_{x\in X}$. These two information structures are called "Blackwell-equivalent," and it follows that the cost of these two information structures must be equal, by Condition 3. The left-invertible Markov transition matrices associated with Blackwell-equivalent information structures are called Markov congruent embeddings by Chentsov [1982]. Chentsov [1982] studied tensors and divergences that are invariant to Markov congruent embeddings (we will say "invariant" for brevity).

Let Π be a Markov congruent embedding from $\mathcal{P}(S)$ to $\mathcal{P}(S')$. By Condition 3, all information cost functions satisfying our conditions are invariant to Markov congruent embeddings. It necessarily follows that, for any Markov congruent embedding Π, that

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

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

The key results of Chentsov [1982] are expressed in terms of the Fisher information matrix. In our context, the Fisher information matrix on the simplex is

$$g(r) = Diag(r)^{+} - \iota\iota^{T},$$

where $Diag(r)^{+}$ is the pseudo-inverse of $Diag(r)$ and ι is a vector of ones. Chentsov establishes the following results:[21]

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.[22]

These results allow us to characterize the local properties of rational inattention cost functions, via a Taylor expansion. Hold fixed the signal alphabet S, and consider an information structure

$$p_x(\epsilon, \nu) = r + \epsilon\nu_x + \nu\omega_x.$$

Here, $r \in \mathcal{P}(S)$ and $\nu_x \in \mathbb{R}^{|S|}$ satisfies $\iota^T\nu_x = 0$ for all x, where ι is a vector of ones. We also assume that, for all $s \in S$, $e_s^T\nu_x \neq 0$ only if $e_s^T r > 0$. That is, ν_x is an element of the tangent space of the probability simplex at r. The same properties hold true for ω_x. As a result, for values of the perturbation parameters ϵ and ν sufficiently close to zero, $p_x \in \mathcal{P}(S)$ for all $x \in X$. In other words, the parameters ϵ and ν index a two-parameter family of perturbations of an uninformative information structure (corresponding to $\epsilon = \nu = 0$), in which the perturbed information structures will generally be informative; the ν_x and ω_x specify two directions of perturbation. Each of the perturbed information structures has the property that p_x is absolutely continuous with respect to r.

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

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

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

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

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

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

$$\frac{\partial}{\partial \nu}\frac{\partial}{\partial \epsilon}C(\{p_x(\epsilon,\nu)\}_{x\in X}, q; S)|_{\epsilon=\nu=0} = \sum_{x'\in X}\sum_{x\in X}\omega_{x'}^T \cdot C_{xx'}(\{r\}_{x\in X}, q; S)\cdot \nu_x.$$

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

$$\frac{\partial}{\partial \nu}\frac{\partial}{\partial \epsilon}C(\{p(\cdot|\cdot;\epsilon,\nu)\}, q)|_{\epsilon=\nu=0} = \sum_{x'\in X}\sum_{x\in X}k_{x,x'}(q)\omega_{x'}^T g(r)\nu_x,$$

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-valued function $k(q)$ can depend on the prior q, but cannot depend on the unconditional distribution of signals, r; otherwise, invariance would not hold.

We begin by considering perturbations that preserve the support of the signal structure. As a result, this theorem should be interpreted as applying to "frequent but not very informative" signals, as opposed to "rare but informative" signals. We will discuss the latter type of signals shortly. Note that the pseudo-inverse of the Fisher information matrix is

$$g^+(q) \;=\; Diag(q) - qq^T.$$

Lemma 11. *Suppose that a sequence of information structures p_m, with signal alphabet S, is described by the equation*

$$p_{m,s,x} = \Delta_m^{\alpha(s)} r_s + \Delta_m^{\frac{1}{2}(1+\alpha(s))}\nu_{s,x} + o(\Delta_m^{\frac{1}{2}(1+\alpha(s))}),$$

where, for all $s \in S$, $x \in X$, and $\Delta_m \geq 0$, $p_{m,x,s} \neq 0 \Rightarrow r_s > 0$, $\alpha(s) \in [0, 1)$, and $\sum_{x\in X} v_{s,x} = 0$. Let $C(\cdot)$ be an information cost function that satisfies Conditions 1-4.

There exists a matrix valued function $k(q)$ such that

$$C(p_m; q; S) = \frac{1}{2}\Delta_m \sum_{x'\in X}\sum_{x\in X} k_{x,x'}(q)\nu_{x'}^T g(r)\nu_x + o(\Delta_m).$$

For all q, the matrix-valued function $k(q)$ is continuous, positive semi-definite and symmetric, and satisfies $v^T k(q)v = 0$ for any vector $z \in \mathbb{R}^{|X|}$ that is constant in the support of q.

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

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

In the case of the mutual-information cost function, the matrix $k(q)$ is itself the pseudo-inverse of the Fisher information matrix.

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 and the conditional probabilities of signals. The corollary below expresses the results of Lemma 11 in terms of posterior beliefs.

Corollary 4. *Under the assumptions of Lemma 11, the posterior beliefs can be written, for any $s \in S$ such that $r_s > 0$, as*

$$q_{s,x}(p_m, q) = q_x + \Delta_m^{\frac{1}{2}(1-\alpha(s))} q_x \frac{\nu_{s,x} - \sum_{x' \in X} q_{x'} \nu_{s,x'}}{r_s} + o(\Delta_m^{\frac{1}{2}(1-\alpha(s))}).$$

The cost function can be written as

$$C(p_m, q; S) = \frac{1}{2} \sum_{s \in S: r_s > 0} \pi_s(p_m, q)(q_{s,x}(p_m, q) - q)^T \bar{k}(q)(q_{s,x}(p_m, q) - q) + o(\Delta_m),$$

where $\bar{k}(q) = Diag(q)^+ k(q) Diag(q)^+$.

Proof. See the appendix, section B.4. □

There are, in effect, two ways for a signal to be contain a small amount of information, and different costs associated with these different types of signals. The results of Lemma 11 characterize, for any rational inattention cost function satisfying our conditions, the cost of receiving frequently, but relatively uninformative, signals. We next consider the cost of receiving a rare but informative signal.

Lemma 12. *Under the assumptions of Lemma 11, define the signal structure*

$$\hat{p}_m = p_m + \Delta_m \omega,$$

where p_m is a signal structure of the type described in Lemma 11, with $\sum_{s \in S} \omega_x = 0$ for all $x \in X$ and with $\omega_{s,x} \geq 0$ for all $s \in S$ such that $p_{m,s,x} = 0$.

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

$$C(\hat{p}_m; q; S) = \frac{1}{2} \sum_{s \in S: r_s > 0} \pi_s(\hat{p}_m, q)(q_{s,x}(\hat{p}_m, q) - q)^T \bar{k}(q)(q_{s,x}(\hat{p}_m, q) - q) + \sum_{s \in S: r_s = 0} \pi_s(\hat{p}_m, q) D^*(q_{s,x}(\hat{p}_m, q) || q) + o(\Delta_m), \quad (26)$$

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

$$\frac{\partial^2 D^*(r||q)}{\partial r^i \partial r^j}|_{r=q} \;=\; \bar{k}(q). \tag{27}$$

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

The divergence D^* represents the cost of acquiring an infrequent, but potentially informative, signal. Naturally, if the signal is in fact not very informative, this cost must be closely related to the costs of other uninformative signals, which gives rise to the condition on the Hessian of the divergence. Note that the lemma demonstrates that the cost is additive with respect to the other signals being received (at least up to order Δ). The result follows from the directional differentiability of the cost function with respect to signals that occur with zero probability and the continuity of that directional derivative (Condition 4) and invariance.

To conclude the proof, observe that by assumption,

$$\pi_s(p_m, q)||q_s(p_m, q) - q||_X^2 \leq B\Delta_m$$

for all $m \in \mathbb{N}$ and $s \in S$. Consequently, for all convergent subsequences of m (denote them by n), either

$$\lim_{n\to\infty} \frac{\pi_s(p_n, q)}{\Delta_n^{\alpha(s)}} \leq B$$

for some $\alpha(s) \in [0, 1)$, or $\pi_s(p_n, q) = O(\Delta_n)$.

In the first case, we must have

$$||q_s(p_n, q) - q||_X^2 = O(\Delta_n^{1-\alpha(s)}).$$

In this case, it follows by Taylor's theorem that

$$\frac{1}{2}(q_{s,x}(p_m, q) - q)^T \bar{k}(q)(q_{s,x}(p_m, q) - q) = D^*(q_s(p_m, q)||q) + o(\Delta_m^{1-\alpha(s)}).$$

Defining

$$r_s = \lim_{n\to\infty} \frac{\pi_s(p_n, q)}{\Delta_n^{\alpha(s)}}$$

and

$$v_{s,x} = \lim_{n\to\infty} \frac{q_{s,x}(p_m, q) - q_x}{q_x \pi_s(p_n, q)} \Delta_n^{-\frac{1}{2}(1+\alpha(s))},$$

we can apply Lemma 11.

In the second case, defining

$$\omega_{s,x} = \lim_{n\to\infty} \frac{q_{s,x}(p_n, q)}{q_x \Delta_n} \pi_s(p_n, q)$$

allows us to apply Lemma 12. It follows that, for all convergent subsequences (therefore by boundedness for all m),

$$C(p_m;q;S)=\sum_{s\in S}\pi_s(p_m,q)D^*(q_s(p_m,q)||q)+o(\Delta_m).$$

The claimed properties of the divergence and its Hessian follow from the two lemmas.

B.3 Proof of Lemma 11

Consider a Taylor expansion of $C(p_m,q;S)$ around the value of

$$r_{m,s}=\Delta_m^{\alpha(s)}r_s.$$

We have

$$p_{m,s,x}-r_{m,s}=\Delta_m^{\frac{1}{2}(1+\alpha(s))}\nu_{s,x}+o(\Delta_m^{\frac{1}{2}(1+\alpha(s))}),$$

and therefore by Taylor's theorem we have

$$C(p_m,q;S)=\frac{1}{2}\sum_{x'\in X}\sum_{x\in X}k_{x,x'}(q)(p_{m,x}-r_m)^T g(r_m)(p_{m,x}-r_m)+o(\sum_{s\in S}|p_{m,x,s}-r_{m,s}|^2).$$

By construction,

$$o(\sum_{s\in S}|p_{m,x,s}-r_{m,s}|^2)=o(\Delta_m).$$

Observing that

$$\sum_{x\in X}v_{s,x}=0,$$

that $p_{m,s,x}-r_{m,s}\neq 0$ if and only if $r_{m,s}>0$, and using the definition

$$g(r_m)=Diag(r_m)^+-\iota\iota^T,$$

we have

$$(p_{m,x}-r_m)^T g(r_m)(p_{m,x}-r_m)=\Delta_m^{(1+\alpha(s))}\nu_x^T Diag(r_m)^+\nu_x+o(\Delta_m),$$

which is

$$(p_{m,x}-r_m)^T g(r_m)(p_{m,x}-r_m)=\Delta_m\nu_x^T Diag(r)^+\nu_x+o(\Delta_m).$$

It therefore follows that

$$C(p_m;q;S)=\frac{1}{2}\Delta_m\sum_{x'\in X}\sum_{x\in X}k_{x,x'}(q)\nu_{x'}^T g(r)\nu_x+o(\Delta_m).$$

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

Now consider a particular sequence of information structures for which $\nu_{s,x}=\phi_s v_x$, where $v\in\mathbb{R}^{|X|}$ and $\phi\in\mathbb{R}^{|S|}$, with $\sum_{s\in S}e_s^T\phi=0$, and $\alpha(s)=0$ for all $s\in S$. Suppose that both v and ϕ

are not zero. For this sequence of information structures,

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

where $\phi^T g(r)\phi = \bar{g} > 0$. Suppose the information structure is uninformative for all Δ_m. This would be the case if v is proportional to ι, or any vector constant in the support of q, because such a signal structure has the same distribution of signals in each state in the support of q. Therefore, for such a v,

$$v^T k(q) v = 0$$

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

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

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

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

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

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

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

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

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

$$q_s(p_m, q) - q = \frac{(Diag(q) - qq^T) p_{m,s}^T}{q^T p_{m,s}^T}.$$

By convention, $q_s = q$ for any s such that $\pi_s(p, q) = 0$.

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

$$C(p_m, q; S) \geq \frac{m_g}{2} \sum_{s \in S} \pi_s(p, q) (q_s(p_m, q) - q)^T Diag^+(q) (q_s(p_m, q) - q)$$

for some constant $m_g > 0$.

For sufficiently small Δ_m, $\pi_s(p, q) > 0$ if $r_s > 0$, and therefore

$$C(p_m, q; S) \geq \frac{m_g}{2} \sum_{s \in S : r_s > 0} \frac{(p_{m,s}(Diag(q) - qq^T) Diag^+(q) (Diag(q) - qq^T) p_{m,s}^T)}{\pi_s(p, q)},$$

or, for the particular sequence defined by the vectors ϕ and v,

$$C(p_m, q; S) \geq \frac{m_g}{2}\Delta_m \sum_{s \in S: r_s > 0} (\phi_s)^2 \frac{v^T(Diag(q) - qq^T)Diag^+(q)(Diag(q) - qq^T)v}{(r_s)} + o(\Delta).$$

Noting that

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

and that

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

we have

$$C(p_m, q; S) \geq \frac{m_g}{2}\Delta_m \bar{g} v^T g^+(q) v + o(\Delta_m).$$

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.

