

*Explosive Convexity of the Value Function in Learning Models**

Axel Anderson[†]

Lones Smith[‡]

Georgetown University

University of Michigan

Economics Department

Economics Department

Washington, DC 20057

Ann Arbor, MI 48109-1220

this version: October 25, 2004

(VERY preliminary)

Abstract

We establish the asymptotic behavior of value functions in two state experimentation problems. In particular, we show that under quite general assumptions $\phi(x) \equiv v(x) - v(0) - v'(0)x \sim c_0 x^{2-\alpha(0)}$ near $x = 0$ and $\phi(x) \sim c_1(1-x)^{2-\alpha(1)}$ near $x = 1$, where $c_0, c_1 > 0$ and $\alpha(0), \alpha(1) \in (0, 1)$. We provide an exact formula for $\alpha(0)$ and $\alpha(1)$ that do not depend on the decision maker's payoff function. Thus, fixing the payoff function, α ranks the value of decision problems near the extremes $x = 0$ and $x = 1$. For high δ , α is increasing in the Kullback-Liebler distance between the two signal densities. We use our approximation to provide an explicit lower bound for δ above which complete learning obtains, and to explicitly solve for the value of any small signal, which turns out to be of the order $x^{2-\alpha(0)}$ and $(1-x)^{2-\alpha(1)}$ near $x = 0$ and $x = 1$. Our result also implies that if $v''(x)$ exists then it must be unbounded near the extremes $x = 0$ and $x = 1$.

*This paper is in restricted circulation for the sole purpose of having an available reference for our paper "Assortative Matching and Reputation". Our preference is to hide such infant papers until finished. Lones wishes to thank the NSF for continual funding.

[†]Email: aza@georgetown.edu

[‡]Email: lones@umich.edu; web: www.umich.edu/~lones

1 Introduction

We investigate the behavior of value functions in infinite horizon experimentation problems in a two state world (H and L). We find that if the discount rate is above a certain computed threshold (which may be quite low), the value function must be asymptotically equal to a linear function plus $c(x - x^*)^{2-\alpha(x^*)}$ for some $c > 0$ and $\alpha(x^*) \in (0, 1)$ near any fixed point x^* , where x is the probability that the state is H . Specifically, $v(x) - v(x^*) - v'(x^*)x \sim c(x - x^*)^{2-\alpha(x^*)}$, where v is the value function.

We then use this result to show that interior fixed points cannot exist for δ above our computed threshold. Our proof proceeds by contradiction, we assume x^* is an interior fixed point and then show that the marginal gain to making a small ε change from the statically optimal policy implied by our approximation to the value function must be positive. Thus, our computed δ provides a lower bound such that complete learning obtains above this δ . We hope this alternative proof that complete learning obtains for high δ will provide some intuition for the result.

Having characterized the value function near the extreme points in the belief space, we investigate the implications for the value of information. We show that $v(0)$ and $v'(0)$ only depend on the decision maker's payoff function (not the signal structure). On the other hand, α is an easily computed function of the decision maker's discount factor and the signal (information) structure alone. Thus if we fix the payoff function, then for x sufficiently close to 0, $v(x)$ will be larger *iff* α is larger. On corollary of this is that for a fixed the discount factor, α provides a complete ordering of information structures independent of payoffs near $x = 0$ and $x = 1$.

We are aware of no function in the literature similar to our α function itself. However, we can link this to the statistics literature for high δ . It turns out that for high enough δ , α is increasing in the Kullback-Liebler distance between the probability densities in the high and low states, where the Kullback-Leibler distance between probability density vectors a and b is denoted as:

$$KL(a, b) = \int a(\sigma) \log(a(\sigma)/b(\sigma)).$$

We then ask, if we stopped the dynamic problem for an instant and offered the decision maker a one time signal, what would be the value of that signal to the decision maker. We show that our approximation implies that for relatively weak signals, the

value of the signal is of the order $x^{2-\alpha(0)}$ for x near 0 and $(1-x)^{2-\alpha(1)}$ for x near 1.

Naively differentiating our asymptotic approximation yields $v''(x) \propto x^{-\alpha}$. Why is convexity unbounded as $x \rightarrow 0$? We show that total convexity $\int v''(x)dx$ is a constant in the discount factor. But, as is well known, the value function flattens out (becomes linear) as the discount factor converges to 1. To resolve this apparent conflict, convexity must accumulate near $x = 0$ and $x = 1$. In fact, we show that if the policy function is C^2 then this naive differentiation is justified and the asymptotic approximation for the second derivative of the value function is $v''(x) \sim \hat{c}x^{-\alpha}$ near $x = 0$ ($\hat{c} > 0$).

As far as we know this is the first time that anyone has noticed that convexity accumulates around extreme points in experimentation problems, let alone characterized the rate of accumulation. Therefore, we think our results have significant theoretical economic content for the learning literature. However, we feel this paper contributes on more than this level. Our main characterization result should prove useful to applied theorists studying problems with some element of experimentation (as they already have to us in (Anderson and Smith 2002)). Discrete time models with learning often do not have a closed form solution. We provide a very simple parameterized solution in the neighborhood of a fixed point.

In a model in which the decision maker observes the outcome of a single experiment prior to taking an action, Blackwell (1953) established that one information structure must be *sufficient* for another to be preferred for all preferences and priors. The ordering of information structures implied by sufficiency is partial, prompting others to investigate restrictions on the class of preferences, priors, and/or information structures and their associated orders on information structures, while retaining the static framework of Blackwell. Ideally one would like to make minimal restrictions, and yet obtain a complete order: Athey and Levin (2000) and Moscarini and Smith (2002) both obtain complete orders with reasonable assumptions.

We investigate the implications of our asymptotic approximation for the ordering of information structures in a *dynamic* context. Decision makers are described by their payoff function, discount factor, and prior. We provide an index that completely orders all information structures given a fixed discount factor and prior, i.e. *regardless* of the payoff function.

In addition to our substantive economic contributions, we feel that the methodology

used in our main proof may prove useful in generating similar asymptotic approximations outside of the experimentation literature. One (widely used) methodology for characterizing value functions is to define a map on a closed metric space; show that this map is a contraction; establish that the subspace defined by the desired property is closed in the defined metric; and finally, show that the defined map preserves the desired property. In learning problems the natural map is the Bellman equation, while the natural norm is the sup-norm. It is generally straightforward to show (see Easley and Kiefer (1988)) for a general treatment) that the Bellman equation yields a contraction mapping in the sup-norm. However, this does not provide any information about the asymptotic form of the value function, because the sup-norm cannot nuance between asymptotic expressions. Put another way, how does one evaluate $\sup_x o(1)$?

