Coordination and Continuous Choice*

Stephen Morris and Ming Yang
Princeton University and Duke University

December 2016

Abstract

We study a coordination game where players choose what information to acquire about payoffs prior to the play of the game. We allow general information acquisition technologies, modeled by a cost functional defined on information structures.

A cost functional satisfies *continuous choice* if players choose a continuous decision rule even in a decision problem with discontinuous payoffs. If continuous choice holds, there is a unique equilibrium in the coordination game; if continuous choice fails, there are multiple equilibria. We show how continuous choice captures the idea that it is sufficiently harder to distinguish states that are close to each other relative to far away states.

JEL: C72 D82

KEYWORDS: coordination, endogenous information acquisition, continuous choice, higher order beliefs

1 Introduction

Situations where players must coordinate their actions are ubiquitous. Under complete information, the resulting coordination game will have multiple equilibria. But what if there is incomplete information? And what if the information structure is chosen endogenously? We will show that if it is particularly difficult for players to distinguish nearby states, there will be uniqueness. Otherwise, there will be multiple equilibria.

Our results will come in two parts. We will first show that continuous choice, a property of choice in (non-strategic) decision problems, implies uniqueness in coordination games. We

*An earlier version of this paper, Morris and Yang (2016), was circulated under the title "Coordination and the Relative Cost of Distinguishing Nearby States." We are grateful for the comments of Muhamet Yildiz who discussed the paper at both the "Global Games in Ames" conference in April 2016 and the Cowles Economic Theory conference in June 2016. We are also grateful for discussions with Ben Hebert, Filip Matejka, Chris Sims, Jakub Steiner, Philipp Strack, Tomasz Strzalecki and Mike Woodford.

then argue that the continuous choice property reflects difficulty in distinguishing nearby states.

For the first part, consider a player who must acquire information (an experiment) about a real-valued state of the world and then make a binary choice: "invest" or "not invest". There is a cost $0 < t < 1$ to investing. If the state is positive, there is a return of 1 to investing. If the state is negative, there is no return to investing. Thus it is optimal to invest if the state is positive and not invest if the state is negative. The information choice and binary decision rule will together generate an *effective strategy*: a probability of investing in each state of the world. Only a player's effective strategy - not the experiment and decision rule giving rise to it - will matter in evaluating payoffs. If information were costless, the optimal effective strategy would be a step function (invest exactly when the state is greater than 0, and not otherwise). We will say that an information cost functional - mapping experiments to costs - satisfies continuous choice if the optimal effective strategy is continuous in the state in such discontinuous decision problems.

Now consider a continuum player coordination game where investment gives the return of 1 only if the proportion of others investing exceeds a threshold that is decreasing in the state (a "regime change" game). For any profile of effective strategies, there will be a critical point at which enough players invest in order for there to be an investment return of 1. But any critical point will give rise to a profile of effective strategies that is a best response given that critical point. Combining these two steps, we have a best response mapping from effective strategy profiles to effective strategy profiles, with equilibria corresponding to fixed points. Now assume that the cost functional is *translation insensitive*: translating an effective strategy does not change its cost too fast. Suppose that we start at an equilibrium, and translate the equilibrium effective strategies - say, to the right. The implied critical point will also move to the right, but - if continuous choice and translation insensitivity are satisfied - less than the translation of the strategies, because the threshold is decreasing in the state. This in turn implies that the best response will also move to the right less than the original translation. Based on this logic, one can show that the mapping from critical points to critical points is a contraction and the equilibrium that we started with is unique. This argument breaks if continuous choice fails, because the translation of discontinuous effective strategy profiles is consistent with the implied critical point moving one for one with the translation.