B.4 Proof of Corollary 4

Under the stated assumptions,

$$p_{m,s,x} = \Delta_m^{\alpha(s)} r_s + \Delta_m^{\frac{1}{2}(1+\alpha(s))}\nu_{s,x} + o(\Delta_m^{\frac{1}{2}(1+\alpha(s))}).$$

By Bayes' rule, for any $s \in S$ such that $\pi_s(p_m, q) > 0$,

$$q_s(p_m, q) = \frac{Diag(q)p_{m,s}^T}{q^T p_{m,s}^T}.$$

It follows immediately that

$$\lim_{\Delta \to 0^+} q_s(p_m, q) = Diag(q)\frac{\iota r_s}{r_s} = q.$$

Next, using the notation $v_s \in \mathbb{R}^{|X|}$ to denote the vector of $\{v_{s,x}\}_{x \in X}$,

$$\begin{aligned}\Delta_m^{-\frac{1}{2}(1-\alpha(s))}(q_s(p_m, q) - q) &= \frac{\Delta_m^{-\frac{1}{2}(1+\alpha(s))}}{\Delta_m^{-\alpha(s)}} \frac{(Diag(q) - qq^T)p_s^T}{\pi_s(p_m, q)} \\ &= Diag(q)\frac{\nu_s - \iota q^T \nu_s + o(1)}{\Delta_m^{-\alpha(s)}\pi_s(p_m, q)}.\end{aligned}$$

For any s such that $r_s > 0$,

$$\lim_{\Delta \to 0^+} \Delta_m^{-\frac{1}{2}(1-\alpha(s))}(q_s(p_m, q) - q) = Diag(q)\frac{\nu_s - \iota q^T \nu_s}{r_s}.$$

By Lemma 11,

$$C(p_m, q; S) = \frac{1}{2}\Delta_m \sum_{x' \in X} \sum_{x \in X} k_{x,x'}(q)\nu_{x'}^T g(r)\nu_x + o(\Delta_m).$$

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

$$\nu_{x'}^T g(r)\nu_x = \sum_{s \in S: r_s > 0} r_s \frac{\nu_{x',s}}{r_s}\frac{\nu_{x,s}}{r_s}.$$

Substituting in the result regarding the posterior,

$$C(p_m, q; S) = \frac{1}{2} \sum_{s \in S: r_s > 0} \pi_s(p, q)(q_s(p_m, q) - q)^T Diag^+(q)k(q)Diag^+(q)(q_s(p_m, q) - q) + o(\Delta_m),$$

which is the result, observing that $q^T Diag^+(q)$ is constant in the support of q and applying Lemma 11.

B.5 Proof of Corollary 12

By directional differentiability and the continuity of the directional derivatives, there exists a function

$$f(\omega, r, q; S) = \lim_{m \to \infty} \frac{C(p_m + \Delta_m \omega, q; S) - C(p_m, q; S)}{\Delta_m}.$$

Observe that, if ωe_x is in the support of r for all x in the support of q, we must have $f(\omega, \bar{p}, q; S) = 0$, by the results of Lemma 11. Relatedly, if ω and ω' differ only with respect to the frequency of signals in the support of r for all x in the support of q, we must have

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

Assuming there are signals with $\pi_s(p_m, q) = 0$, we can write $\omega = \omega_1 + \omega_2 + \ldots$, where each ω_i is a perturbation that contains only one signal with $\pi_s(p_m, q) = 0$. Let $N \leq |S|$ denote the number of these perturbations. We can define

$$f_i(\omega_i, r, q; S) = \lim_{m \to \infty} \frac{C(p_{i-1,m} + \Delta\omega_i, q; S) - C(p_{i-1,m}, q; S)}{\Delta_m},$$

where $p_{i-1,m} = p_m + \Delta \sum_{j=1}^{i-1} \omega_i$. By the assumption of the continuity of the directional derivatives,

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

It follows that

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

By invariance, the function $f(\omega_i, r, q; S)$ does not depend on r or S. By the argument above, it is only a function of $\omega_{i,s_i} \in \mathbb{R}^{|X|}$, where $s_i \in S$ is the unique signal in ω_i with $r_{s_i} = 0$. By Bayes' rule, if the prior q has full support,

$$\omega_{i,s_i} = (q^T \omega_{i,s_i}) Diag(q)^+ q_{s_i},$$

where q_{s_i} is the posterior associated with signal s_i. If not, by Condition 1, it is without loss of generality to assume $\omega_{i,s_i} e_x = r$ for all x not the support of q, and hence that this equation holds for all q. By the homogeneity of the directional derivative, we can rewrite this as

$$f(\omega_i, r, q; S) = (q^T \omega_{i,s_i}) F(q_{s_i}, q).$$

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

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

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

for all $q' \neq q$. Therefore, F is a divergence, which we write $D^*(q'||q)$. The finiteness and continuity of $D^*(q'||q)$ is implied by the existence and continuity of the directional derivative. The approximation of the cost function follows from this result and Corollary 4, observing that $\pi_{s_i}(\hat{p}_m, q) = q^T \omega_{i,s_i}$.

By invariance, there exists a Markov congruent embedding that splits each signal in S into $M > 1$ distinct signals in S'. As M becomes arbitrarily large, the probability of each signal becomes small — and in particular, can be of order Δ. It follows for all $s \in S'$ such that $||q_s - q|| = O(\Delta_m^{\frac{1}{2}(1-\alpha(s))})$ (e.g. the signals described in Corollary 4), we must have

$$D^*(q_s||q) = \frac{1}{2} \Delta_m^{(1-\alpha(s))} (q_s^T - q) \bar{k}(q) (q_s - q) + o(\Delta_m^{(1-\alpha(s))}),$$

and therefore $D^*(q'||q)$ must be twice differentiable in q' evaluated at q.

Lastly, we prove that D^* is convex in its first argument. By the convexity of $C(p, q; S)$,

$$C(p_m, q; S) \geq \sum_{s \in S} \pi_s(p_m, q) D^*(q_s(p_m, q)||q).$$

Therefore, for all signal structures p_m^1 and p_m^2 satisfying the conditions of the lemma and all $\lambda \in (0, 1)$, letting $p_m = \lambda p_m^1 + (1-\lambda) p_m^2$, by convexity

$$\lambda C(p_m^1, q; S) + (1-\lambda) C(p_m^2, q; S) \geq \sum_{s \in S} \pi_s(p_m, q) D^*(q_s(p_m, q)||q).$$

Taking the limit as $m \to \infty$, we must have

$$\lim_{m\to\infty} \lambda \sum_{s \in S} \pi_s(p_m^1, q) D^*(q_s(p_m^1, q)||q) +$$
$$\lim_{m\to\infty} (1-\lambda) \sum_{s \in S} \pi_s(p_m^2, q) D^*(q_s(p_m^2, q)||q) \geq \lim_{m\to\infty} \sum_{s \in S} \pi_s(p_m, q) D^*(q_s(p_m, q)||q).$$

For any $q^1, q^2 \in \mathcal{P}(X)$ absolutely continuous with respect to q, define

$$p_m^1 = r + \Delta_m \omega^1,$$
$$p_m^2 = r + \Delta_m \omega^2,$$

where ω^1 and ω^2 are both non-zero only for some $s \in S$ with $r_s = 0$, and satisfy $q^1 = q_s(p_m^1, q)$, $q^2 = q_s(p_m^2, q)$, and $\pi_s(p_m^1, q) = \pi_s(p_m^2, q)$ (which can be achieved by Bayes' rule). It follows that

$$q_s(p_m, q) = \lambda q^1 + (1-\lambda) q^2,$$

and therefore that D^* is convex in its first argument.

B.6 Proof of Theorem 2

We begin by describing three lemmas that we will employ to prove the convergence result. Our first lemma shows that the dual discrete time value function $W(q_t, \lambda; \Delta)$ is well-behaved:

Lemma 13. *If $\lambda \in (0, \kappa c^{-\rho})$ and $\beta = 1$, or if $\beta < 1$, for all $\Delta \leq 1$ the value function $W(q_t, \lambda; \Delta)$ is bounded above on $q_t \in \mathcal{P}(X)$ by a constant $\bar{W}$, bounded below by zero, and is convex in q. Moreover, for all $\Delta \leq 1$,*

$$\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W} > 0.$$

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

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

Lemma 14. *Let $n \in \mathbb{N}$ denote a sequence such that $\Delta_n \leq 1$ and $\lim_{n\to\infty} \Delta_n = 0$. Under the assumptions of Lemma 13, any associated sequence of optimal policies $p_{t,n}^*$ satisfies, for all elements of the sequence,*

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

where $\theta = \lambda(\rho \frac{\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W}}{\lambda(\rho-1)})^{\frac{\rho-1}{\rho}}$ and $\bar{W}$ is the upper bound of Lemma 13.

Proof. See appendix, section B.8. □

Our next lemma discuss the convergence of an arbitrary sequence of stochastic processes for beliefs (denoted $q_{t,m}$) and of stopping times (denoted τ_m) to their continuous-time limits, under the assumption that the policies generating them satisfy the bound in Lemma 14 and a bound on expected stopping times. This lemma applies to a sequence of optimal policies, but also to sequences of sub-optimal policies. The lemma describes the convergence of the beliefs process to a martingale, which is not necessarily a diffusion (it may have jumps, or even be a semi-martingale that is not a jump-diffusion).

Lemma 15. *Let Δ_m, $m \in \mathbb{N}$, denote a sequence such that $\lim_{m\to\infty} \Delta_m = 0$. Let $p_m(q)$ denote a sequence of Markov policies satisfying the bound in Lemma 14. Let $q_{t,m}$ denote the stochastic process for the DM's beliefs at time t, under such a policy, and let τ_m be a sequence of stopping policies such that $E_0[\tau_m] \leq \bar{\tau}$.*

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

1. *The beliefs $q_{t,n}$ and the stopping time τ_n converge almost surely to a martingale q_t and a stopping time τ.*

2. *The martingale q_t can be represented in terms of its semi-martingale characteristics,*

$$B_t = -\int_0^t (\int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) z dz) dA_s$$

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

$$\nu_t(z) = dA_t \psi_t(x),$$

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

3. *For all stopping times T,*

$$E_t[\int_t^T \beta^{s-t}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z||q_{s^-})dz\} dA_s] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \frac{1-\beta^{T-t}}{-\ln(\beta)}.$$

4. *The limit of the cumulative information cost is bounded below,*

$$\lim_{n\to\infty} E_0[\Delta_n^{1-\rho} \sum_{j=0}^{\tau_n\Delta_n^{-1}-1} \beta^{\Delta_n j} C(p_n(q_{\Delta_n j,n}), q_{\Delta_n j,n}; S)^\rho] \geq E_t[\int_0^\tau \beta^s \{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z||q_{s^-})dz\}^\rho (\frac{dA_s}{ds})^\rho ds].$$

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

Having described these three lemmas (all proven below), we now proceed to the main proof. Assume that $\lambda \in (0, \kappa c^{-\rho})$ if $\beta = 1$, $\lambda > 0$ if $\beta < 1$. Under this assumption, lemmas 13, 14, and 15 apply.

Let m index a sequence of Markov optimal policies, $p_m^*(q)$, and of stopping times τ_m^*. Let $q_{t,n}^*$ denote the associated process for beliefs. By the uniform boundedness and convexity of the family of value functions $W(q, \lambda; \Delta_m)$, a uniformly convergent sub-sequence exists. Rockafellar [1970] Theorem 10.9 demonstrates that a uniformly convergent sub-sequence exists on the relative interior of the simplex, and Rockafellar [1970] Theorem 10.3 demonstrates that there is a unique extension to a convex and continuous function on the boundary of the simplex. Therefore, by Lemma 13, that proving the discrete time value function converges to the stated continuous time limit also proves that the continuous time limit is bounded and convex.

Pass to this sub-sequence, which (for simplicity) we also index by m, and let $W(q,\lambda)$ denote its limit. Let $W^{+}(q,\lambda)$ denote the continuous time problem defined in Definition 1. We will prove that $W(q,\lambda) = W^{+}(q,\lambda)$.

By Lemmas 13 and 14, the sequence of optimal policies and stopping time satisfies the conditions of Lemma 15. It follows by that lemma that

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

where q_s^* is the limiting stochastic process and $\sigma_s^*,\psi_s^*,dA_s^*$ are associated with the characteristics of the martingale q_s^*.

We also have, by weak convergence,

$$\lim_{n\to\infty} E_0[\beta^{\tau_n}\hat{u}(q_{\tau_n^*,n}) - \Delta_n\frac{1-\beta^{\tau_n^*}}{1-\beta^{\Delta_n}}(\kappa-\lambda c^{\rho}))] = E_0[\beta^{\tau^*}\hat{u}(q_{\tau^*}) - \frac{1}{-\ln(\beta)}(1-\beta^{\tau^*})(\kappa-\lambda c^{\rho}))].$$

Recall also the bound, for any stopping time T measurable with respect filtration generated by q_s^* ,

$$E_t[\int_t^{T} \beta^{s}\{\frac{1}{2}tr[\sigma_s^*\sigma_s^{*T}k(q_{s-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^*(z)D^*(q_{s-}^*+z||q_{s-}^*)dz\}dA_s^*] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}E_t[\frac{1-\beta^{(T-t)}}{-\ln(\beta)}].$$

It follows that

$$W(q,\lambda) \leq W^{+}(q,\lambda)$$

for all $q \in \mathcal{P}(X)$, where

$$W^{+}(q_t,\lambda) = \sup_{\{\sigma_s,\psi_s,dA_s,\tau\}} E_t[\beta^{(\tau^*-t)}\hat{u}(q_{\tau^*}) - \frac{1}{-\ln(\beta)}(1-\beta^{(\tau^*-t)})(\kappa-\lambda c^{\rho})]- -\frac{\lambda}{\rho}E_t[\int_t^{\tau}\beta^{(s-t)}\{\frac{1}{2}tr[\sigma_s\sigma_s^{T}k(q_s)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z)D^*(q_{s-}+z||q_{s-})dz\}^{\rho}(\frac{dA_s}{ds})^{\rho}ds],$$

subject to the constraints, for all stopping times T measurable with respect filtration generated by

q_s^* ,

$$E_t[\int_t^T \beta^{(s-t)}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_t)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z)D^*(q_{s^-}+z||q_{s^-})dz\}dA_s] \leq$$
$$(\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}E_t[\frac{1-\beta^{(T-t)}}{-\ln(\beta)}].$$

and the evolution of beliefs as implied by the characteristics derived from σ_s, ψ_s, dA_s. Observe, by the arguments in the proof of Lemma 13, that $W^+(q,\lambda)$ is bounded, and convex in q, and satisfies $\kappa - \lambda c^\rho - \ln(\beta)W^+(q_t,\lambda) > 0$.