The key to our characterization is the *monotonicity* of the Bellman equation. We construct two functions, both of which have the correct asymptotic expression. One of these functions the defined map makes smaller, and the other the defined map makes larger. Thus, since the map is a monotonic contraction we know that the unique fixed point (our value function) must lay between these two constructed function, which implies that it has the correct asymptotic form.

We prove our results for a special case of Easley and Kiefer (1988). One of the keys to our results are the properties of the first and second derivative of expected functions of martingales that follow from the martingale property of beliefs. Smith and Sorensen (2000) use the same first derivative characteristics to prove exponential stochastic convergence of beliefs in their model of observational learning.

In § 2 we provide the basic model. The main proof turns out to be somewhat involved, so rather than present it immediately we first try to develop some of the basic mathematical intuition in a related (although much simplified) model in § 3. In § 4.1 we characterize the asymptotic form of the value function near any fixed point. In § 4.2 we present our complete learning result. In § 4.3 we relate α to the Kullback-Liebler distance between the two state contingent probability densities. In § 4.4 we determine the value of a ‘small’ signal. In § 4.5 we show that our approximation implies that the second derivative (when it exists) must be unbounded near the extremes. In § 5 we conclude.

2 The Model

2.1 The Basics

Our model is a special case of Easley and Kiefer (1988).¹ There are two states of the world $\Theta = H$ and $\Theta = L$, and x is the chance that the state of the world is H . In each of an infinite number of periods a decision maker chooses an action a in a compact set A . At the end of the period he receives a signal $\sigma \in \Sigma$ a compact set, and an expected reward $\pi(x, a)$. Let $f(\sigma, x, a) = f^H(\sigma, a)x + f^L(\sigma, a)(1 - x)$ be the density over Σ given the state of the world and the action taken. For simplicity, we assume this density is C^2 in x and a . Upon observing the signal realization the decision maker updates his beliefs according to Bayes' rule, which implies that the updated belief that $\Theta = H$ conditional on the current belief, signal, and the action taken is $z(\sigma, x, a) \equiv f^H(\sigma, a)x/f(\sigma, x, a)$.

Let v_δ be the decision maker's value function. This value function solves:

$$v_\delta(x) = \max_a \{ (1 - \delta)\pi(x, a) + \delta \int v_\delta(z(\sigma', x, a))f(\sigma', x, a)d\sigma' \} \quad (1)$$

Note that v_δ must be convex by Easley and Kiefer (1988).

3 A Motivational Toy Example

In order to gain some insight into the asymptotic behavior near $x = 0$, we consider an approximation to the Bellman equation (1) near $x = 0$. Fix a policy function $a(x)$, and then set $p(\sigma) \equiv p(\sigma, 0, a(0))$ and $z(\sigma) \equiv z(\sigma, x, a(x))/x$ evaluated at $x = 0$. Further assume for simplicity that $\pi(x, a(x)) = x^2$. These substitutions yield the functional equation:

$$g(x) = (1 - \delta)x^2 + \delta \int g(z(\sigma)x)p(\sigma)d\sigma \quad (2)$$

For this to make sense as a functional equation we must expand the domain of g to \mathbb{R}_+ . This is purely a pedagogical tool - we will return to the correct formulation. Now we ask, how do twice differentiable solutions to (2) behave near $x = 0$?

We evaluate (2) at $x = 0$ to get:

$$g(0) = \delta \int g(0)p(\sigma)d\sigma$$

¹Our main substantive restriction is to a two state world, a prototypical special case in many experimentation models. See for example, McLennan (1984) and Sørensen and Smith (2000)

and thus $g(0) = 0$ (if it is finite). If we differentiate (2) we find

$$g'(x) = 2(1 - \delta)x + \delta \int z(\sigma)g'(z(\sigma)x)p(\sigma)d\sigma$$

Evaluating at $x = 0$ yields:

$$g'(0) = \delta \int z(\sigma)g'(0)p(\sigma)d\sigma$$

Since $\sum z(\sigma)p(\sigma) = 1$, $g'(0) = 0$ (again if it is finite).

Differentiate (2) a second time, and set $x = 0$ to get:

$$g''(0) = 2(1 - \delta) + \delta \int z_i(\sigma)^2 g''(0)p(\sigma)d\sigma$$

Rearrange to:

$$g''(0) = \frac{2(1 - \delta)}{1 - \delta \int z(\sigma)^2 p(\sigma)d\sigma}$$

We claim that $\int z(\sigma)^2 p(\sigma)d\sigma > 1$. Thus, for large δ , $\delta \int z(\sigma)^2 p(\sigma)d\sigma > 1$, which implies $g''(0) < 0$ (again if it is finite - this assumption is *not* innocuous for the second derivative). To show that $\int z(\sigma)^2 p(\sigma)d\sigma > 1$, note that the Power Mean Inequality, or Theorem 96 in Hardy, Littlewood, and Polya (1952) implies that $(\int z(\sigma)^\theta p(\sigma))^{1/\theta}$ is increasing in θ . Thus,

$$\left(\int z(\sigma)^2 p(\sigma)d\sigma \right)^{1/2} > \int z(\sigma)p(\sigma)d\sigma = 1.$$

So, $\delta \int z(\sigma)^2 p(\sigma) > 1$ iff $\delta > \delta^* \equiv (\int z(\sigma)^2 p(\sigma)d\sigma)^{-1}$.

Should we be concerned about our assumption that $g''(0)$ is finite? The answer is that we should be very concerned if we hope that $g''(x)$ is a good approximation to $v_\delta''(x)$ near $x = 0$. Recall that $v_\delta(x)$ must be convex, but we just argued that if $g''(0)$ is finite then it must be negative for $\delta > \delta^*$. Thus, our assumption that $g''(0)$ is finite conflicts with convexity for large δ !²

Now let us see if we can construct a solution which behaves like our proposed approximation. In fact, we claim that there exists $b > 0$ (which we shall construct) such that if $\delta > \delta^*$ then $f''(x) = cx^{-\alpha_\delta} - b$ solves (2) for all c , where α_δ solves:

$$1 = \delta \int z(\sigma)^{2-\alpha_\delta} p(\sigma)d\sigma$$

²Note that we do not claim that there exists a solution to (2) that approximates v_δ , but rather that to have any hope that a solution to (2) approximates v_δ one must consider g with unbounded second derivatives near $x = 0$.

Twice differentiating (2) and substituting $f''(x) = cx^{-\alpha_\delta} - b$ on the right hand side yields:

$$\begin{aligned} f''(x) &= 2(1 - \delta) - b\delta \int z(\sigma)^2 p(\sigma) + c\delta \int z(\sigma)^{2-\alpha_\delta} p(\sigma) x^{-\alpha_\delta} d\sigma \\ &= 2(1 - \delta) - b\delta \int z(\sigma)^2 p(\sigma) d\sigma + cx^{-\alpha_\delta} \end{aligned}$$

So if we set:

$$b = \frac{2(1 - \delta)}{\delta \int z(\sigma)^2 p(\sigma) d\sigma - 1}$$

then the resulting g solves (2).