Our results offer a novel perspective on recent work on endogenous information acquisition in coordination games. Szkup and Trevino (2015) and Yang (2015) have considered the case where players can (simultaneously) choose the precision of noisy signals about the state, with the cost increasing in the precision. In this case, a low cost of information will imply that players will acquire signals with high precision. Carlsson and Damme (1993) have shown that in such "global game" environments, a unique equilibrium must then be played. This information structure does give rise to continuous choice and thus we are generalizing these global game results. But in the global game setting, players are restricted to a particular one dimensional class of possible experiments, parameterized by the precision of private information. Conceptually, this gives rise to a couple of problems. First, we do not allow players to decide *where* to pay attention, a margin that we should expect to be of first order importance in this and other economic settings. Second, it is infeasible to choose discontinuous effective strategies. In this sense, uniqueness is assumed rather than implied by optimal information acquisition. In response to this, Yang (2015) considered *flexible* information acquisition, where players can acquire *any* information, and it is always strictly cheaper to acquire information that is strictly less informative in the sense of Blackwell (1953). Yang (2015) used entropy reduction as a flexible cost functional for information, and showed that there are multiple equilibria. However, the entropy reduction cost function fails continuous choice, because the effective strategy must be continuous in the payoff gain regardless of the mapping from states to payoff gains. Thus (under the entropy reduction cost functional) discontinuous decision problems must imply discontinuous choice. This paper incorporates cost functionals which are flexible in the sense described above but are allowed to depend on the distance between states and not just payoffs at those states. Both uniqueness and multiplicity are consistent with flexible information acquisition in our setting, and there is a natural interpretation of which cost functionals give rise to continuous choice and thus uniqueness.

The second part of our results concerns the interpretation of the continuous choice property. We are interested in primitive conditions on the underlying cost functionals that give rise to continuous choice (or not). Given the richness of the set of cost functionals, we do not have a single criterion that characterizes continuous choice. Rather, we report a number of sufficient conditions at different levels of abstraction - all of which share the feature that it is more difficult to distinguish nearby states relative to distant states.

At the most abstract level, we say that a discontinuous effective strategy has a cheap continuous approximation if there is a sequence of continuous effective strategies that approach the discontinuous effective strategy, such that the continuous effective strategies are cheaper than the given discontinuous effective strategy, and that the cost saving from using the approximating continuous effective strategies is large relative to the distance from the discontinuous effective strategy. Now a sufficient condition for continuous choice is that every discontinuous effective strategy has a cheap continuous approximation. This observation gives a natural interpretation of continuous choice in terms of local distinguishability: there is a significant saving from approximating a discontinuous effective strategy with a continuous effective strategy because it is hard to distinguish states at the discontinuity.

This high level characterization is general and formalizes the idea that continuous choice reflects a high cost of distinguishing nearby states. However, the characterization is close to the conclusion. For an alternative approach, we use the slope of an effective strategy as a natural cost of distinguishing states in a small neighborhood. For a more concrete characterization of continuous choice, we then allow a *local cost* to depend in an arbitrary way on both this derivative but also location, because there may be reasons why it is more difficult to distinguish states at different locations; we can allow the local costs to be aggregated in any way (ranging from the average local cost to the maximum local cost); and we can allow costs to be an arbitrary increasing function of this measure. Within this class, we can give tight conditions for continuous choice that depend on three interpretable statistics of the cost functional. Cheap continuous approximation and thus continuous choice will now arise if cost increases slowly enough in the maximum slope.

Another approach to characterizing continuous choice is to take a position on how information is acquired. One hybrid approach is to suppose that there is both a perception cost (a cost generating signals of a certain accuracy) and an attention cost associated with processing those signals. The perception cost guarantees that the cost functional satisfies continuous choice (because no discontinuous effective strategies are feasible). But it will also have the property that players do not acquire any information that they do not use, so that it is flexible in the sense described above. Another model of how information is acquired is that players observe a diffusion whose drift is given by the states and can choose a stopping rule. This gives rise to an implied cost functional (the expected time to stopping), and this gives rise to continuous choice under very weak assumptions.

A final contribution of the paper is to formalize an idea alluded to above: there is a distinction between *assuming* continuous choice, by setting the cost of all discontinuous choice functions to infinity, as implicitly assumed in the global games literature; and establishing continuous choice and thus uniqueness in games in settings where all effective strategies are feasible (i.e., have finite cost) but players nonetheless choose continuous ones. We formalize this by giving a characterization - based on players' higher-order beliefs about the state of the world - of equilibria of the regime change game for a fixed information structure. We can then identify when optimal information acquisition of information implies unique equilibrium even though multiple equilibria could arise under feasible information structures.