Also note that, for W^+, it is without loss of generality to set $dA_s = ds$. Scaling dA_s up and scaling $\sigma_s\sigma_s^T$ and ψ_s down, or vice versa, does not change the constraint, and setting $dA_s = 0$ is clearly sub-optimal by the result that $\kappa - \lambda c^\rho - \ln(\beta)W^+(q_t,\lambda) > 0$. Note also that there is a version of the optimal policies which satisfy the constraint everywhere:

$$\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z)D^*(q_{s^-}+z||q_{s^-})dz \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}.$$

Next, observe that increasing $\sigma_s\sigma_s^T$ by a quantity ϵzz^T results in a first order condition, anywhere W^+ is twice-differentiable, of

$$\lambda\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z'')D^*(q_{s^-}+z''||q_{s^-})dz''\}^{\rho-1}\frac{1}{2}tr[zz^T k(q_{s^-})] \geq$$
$$\frac{1}{2}z^T W_{qq}^+(q_{s^-})z,$$

with equality if the diffusion terms are non-zero in that direction. Note that the bound that the optimal policies satisfy implies that $W_{qq}^+(q_{s^-})$, interpreted in a distributional sense, is finite and hence that W^+ is differentiable.

Similarly, for any z such that $\psi_s^+(z) > 0$, the first-order condition requires that

$$\lambda\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z'')D^*(q_{s^-}+z''||q_{s^-})dz''\}^{\rho-1}D^*(q_{s^-}+z||q_{s^-}) =$$
$$W^+(q_{s^-}+x,\lambda) - W^+(q_{s^-},\lambda) - x^T \cdot W_q^+(q_{s^-},\lambda), \tag{28}$$

where the differentiability of W^+ in the continuation region follows from the envelope theorem.

Combining these two first order conditions, consider a perturbation that decreases $\sigma_s\sigma_s^T$ by ϵzz^T and increases $\psi_s(\nu z)$ and $\psi_s(-\nu z)$ by $\frac{1}{2}\frac{\epsilon}{\nu^2}$. The first-order conditions for this perturbation is

$$\lambda\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z'')D^*(q_{s^-}+z''||q_{s^-})dz''\}^{\rho-1}\times$$
$$\{\frac{1}{2\nu^2}D^*(q_{s^-}+\nu z||q_{s^-}) + \frac{1}{2\nu^2}D^*(q_{s^-}-\nu z||q_{s^-}) - \frac{1}{2}tr[zz^T k(q_{s^-})]\} =$$
$$\frac{1}{2\nu^2}(W^+(q_{s^-}+\nu z,\lambda) + W^+(q_{s^-}-\nu z,\lambda) - 2W^+(q_{s^-},\lambda)) - \frac{1}{2}z^T W_{qq}^+(q_{s^-})z.$$

In the limit as $\nu \to 0^+$, this equation is always satisfied, and therefore it is without loss of generality to suppose that the diffusion term is zero.

Lastly, if there exists an z, z' with $\psi_s^+(z) > 0$ and $\psi_s^+(z') > 0$, an alternative policy that sets $\tilde{\psi}_s^+(z) = \psi_s^+(z) + \frac{D^*(q_{s-}+z'||q_{s-})}{D^*(q_{s-}+z||q_{s-})}\psi_s^+(z')$ and $\tilde{\psi}_s^+(z') = 0$ generates the same cost, and changes utility by

$$\frac{D^*(q_{s^-}+z'||q_{s^-})}{D^*(q_{s^-}+z||q_{s^-})}(W^+(q_{s^-}+z,\lambda) - W^+(q_{s^-},\lambda) - z^T \cdot W_q^+(q_{s^-},\lambda))\psi_s^+(z') - \\ (W^+(q_{s^-}+z',\lambda) - W^+(q_{s^-},\lambda) - z'^T \cdot W_q^+(q_{s^-},\lambda))\psi_s^+(z') = 0.$$

It follows that it is without loss of generality to assume that $\psi_s^+(z) > 0$ for at most one value of z. Recalling that the optimal policies are Markov, let $\sigma^+(q_s)$ denote the optimal policy for the diffusion, let $\bar{\psi}^+(q)$ denote the optimal jump intensity, and let $z^+(q)$ denote the Markov optimal jump direction. Any semi-martingale with these characteristics generates a law that is identical to the jump-diffusion process described in Lemma 15.

Noting that $W^+(q,\lambda) \geq W(q,\lambda)$, it follows that if there exists a sequence of policies that converge to the stochastic process q_t^+, characterized by σ^+, $\bar{\psi}^+$, z^+, and whose cumulative information costs $\Delta_n^{-1}C(\cdot)$ converge to the total information costs in definition 1, then such a sequence of policies achieves, in the limit, at least as much utility as any other sequence of policies. It would then be the case that there must be sequence of optimal policies that converges a.s. (as in Lemma 15) to some optimal policy of W^+ (not necessarily the policies that generate q_t^+). Note also by the result above that it is without loss of generality to suppose $\sigma^+ = 0$.

We can rewrite our controls in terms of the jump destination, $q^+(q_s) = q_s + z^+(q_s)$. To construct such a sequence of convergent policies, consider the "constant control" described in chapter 13.2 of Kushner and Dupuis [2013] ("constant controls", in this context, being a constant $q^+, \bar{\psi}_t^+$ pair over the interval $[t, t+\Delta_n)$, switching to $\psi_t = 0$ after the first jump). By theorem 2.3 of that chapter, there exists a sequence of constant controls that converge (weakly) to the optimal policies of W^+. Moreover, these controls result, of the intervals $[t, t+\Delta_n)$, in a two-point distribution, with support on $q^+(q_t)$ and $q_t - \bar{\psi}_t^+(q^+(q_t) - q_t)\Delta_n$ for the left limit of the process at time $t + \Delta_n$.

Define the constant

$$\theta^+ = \frac{E_t[\int_t^\tau \beta^{(s-t)}\bar{\psi}_s^+(q_{s^-})D^*(q_{s^-}+z^+(q_{s^-})||q_{s^-})ds]}{E_t[\frac{1-\beta^{(\tau-t)}}{-\ln(\beta)}]}.$$

Now consider a modification of these constant control policies, which scale the intensity $\bar{\psi}_t^+$ by the quantity $\alpha_n(q_t)$, so that, for the modified policy,

$$\Delta_n C(\cdot) = \theta^+.$$

By the first-order condition with respect to $\bar{\psi}_t$, and the Bellman equation,

$$\kappa - \lambda c^\rho - \ln(\beta) W^+(q_t, \lambda) = \lambda(1 - \frac{1}{\rho})(\bar{\psi}_t^+(q_{t^-}) D^*(q_{t^-} + z^+(q_{t^-})||q_{t^-}))^\rho.$$

By the convexity of $C(\cdot)$,

$$C(\cdot) \geq \alpha_n(q_t)\bar{\psi}_t^+(q_{t^-}) D^*(q_{t^-} + z^+(q_{t^-})||q_{t^-}).$$

Observe that the lower bound on the value function that

$$\kappa - \lambda c^\rho - \ln(\beta) W^+(q_t, \lambda) > 0$$

for all q_t. It follows that

$$\alpha_n(q_t) \in [0, \frac{\theta^+}{\kappa - \lambda c^\rho - \ln(\beta)\underline{\mathrm{W}}}],$$

where $\underline{\mathrm{W}} = \min_{q \in \mathcal{P}(X)} W^+(q, \lambda)$, and that

$$\lim_{n \to \infty} \alpha_n(q_t) = 1.$$

Therefore, by this uniform bound, the modified policies converge weakly to the same limit as the constant control policies, and hence to an optimal policy of W^+. Moreover, by construction, the costs converge, and hence the dual value function W^+ is achievable and therefore $W(q, \lambda) = W^+(q, \lambda)$.

We next demonstrate equality of the primal and dual. The associated Bellman equation for the dual value function W^+, in the continuation region, is

$$\begin{aligned} -\ln(\beta) W^+(q_s, \lambda) = \max_{\sigma_s, \psi_s} E[dW^+(q_s, \lambda)] - (\kappa - \lambda c^\rho) ds \\ - \frac{\lambda}{\rho}\{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(z) D^*(q_{s^-} + z||q_{s^-}) dz\}^\rho. \end{aligned}$$

Consider a perturbation which scales $\sigma_s^+ \sigma_s^{+T}$ and ψ_s^+ be some constant $(1 + \epsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1 + \epsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $-\ln(\beta) W^+(q_s, \lambda) + \kappa - \lambda c^\rho > 0$. The first order condition for this perturbation is

$$\begin{aligned} -\ln(\beta) W^+(q_s, \lambda) + (\kappa - \lambda c^\rho) + \frac{\lambda}{\rho}\{\frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \\ \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(z) D^*(q_{s^-} + z||q_{s^-}) dz\}^\rho = \\ \lambda\{\frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(z) D^*(q_{s^-} + z||q_{s^-}) dz\}^\rho, \end{aligned}$$

which must hold at the optimal policies for this problem. We can rewrite the Bellman equation as

$$-\ln(\beta)W^{+}(q_s,\lambda)ds + (\kappa - \lambda c^{\rho})ds =$$
$$E[dW^{+}(q_s,\lambda)] - \frac{\lambda}{\rho}(\rho\frac{rW^{+}(q_s,\lambda) + (\kappa - \lambda c^{\rho})}{\lambda(\rho-1)})ds,$$

or

$$(-\ln(\beta)W^{+}(q_s,\lambda) + (\kappa - \lambda c^{\rho}))\frac{\rho}{\rho-1} = E[dW^{+}(q_s,\lambda)].$$

Solving this equation,

$$W^{+}(q_s,\lambda) = E_s[\beta^{\frac{\rho}{\rho-1}(\tau^{*}-s)}\hat{u}(q_{\tau^{*}}) - \frac{\rho}{\rho-1}(\kappa - \lambda c^{\rho})\int_s^{\tau^{*}} \beta^{\frac{\rho}{\rho-1}(l-s)}dl].$$

Define λ^{*} by

$$E_s[\beta^{\frac{\rho}{\rho-1}(\tau^{*}-s)}\hat{u}(q_{\tau^{*}}) - \frac{\rho}{\rho-1}(\kappa - \lambda^{*} c^{\rho})\int_s^{\tau^{*}} \beta^{\frac{\rho}{\rho-1}(l-s)}dl] =$$
$$E_0[\beta^{(\tau^{*}-s)}\hat{u}(q_{\tau^{*}}) - \kappa\int_s^{\tau^{*}} \beta^{(l-s)}dl].$$

We can rewrite this as

$$(\frac{1}{\rho-1}\kappa - \frac{\rho}{\rho-1}\lambda^{*}c^{\rho})E_0[\int_s^{\tau^{*}} \beta^{\frac{\rho}{\rho-1}(l-s)}dl] =$$
$$E_0[\beta^{\frac{\rho}{\rho-1}(\tau^{*}-s)}\hat{u}(q_{\tau^{*}}) - \kappa\int_s^{\tau^{*}} \beta^{\frac{\rho}{\rho-1}(l-s)}dl] -$$
$$E_0[\beta^{(\tau^{*}-s)}\hat{u}(q_{\tau^{*}}) - \kappa\int_0^{\tau^{*}} \beta^{(l-s)}dl].$$

The right-hand side is weakly negative, and zero if $\beta = 1$. Consequently, $\lambda^{*} > 0$, and $\lambda^{*} = \frac{\kappa}{\rho c^{\rho}} < \kappa c^{-\rho}$ if $\beta = 1$.

Consider a convergent sub-sequence of $V(q_0;\Delta_n)$ (which exists by the uniform boundedness and convexity of the problem), and denote its limit $V(q_0)$ (again, we will index this sequence by n). By the standard duality inequalities, for all λ,

$$V(q_0;\Delta_n) \leq W(q_0,\lambda;\Delta_n),$$

for all n, and therefore

$$V(q_0) \leq W^{+}(q_0,\lambda^{*}).$$

Consider the value function $\tilde{V}(q_0)$, which is the value function under the feasible optimal policies for $W^{+}(q_0,\lambda^{*})$. It follows that $\tilde{V}(q_0) = W(q_0,\lambda^{*})$, and $\tilde{V}(q_0) \leq V(q_0)$, and therefore $V(q_0) = W(q_0,\lambda^{*})$.

Note that every convergent sub-sequence of $V(q_0;\Delta_n)$ converges to the same function. It follows that

$$V(q_0) = \lim_{\Delta \to 0^+} V(q_0; \Delta).$$
$$= E_0[\beta^{\tau^*} \hat{u}(q_{\tau^*}) - \kappa \int_0^{\tau^*} \beta^l dl].$$

By the definition of λ^* and the Bellman equation,

$$E_0[\int_0^{\tau^*} \beta^s \frac{1}{\rho} \{\frac{1}{2} tr[\sigma_s^* \sigma_s^{*T} k(q_{s-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^*(z) D^*(q_{s^-} + z || q_{s-}) dz\}^\rho ds] \leq$$
$$c^\rho E_0[\int_0^{\tau^*} \beta^l dl],$$

as required. It follows that the value function is the maximized over all policies satisfying the above constraint (which is the limiting constraint, by the dominated convergence theorem), concluding the proof.

B.7 Proof of Lemma 13

Write the value function in sequence-problem form, for the $\beta < 1$ case:

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

Define

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

By the weak positivity of the cost function $C(\cdot)$, it follows that

$$W(q_0, \lambda; \Delta) \leq \bar{u} + \Delta E_0[\frac{1-\beta^\tau}{1-\beta^\Delta}](\lambda c^\rho - \kappa).$$

If $\lambda \in [0, \kappa c^{-\rho}]$, the value function is bounded above by $\bar{u}$. If $\lambda > \kappa c^{-\rho}$,

$$W(q_0, \lambda; \Delta) \leq \bar{u} + \frac{\Delta}{1-\beta^\Delta}(\lambda c^\rho - \kappa),$$

and

$$1 - \beta^\Delta > \frac{-\Delta \ln(\beta)}{1 - \Delta \ln(\beta)},$$

implying

$$\frac{\Delta}{1-\beta^{\Delta}} < \frac{1-\ln(\beta)}{-\ln(\beta)}$$

for all $\Delta \le 1$. Therefore,

$$\bar{W} = \bar{u} + \frac{1-\ln(\beta)}{-\ln(\beta)} \max\{\lambda c^{\rho} - \kappa, 0\}.$$

It follows immediately that

$$\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W} = \begin{cases} -\ln(\beta)\bar{u} + \kappa - \lambda c^{\rho} & \kappa \ge \lambda c^{\rho} \\ -\ln(\beta)\bar{u} & \kappa < \lambda c^{\rho}, \end{cases}$$

and therefore

$$\kappa - \lambda c^{\rho} - \ln(\beta)\bar{W} > 0.$$

For the $\beta = 1$ case, by the assumption that $\lambda c^{\rho} \le \kappa$, $W(q_0, \lambda; \Delta) \le \bar{u} = \bar{W}$, and the result holds immediately.