The key to this construction was choosing an exponent $-\alpha_\delta$ such that, the resulting sum of the coefficients ($\delta \int z(\sigma)^{2-\alpha_\delta} p(\sigma) d\sigma$) was exactly equal to 1.

4 The Value Function Near a Fixed Point

Let $a^*(x)$ be an optimal policy, and x^* be a fixed point given this policy ($z(\sigma, x^*, a^*(x^*)) = x^* \forall \sigma$). In this section we will show that as long as δ is above a easily computed threshold, $\delta^*(x^*)$, the value function, v_δ is of the form³:

$$a + bx + c(x - x^*)^{2-\alpha}(1 + o(1)) \quad (x \rightarrow x^*),$$

for some $c_\delta > 0$, and $\alpha_\delta \in (0, 1)$. This immediately establishes that the value function has this asymptotic form near the extremes $x = 0$ and $x = 1$.

We then show that if the value function satisfies this asymptotic form, then $x^* \notin (0, 1)$. Thus, we offer an alternative proof of the the complete learning result of Easley and Kiefer (1988). Note that we establish not only the existence of a lower bound for δ such that complete learning obtains above this threshold, but we provide a simple formula for this threshold. We hope this alternative proof provides some intuition for the complete learning result.

Having shown that our asymptotic form rules out interior fixed points, we show the implications for the value of information near the extremal (0 and 1) fixed points.

³Notational reminder: $f(x) = g(x)(1 + o(1)) \quad (x \rightarrow x^*)$ iff $f(x) \sim g(x)$ near x^* iff $\lim_{x \rightarrow x^*} f(x)/g(x) = 1$.

4.1 Characterization Near a Fixed Point

Rather than jump directly to the formal analysis, let us assume that v_δ has the asymptotic form we claim and see what the Bellman Equation then requires. If $a^*(x)$ is an optimal policy then the value function must satisfy the following policy equation:

$$v_\delta(x) = (1 - \delta)\pi(x, a^*(x)) + \delta \int v_\delta(z(\sigma', x, a^*(x)))f(\sigma', x, a^*(x))d\sigma'$$

Substituting our functional form into this policy equation, we find that this policy equation can only be satisfied if $a + bx + c(x - x^*)^{2-\alpha}$

$$\begin{aligned} &= (1 - \delta)\pi(x, a^*(x)) + \delta \int [a + bz(\sigma, x, a^*(x)) + c(z(\sigma, x, a^*(x)) - x^*)^{2-\alpha}] f(\sigma, x, a^*(x))d\sigma \\ &= (1 - \delta)\pi(x, a^*(x)) + \delta \left[a + bx + \int c(z(\sigma, x, a^*(x)) - x^*)^{2-\alpha} f(\sigma, x, a^*(x))d\sigma \right] \end{aligned}$$

Since π is C^2 we may take a first order Taylor Series approximation,

$$\pi(x, a^*(x)) = \pi(x^*, a^*(x^*)) + \pi_1(x^*, a^*(x^*))(x - x^*) + O(x^2)$$

Thus, if we set $a = \pi(x^*, a^*(x^*)) - \pi_1(x^*, a^*(x^*))x^*$ and $b = \pi_1(x^*, a^*(x^*))$ our approximation can only satisfy the policy equation if

$$c(x - x^*)^{2-\alpha} = \left[\delta \int c(z(\sigma, x, a^*(x)) - x^*)^{2-\alpha} f(\sigma, x, a^*(x))d\sigma \right] (1 + o(1)) \quad (x \rightarrow x^*)$$

which can only be satisfied if

$$1 = \delta\Phi(x^*, \alpha)$$

where

$$\Phi(x^*, \alpha) \equiv \lim_{x \rightarrow x^*} \left[\int \left(\frac{z(\sigma, x, a^*(x)) - x^*}{x - x^*} \right)^{2-\alpha} f(\sigma, x, a^*(x))d\sigma \right]$$

Assume $\Phi(x^*, 0) > 1$, define $\delta^*(x^*) \equiv \Phi(x^*, 0)^{-1}$, and let $\alpha_\delta(x^*)$ solve: $\delta\Phi(x^*, \alpha_\delta(x^*)) \equiv 1$.

Lemma 1 *If $\delta^*(x^*) < \delta < 1$ then $\alpha_\delta(x^*) \in (0, 1)$, $\lim_{\delta \rightarrow 1} \alpha_\delta(x^*) = 1$, and $\alpha_\delta(x^*)$ is increasing in δ .*

Proof: $\Phi(x^*, \alpha)$ is continuous in α . By assumption, $\Phi(x^*, 0) > 1$, and straightforward computation yields $\Phi(x^*, 1) = 1$. Thus, there exists a candidate $\alpha_\delta(x^*) \in (0, 1)$ and $\lim_{\delta \rightarrow 1} \alpha_\delta(x^*) = 1$. To show that the the solution to $\delta\Phi(x^*, \alpha) = 1$ is unique and that

$\alpha_\delta(x^*)$ is increasing in δ , it suffices to show that $\Phi(x^*, \alpha)$ decreases in α . By the Power Mean Inequality $\mathcal{M}_\theta(x) = (\int \omega(\sigma)y(\sigma)^\theta d\sigma)^{1/\theta}$ is rising in θ if $\int \omega(\sigma)d\sigma = 1$. Setting $\theta = 2 - \alpha$, $\omega(\sigma) = f(\sigma, x^*, a^*(x^*))$ and $y(\sigma) = [z(\sigma, x^*, a^*(x^*)) - x^*]/[x - x^*]$ yields the desired result. \square

In order for this Lemma to be meaningful, it must be the case that $\delta^*(x^*) < 1$, which requires $\Phi(x^*, 0) > 1$. So, critical in our analysis is that $\Phi(x^*, 0) > 1$. In fact, this is required by our formal proof in addition to the intuitive derivation above. It turns out that the analysis that $\Phi(x^*, 0) > 1$ differs for extremal ($x^* \in \{0, 1\}$) and interior fixed point ($x^* \in (0, 1)$). The latter requires additional assumptions.