Our main result has a partial converse. A cost functional is *Lipschitz* if the cost of changing the effective strategy in a small neighborhood is of the same order of magnitude as that neighborhood. This condition implies discontinuous choices and can be used to establish multiplicity. Our results also extend from regime change games - used for our main result and previously studied by Morris and Shin (1998) among others - to general coordination games.

Our analysis focusses on the case where we fix a cost functional but ask what happens as we let a weight on that functional go to 0. In this case, the cost of all information goes to zero, including the cost of distinguishing nearby states. Thus when we say - in this introduction and in the body of the paper - that it is "particularly difficult to distinguish nearby states," what matters is how difficult it is to distinguish nearby states relative to the cost of far away states: even as the absolute cost of distinguishing nearby states goes to zero, there will be continuous choice and thus equilibrium uniqueness as long as the absolute cost of distinguishing far away states converges to zero faster.

Our results have implications for modelling information acquisition more broadly. Sims (2003) suggested that the ability to process information is a binding constraint, which implies - via results in information theory - that there is a bound on feasible entropy reduction. If information capacity can be bought, this suggests a cost functional that is an increasing function of entropy reduction. But because of its purely information theoretic foundations, this cost function is not sensitive to the labelling of states, and thus it is built in that it is as easy to distinguish nearby states as distant states. Because entropy reduction has a tractable functional form for the cost of information, it has been widely used in economic settings where it does not reflect information processing costs and where the insensitivity to the distance between states does not make sense. While this may not be important in single person decision making, this paper contains a warning about use of entropy as a cost of information in strategic settings.

Our results also have implications for a debate about equilibrium uniqueness without common knowledge. Weinstein and Yildiz (2007) have emphasized that equilibrium selection arguments in the global games literature rely on a particular relaxation of common knowledge (noisy signals of payoffs) and do not go through under other local perturbations from common knowledge. We show that endogenous information acquisition gives rise to uniqueness under some reasonable assumptions about the cost functional.

Our uniqueness result generalizes a result of Carlsson and Damme (1993) on binary action games with exogenous noisy information structures. However, we cannot appeal to the arguments in that and later papers on binary action games because the relevant space of effective strategies cannot be characterized by a threshold. Our results are closer to the argument for uniqueness in general supermodular games in Frankel, Morris, and Pauzner (2003). Here too, translation insensitivity has a crucial role, with contraction like properties giving rise to uniqueness (Mathevet (2008) showed an exact relation to the contraction mapping theorem under slightly stronger assumptions). Mathevet and Steiner (2013) highlighted the role of translation insensitivity in obtaining uniqueness results. All these papers assume noisy information structures and hinge their analysis on the built-in translation insensitivity without noting that continuous choice is also obtained for free in these environments. They studied how translation insensitivity helps pin down the equilibrium strategy while the role of continuous choice is not the focus. In contrast, we show that translation insensitivity leads to limit uniqueness (multiplicity) if continuous choice is satisfied (violated), and thus highlight continuous choice as the essential property that leads to the equilibrium uniqueness.

We proceed as follows. Section 2 sets up the model. Section 3 presents a leading example to illustrate continuous choice, how it depends on the difficulty of distinguishing nearby states and its impact on equilibrium outcomes. Section 4 contains our main result, showing that continuous choice implies uniqueness. In Section 5, we show the relationship between continuous choice and the cost of distinguishing nearby states. Section 6 highlights the difference between establishing uniqueness by restricting attention to information structures implying continuous choice (as in the global games literature) and establishing uniqueness via the optimal choice of information. Section 7 extends the results to coordination games with general payoffs under weaker conditions on the information cost functional, as well as giving a converse to our main result. Section 8 discusses the relation between our general information cost and the entropy-based information cost used in similar settings, allowing players to observe others' actions and evidence on continuous choice. Long proofs are relegated to the appendix.