There is a smallest possible decision utility which is strictly positive, and because stopping now and deciding is always feasible,

$$W(q_0, \lambda; \Delta) \ge 0.$$

We can define the "state-specific" value function, $W(q_t, \lambda; \Delta, x)$, which is the value function conditional on the true state being x. The state-specific value function has a recursive representation, in the region in which the DM continues to gather information:

$$\begin{aligned} W(q_t, \lambda; \Delta, x) \quad &= \quad -\kappa\Delta + \lambda\Delta^{1-\rho}(\Delta^{\rho} c^{\rho} - \frac{1}{\rho} C(\cdot)^{\rho}) + \\ & \quad \beta^{\Delta} \sum_{s \in S:\, e_s^T p_t e_x > 0} (e_s^T p_t^* e_x) W(q_{t+\Delta,s}^*, \lambda; \Delta, x). \end{aligned}$$

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

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

By the optimality of the policies, we have

$$W(q_t, \lambda; \Delta) \ge \sum_{x \in X} q_{t,x} W(q', \lambda; \Delta, x),$$

for any q' in $\mathcal{P}(X)$. Suppose not; then the DM could simply adopt the information structure associated with beliefs q' and achieve higher utility, contradicting the optimality of the policy.

The convexity of the value function follows from the observation that

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

and using the inequality above,

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

B.8 Proof of Lemma 14

Consider an alternative policy that mixes (in the sense of Condition 2) the optimal signal structure and an uninformative signal, with probabilities $1-a$ and a, respectively. We must have

$$-\beta^{\Delta_n} \sum_{s \in S} (e_s^T r_{t,n}^*)(W(q_{t,n,s}^*, \lambda; \Delta_n) - W(q_{t,n}, \lambda; \Delta_n)) - \lambda \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^{\rho-1} \frac{\partial C(p_{t,n}(a), q_{t,n})}{\partial a}|_{a=0^+} \leq 0,$$

which is the first-order condition at the optimal policy in the direction of adding a little bit of the uninformative signal (decreasing a). By the convexity of $C(\cdot)$ and Condition 1,

$$C(p_{t,n}^*, q_{t,n}) + \frac{\partial C(p_{t,n}(a), q_{t,n})}{\partial a}|_{a=0^+} \leq 0,$$

and therefore we must have

$$\beta^{\Delta_n} \sum_{s \in S} (e_s^T r_{t,n}^*)(W(q_{t,n,s}^*, \lambda; \Delta_n) - W(q_{t,n}, \lambda; \Delta_n)) \geq \lambda \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^{\rho}.$$

Applying the Bellman equation in the continuation region,

$$(1-\beta^{\Delta_n})W(q_{t,n}, \lambda; \Delta_n) + (\kappa - \lambda c^{\rho})\Delta_n + \frac{\lambda}{\rho}\Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^{\rho} \geq \lambda \Delta_n^{1-\rho} C(p_{t,n}^*, q_{t,n})^{\rho}.$$

Therefore,

$$\lambda(1-\frac{1}{\rho})\Delta_n^{-\rho} C(p_{t,n}^*, q_{t,n})^{\rho} \leq (\kappa - \lambda c^{\rho}) + \frac{(1-\beta^{\Delta_n})}{\Delta_n} W(q_{t,n}, \lambda; \Delta_n).$$

If $\beta = 1$, then

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

for the constant $\theta = \lambda(\rho \frac{\kappa - \lambda c^{\rho}}{\lambda(\rho-1)})^{\frac{\rho-1}{\rho}} > 0$.

If $\beta < 1$, note that

$$\frac{(1-\beta^{\Delta_n})}{\Delta_n} < -\ln(\beta).$$

Let $\bar{W}$ denote the upper bound on $W(q_{t,n}, \lambda; \Delta_n)$, which exists by Lemma 13. We have

$$C(p^*_{t,n}, q_{t,n}) \leq \Delta_n (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}},$$

where

$$\theta = \lambda(\rho \frac{\kappa - \lambda c^\rho - \ln(\beta)\bar{W}}{\lambda(\rho-1)})^{\frac{\rho-1}{\rho}}.$$

The constant θ is positive by (13). Note that this generalizes the formula of the $\beta = 1$ case.

B.9 Proof of Lemma 15

We begin by discussing the convergence of stopping times. Let $\bar{W}$ denote the upper bound on $W(q_{t,n}, \lambda; \Delta_n)$, which exists by Lemma 13. Suppose that under an optimal policy,

$$\lim_{T\to\infty} Pr\{\tau_n < T\} = 1 - \alpha < 1.$$

The value function at time T must be bounded above by

$$W(q_T, \lambda; \Delta) \leq (1 - \alpha)\bar{W},$$

as the payoff conditional on never stopping is negative. Now consider an alternative policy that follows the optimal policy until time T, and then stops. The difference in the initial value functions is bounded above by the possibility of making the best possible decision under the optimal policy vs. the worst possible decision under the alternative policy, with utility $\underline{\mathrm{u}} > 0$:

$$(1 - \alpha - Pr\{\tau_n < T\})\beta^T \bar{W} \geq (1 - Pr\{\tau_n < T\})\beta^T \underline{\mathrm{u}}.$$

This inequality cannot hold in the limit as $T \to \infty$. Therefore, by the positivity of τ_n, the laws of τ_n are tight, and therefore there exists a sub-sequence that converges in measure. Pass to this sub-sequence (which we will also index by n), and let τ denote the limit of this sub-sequence.

The beliefs $q_{t,n}$ are a family of $\mathbb{R}^{|X|}$-valued stochastic processes, with $q_{t,n} \in \mathcal{P}(X)$ for all $t \in [0, \infty)$ and $n \in \mathbb{N}$. Construct them as RCLL processes by assuming that $q_{\Delta_n j+\epsilon,n} = q_{\Delta_n j,n}$ for all m, $\epsilon \in [0, \Delta_n)$, and $j \in \mathbb{N}$. We next establish that the laws of $q_{t,n}$ are tight. By Condition 5 and Lemma 14,

$$\frac{m}{2} \sum_{s\in S} (e_s^T p_n(q_{t,n}) q_{t,n}) ||q_{s,n}(q_{t,n}) - q_{t,n}||_2^2 \leq C(p_n(q_{t,n}), q_{t,n}; S) \leq \Delta_n (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}},$$

where $q_{s,n}(q)$ is defined by $p_n(q)$ and Bayes' rule. It follows that, for any $\epsilon > 0$, there exists an N_ϵ such that, for all $n > N_\epsilon$,

$$P(||q_{t+\Delta_n,n} - q_{t,n}|| > \epsilon) \leq K_\epsilon \Delta_n,$$

for the constant $K_\epsilon = 2m^{-1}\epsilon^{-2}\theta^{\frac{1}{\rho-1}}$. By Theorem 3.21 in chapter 6 of Jacod and Shiryaev [2013], and the boundedness of $q_{t,n}$, it follows that the laws of $q_{t,n}$ are tight. By Prokhorov's theorem (Theorem 3.9 in chapter 6 of Jacod and Shiryaev [2013]), it follows that there exists a convergent sub-sequence. Pass to this sub-sequence, and let q_t denote the limiting stochastic process. By Proposition 1.1 in chapter 9 of Jacod and Shiryaev [2013], q_t is a martingale with respect to the filtration it generates. By Skorohod's representation theorem, there exists a probability space and random variables (which we will also denote with $q_{t,n}$ and q_t) such the convergence is almost sure. We refer to this probability space and these random variables in what follows.

Note that, by Bayes' rule, if $e_x^T q_{t,n} = 0$ for some $x \in X$ and time t, then $e_x^T q_{s,n} = 0$ for all $s > t$. By Proposition 2.9 and Corollary 2.38 in chapter 2 of Jacod and Shiryaev [2013], we can write the "good" version of the martingale with characteristics

$$B = -\int_0^t (\int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) z dz) dA_s$$

$$C = \int_0^t \Sigma_s dA_s$$

$$\nu = dA_s \psi_s(x).$$

Because beliefs remain in the simplex, $\psi_s(z)$ has support only on z such that $q_s + z \in \mathcal{P}(X)$ and $q_s + z \ll q_s$. Relatedly, $\iota^T \Sigma_s = 0$, and Σ_s can be decomposed as $\Sigma_s = D(q_{s^-})\sigma_s \sigma_s^T D(q_{s^-})$.

By the convexity of the cost function and Theorem 1,

$$C(p_n(q_{t,n}), q_{t,n}; S) \geq \sum_{s \in S} (e_s^T p_n(q_{t,n}) q_{t,n}) D^*(q_{s,n}(q_{t,n}) || q_{t,n}).$$

Defining the process, for arbitrary stopping time T,

$$D_{s,n} = \lim_{\epsilon \to 0^+} D^*(q_{s^- + \epsilon, n} || q_{s^-,n})$$

and

$$D_{t,T,n} = E_t[\int_t^T \beta^{\Delta_n \lfloor \Delta_n^{-1}(s-t) \rfloor} D_{s,n} ds] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \Delta_n E_t[\sum_{j=0}^{\lfloor \Delta_n^{-1}(s-t) \rfloor} \beta^{j\Delta_n}],$$

we have by Ito's lemma, almost sure convergence, and the dominated convergence theorem,

$$D_{t,T} = \lim_{n \to \infty} D_{t,T,n} = E_t[\int_t^T \beta^{s-t} \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z || q_{s^-}) dz\} dA_s].$$

Hence, for all such stopping times T,

$$E_t[\int_t^T \beta^{s-t} \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z || q_{s^-}) dz\} dA_s] \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}} \frac{1 - \beta^{T-t}}{-\ln(\beta)}.$$

Note also by this argument that

$$\lim_{n\to\infty} E_0[\Delta_n^{1-\rho} \sum_{j=0}^{\tau_n \Delta_n^{-1}-1} \beta^{\Delta_n j} C(p_n(q_{\Delta_n j,n}), q_{\Delta_n j,n}; S)^\rho]$$
$$= \lim_{n\to\infty} E_0[\int_0^{\tau_n} \beta^{\Delta_n \lfloor \Delta_n^{-1} t \rfloor} \Delta_n^{-\rho} C(p_n(q_{t,n}), q_{t,n}; S)^\rho dt]$$
$$\geq E_t[\int_0^\tau \beta^s \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z || q_{s^-}) dz\}^\rho (\frac{dA_s}{ds})^\rho ds]$$

B.10 Proof of Corollary 1

With $\beta = 1$, the associated Bellman equation for the candidate value function W^+, in the continuation region, is

$$0 = \max_{\sigma_s, \psi_s} E[dW^+(q_s, \lambda)] - (\kappa - \lambda c^\rho) ds$$
$$- \frac{\lambda}{\rho} \{\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_s)] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z || q_{s^-}) dz\}^\rho.$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem. Consider a perturbation which scales $\sigma_s^+ \sigma_s^{+T}$ and ψ_s^+ be some constant $(1+\epsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1+\epsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $\kappa - \lambda c^\rho > 0$. The first order condition for this perturbation is

$$(\kappa - \lambda c^\rho) + \frac{\lambda}{\rho} \{\frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z) D^*(q_{s^-} + z || q_{s^-}) dz\}^\rho =$$
$$\lambda \{\frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z) D^*(q_{s^-} + z || q_{s^-}) dz\}^\rho,$$

which must hold at the optimal policies for this problem. It follows by the definition of θ in the $\beta = 1$ case (see the proof of Lemma 14),

$$\theta = \lambda(\rho \frac{\kappa - \lambda c^\rho}{\lambda(\rho - 1)})^{\frac{\rho-1}{\rho}},$$

that the constraint

$$\frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \bar{\psi}_s D^*(q_{s^-} + z_s || q_{s^-}) \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}$$

binds with equality everywhere, where we have used the result in the proof of Theorem 2 that it is without loss of generality to suppose $\psi_s(z)$ has support on at most one value of z, which we denote z_s.

Consequently, the Bellman equation can be rewritten as

$$\max_{\sigma_s, \bar{\psi}_s, z_s} E[dW^+(q_s, \lambda)] - (\kappa - \lambda c^\rho) ds - \frac{\lambda}{\rho} (\frac{\theta}{\lambda})^{\frac{\rho}{\rho-1}} ds$$

subject to

$$\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \bar{\psi}_s D^*(q_{s^-} + z_s||q_{s^-}) \leq (\frac{\theta}{\lambda})^{\frac{1}{\rho-1}}.$$

Defining

$$\chi(\lambda) = (\rho\frac{\kappa - \lambda c^\rho}{\lambda(\rho-1)})^{\frac{1}{\rho}}$$

and observing that

$$\begin{aligned}\kappa - \lambda c^\rho + \frac{\lambda}{\rho}(\frac{\theta}{\lambda})^{\frac{\rho}{\rho-1}} &= \kappa - \lambda c^\rho + \frac{\kappa - \lambda c^\rho}{\rho - 1} \\ &= (\kappa - \lambda c^\rho)\frac{\rho}{\rho-1},\end{aligned}$$

the result follows, noting from the proof of Theorem 2 that $\lambda^* = \frac{1}{\rho}\kappa c^{-\rho}$ when $\beta = 1$, and therefore

$$\chi(\lambda^*) = c\rho^{\frac{1}{\rho}}.$$

B.11 Proof of Lemma 2

Consider a two-signal alphabet, $s \in \{s_1, s_2\}$, with $\pi_{s_1} = \pi_{s_2}$, and $q_{s_1} = (1+\epsilon)q' - \epsilon q$ and $q_{s_2} = (1-\epsilon)q' + \epsilon q$. Applying the "chain rule" inequality,

$$\begin{aligned}D^*(q'||q) + \frac{1}{2}D^*(q' + \epsilon(q'-q)||q') + \frac{1}{2}D^*(q' - \epsilon(q'-q)||q') \\ \leq \frac{1}{2}D^*(q' + \epsilon(q'-q)||q) + \frac{1}{2}D^*(q' - \epsilon(q'-q)||q).\end{aligned}$$

Dividing by ϵ^2 and taking the limit as $\epsilon \to 0^+$,

$$(q'-q)^T \cdot \bar{k}(q') \cdot (q'-q) \leq \frac{d^2}{d\epsilon^2}D^*(q' + \epsilon(q'-q)||q)|_{\epsilon=0}.$$

Since this must hold for all $q' \ll q$, it holds for $q' = q + t(q''-q)$, with some arbitrary $q'' \ll q$ and $t \in [0,1]$. Therefore,

$$\frac{d^2}{dt^2}D^*(q + t(q''-q)||q)|_{t=0} \geq (q''-q)^T \cdot \bar{k}(q + t(q''-q)) \cdot (q''-q).$$

Integrating,

$$D^*(q''||q) \geq \int_0^1 \int_0^s (q''-q)^T \cdot \bar{k}(q + t(q''-q)) \cdot (q''-q)dtds,$$

which is

$$D^*(q''||q) \geq \int_0^1 (1-t)(q''-q)^T \cdot \bar{k}(q + t(q''-q)) \cdot (q''-q)dt.$$

B.12 Proof of Theorem 3

Conjecture that $\lambda \in (0, \kappa c^{-\rho})$. Under this conjecture, lemmas 13, 14, 15, and 2 apply.