Lemma 2 *Let x^* be a fixed point. Then*

1. *if $x^* \in \{0, 1\}$ then $\Phi(x^*, 0) > 1$.*
2. *if $x^* \in (0, 1)$ and the policy function a^* is differentiable at x^* , then $\Phi(x^*, 0) > 1$.*

Proof that $\Phi(0, 0) > 1$ and $\Phi(1, 0) > 1$:

$$\begin{aligned} \Phi(0, 0) &= \lim_{x \rightarrow 0} \int \frac{z(\sigma, x, a^*(x))^2}{x^2} f(\sigma, x, a^*(x)) d\sigma \\ &= \lim_{x \rightarrow 0} \int \frac{f^H(\sigma, a^*(x))^2 x^2}{f(\sigma, x, a^*(x)) x^2} f(\sigma, x, a^*(x)) d\sigma \\ &= \lim_{x \rightarrow 0} \int \frac{f^H(\sigma, a^*(x))^2}{f(\sigma, x, a^*(x))} d\sigma \\ &= \int \frac{f^H(\sigma, a^*(0))^2}{f^L(\sigma, a^*(0))} d\sigma > 1 \end{aligned}$$

To see this last inequality, write

$$\int \frac{f^H(\sigma)^2}{f^L(\sigma)} d\sigma - 1 = \int \frac{f^H(\sigma)}{f^L(\sigma)} [f^H(\sigma) - f^L(\sigma)] d\sigma > 0$$

since $f^H \neq f^L$, $f^H, f^L \geq 0$ and $\int f^H(\sigma) d\sigma = \int f^L(\sigma) d\sigma = 1$, where we have abused notation and set $f^\theta(\sigma) = f^\theta(\sigma, a^*(0))$.

$$\begin{aligned} \Phi(1, 0) &= \lim_{x \rightarrow 1} \int \frac{(1 - z(\sigma, x, a^*(x)))^2}{(1 - x)^2} f(\sigma, x, a^*(x)) d\sigma \\ &= \lim_{x \rightarrow 1} \int \frac{(1 - \frac{f^H(\sigma, a^*(x))x}{f(\sigma, x, a^*(x))})^2}{(1 - x)^2} f(\sigma, x, a^*(x)) d\sigma \end{aligned}$$

$$\begin{aligned}
&= \lim_{x \rightarrow 1} \int \frac{f^L(\sigma, a^*(x))^2 (1-x)^2}{(1-x)^2 f(\sigma, x, a^*(x))} d\sigma \\
&= \int \frac{f^L(\sigma, a^*(1))^2}{f^H(\sigma, a^*(1))} d\sigma > 1
\end{aligned}$$

where the last inequality is established similarly to the same step above.

Since (assumptions from Real Analysis text), we have that:

$$\Phi(x^*, 0) = \int \left(\lim_{x \rightarrow x^*} \frac{z(\sigma, x, a^*(x)) - x^*}{x - x^*} \right)^2 f(\sigma, x^*, a^*(x^*)) d\sigma$$

We have:

$$\begin{aligned}
\frac{z(\sigma, x, a^*(x)) - x^*}{x - x^*} &= \frac{f^H(\sigma, a^*(x))x - f(\sigma, x, a^*(x))x^*}{f(\sigma, x, a^*(x))(x - x^*)} \\
&= \frac{f(\sigma, x, a^*(x))(x - x^*) + f^H(\sigma, a^*(x))x - f(\sigma, x, a^*(x))x}{f(\sigma, x, a^*(x))(x - x^*)} \\
&= 1 + \frac{x(f^H(\sigma, a^*(x)) - f(\sigma, x, a^*(x)))}{f(\sigma, x, a^*(x))(x - x^*)} \\
&= 1 + \frac{x(f^H(\sigma, a^*(x)) - x f^H(\sigma, a^*(x)) - (1-x)f^L(\sigma, a^*(x)))}{f(\sigma, x, a^*(x))(x - x^*)} \\
&= 1 + \frac{x(1-x)(f^H(\sigma, a^*(x)) - f^L(\sigma, a^*(x)))}{f(\sigma, x, a^*(x))(x - x^*)}
\end{aligned}$$

Thus,

$$\lim_{x \rightarrow x^*} \frac{z(\sigma, x, a^*(x)) - x^*}{x - x^*} = 1 + \frac{x^*(1-x^*)}{f(\sigma, x^*, a^*(x^*))} \lim_{x \rightarrow x^*} \frac{f^H(\sigma, a^*(x)) - f^L(\sigma, a^*(x))}{x - x^*}$$

Note that $f^H(\sigma, a^*(x^*)) = f^L(\sigma, a^*(x^*))$, thus we need to use L'Hopital to determine this limit. However, to do so requires that we take a derivative of $a^*(x)$ at x^* , which is precisely why we require additional assumptions for the interior fixed point analysis:

$$\lim_{x \rightarrow x^*} \frac{f^H(\sigma, a^*(x)) - f^L(\sigma, a^*(x))}{x - x^*} = (f_a^H(\sigma, a^*(x^*)) - f_a^L(\sigma, a^*(x^*))) (a^*)'(x^*)$$

and thus,

$$\begin{aligned}
\Phi(x^*, 0) &= \int \left[1 + \frac{x^*(1-x^*)}{f(\sigma, x^*, a^*(x^*))} (f_a^H(\sigma, a^*(x^*)) - f_a^L(\sigma, a^*(x^*))) (a^*)'(x^*) \right]^2 f(\sigma, x^*, a^*(x^*)) d\sigma \\
&= 1 + [x^*(1-x^*)(a^*)'(x^*)]^2 \int \frac{[f_a^H(\sigma, a^*(x^*)) - f_a^L(\sigma, a^*(x^*))]^2}{f(\sigma, x^*, a^*(x^*))} > 1
\end{aligned}$$

We have shown that if v_δ is of the specified form it can only obey the policy equation if $\alpha_\delta(x^*)$ is as given. Although the analysis above should hopefully provide some intuition for the following result, it clearly does not prove it.

Proposition 1 *If x^* is a fixed point and the assumptions in Lemma 2 hold then we have:*

$$v_\delta(x) = \pi(x^*, a^*(x^*)) + \pi_1(x^*, a^*(x^*))(x - x^*) + c_\delta(x - x^*)^{2 - \alpha_\delta(x^*)}(1 + o(1)) \quad (x \rightarrow x^*) \quad (3)$$

for some $c_\delta > 0$.

We prove this Proposition in Appendix A.

4.2 Complete Learning Threshold

Now that we have a sharp characterization of the value function in the neighborhood of a fixed point we can use this characterization to ask when interior fixed points exist. We say *complete learning* occurs if the only fixed points of an optimal policy are 0 and 1. Specifically, we will use our characterization result to establish a lower bound such that complete learning must occur for all δ above this lower bound.

Following Easley and Kiefer (1988) let $a^0(x)$ be the statically optimal policy, i.e. $a^0(x) \equiv \operatorname{argmax} \pi(x, a)$, and then define the set of potentially confounding beliefs as:

$$S \equiv \{x : x \in (0, 1) \text{ and } z(\sigma, x, a^0(x)) = x \forall \sigma\}.$$

Let

$$\Phi^0(x) \equiv \lim_{s \rightarrow x} \int \left(\frac{z(\sigma, s, a^0(s)) - x}{s - x} \right)^2 f(\sigma, s, a^0(s)) d\sigma$$

and $\delta^0(x) \equiv \Phi^0(x)^{-1}$. Notice that if x is a fixed point of the optimal policy then $\delta^0(x) = \delta^*(x)$.