Consider a possibly sub-optimal policy which sets $\psi_s(z) = 0$ for all z and satisfies the constraint. The above FOC applies, and therefore we must have

$$tr[\tilde{\sigma}_s \tilde{\sigma}_s^T (D(q_{s^-}) W_{qq}^+(q_{s^-}, \lambda) D(q_{s^-}) - \theta k(q_{s^-}))] \leq 0,$$

where W_{qq}^+ is understood in a distributional sense. It follows that, for all feasible z,

$$W^+(q_{s^-} + z, \lambda) - W^+(q_{s^-}, \lambda) - z^T W_q^+(q_{s^-}, \lambda) \leq \int_0^1 \int_0^s z^T \bar{k}(q_{s^-} + lz) z dl ds.$$

By our assumption of gradual learning (definition 3), this implies that

$$W^+(q_{s^-} + z, \lambda) - W^+(q_{s^-}, \lambda) - z^T W_q^+(q_{s^-}, \lambda) \leq \theta D^*(q_{s^-} + z || q_{s^-}).$$

Hence, it is without loss of generality to assume that $\psi_s^+(z) = 0$ for all z. Note that, if there is a strict preference for gradual learning, the above inequality is strict for all non-zero z. As a result, in this case we must have $\psi_s^+(z) = 0$ for all z. Note also that our control problem involves direct control of the diffusion coefficients, and hence satisfies the standard requirements for the existence and uniqueness of a strong solution to the resulting SDE (Pham [2009] sections 1.3 and 3.2).

B.13 Proof of Theorem 4

The associated Bellman equation, in the continuation region, is (letting $W^+(q, \lambda)$ denote the continuous time value function of Definition 1, generalized to allow multiple jumps)

$$\begin{aligned} 0 = \max_{\sigma_s, \psi_s} & E[dW^+(q_s, \lambda)] + \ln(\beta) W^+(q_{s^-}, \lambda) ds - (\kappa - \lambda c^\rho) ds \\ & - \frac{\lambda}{\rho} \{ \frac{1}{2} tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s(z) D^*(q_{s^-} + z || q_{s^-}) dz \}^\rho ds. \end{aligned} \tag{29}$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem. Suppose that the constraint does not bind, and consider a perturbation which scales $\sigma_s^+ \sigma_s^{+T}$ and ψ_s^+ be some constant $(1 + \epsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1 + \epsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $-\ln(\beta) W^+(q_s, \lambda) + \kappa - \lambda c^\rho > 0$. The first order condition for this perturbation is

$$\begin{aligned} - \ln(\beta) W^+(q_{s^-}, \lambda) + (\kappa - \lambda c^\rho) + \frac{\lambda}{\rho} \{ \frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \\ \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(z) D^*(q_{s^-} + z || q_{s^-}) dz \}^\rho = \\ \lambda \{ \frac{1}{2} tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})] + \int_{\mathbb{R}^{|X|} \setminus \{0\}} \psi_s^+(z) D^*(q_{s^-} + z || q_{s^-}) dz \}^\rho, \end{aligned}$$

which must hold at the optimal policies for this problem.

Define

$$\theta(q_{s^-}) = \lambda\Big(\frac{-\ln(\beta)W^+(q_{s^-},\lambda) + (\kappa - \lambda c^\rho)}{\lambda(1-\frac{1}{\rho})}\Big)^{\frac{\rho-1}{\rho}}$$

Observe by Lemma 13 that $\theta(q_{s^-}) > 0$.

For any feasible z, define

$$\tilde{\theta}(q_{s^-}, z) = \min_{\alpha\in[0,1]} \theta(q_{s^-} + \alpha z)$$

and let $\alpha^*(q_{s^-}, z)$ denote the minimizer.

Consider a sub-optimal policy $\tilde{\sigma}_s$ which sets $\psi_s(z) = 0$ and satisfies

$$-\ln(\beta)W^+(q_{s^-} + \alpha^*(q_{s^-}, z)z, \lambda) + (\kappa - \lambda c^\rho) = \lambda(1 - \frac{1}{\rho})\{\frac{1}{2}tr[\sigma_s^+ \sigma_s^{+T} k(q_{s^-})]\}^\rho,$$

which is

$$\frac{1}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T k(q_{s^-})] = (\frac{\tilde{\theta}(q_{s^-}, z)}{\lambda})^{\frac{1}{\rho-1}}.$$

For such a policy, the Bellman equation must be an inequality,

$$\frac{1}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T(Diag(q_{s^-})W_{qq}^+(q_{s^-},\lambda)Diag(q_{s^-})]ds \leq$$
$$-\ln(\beta)W^+(q_{s^-},\lambda)ds + (\kappa - \lambda c^\rho)ds + \frac{\lambda}{\rho}\{\frac{1}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T k(q_{s^-})]\}^\rho ds,$$

where W_{qq}^+ is understood in a distributional sense. We simplify this expression to

$$\frac{1}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T(Diag(q_{s^-})W_{qq}^+(q_{s^-},\lambda)Diag(q_{s^-})] \leq -\ln(\beta)[W^+(q_{s^-},\lambda) - W^+(q_{s^-} + \alpha^*(q_{s^-}, z)z, \lambda)]$$
$$+ \frac{\tilde{\theta}(q_{s^-}, z)}{2}tr[\tilde{\sigma}_s\tilde{\sigma}_s^T k(q_{s^-})].$$

This inequality must hold for all $\tilde{\sigma}_s$ with optimal scale. It follows that, integrating along a line (which must lie in the continuation region) and using the positivity of the value function, that

$$W^+(q_{s^-} + z, \lambda) - W^+(q_{s^-}, \lambda) - z^T W_q^+(q_{s^-}, \lambda) \leq \tilde{\theta}(q_{s^-}, z)\int_0^1\int_0^s z^T\bar{k}(q_{s^-} + lz)z\,dl\,ds$$
$$- 2\ln(\beta)\int_0^1\int_0^t W^+(q_{s^-} + lz)\,dl\,dt,$$

where $W_q^+(q_{s^-}, \lambda)$, the derivative, exists by Theorem 2.

By the strong preference for gradual learning and the upper bound on utility,

$$W^+(q_{s^-} + z, \lambda) - W^+(q_{s^-}, \lambda) - z^T W_q^+(q_{s^-}, \lambda) - \tilde{\theta}(q_{s^-}, z)D^*(q_{s^-} + z||q_{s^-}) \leq \quad (30)$$
$$-\ln(\beta)\bar{u}||z||_2^2 - m||z||_2^{2+\delta}.$$

If a jump is optimal, we must have (by the first-order condition)

$$W^{+}(q_{s^-}+z,\lambda)-W^{+}(q_{s^-},\lambda)-z^{T}W^{+}_{q}(q_{s^-},\lambda)=$$
$$\lambda\{\psi^{+}_{s}(z)D^{*}(q_{s^-}+z||q_{s^-})\}^{\rho-1}D^{*}(q_{s^-}+z||q_{s^-}),$$

and

$$\lambda(1-\rho^{-1})\{\psi^{+}_{s}(z)D^{*}(q_{s^-}+z||q_{s^-})\}^{\rho}=-\ln(\beta)W^{+}(q_{s^-},\lambda)+(\kappa-\lambda c^{\rho}).$$

Therefore, by the monotone relationship between $\theta(q_{s^-})$ and $W^{+}(q_{s^-},\lambda)$,

$$\lambda\{\psi^{+}_{s}(x)D^{*}(q_{s^-}+z||q_{s^-})\}^{\rho-1}\geq\lambda\Big(\frac{-\ln(\beta)W^{+}(q_{s^-}+\alpha^{*}(q_{s^-},z)z,\lambda)+(\kappa-\lambda c^{\rho})}{\lambda(1-\rho^{-1})}\Big)^{\frac{\rho-1}{\rho}},$$

which implies

$$W^{+}(q_{s^-}+z,\lambda)-W^{+}(q_{s^-},\lambda)-z^{T}W^{+}_{q}(q_{s^-},\lambda)\geq\tilde{\theta}(q_{s^-},z)D^{*}(q_{s^-}+z||q_{s^-}).$$

Using equation (30) above,

$$m||z||_{2}^{\delta}\leq-\ln(\beta)\bar{u},$$

which is

$$||z||_{2}\leq\Big(-\frac{\bar{u}\ln(\beta)}{m}\Big)^{\delta^{-1}}.$$

Now suppose that the jump reduces the value function,

$$W^{+}(q_{s^-}+z,\lambda)\leq W^{+}(q_{s^-},\lambda).$$

Consider again a sub-optimal diffusion policy, but with (for all q)

$$-\ln(\beta)W^{+}(q,\lambda)+(\kappa-\lambda c^{\rho})=\lambda\Big(1-\frac{1}{\rho}\Big)\Big\{\frac{1}{2}tr[\sigma^{+}_{s}\sigma^{+T}_{s}k(q)]\Big\}^{\rho},$$

which is

$$\frac{1}{2}tr[\tilde{\sigma}_{s}\tilde{\sigma}^{T}_{s}k(q)]=\Big(\frac{\theta(q)}{\lambda}\Big)^{\frac{1}{\rho-1}}.$$

The Bellman inequality in this case simplifies to

$$\frac{1}{2}tr[\tilde{\sigma}_{s}\tilde{\sigma}^{T}_{s}(Diag(q)W^{+}_{qq}(q,\lambda)Diag(q)]\leq\frac{\theta(q)}{2}tr[\tilde{\sigma}_{s}\tilde{\sigma}^{T}_{s}k(q)].$$

Observe by Lemma 13 and Theorem 5 that W^{+} is the limit of a sequence of bounded and convex functions, and hence convex. By the convexity of W^{+}, for all $\alpha\in[0,1)$,

$$W^{+}(q_{s^-}+\alpha z,\lambda)\leq W^{+}(q_{s^-},\lambda),$$

and therefore (by the definition of $\theta(q)$) $\theta(q_{s^-}+\alpha z)<\theta(q_{s^-})$. Consequently, integrating, we have

$$W^+(q_{s^-}+z,\lambda)-W^+(q_{s^-},\lambda)-z^T W_q^+(q_{s^-},\lambda)\leq \theta(q_{s^-})\int_0^1\int_0^s z^T\bar{k}(q_{s^-}+lz)zdlds,$$

and by the definition of a strong (and hence strict) preference for gradual learning,

$$W^+(q_{s^-}+z,\lambda)-W^+(q_{s^-},\lambda)-z^T W_q^+(q_{s^-},\lambda)<\theta(q_{s^-})D^*(q_{s^-}+x||q_{s^-}).$$

However, if a jump downwards is optimal, we must have (as argued above)

$$W^+(q_{s^-}+z,\lambda)-W^+(q_{s^-},\lambda)-z^T W_q^+(q_{s^-},\lambda)=\theta(q_{s^-})D^*(q_{s^-}+z||q_{s^-}),$$

and therefore downwards jumps are never optimal.

B.14 Proof of Lemma 3

Suppose the cost function satisfies a preference for discrete learning. Consider a two-signal alphabet, $s\in\{s_1,s_2\}$, with $\pi_{s_1}=\pi_{s_2}$, and $q_{s_1}=(1+\epsilon)q'-\epsilon q$ and $q_{s_2}=(1-\epsilon)q'+\epsilon q$. Applying the "chain rule" inequality,

$$\begin{aligned}D^*(q'||q)&+\frac{1}{2}D^*(q'+\epsilon(q'-q)||q')+\frac{1}{2}D^*(q'-\epsilon(q'-q)||q')\\&\geq\frac{1}{2}D^*(q'+\epsilon(q'-q)||q)+\frac{1}{2}D^*(q'-\epsilon(q'-q)||q),\end{aligned}\tag{31}$$

strictly if the preference is strict and $q'\neq q$. Dividing by ϵ^2 and taking the limit as $\epsilon\to 0^+$,

$$(q'-q)^T\cdot\bar{k}(q')\cdot(q'-q)\geq\frac{d^2}{d\epsilon^2}D^*(q'+\epsilon(q'-q)||q)|_{\epsilon=0}.$$

Since this must hold for all $q'\ll q$, it holds for $q'=q+t(q''-q)$, with some arbitrary $q''\ll q$ and $t\in[0,1]$. Therefore,

$$\frac{d^2}{dt^2}D^*(q+t(q''-q)||q)|_{\epsilon=0}\leq(q''-q)^T\cdot\bar{k}(q+t(q''-q))\cdot(q''-q).$$

Integrating,

$$D^*(q''||q)\leq\int_0^1\int_0^s(q''-q)^T\cdot\bar{k}(q+t(q''-q))\cdot(q''-q)dtds,$$

which is

$$D^*(q''||q)\leq\int_0^1(1-t)(q''-q)^T\cdot\bar{k}(q+t(q''-q))\cdot(q''-q)dt.$$

It follows that equality in this equation must hold if the cost function satisfies both a preference for discrete learning and a preference for gradual learning. Consequently, a strict preference for gradual learning is incompatible with a preference for discrete learning. Moreover, in this case equation (31) cannot hold strictly, and therefore a strict preference for discrete learning implies no preference for gradual learning.

B.15 Proof of Theorem 5

The problem described in Corollary 1, using the fact that it is without loss of generality to assume a pure jump process, is

$$W^+(q_t,\lambda)=\sup_{\{\bar{\psi}_s,z_s\},\tau} E_t[\hat{u}(q_{\tau^*})-\tau\frac{\rho}{\rho-1}(\kappa-\lambda c^{\rho})]$$

subject to

$$\bar{\psi}_s D^*(q_{s^-}+z_s||q_{s^-})\leq\chi(\lambda).$$

Suppose that the theorem is false– that for some q_{t^-} and z_t^*, $q_{t^-}+z_t^*=q'$ is in the continuation region. The first-order condition (see equation (28) in the proof of Theorem 2, which proves differentiability) can be written as

$$W^+(q_{t^-}+z_t^*,\lambda)-W^+(q_{t^-},\lambda)-z_t^{*T}\cdot W_q^+(q_{t^-},\lambda)=\theta D^*(q'||q_{t^-}).$$

If q' is in the continuation region, there must be some $q''\ll q'$ such that

$$W^+(q'',\lambda)-W^+(q',\lambda)-(q''-q')^T\cdot W_q^+(q',\lambda)=\theta D^*(q''||q').$$

Adding these two equations together and re-arranging,

$$\begin{aligned}W^+(q'',\lambda)-W^+(q_{t^-},\lambda)-(q''-q_{t^-})^T\cdot W_q^+(q_{t^-},\lambda)=\\ \theta D^*(q'||q_{t^-})+\theta D^*(q''||q')+(q''-q')^T\cdot(W_q^+(q',\lambda)-W_q^+(q_{t^-},\lambda)).\end{aligned}$$

Observe that, because $q''\ll q'$, there exists an $\bar{\epsilon}>0$ such that, for all $\epsilon\in[0,\bar{\epsilon}]$, $q'-\frac{\epsilon}{1-\epsilon}(q''-q')$ remains in the the simplex. By the fact that z_t^* is optimal, for all such ϵ,

$$\begin{aligned}W^+(q',\lambda)-W^+(q_{t^-},\lambda)-(q'-q_{t^-})^T\cdot W_q^+(q_{t^-},\lambda)-\theta D^*(q'||q_{t^-})\geq\\ W^+(q'-\frac{\epsilon}{1-\epsilon}(q''-q'),\lambda)-W^+(q_{t^-},\lambda)-(q'-\frac{\epsilon}{1-\epsilon}(q''-q')-q_{t^-})^T\cdot W_q^+(q_{t^-},\lambda)-\\ \theta D^*(q'-\frac{\epsilon}{1-\epsilon}(q''-q')||q_{t^-}),\end{aligned}$$

or

$$\begin{aligned}\theta D^*(q'||q_{t^-})-\theta D^*(q'-\frac{\epsilon}{1-\epsilon}(q''-q')||q_{t^-})\leq\\ W^+(q',\lambda)-W^+(q'-\frac{\epsilon}{1-\epsilon}(q''-q'),\lambda)-\frac{\epsilon}{1-\epsilon}(q''-q')^T W_q^+(q_{t^-},\lambda).\end{aligned}$$

Now consider the chain rule inequality, supposing that $s\in\{s_1,s_2\}$ with $\pi_{s_1}=\epsilon$, $\pi_{s_2}=1-\epsilon$,

$q_{s_1} = q''$, and $q_{s_2} = q' - \frac{\epsilon}{1-\epsilon}(q'' - q')$,

$$D^*(q'||q_{t^-}) + \epsilon D^*(q''||q') + (1-\epsilon)D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q') \geq$$
$$\epsilon D^*(q''||q_{t^-}) + (1-\epsilon)D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q_{t^-}).$$

We have, for all $\epsilon \in [0, \bar{\epsilon}]$,

$$W^+(q', \lambda) - W^+(q' - \frac{\epsilon}{1-\epsilon}(q''-q'), \lambda) - \frac{\epsilon}{1-\epsilon}(q''-q')^T W_q^+(q_{t^-}, \lambda) +$$
$$\epsilon\theta D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q_{t^-}) + \epsilon\theta D^*(q''||q') \geq$$
$$\epsilon\theta D^*(q''||q_{t^-}) - (1-\epsilon)\theta D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q').$$

Dividing by ϵ and taking limits,

$$(q''-q')^T \cdot (W_q^+(q', \lambda) - W_q^+(q_{t^-}, \lambda)) +$$
$$\theta D^*(q'||q_{t^-}) + \theta D^*(q''||q') \geq \theta D^*(q''||q_{t^-}).$$

Consequently,

$$W^+(q'', \lambda) - W^+(q_{t^-}, \lambda) - (q'' - q_{t^-})^T \cdot W_q^+(q_{t^-}, \lambda) \geq \theta D^*(q''||q_{t^-}),$$

meaning that it is without loss of generality to suppose that beliefs jump directly to q'' instead of to q'. Therefore, it is without loss of generality to suppose beliefs jump directly to the stopping region.

B.16 Proof of Theorem 6

The associated Bellman equation, in the continuation region, is (letting $W^+(q, \lambda)$ denote the continuous time value function of Definition 1, generalized to allow multiple jumps)

$$0 = \max_{\sigma_s, \psi_s} E[dW^+(q_s, \lambda)] + \ln(\beta)W^+(q_{s^-}, \lambda)ds - (\kappa - \lambda c^\rho)ds$$
$$- \frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s \sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z) D^*(q_{s^-} + z||q_{s^-})dz\}^\rho ds.$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem. Suppose that the constraint does not bind, and consider a perturbation which scales $\sigma_s^+ \sigma_s^{+T}$ and ψ_s^+ be some constant $(1+\epsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1+\epsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $-\ln(\beta)W^+(q_s, \lambda) + \kappa - \lambda c^\rho > 0$. The first order condition for this perturbation is

$$-\ln(\beta)W^{+}(q_{s^-},\lambda)+(\kappa-\lambda c^{\rho})+\frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s^{+}\sigma_s^{+T}k(q_{s^-})]+ \int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^{+}(z)D^{*}(q_{s^-}+z||q_{s^-})dz\}^{\rho}=$$
$$\lambda\{\frac{1}{2}tr[\sigma_s^{+}\sigma_s^{+T}k(q_{s^-})]+\int_{\mathbb{R}^{|X|}\setminus\{0\}}\psi_s^{+}(z)D^{*}(q_{s^-}+z||q_{s^-})dz\}^{\rho},$$

which must hold at the optimal policies for this problem.

Define

$$\theta(q_{s^-})=\lambda(\frac{-\ln(\beta)W^{+}(q_{s^-},\lambda)+(\kappa-\lambda c^{\rho})}{\lambda(1-\frac{1}{\rho})})^{\frac{\rho-1}{\rho}}$$

and observe that it is strictly positive by Theorem 2.

If a jump is optimal, we must have (by the above first-order condition)

$$W^{+}(q_{s^-}+z_s^{*},\lambda)-W^{+}(q_{s^-},\lambda)-z_s^{*T}W_q^{+}(q_{s^-},\lambda)=\theta(q_{s^-})D^{*}(q_{s^-}+z_s^{*}||q_{s^-}),$$

where $W_q^{+}(q_{s^-},\lambda)$ is the derivative that exists by Theorem 2.

Suppose that for some q_{t^-} and z_t^{*}, $q_{t^-}+z_t^{*}=q'$ is in the continuation region and that $W^{+}(q',\lambda)\geq W^{+}(q_{t^-},\lambda)$. Then we have

$$\theta(q')\geq\theta(q_{t^-})$$

and, for some $q''\ll q'$,

$$W^{+}(q'',\lambda)-W^{+}(q',\lambda)-(q''-q')^{T}\cdot W_q^{+}(q',\lambda)=\theta(q')D^{*}(q''||q'),$$

and therefore

$$W^{+}(q'',\lambda)-W^{+}(q',\lambda)-(q''-q')^{T}\cdot W_q^{+}(q',\lambda)\geq\theta(q_{t^-})D^{*}(q''||q').$$

We also have the first order condition

$$W^{+}(q_{t^-}+z_t^{*},\lambda)-W^{+}(q_{t^-},\lambda)-z_t^{*T}\cdot W_q^{+}(q_{t^-},\lambda)=\theta(q_{t^-})D^{*}(q'||q_{t^-})$$

and, putting these two equations together,

$$W^{+}(q'',\lambda)-W^{+}(q_{t^-},\lambda)-(q''-q_{t^-})^{T}\cdot W_q^{+}(q_{t^-},\lambda)\geq$$
$$\theta(q_{t^-})D^{*}(q'||q_{t^-})+\theta(q_{t^-})D^{*}(q''||q')+(q''-q')^{T}\cdot(W_q^{+}(q',\lambda)-W_q^{+}(q_{t^-},\lambda)).$$

Observe that, because $q''\ll q'$, there exists an $\bar{\epsilon}>0$ such that, for all $\epsilon\in[0,\bar{\epsilon}]$, $q'-\frac{\epsilon}{1-\epsilon}(q''-q')$

remains in the the simplex. By the fact that z_t^* is optimal, for all such ϵ,

$$W^+(q',\lambda) - W^+(q_{t-},\lambda) - (q'-q_{t-})^T \cdot W_q^+(q_{t-},\lambda) - \theta(q_{t-})D^*(q'||q_{t-}) \geq$$
$$W^+(q' - \frac{\epsilon}{1-\epsilon}(q''-q'),\lambda) - W^+(q_{t-},\lambda) - (q' - \frac{\epsilon}{1-\epsilon}(q''-q') - q_{t-})^T \cdot W_q^+(q_{t-},\lambda) -$$
$$\theta(q_{t-})D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q_{t-}),$$

or

$$\theta(q_{t-})D^*(q'||q_{t-}) - \theta(q_{t-})D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q_{t-}) \leq$$
$$W^+(q',\lambda) - W^+(q' - \frac{\epsilon}{1-\epsilon}(q''-q'),\lambda) - \frac{\epsilon}{1-\epsilon}(q''-q')^T W_q^+(q_{t-},\lambda).$$

Now consider the chain rule inequality, supposing that $s \in \{s_1, s_2\}$ with $\pi_{s_1} = \epsilon$, $\pi_{s_2} = 1-\epsilon$, $q_{s_1} = q''$, and $q_{s_2} = q' - \frac{\epsilon}{1-\epsilon}(q''-q')$,

$$D^*(q'||q_{t-}) + \epsilon D^*(q''||q') + (1-\epsilon)D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q') \geq$$
$$\epsilon D^*(q''||q_{t-}) + (1-\epsilon)D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q_{t-}).$$

We have, for all $\epsilon \in [0,\bar{\epsilon}]$,

$$W^+(q',\lambda) - W^+(q' - \frac{\epsilon}{1-\epsilon}(q''-q'),\lambda) - \frac{\epsilon}{1-\epsilon}(q''-q')^T W_q^+(q_{t-},\lambda) +$$
$$\epsilon\theta(q_{t-})D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q_{t-}) + \epsilon\theta(q_{t-})D^*(q''||q') \geq$$
$$\epsilon\theta(q_{t-})D^*(q''||q_{t-}) - (1-\epsilon)\theta(q_{t-})D^*(q' - \frac{\epsilon}{1-\epsilon}(q''-q')||q').$$

Dividing by ϵ and taking limits,

$$(q''-q')^T \cdot (W_q^+(q',\lambda) - W_q^+(q_{t-},\lambda)) +$$
$$\theta(q_{t-})D^*(q'||q_{t-}) + \theta(q_{t-})D^*(q''||q') \geq \theta(q_{t-})D^*(q''||q_{t-}).$$

Consequently,

$$W^+(q'',\lambda) - W^+(q_{t-},\lambda) - (q''-q_{t-})^T \cdot W_q^+(q_{t-},\lambda) \geq \theta(q_{t-})D^*(q''||q_{t-}),$$

and therefore it is without loss of generality to suppose that beliefs jump directly to q'' instead of to q'. Note that this inequality is strict if $W^+(q',\lambda) > W^+(q_{t-},\lambda)$. Hence it follows that if beliefs jump in such a way that increases the value function, they must jump to the stopping region.

Observe by Lemma 13 and Theorem 2 that W^+ is the limit of a sequence of bounded and convex functions, and hence convex. It follows that if $W^+(q',\lambda) = W^+(q_{t-},\lambda)$, we must have (by the mean value theorem) $(q'-q_{t-})^T \cdot W_q^+(\alpha q_{t-} + (1-\alpha)q',\lambda) = 0$ for some $\alpha \in (0,1)$, and therefore by convexity $(q'-q_{t-})^T \cdot W_q^+(q'_{t-},\lambda) \leq 0$. If such a jump were optimal, we would require $\theta(q_{s-})D^*(q'||q_{t-}) \leq 0$, which cannot hold. Therefore, jumps to the same level of the value function do not occur.

Now suppose that

$$W^+(q', \lambda) < W^+(q_{t^-}, \lambda)$$

and therefore $\theta(q') < \theta(q_{t^-})$. Define

$$q'' = \alpha q' + (1-\alpha) q_{t^-}$$

for some $\alpha \in (0,1)$.