Proposition 2 *If $\delta > \max_{x \in S} \delta^0(x)$ then complete learning occurs.*

Proof: The proof is by contradiction: first assume that an interior x is a fixed point. Then $\delta^0(x) < \delta$ means we can use our approximation to the value function in the neighborhood of this fixed point. We then show that an ϵ deviation from the optimal policy increases the total expected payoff, and thus show that the policy could not have been optimal.

Proof of Proposition 2:

Let $x^* \in S$ be a fixed point of an optimal policy a^* , and consider an ε deviation from a^* at $x = x^*$. The static losses from such a deviation are:

$$L(\varepsilon) \equiv (1 - \delta) [\pi(x^*, a^*(x^*) + \varepsilon) - \pi(x^*, a^*(x^*))]$$

Since π is C^2 in a , we have:

$$\begin{aligned} L(\varepsilon) &= (1 - \delta) \left[\pi_2(x^*, a^*(x^*))\varepsilon + \frac{1}{2}\pi_{22}(x^*, a^*(x^*))\varepsilon^2 \right] + o(\varepsilon^2) \quad (\varepsilon \rightarrow 0) \\ &= \frac{1}{2}(1 - \delta)\pi_{22}(x^*, a^*(x^*))\varepsilon^2 + o(\varepsilon^2) \quad (\varepsilon \rightarrow 0) \end{aligned}$$

since $a^*(x^*)$ solves $\max_a \pi(x^*, a)$. Thus,

$$\lim_{\varepsilon \rightarrow 0} \frac{L(\varepsilon)}{\varepsilon^2} = \frac{1}{2}(1 - \delta)\pi_{22}(x^*, a^*(x^*))$$

The expected future change in value from this ε change in policy is:

$$\begin{aligned} G(\varepsilon) &\equiv \delta \lim_{x \rightarrow x^*} \left[\int v_\delta(z(\sigma, x, a^*(x^*) + \varepsilon))f(\sigma, x, a^*(x^*) + \varepsilon)d\sigma \right] \\ &\quad - \int v_\delta(z(\sigma, x, a^*(x^*)))f(\sigma, x, a^*(x^*))d\sigma \end{aligned}$$

We will show that $\lim_{\varepsilon \rightarrow 0} G(\varepsilon)/\varepsilon^2 = \infty$, so that $\lim_{\varepsilon \rightarrow 0} (L(\varepsilon) + G(\varepsilon))/\varepsilon^2 = \infty > 0$. Thus, a^* could not have been an optimal policy as assumed, so that x^* cannot be a fixed point.

Since $\delta > \delta^*(x^*)$, Proposition 1 applies and

$$\lim_{x \rightarrow x^*} \int v_\delta(z(\sigma, x, a^*(x^*)))f(\sigma, x, a^*(x^*))d\sigma = \pi(x^*, a^*(x^*))$$

so,

$$\begin{aligned} G(\varepsilon) &= \lim_{x \rightarrow x^*} \int v_\delta(z(\sigma, x, a^*(x^*) + \varepsilon))f(\sigma, x, a^*(x^*) + \varepsilon)d\sigma - \pi(x^*, a^*(x^*)) \\ &= c_\delta \lim_{x \rightarrow x^*} \int (z(\sigma, x, a^*(x^*) + \varepsilon) - x^*)^{2-\alpha_\delta(x^*)} f(\sigma, x, a^*(x^*) + \varepsilon)d\sigma \\ &\equiv c_\delta \lim_{x \rightarrow x^*} g(x, \varepsilon) \\ &= c_\delta \lim_{x \rightarrow x^*} \left[g(x, 0) + g_2(x, 0)\varepsilon + \frac{1}{2}g_{22}(x, 0)\varepsilon^2 \right] + o(\varepsilon^2) \end{aligned}$$

Straightforward computation yields:

$$\lim_{x \rightarrow x^*} g(x, 0) = \lim_{x \rightarrow x^*} g_2(x, 0) = 0$$

and $\lim_{x \rightarrow x^*} g_{22}(x, 0) =$

$$\lim_{x \rightarrow x^*} (2 - \alpha_\delta(x^*)) (1 - \alpha_\delta(x^*)) \int (z(\sigma, x, a^*(x^*)) - x^*)^{-\alpha_\delta(x^*)} z_3(\sigma, x, a^*(x^*))^2 f(\sigma, x, a^*(x^*)) d\sigma$$

Clearly the limit is unbounded. To see that it is positive, again use the Power Mean Inequality with $w(\sigma) \equiv f(\sigma, x, a^*(x^*))$ and $y(\sigma) \equiv (z(\sigma, x, a^*(x^*)) - x^*)^{-1}$, then the Power Mean Inequality implies

$$\left(\int y(\sigma)^{\alpha_\delta(x^*)} w(\sigma) d\sigma \right)^{1/\alpha_\delta(x^*)} > \lim_{\theta \rightarrow 0} \left(\int y(\sigma)^\theta w(\sigma) d\sigma \right)^{1/\theta} = \lim_{\theta \rightarrow 0} 1^{1/\theta} = 1.$$

Thus, $\lim_{\varepsilon \rightarrow 0} G(\varepsilon)/\varepsilon^2 = \infty$, and $\lim_{\varepsilon \rightarrow 0} (L(\varepsilon) + G(\varepsilon))/\varepsilon^2 = \infty > 0$. \square .

4.3 A Statistical Interpretation for α

We have established that the local curvature of the value function is determined by $\alpha_\delta(x^*)$ near extremal fixed points. As we saw above, for δ above our threshold interior fixed points cannot exist, so we focus now on $x^* \in \{0, 1\}$. In fact, in this section we assume $x^* = 0$ to conserve space. We have shown that $\alpha_\delta(0)$ solves:

$$\delta \int \frac{f^H(\sigma, a^*(0))^{2-\alpha_\delta(0)}}{f^L(\sigma, a^*(0))} d\sigma \equiv 1,$$

but have offered no interpretation for this formula. We can relate this formula to the information literature when δ is near 1. For simplicity we suppress notation and let $f^\theta(\sigma) \equiv f^\theta(\sigma, a^*(0))$, and $\alpha_\delta \equiv \alpha_\delta(0)$.

Recall, the Kullback-Leibler distance between probability densities a and b is:

$$KL(a, b) = \int a(\sigma) \log(a(\sigma)/b(\sigma)).$$

This is the most commonly used (albeit *asymmetric*) *distance* measure between two conditional signal distributions in the information literature outside of economics. It obeys $KL(a, b) \geq 0$ with $KL(a, b) = 0$ iff $a = b$, but it is not formally a distance measure as it does not obey the triangle inequality.