By the convexity of W^+, for all $\alpha \in [0,1)$, $W^+(q'', \lambda) < W^+(q_{t^-}, \lambda)$, and therefore

$$W^+(q', \lambda) - W^+(q'', \lambda) - (q' - q'')^T \cdot W_q^+(q'', \lambda) \geq \theta(q_{t^-}) D^*(q'||q'')$$

and

$$W^+(q', \lambda) - W^+(q'', \lambda) - (q' - q'')^T \cdot W_q^+(q'', \lambda) > \theta(q'') D^*(q'||q'').$$

Define $\bar{\psi}'$ by

$$\lambda(\bar{\psi}' D^*(q'||q''))^{\rho-1} D^*(q'||q'') = W^+(q', \lambda) - W^+(q'', \lambda) - (q' - q'')^T \cdot W_q^+(q', \lambda).$$

We have

$$\lambda(\bar{\psi}' D^*(q'||q''))^{\rho-1} > \lambda\left(\frac{-\ln(\beta) W^+(q'', \lambda) + (\kappa - \lambda c^\rho)}{\lambda(1-\frac{1}{\rho})}\right)^{\frac{\rho-1}{\rho}}$$

and therefore

$$-\ln(\beta) W^+(q'', \lambda) + (\kappa - \lambda c^\rho) < \lambda(1 - \frac{1}{\rho})(\bar{\psi}' D^*(q'||q''))^\rho,$$

which is

$$-\ln(\beta) W^+(q'', \lambda) + (\kappa - \lambda c^\rho) < -\frac{\lambda}{\rho}(\bar{\psi}' D^*(q'||q''))^\rho + \bar{\psi}'[W^+(q', \lambda) - W^+(q'', \lambda) - (q' - q'')^T \cdot W_q^+(q', \lambda)].$$

It follows that the policy $z' = (q' - q'')$ and $\bar{\psi}'$ violates the HJB equation at q'', and therefore jumps downward never occur.

Hence we conclude that only upward jumps in the value function occur, and only to the stopping region.

B.17 Proof of Theorem 7

The associated Bellman equation, in the continuation region, is (letting $W^{+}(q,\lambda)$ denote the continuous time value function of Definition 1, generalized to allow multiple jumps)

$$0 = \max_{\sigma_s,\psi_s} E[dW^{+}(q_s,\lambda)] + \ln(\beta)W^{+}(q_{s^-},\lambda)ds - (\kappa - \lambda c^{\rho})ds \\ - \frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s\sigma_s^T k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s(z)D^*(q_{s^-}+z||q_{s^-})dz\}^{\rho}ds.$$

Let σ_s^+ and ψ_s^+ denote optimal policies for this problem. Suppose that the constraint does not bind, and consider a perturbation which scales $\sigma_s^+\sigma_s^{+T}$ and ψ_s^+ be some constant $(1+\epsilon)$. Note that such a perturbation would also scale $E[dW^+]$ by $(1+\epsilon)$, and that at least one of σ_s^+ and ψ_s^+ must be non-zero by the assumption that $-\ln(\beta)W^{+}(q_s,\lambda)+\kappa-\lambda c^{\rho} > 0$. The first order condition for this perturbation is

$$-\ln(\beta)W^{+}(q_{s^-},\lambda) + (\kappa - \lambda c^{\rho}) + \frac{\lambda}{\rho}\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})] + \\ \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z)D^*(q_{s^-}+z||q_{s^-})dz\}^{\rho} = \\ \lambda\{\frac{1}{2}tr[\sigma_s^+\sigma_s^{+T}k(q_{s^-})] + \int_{\mathbb{R}^{|X|}\setminus\{0\}} \psi_s^+(z)D^*(q_{s^-}+z||q_{s^-})dz\}^{\rho},$$

which must hold at the optimal policies for this problem.

Define

$$\theta(q_{s^-}) = \lambda(\frac{-\ln(\beta)W^{+}(q_{s^-},\lambda)+(\kappa-\lambda c^{\rho})}{\lambda(1-\frac{1}{\rho})})^{\frac{\rho-1}{\rho}}$$

and observe that it is strictly positive by Theorem 2.

Because a jump is optimal, we must have (by the above first-order condition)

$$W^{+}(q_{s^-}+z_s^*,\lambda) - W^{+}(q_{s^-},\lambda) - z_s^{*T}W_q^{+}(q_{s^-},\lambda) = \theta(q_{s^-})D^*(q_{s^-}+z_s^*||q_{s^-}),$$

where $W_q^{+}(q_{s^-},\lambda)$ is the derivative that exists by Theorem 2, and for all feasible jumps,

$$W^{+}(q_{s^-}+z,\lambda) - W^{+}(q_{s^-},\lambda) - z^{T}W_q^{+}(q_{s^-},\lambda) \leq \theta(q_{s^-})D^*(q_{s^-}+z||q_{s^-}).$$

We begin by proving that a preference for gradual learning exists for two-signal alphabets, and assuming that all of the relevant elements of the simplex are interior. We then extend the result to prove the full preference for discrete learning.

Proof by contradiction: suppose there exists an interior $q, q', q_1, q_2 \in \mathcal{P}(X)$ and $\pi \in (0,1)$ such that

$$\pi q_1 + (1-\pi)q_2 = q'$$

and

$$D^*(q'|q) + \pi D^*(q_1|q') + (1-\pi)D^*(q_2||q') < \pi D^*(q_1|q) + (1-\pi)D^*(q_2||q).$$

Now suppose that there exists a utility function such that $z = q_1 - q$ and $z = q_2 - q$ are both optimal policies from q, and for which

$$W^+(q', \lambda) \leq W^+(q, \lambda).$$

Then we must have, for $i \in \{1, 2\}$,

$$W^+(q_i, \lambda) - W^+(q, \lambda) - (q_i - q)^T \cdot W_q^+(q, \lambda) = \theta(q) D^*(q_i || q),$$

$$W^+(q', \lambda) - W^+(q, \lambda) - (q' - q)^T \cdot W_q^+(q, \lambda) \leq \theta(q) D^*(q' || q),$$

$$W^+(q_i, \lambda) - W^+(q', \lambda) - (q_i - q')^T \cdot W_q^+(q', \lambda) \leq \theta(q') D^*(q_i || q') \leq \theta(q) D^*(q_i || q'),$$

where $\theta(q') \leq \theta(q)$ by the definition of $\theta(\cdot)$ and $W^+(q', \lambda) \leq W^+(q, \lambda)$. Putting these together,

$$\theta(q) D^*(q' || q) + \theta(q) D^*(q_i || q') - \theta(q) D^*(q_i || q) \geq -(q_i - q')^T \cdot [W_q^+(q', \lambda) - W_q^+(q, \lambda)].$$

It would follow in this case that

$$D^*(q' | q) + \pi D^*(q_1 | q') + (1 - \pi) D^*(q_2 || q') \geq \pi D^*(q_1 | q) + (1 - \pi) D^*(q_2 || q),$$

a contradiction. To prove the result, we construct such a utility function. Note that our construction below will assume there are three actions; when applying this proof to the case of $|X| = 2$, one of the actions will be redundant.

Define, for some $\mu = (0, 1)$, a q_3 such that

$$\mu q_3 + (1 - \mu) q' = q.$$

Note that such a q_3 exists by the assumption that q is in the interior of the simplex.

Let $v \in \mathbb{R}^{|X|}$ be a vector and let k_1, k_2, k_3, K be constants. Suppose there are three actions, and let their utilities satisfy

$$u_i \in \theta(q) \partial D^*(q_i || q) + v + \iota k_i,$$

where $\partial D^*(q_i || q)$ denotes the sub-gradient with respect to the first argument. This sub-gradient exists by the convexity of D^* in its first argument and the assumption that q_i is interior. Define

$$k_i = \theta(q) D^*(q_i || q) - q_i^T \cdot \theta(q) \partial D^*(q_i || q)] + K - q^T v$$

so that

$$\theta(q) D^*(q_i || q) = q_i^T \cdot u_i - K + (q - q_i)^T v.$$

Observe by convexity that

$$\begin{aligned} q_i^T(u_i - u_j) = \theta(q)D^*(q_i||q) + K - (q - q_i)^T v \\ - \theta(q)D^*(q_j||q) - K + (q - q_j)^T v \\ - (q_i - q_j)^T \cdot u_j, \end{aligned}$$

and by the definition of the sub-gradient this yields

$$q_i^T(u_i - u_j) \geq 0.$$

By sub-optimality, for any $q'' \ll q$ and any $i \in \{1, 2, 3\}$,

$$\theta(q)D^*(q''||q) \geq (q'')^T u_i - W^+(q, \lambda) - (q'' - q)^T \cdot W_q^+(q, \lambda).$$

Therefore, for all $i \in \{1, 2, 3\}$,

$$(q_i - q)^T \cdot (W_q^+(q, \lambda) - v) \geq K - V(q)$$

Since this must hold for all q_i, we must have

$$W^+(q, \lambda) = K$$

and

$$(q_i - q)^T \cdot (W_q^+(q, \lambda) - v) = 0.$$

Hence it follows that

$$(q' - q)^T \cdot W_q^+(q, \lambda) = (q' - q)v.$$

By sub-optimality,

$$\theta(q)D^*(q'||q) + (q' - q)^T \cdot v \geq W^+(q', \lambda) - W^+(q, \lambda).$$

Setting

$$v \in -\theta(q)\partial D^*(q'||q)$$

ensures by convexity that $W^+(q, \lambda) \geq W^+(q', \lambda)$.

Observe that

$$\begin{aligned} \theta(q)D^*(q_i||q) &= q_i^T \cdot u_i - K + (q - q_i)^T v \\ &= q_i^T \cdot u_i - W^+(q', \lambda) + (q - q_i)^T W_q^+(q, \lambda) \end{aligned}$$

and therefore that jumps to the points q_i are optimal. Observe also that for any other q'', by the definition of the sub-gradient,

$$\theta(q)D^*(q''||q) - \theta(q)D^*(q_i||q) \geq (q'' - q_i)^T(u_i - v)$$

and therefore

$$\theta(q)D^*(q''||q) \geq (q'')^T u_i - W^+(q', \lambda) + (q - q'')^T W_q^+(q, \lambda)$$

as required. Therefore, the stationary policy of jumping to $\{q_1, q_2, q_3\}$ in proportions $\pi(1-\mu), (1-\pi)(1-\mu), \mu$ is optimal.

We conclude that for all interior pairs,

$$D^*(q'|q) + \pi D^*(q_1|q') + (1-\pi)D^*(q_2||q') \geq \pi D^*(q_1|q) + (1-\pi)D^*(q_2||q).$$

The result extends immediately to more than two $\{q_s\}$ by adding this expression for different pairs. The result extends to the boundary of the simplex by continuity.

B.18 Proof of Lemma 4

Recall the definition of a preference for discrete learning: for all $q, q', \{q_s\}_{s\in S}$ with $q' \ll q$ and $\sum_{s\in S} \pi_s q_s = q'$,

$$D^*(q'||q) + \sum_{s\in S} \pi_s D^*(q_s||q') \geq \sum_{s\in S} \pi_s D^*(q_s||q)$$

Therefore, for all $z \in \mathbb{R}^{|X|}$ with support on the support of q' and ϵ sufficiently small,

$$D^*(q'||q' + \epsilon z) + \sum_{s\in S} \pi_s D^*(q_s||q') \geq \sum_{s\in S} \pi_s D^*(q_s||q' + \epsilon z).$$

It follows immediately by the differentiability assumption that

$$\sum_{s\in S} \pi_s \frac{\partial}{\partial \epsilon} D^*(q_s||q' + \epsilon z)|_{\epsilon=0} = 0.$$

By step 1 in the proof of theorem 4 of Banerjee et al. [2005], it follows immediately that

$$D^*(q'||q) = H(q') - H(q) - (q' - q)^T H_q(q)$$

for some convex function H, where H_q denotes the gradient. Note that theorem 4 of Banerjee et al. [2005] is stated as requiring that

$$\sum_{s\in S} \pi_s D^*(q_s||q' + \epsilon z)$$

be minimized at $\epsilon = 0$ for all z, but step 1 of the proof in fact only requires that $\epsilon = 0$ correspond to a critical value for all z. Step 2 of the proof relaxes slightly the regularity conditions, but we have simply assumed these. Minimization is only required to establish the last step of the proof, step 3, which proves strict convexity of H. Strict convexity of $H(q)$ on the support of q follows in our setting immediately from the properties of the $k(q)$ matrix (Theorem 1).

B.19 Proof of Lemma 5

Define $M(q_t)$ as the set of $|X| \times |X|$ matrices such that, for all $\sigma \in M(q_t)$, $q_t^T \sigma = \vec{0}$.

In the continuation region, everywhere the value function is twice differentiable,

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

subject to

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

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

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

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

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

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

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

and $z^T k(q_t) z = 0$, but the null space of $k(q_t)$ consists only of vectors whose elements are constant over the support of q_t by Theorem 1, and therefore satisfy $q^T z \neq 0$, implying that $zz^T \notin M(q_t)$. Therefore, the constraint binds, and an optimal policy exists.

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

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

B.20 Proof of Theorem 8

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

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

and therefore converse of Euler's homogenous function theorem applies. That is, $H_q(q_t)$ is homogenous of degree zero, and $H(q_t)$ is homogeneous of degree one.

We start by showing that the function $\phi(q_t)$ is twice-differentiable in certain directions. Substituting the definition of the divergence into the statement of theorem,

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

subject to the same constraint. Define a new choice variable, $\hat{q}_a = \pi(a) q_a$. By definition, $\hat{q}_a \in \mathbb{R}_+^{|X|}$, and the constraint is $\sum_{a \in A} \hat{q}_a = q_0$. By the homogeneity of H, the objective is

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

Any choice of $\hat{q}_a$ satisfying the constraint can be implemented by some choice of π and q_a in the following way: set $\pi(a) = \iota^T \hat{q}_a$, and (if $\pi(a) > 0$) set

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

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

Rewriting the problem in Lagrangian form,

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

Observe that $\phi(q_0)$ is convex in q_0. Suppose not: for some $q = \lambda q_0 + (1-\lambda) q_1$, with $\lambda \in (0,1)$, $\phi(q) < \lambda \phi(q_0) + (1-\lambda)\phi(q_1)$. Consider a relaxed version of the problem in which the DM is allowed to choose two different $\hat{q}_a$ for each a. Because of the convexity of H, even with this option, the DM will set both of the $\hat{q}_a$ to the same value, and therefore the relaxed problem reaches the same value as the original problem. However, in the relaxed problem, choosing the optimal policies for q_0 and q_1 in the original problem, scaled by λ and $(1-\lambda)$ respectively, is feasible. It follows that $\phi(q) \geq \lambda \phi(q_0) + (1-\lambda)\phi(q_1)$. Note also that $\phi(q_0)$ is bounded on the interior of the simplex. It follows by Alexandrov's theorem that 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. Anywhere 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 $a \in A$ with $\hat{q}_a \neq \vec{0}$),

$$u_a - \theta H_q(\hat{q}_a) - \kappa + \nu_a = 0. \tag{32}$$

If $\hat{q}_a = \vec{0}$, we must have $q^T(u_a - \kappa) \leq \theta H(q)$ for all q, meaning that $u_a - \kappa$ is a sub-gradient of $H(q)$ at $q = 0$. In this case, we can define $\nu_a = \vec{0}$ and observe that the first-order condition holds for an appropriately-chosen sub-gradient. 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.