Consider $(f^H, f^L) \neq (\hat{f}^H, \hat{f}^L)$, and let α_δ be associated with the former and $\hat{\alpha}_\delta$ be associated with the latter, then we have the following proposition.

Proposition 3 *There exists δ^\dagger such that $KL(f^H, f^L) > KL(\hat{f}^H, \hat{f}^L) \Rightarrow \alpha_\delta > \hat{\alpha}_\delta$ for all $\delta > \delta^\dagger$.*

Proof: Take the derivative of α_δ in δ evaluated at $\delta = 1$ we find:

$$\left. \frac{\partial \alpha_\delta}{\partial \delta} \right|_{\delta=1} = \frac{1}{KL(f^H, f^L)}$$

Also α_δ is continuous in δ . Thus, for δ near 1, higher KL implies α_δ falls more slowly as we reduce δ . This in turn implies that a higher KL implies a higher α_δ for high enough δ . \square .

Notice that our approximation for v_δ involves three terms. The linear portion is only a function of the flow profit function. For x close enough to 0, the final term will be larger *iff* α_δ is larger, regardless of the value of c_δ . Now consider two decision makers, both with the same δ and flow payoff function π , but differing signal structures $(f^H, f^L) \neq (\hat{f}^H, \hat{f}^L)$. Let v_δ be the value function associated with former and \hat{v}_δ be the value function associated with the latter. We then have the following Corollary to Proposition 3.

Corollary 1 *There exists δ^\dagger such that $KL(f^H, f^L) > KL(\hat{f}^H, \hat{f}^L) \Rightarrow v_\delta > \hat{v}_\delta$ for all $\delta > \delta^\dagger$.*

4.4 The Value of Marginal Signals

Assume the decision maker facing our learning problem is given the chance to observe an extra signal with densities $g^\theta(\sigma, \varepsilon)$ for $\theta \in \{H, L\}$, with $\varepsilon \in \mathbb{R}_+$. Assume g^θ is C^2 in ε and that $\exists \sigma$ such that $g^H(\sigma, \varepsilon) \neq g^L(\sigma, \varepsilon)$ for all $\varepsilon > 0$. Thus, we have an informative signal for all non zero ε , but the signal becomes uninformative as $\varepsilon \rightarrow 0$. Let $g(\sigma, x, \varepsilon) = xg^H(\sigma, \varepsilon) + (1-x)g^L(\sigma, \varepsilon)$. Overuse notation and let $z(\sigma, x, \varepsilon)$ be the updated value of x upon observing signal σ generated by g , so that $z(\sigma, x, \varepsilon) = (g^H(\sigma, \varepsilon)x)/g(\sigma, x, \varepsilon)$.

Assume x near 0 and $\delta > \delta^*(0)$ so that Proposition 1 obtains. We then ask, what is the value of this signal? More specifically, we have the following Proposition.

Proposition 4 *Assume $\delta > \max\{\delta^*(0), \delta^*(1)\}$, then*

1. *near $x = 0$:*

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{V}(x, \varepsilon)}{\varepsilon^2} = \frac{1}{2}(2 - \alpha)(1 - \alpha)x^{2-\alpha} \int \frac{[g_2^H(\sigma, 0) - g_2^L(\sigma, 0)]^2}{g(\sigma, 0, 0)} d\sigma + o(x^{2-\alpha})$$

2. near $x = 1$:

$$\lim_{\varepsilon \rightarrow 0} \frac{\mathcal{V}(x, \varepsilon)}{\varepsilon^2} = \frac{1}{2}(2-\alpha)(1-\alpha)(1-x)^{2-\alpha} \int \frac{[g_2^H(\sigma, 1) - g_2^L(\sigma, 1)]^2}{g(\sigma, 1, 0)} d\sigma + o((1-x)^{2-\alpha})$$

Given ε and value function v the value of a one time signal is:

$$\mathcal{V}(x, \varepsilon) \equiv \int v(z(\sigma, x, \varepsilon))g(\sigma, x, \varepsilon)d\sigma - v(x)$$

For x near 0 our asymptotic approximation is linear plus $c_\delta x^{2-\alpha}$. Thus our approximation implies:

$$\mathcal{V}(x, \varepsilon) = c_\delta \left[\int z(\sigma, x, \varepsilon)^{2-\alpha} g(\sigma, x, \varepsilon) d\sigma - x^{2-\alpha} \right] + o(x^{2-\alpha}) = h(x, \varepsilon) + o(x^{2-\alpha})$$

We are interested in $h(x, \varepsilon)$ for small ε . Taking a second order Taylor series, we note:

$$h(x, \varepsilon) = h(x, 0) + h_\varepsilon(x, 0)\varepsilon + \frac{1}{2}h_{\varepsilon\varepsilon}(x, 0)\varepsilon^2$$

Routine (by now) algebra establishes that $h(x, 0) = h_\varepsilon(x, 0) = 0$, and simplification yields:

$$h_{\varepsilon\varepsilon}(x, 0) = c_\delta(2-\alpha)(1-\alpha)x^{-\alpha} \int z_3(\sigma, x, 0)^2 g(\sigma, x, 0) d\sigma$$

Substituting we have:

$$\int z_3(\sigma, x, 0)^2 g(\sigma, x, 0) d\sigma = x^2(1-x)^2 \int \frac{[g_2^H(\sigma, 0) - g_2^L(\sigma, 0)]^2}{g(\sigma, 0, 0)} d\sigma$$

Thus,

$$\frac{\mathcal{V}(x, \varepsilon)}{\varepsilon^2} = \frac{1}{2}c_\delta(2-\alpha)(1-\alpha)x^{2-\alpha} \int \frac{[g_2^H(\sigma, 0) - g_2^L(\sigma, 0)]^2}{g(\sigma, 0, 0)} d\sigma + o(x^{2-\alpha})$$

The result for x near 1 is established similarly. \square

Note that changes α_δ impact the value of a marginal signal in two ways: via the exponent $2-\alpha$, and via the coefficient $(2-\alpha_\delta)(1-\alpha_\delta)$, thus changes in the information structure that increase α_δ may increase or decrease the value of a marginal signal depending on the value of x . However, for low or high enough x , increasing α_δ increases the value of a marginal signal. Thus, α_δ fully orders the value of a marginal signal for decision makers who are almost sure about the state of the world.

4.5 Unbounded Extremal Convexity

We have found the tail asymptotic order on the value function. One might be tempted to conclude that if v is twice differentiable, we must have $v''_\delta(x) \propto x^{2-\alpha_\delta(0)}$. However, while integrating asymptotic relations is a fully valid exercise, differentiation requires regularity conditions on the derivative (results known as ‘Tauberian Theorems’). Thus, we cannot make such a statement without making stronger assumptions (a C^2 optimal policy function for starters). Instead we show that our approximation necessarily implies that the second derivative (where it exists) must be unbounded.

Proposition 5 *Assume $\delta > \delta^*(0)$ and fix $t^* > 0$, then we cannot have $v''_\delta(t) < [c_\delta/(2 - \alpha_\delta(0))(1 - \alpha_\delta(0))]t^{-\alpha_\delta(0)}(1 + o(1))$ for all $t < t^*$.*

Proof: v_δ is convex and thus monotonic. The derivative of monotonic functions exists almost everywhere, so that we may write:

$$v_\delta(x) = v_\delta(0) + v'_\delta(0)x + \int_0^x \int_0^s v''_\delta(t) dt ds$$

Which combined with our asymptotic approximation implies that:

$$\int_0^x \int_0^s v''_\delta(t) dt ds = c_\delta x^{2-\alpha_\delta(0)}(1 + o(1))$$

Now assume $v''_\delta(t) < [c_\delta/(2 - \alpha_\delta(0))(1 - \alpha_\delta(0))]t^{-\alpha_\delta(0)}(1 + o(1))$ for all $t < t^*$. Which if we integrate implies:

$$\int_0^x \int_0^s v''_\delta(t) dt ds < c_\delta x^{2-\alpha_\delta(0)}(1 + o(1))$$

a contradiction. □

5 Conclusion

We have shown that in two state infinite horizon learning problems with sufficient patience the value function had to be linear plus $c(x^*)x^{2-\alpha(x^*)}$ near any fixed point of the learning dynamics x^* . We then used this result to provide an alternative proof that complete learning obtains for high enough δ .

The exponent $\alpha(x^*)$ was only a function of the discount rate and the signal structure, not the flow profit function. Thus, close to the extremes, the value of alternative

learning problems is determined by δ and the signal structure. For high enough δ , α is determined by the Kullback-Liebler distance between the probability densities in the two states.

We also used our approximation to approximate the value of a marginal signal near $x = 0$ and $x = 1$.

A Proof of Proposition 1.

Subtract $\pi(0, a^*(0)) + \pi_1(0, a^*(0))x$ from both sides of the Bellman equation (1) and use

$$\pi(0, a^*(0)) = \int \pi(0, a^*(0))f(\sigma, x, a^*(x))$$

and

$$\pi_1(0, a^*(0))x = \int \pi_1(0, a^*(0))z(\sigma, x, a^*(x))f(\sigma, x, a^*(x))d\sigma$$

to get:

$$\phi_\delta(x) = \max_a \left\{ (1 - \delta)\Delta\pi(x, a) + \delta \int \phi_\delta(z(\sigma', x, a))f(\sigma', x, a)d\sigma' \right\} \quad (4)$$

where $\Delta\pi(x, a) \equiv \pi(x, a) - \pi(0, a^*(0)) - \pi_1(0, a^*(0))x$ and $\phi_\delta(x) \equiv v_\delta(x) - \pi(0, a^*(0)) - \pi_1(0, a^*(0))x$.

Now form the mapping $T : C[0, 1] \rightarrow C[0, 1]$ defined by (4). Notice that T is a monotonic contraction in the sup-norm. Thus, there exists a unique solution to (4) by Banach's fixed point theorem. We prove our result by constructing two functions: $\underline{\phi}$ and $\bar{\phi}$ both of which have the correct asymptotic form, that satisfy: $\underline{\phi} \leq T\underline{\phi}$ and $T\bar{\phi} \leq \bar{\phi}$. Thus, since T monotonic and the sequence $T^N\underline{\phi}$ converges to the unique fixed point for any ϕ , we have $\underline{\phi} \leq \phi_\delta \leq \bar{\phi}$, and since both functions have the same asymptotic form, ϕ_δ must have the same asymptotic form.

Step 1: There exists an $a > 0$ such that $\underline{\phi}(x) = a(x - x^*)^{2-\alpha_\delta(x^*)} \Rightarrow \underline{\phi} \leq T\underline{\phi}$. Let $a^*(x)$ be an optimal policy function, we have:

$$\begin{aligned} T\underline{\phi}(x) &= \max_a \left\{ (1 - \delta)\Delta\pi(x, a) + a(x - x^*)^{2-\alpha_\delta(x^*)} \int \left[\frac{z(\sigma, x, a)}{x - x^*} \right]^{2-\alpha_\delta(x^*)} f(\sigma, x, a)d\sigma \right\} \\ &\geq (1 - \delta)\Delta\pi(x, a^*(x)) + a(x - x^*)^{2-\alpha_\delta(x^*)}\Upsilon(x) \end{aligned}$$

where $\Upsilon(x) \equiv \delta \int [(z(\sigma, x, a^*(x)))/(x - x^*)]^{2-\alpha_\delta(x^*)} f(\sigma, x, a(x)) d\sigma$. So to establish the claim, we need:

$$(1 - \delta)\Delta\pi(x, a^*(x)) + a(x - x^*)^{2-\alpha_\delta(x^*)}\delta\Upsilon(x) \geq a(x - x^*)^{2-\alpha_\delta(x^*)}$$

or

$$(1 - \delta)\Delta\pi(x, a^*(x)) \geq a(x - x^*)^{2-\alpha_\delta(x^*)}(1 - \Upsilon(x))$$

Note that $\Delta\pi(x^*, a^*(x^*)) = (1 - \Upsilon(x^*)) = 0$, so that this inequality is satisfied at $x = x^*$. By E-K $\pi(x, a^*(x))$ is a strictly convex function, so that $\Delta\pi(x, a^*(x))$ is strictly positive away from x^* , while Υ is bounded below by 0. So for a low enough this inequality holds boundedly far from $x = x^*$. Finally, $\Delta\pi$ is $O((x - x^*)^2)$, while $(1 - \Upsilon(x))$ is $O(x - x^*)$, so that $(x - x^*)^{2-\alpha_\delta(x^*)}(1 - \Upsilon(x))$ is $O((x - x^*)^{3-\alpha_\delta(x^*)})$. Thus the right hand side is of a lower order than the left hand side. Since π is strictly positive, this inequality holds in a neighborhood of $x = x^*$.

Step 2: There exist constants $b, d > 0$ such that $\bar{\phi}(x) = b(x - x^*)^{2-\alpha_\delta(x^*)} - d(x - x^*)^2 \Rightarrow T\bar{\phi}(x) \leq \bar{\phi}(x)$. We have:

$$T\bar{\phi}(x) = (1 - \delta)\Delta\pi(x, a^*(x)) + b(x - x^*)^{2-\alpha_\delta(x^*)}\Upsilon(x) - d(x - x^*)^2\hat{\Upsilon}(x)$$

where $\hat{\Upsilon}(x) \equiv \delta \int [(z(\sigma, x, a^*(x)) - x^*)/(x - x^*)]^2 f(\sigma, x, a^*(x)) d\sigma$. The analysis used to establish that $\Upsilon(x^*) = 1$ can be used to establish that $\hat{\Upsilon}(x^*) = \Phi(x^*) > 1$, bounded away from 1.