We next prove the "locally invariant posteriors" property described by Caplin et al. [2019]. 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 $\nu_a(\tilde{q}_0) = \nu_a(q_0)$. By the homogeneity property,

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

and therefore the first-order conditions are satisfied. By construction, the constraint is satisfied, the complementary slackness conditions are satisfied, and $\hat{q}_a$ and 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

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

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

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

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

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

as required for the stopping region. Within the continuation region, the strict convexity of $H(q_0)$ in all directions orthogonal to q_0 implies that, as required,

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

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

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

It follows that

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

By the first-order condition,

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

Consider the term

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

By the complementary slackness condition,

$$(\nu_a^T(q_0(\epsilon; z)) - \nu_a^T(q_0))(\hat{q}_a(\epsilon; z) - \hat{q}_a(q_0)) = -\nu_a^T(q_0(\epsilon; z))\hat{q}_a(q_0) - \nu_a^T(q_0)\hat{q}_a(\epsilon; z) \leq 0.$$

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,

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

Thus, anywhere ϕ is twice differentiable (almost everywhere on the interior of the simplex),

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

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

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

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

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

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

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

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

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

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

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

Therefore, we must have

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

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

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

and $d\pi_t = \sigma_{\pi,t} dB_t$. Define $\tilde{\phi}(\pi_t) = \phi(q_t)$. As shown above,

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

exists everywhere in Ω, and therefore

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

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

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

This can be rewritten as

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

where $Diag^+$ denotes the pseudo-inverse of the diagonal matrix.

By the positive-definiteness of k in all directions except those constant in the support of $Q(q_0)\pi_t$, 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 ϕ, $\tilde{\phi}(\pi_\tau) = \hat{u}(Q(q_0)\pi_\tau)$. It follows that, as required,

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

Finally, we verify that alternative policies are sub-optimal. Consider an arbitrary control process σ_t and stopping rule described by the stopping time τ. We have, by the convexity of ϕ and the generalized Ito formula for convex functions (noting that we have shown that the Hessian measure associated with ϕ_{qq} has no pure point component), interpreting ϕ_{qq} in a distributional sense,

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

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,

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

By inequality $\phi(q_\tau) \geq \hat{u}(q_\tau)$, $\phi(q_0) \geq E_0[\hat{u}(q_\tau) - \mu\tau]$ for all policies, verifying optimality.

B.21 Proof of Corollary 2

We begin by observing that Theorem 8 characterizes the solution to the value function in this case– the proof of Theorem 8 requires only that the problem of Definition 2 be further restricted to have no jumps, not that there be a preference for gradual learning per se.

Now consider in particular utility functions with only two actions, L and R (all other action in A are dominated by those two and hence will never occur with positive probability). Using the first-order conditions for the static problem (equation (22)), we have, assuming interior solutions,

$$u_L - \theta H_q(q_L^*(q_0)) = u_R - \theta H_q(q_R^*(q_0))$$

and

$$\pi_L^*(q_0)q_L^*(q_0) + (1 - \pi_L^*(q_0))q_R^*(q_0) = q_0.$$

Now pick any q_0, q_L, q_R such that $q_0 = \pi q_L + (1 - \pi)q_R$ for some $\pi \in (0, 1)$. Set

$$u_L = \theta H_q(q_L)) - \theta H_q(q_0) + K\iota$$

and

$$u_R = \theta H_q(q_R)) - \theta H_q(q_0) + K\iota$$

for some K such that both u_L and u_R are strictly positive. Observe that if the solution is interior, q_L, q_R, and π are optimal policies.

If the solution is not interior, stopping must be optimal. By the convexity of H,

$$q_L^T \cdot u_L - \theta H(q_L) + \theta H(q_0) + \theta(q_L - q_0)^T H_q(q_0)) - q_0^T \cdot u_L = \\ \theta(q_L - q_0)^T H_q(q_L) - \theta H(q_L) + \theta H(q_0) \geq 0,$$

and likewise for q_R. It follows that the q_0 is in the continuation region, and therefore that (q_L, q_R, π) are indeed optimal policies in the static problem.

By the "locally invariant posteriors" property described by Caplin et al. [2019], it follows that for any $q = \alpha q_L + (1 - \alpha)q_R$ with $\alpha \in [0, 1]$, (q_L, q_R, α) are optimal policies given initial prior q_0.

As in the proof of Theorem 8, this implies that the value function is twice-differentiable on the line segment between q_L and q_R, with

$$(q_L - q_0)^T \cdot W_{qq}(q, \lambda) \cdot (q_L - q_0) = \theta(q_L - q_0)^T \bar{k}(q)(q_L - q_0)$$

for all q on that line segment. Integrating,

$$W(q_L, \lambda) - W(q_0, \lambda) - (q_L - q_0)^T \cdot W_q(q_0, \lambda) = \\ \theta(q_L - q_0)^T \cdot (\int_0^1 (1 - s)\bar{k}(sq_L + (1 - s)q_0)ds) \cdot (q_L - q_0) = \\ \theta H(q_L) - \theta H(q_0) - \theta(q_L - q_0)^T H_q(q_0)).$$

By the sub-optimality of jumping directly from q_0 to q_L, it must be the case that

$$W(q_L, \lambda) - W(q_0, \lambda) - (q_L - q_0)^T \cdot W_q(q_0, \lambda) \leq \theta D^*(q_L||q_0)$$

and therefore a preference for gradual learning holds between the points q_0 and q_L.

This argument can be repeated for all (q_0, q_L) in the relative interior of the simplex. By the convexity of D^* and H, we can extend the result to the entirety of the simplex by continuity, proving that a preference for gradual learning must hold.

B.22 Proof of Lemma 7

Let $\phi(q,t,q_0)$ denote the likelihood that $q_t = q \in [q_L, q_H]$ given the initial beliefs q_0. The forward equation is

$$\phi_t(q,t;q_0) = \frac{\partial^2}{\partial q^2}[\sigma^*(q)^2\phi(q,t;q_0)].$$

By Lemma 5, the constraint binds,

$$\frac{1}{2}tr[\sigma_t^T Diag(q_t)\bar{k}(q_t)Diag(q_t)\sigma_t] = \rho^{\frac{1}{\rho}}c.$$

In the two-state model, with an alpha-divergence, $\bar{k}$ is the Fisher matrix,

$$\bar{k}(q_t) = \begin{bmatrix} \frac{1}{q_t} - 1 & -1 \\ -1 & \frac{1}{1-q_t} - 1 \end{bmatrix}$$

and assuming one-dimension of Brownian motion without loss of generality,

$$Diag(q_t)\sigma_t = \begin{bmatrix} \sigma^*(q_t) & 0 \\ -\sigma^*(q_t) & 0 \end{bmatrix}.$$

Therefore,

$$\frac{1}{2}\sigma^*(q_t)^2(\frac{1}{q_t} + \frac{1}{1-q_t}) = \rho^{\frac{1}{\rho}}c,$$

which is

$$\sigma^*(q_t)^2 = 2\rho^{\frac{1}{\rho}}cq_t(1-q_t),$$

as required.

For the conditional dynamics,

$$\begin{aligned} e_1 Diag(q_t)\sigma_t\sigma_t^T e_1 &= e_1 Diag(q_t)\sigma_t\sigma_t^T Diag(q_t)Diag(q_t)^{-1}e_1 \\ &= \frac{\sigma^*(q)^2}{q}, \end{aligned}$$

and the result follows, and likewise

$$e_1 Diag(q_t)\sigma_t\sigma_t^T e_2 = -\frac{\sigma^*(q)^2}{1-q}.$$

B.23 Proof of Lemma 8

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

$$C(p,q;S) = \sum_{s\in S}\pi_s(p,q)D(q_s(p,q)||q).$$

B.23.1 Condition 1

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

$$q_s = q,$$

and by our convention that $q_s = q$ if $\pi_s(p, q) = 0$, this also holds for zero-probability signals. By the definition of a divergence, $D(q||q) = 0$ for all q, and therefore the cost of an uninformative information structure is zero.

The cost is strictly positive by the definition of a divergence (being strictly positive if $q_s \neq q$) and the fact that probabilities must sum to one.

B.23.2 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,

$$\begin{aligned} C(p_M, q; S) &= \sum_{s \in S} \pi_s(p_M, q) D(q_s(p_M, q)||q) \\ &= \lambda \sum_{s \in S_1} \pi_s(p_1, q) D(q_s(p_1, q)||q) + (1 - \lambda) \sum_{s \in S_2} \pi_s(p_2, q) D(q_s(p_2, q)||q) \\ &= \lambda C(p_1, q; S_1) + (1 - \lambda) C(p_2, q; S_2). \end{aligned}$$

verifying that the condition holds.

B.23.3 Condition 3

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

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

By definition,

$$\pi_{s'}(\Pi p, q) = \sum_{s \in S} \Pi_{s',s} \pi_s(p, q)$$

and by Bayes' rule, treating p as an $|S| \times |X|$ matrix and letting e_s denote a vector with one corresponding to s and zero otherwise,

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

where $q_{s'}$ is the posterior associated with $s' \in S'$. This is

$$q_{s'}(\Pi p, q) = \frac{\sum_{s \in S} \pi_s(p,q) q_s(p,q) \Pi_{s',s}}{\pi_{s'}(\Pi p, q)}$$

It follows by the convexity of D in its first argument and Jensen's inequality that

$$\pi_{s'}(\Pi p, q) D(q_{s'}(\Pi p, q)||q) \leq \sum_{s \in S} \Pi_{s',s} \pi_s(p,q) D(q_s(p,q)||q).$$

It immediately follows that

$$\sum_{s' \in S'} \pi_{s'}(\Pi p, q) D(q_{s'}(\Pi p, q)||q) \leq \sum_{s \in S} \pi_s(p,q) D(q_s(p,q)||q).$$

B.23.4 Condition 4

We begin by showing twice-differentiability with respect to perturbations that do not change the support of the signal structure. By the definition of the cost function and the twice-differentiability of D in its first argument, it is sufficient to show that $\pi_s(p,q)$ and $q_s(p,q)$ are both twice-differentiable with respect to these perturbations, in the neighborhood of an uninformative information structure.

Suppose that

$$p(\epsilon, \nu) = r\iota^T + \epsilon\tau + \nu\omega,$$

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

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

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

Simplifying,

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

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

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

and

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

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, $\frac{\partial}{\partial\nu}q_s(\epsilon,\nu)$ 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 information structure, is twice-differentiable in the directions that do not change the support of the distribution of signals. By construction, $\pi_s(p,q)=(e_s^T pq)$ is twice-differentiable.

Now consider a perturbation that changes the support of the signals,

$$p(\epsilon)=r\iota^T+\epsilon\tau+\epsilon\omega,$$

where $e_s^T\omega=0$ for all s such that $e_s^T r>0$, and greater than or equal to zero otherwise, and the support of τe_x is in the support of r for all $x\in X$. We have

$$\begin{aligned}q_s(\epsilon)=q&\frac{r^T e_s}{r^T e_s+\epsilon q^T\tau^T e_s+\epsilon q^T\omega^T e_s}+\frac{\epsilon D(q)\tau^T e_s}{r^T e_s+\epsilon q^T\tau^T e_s+\epsilon q^T\omega^T e_s}\\ &+\frac{\epsilon D(q)\omega^T e_s}{r^T e_s+\epsilon q^T\tau^T e_s+\epsilon q^T\omega^T e_s}.\end{aligned}$$

For s such that $e_s^T\omega>0$,

$$q_s(\epsilon)=\frac{D(q)\omega^T e_s}{q^T\omega^T e_s},$$

and hence does not depend on ϵ. We also have $(e_s^T pq)=\epsilon q^T\omega^T e_s$ for such s. Directional differentiability, continuous in (ω,τ), follows immediately.

B.23.5 Condition 5

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

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

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

B.24 Proof of Lemma 9

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

$$C(p, q; S) = \sum_{x \in X} q_x D(p_x || \pi(p, q); S).$$

B.24.1 Condition 1

Condition 1 requires that if the information structure is uninformative, the cost is zero, and if it is not, the cost is strictly positive. If the signal is uninformative, $pe_x = pq$ for all $x \in X$, and the result holds by the definition of a divergence. The cost for informative signals is strictly positive by the definition of a divergence.

B.24.2 Condition 2

Mixture feasibility requires that

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

This follows by the convexity of the divergence, the Blackwell condition, and Lemma 1.

B.24.3 Condition 3

The result follows immediately by the Blackwell assumption on the divergence.

B.24.4 Condition 4

Twice differentiability follows by assumption. Directional differentiability, with continuous directional derivatives, follows from convexity (for the existence of directional derivatives) and twice-differentiability in the interior (which ensures continuity), and the assumption of continuity in the limit (as the signal probability reaches zero, and the signal alphabet changes).

B.24.5 Condition 5

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

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

where $|| \cdot ||_X$ is an arbitrary norm on the tangent space of $\mathcal{P}(X)$.

By assumption,

$$D(r' || r; S) \geq m(r' - r)^T g(r)(r' - r),$$

where $g(r)$ is the Fisher information matrix.

Consequently,

$$\sum_{x\in X} q_x D(p_x||\pi(p,q);S) \geq m \sum_{x\in X}\sum_{s\in S} \frac{q_x}{\pi_s(p,q)}(e_x^T - q^T)p^T e_s e_s^T p(q-e_x),$$

which by Bayes' rule is

$$\sum_{x\in X}(e_x^T q)D(pe_x||pq;S) \geq m \sum_{s\in S}\pi_s(p,q)(q_s^T(p,q)-q^T)g(q)(q_s(p,q)-q).$$

Therefore, by $g(q) \succeq I$,

$$\sum_{x\in X} q_x D(p_x||\pi(p,q);S) \geq m \sum_{s\in S}\pi_s(p,q)||q_s(p,q)-q||_2^2$$

where $||\cdot||_2$ denotes the Euclidean norm. The result follows by the equivalence of norms.