We establish step 2 in two parts. First we show that \exists an $\varepsilon > 0$ such that $T\bar{\phi}(x) \leq \bar{\phi}(x)$ for all x in an ε neighborhood of x^* using the asymptotic properties of $\Delta\pi$, Υ and $\hat{\Upsilon}$ plus continuity. We only consider a neighborhood $[x^*, x^* + \varepsilon]$, thus we implicitly assume $x > x^*$. This is done simply to conserve space. Like reasoning holds for $x < x^*$ near x^* .

Then we show that $T\bar{\phi}(x) \leq \bar{\phi}(x)$ for all x outside of this neighborhood via Jensen's inequality, by choosing b and d such that $\hat{\phi}$ is concave over the appropriate range (we will need signals to be bounded for this step).

Step 2-a: There exist $d, \varepsilon > 0$ such that $T\bar{\phi}(x) \leq \bar{\phi}(x) \forall b > 0$ and $x \leq x^* + \varepsilon$.

We have:

$$T\bar{\phi}(x) \leq \bar{\phi}(x)$$

$$\Leftrightarrow (1 - \delta)\Delta(x, a^*, x) \leq b(x - x^*)^{2-\alpha_\delta(x^*)}(1 - \Upsilon(x)) - d(x - x^*)^2(1 - \hat{\Upsilon}(x))$$

At $x = x^*$, both sides of this inequality are zero. Since $\Delta\pi$ is the difference between a C^∞ function and its best linear approximation, it is of the form $t(x - x^*)^2(1 + o(1))$ for some constant t . We know Υ is at least C^2 and $1 - \Upsilon(x^*) = 0$, so $b(x - x^*)^{2-\alpha_\delta(x^*)}(1 - \Upsilon(x)) = O(x^{3-\alpha_\delta(x^*)})$. Since, $\hat{\Upsilon}$ is continuous and C^2 we know $d(x - x^*)^2(\hat{\Upsilon}(x) - 1) = d(\hat{\Upsilon}(x^*) - 1)(1 + o(1))$. Thus, to establish the claim we need to choose a d such that $t \leq d(\hat{\Upsilon}(x^*) - 1)$, which is true for large d .

Step 2-b: Fix ε from Step 2-a, then there exist $b > 0$ and $d > 0$ (large enough for Step 2-a) such that $T\bar{\phi}(x) \leq \bar{\phi}(x) \forall x > x^* + \varepsilon$.

Let $\bar{x}(b, d)$ be the threshold such that $\bar{\phi}$ is convex for all $x \in (x^*, \bar{x}(b, d)]$ and strictly concave for all $x > \bar{x}(b, d)$. We shall choose b such that for all $x > x^* + \varepsilon$, $z(\sigma, x, a^*(x)) > \bar{x}(b, d)$, thus we can use Jensen's inequality in this step of the proof. So, fix $\bar{x}(b, d)$ to solve: $\min_\sigma z(\sigma, x^* + \varepsilon, a^*(x^* + \varepsilon)) = \bar{x}(b, d)$, and then choose b so that $\bar{x}(b, d)$ is the threshold between the convex and concave regions, i.e. :

$$(2 - \alpha_\delta(x^*))(1 - \alpha_\delta(x^*))b(\bar{x}(b, d) - x^*)^{-\alpha_\delta(x^*)} = 2d((\bar{x}(b, d) - x^*))$$

so,

$$\bar{x}(b, d) = x^* + \left[\frac{(2 - \alpha_\delta(x^*))(1 - \alpha_\delta(x^*))b}{2d} \right]^{\frac{1}{1+\alpha_\delta(x^*)}}$$

So, given fixed $\varepsilon > 0$ and $d > 0$ we can choose a $b > 0$ so that $\bar{\phi}$ is strictly concave for all $x > \bar{x}(b, d)$, since $z(\sigma, x, a^*(x)) > \bar{x}(b, d)$ for all σ and $x > x^* + \varepsilon$, we have:

$$\int \bar{\phi}(z(\sigma, x, a^*(x)))f(\sigma, x, a^*(x))d\sigma < \bar{\phi}(x) \quad \forall x > x^* + \varepsilon$$

Further by scaling up b and d (holding $\bar{x}(b, d)$ constant we can make:

$$\bar{\phi}(x) - \delta \int \bar{\phi}(z(\sigma, x, a^*(x)))f(\sigma, x, a^*(x))d\sigma$$

arbitrarily large for any $x > x^* + \varepsilon$. Thus, we can choose b and d such that:

$$T\bar{\phi}(x) = (1-\delta)\Delta\pi(x, a^*(x)) + \delta \int \bar{\phi}(z(\sigma, x, a^*(x)))f(\sigma, x, a^*(x))d\sigma \leq \bar{\phi}(x) \quad \forall x > x^* + \varepsilon$$

Having established an upper and lower bound for ϕ_δ both are of the form $c(x - x^*)^{2-\alpha_\delta(x^*)}$ we know ϕ_δ must be of the same asymptotic form. \square

References

- ANDERSON, A., AND L. SMITH (2002): “Assortative Matching, Reputation, and the Beatles Break-up,” *Working Paper*.
- ATHEY, S., AND J. LEVIN (2000): “The Value of Information in Monotone Decision Problems,” *MIT Working Paper*.
- BLACKWELL, D. (1953): “Equivalent Comparison of Experiments,” *Annals of Mathematics and Statistics*, 24.
- EASLEY, D., AND N. KIEFER (1988): “Controlling a Stochastic Process with Unknown Parameters,” *Econometrica*, 56(5), 1045–1064.
- HARDY, G., J. LITTLEWOOD, AND G. POLYA (1952): *Inequalities*. Cambridge University Press, Cambridge, MA, second edn.
- MCLENNAN, A. (1984): “Price Dispersion and Incomplete Learning in the Long Run,” *Journal of Economic Dynamics and Control*, 7(3), 331–347.
- MOSCARINI, G., AND L. SMITH (2002): “The Law of Large Demand for Information,” *Econometrica*, 70(6), 2351–2366.
- SMITH, L., AND P. SORENSEN (2000): “Pathological Outcomes of Observational Learning,” *Econometrica*, 68(2), 371–398.
- SØRENSEN, P., AND L. SMITH (2000): “Rational Social Learning by Random Sampling,” *Working Paper*, <http://www-personal.umich.edu/~lones/ftp/rational.pdf>.