2 The Model

2.1 Environment

A continuum of players simultaneously choose an action, "not invest" or "invest". The payoff from not investing is normalized to 0. A player's payoff if she invests is $\pi(l, \theta)$, where $l \in [0,1]$ is the proportion of players investing and $\theta \in \mathbb{R}$ is a payoff relevant state. We assume the *regime change* payoffs to be

$$\pi(l, \theta) = \begin{cases} 1 - t, \text{ if } l \geq \beta(\theta) \\ -t, \text{ otherwise} \end{cases}, \tag{1}$$

where $\beta : \mathbb{R} \to \mathbb{R}$ is a continuous and strictly decreasing function; $t \in (0,1)$ is the *investment cost* and $\beta(\cdot)$ is the *threshold function*, so that a player gets a return 1 from investing only if at least proportion $\beta(\theta)$ of other players invest. We assume that $\beta(\theta) > 1$ for small enough θ and $\beta(\theta) < 0$ for large enough θ; without loss of generality, we set $\beta(0) = 1$ and $\beta(1) = 0$. Actions are strategic complements: players are (weakly) more willing to take an action if they expect others to take that action. Players do not know the payoff relevant state θ but do share a *common prior* on θ, denoted by density g. A maintained assumption is that g is continuous and strictly positive on $[0,1]$.[1]

Before selecting an action, players can simultaneously and privately acquire information about θ. We assume that they observe conditionally independent real-valued signals that

[1] Games with this structure of payoffs were studied in Morris and Shin (1998) among many others.

are informative about θ; as always, the labelling of signals does not matter and we are using $\mathbb{R}$ as a signal space to economize on notation. Each player can then pick an *experiment* q, where $q(\cdot|\theta) \in \Delta(\mathbb{R})$ is a probability measure on $\mathbb{R}$ conditional on θ.

We write Q for the space of all experiments. A *cost functional* $C : Q \to [0, \bar{c}]$ maps experiments to a bounded interval of costs. A player incurs a cost $\lambda \cdot C(q)$ if she chooses experiment $q \in Q$. We will hold the cost functional C fixed in our analysis and vary $\lambda \geq 0$, a parameter that represents the difficulty of information acquisition; we will refer to the resulting game as the λ-regime change game. If $\lambda = 0$, the players can choose to observe θ at no cost and the model reduces to a complete information game. We will perturb this complete information game by letting λ be strictly positive but close to zero. Focussing on small but positive λ sharpens the statement and intuition of our results. We will maintain the assumption that the cost functional respects Blackwell's ordering (Blackwell (1953)), so that an experiment that is strictly more informative than another has a (weakly) higher cost.

A player's strategy corresponds to an experiment q together with a decision rule $\sigma : R \to [0,1]$, with $\sigma(x)$ being the probability of investing upon signal realization x. The experiment and decision rule jointly determine the player's *effective strategy*, which is a function $s : \mathbb{R} \to [0,1]$ given by

$$s(\theta) = \int q(x|\theta) \cdot \sigma(x)\, dx \ .$$

That is, $s(\theta)$ is the player's probability of investing conditional on the state being θ.[2] We call s the *effective strategy* since it describes action choices integrating out signal realizations. Any effective strategy s can be viewed as arising from a binary-signal experiment where a player's signal is an action recommendation. Formally, we can identify an effective strategy s with the experiment given by

$$q(x|\theta) = \begin{cases} s(\theta), & \text{if } x = 1 \\ 1 - s(\theta), & \text{if } x = 0 \end{cases}$$

and decision rule σ given by

$$\sigma(x) = \begin{cases} 1, \text{ if } x = 1 \\ 0, \text{ if } x = 0 \end{cases} .$$

Players care only about the effective strategies of other players, but not the experiments or decision rules generating them. Moreover, because the cost functional is weakly increasing in the information content, each player will weakly prefer to acquire a binary-signal experi-

[2] Here, the signal realization x could have full support or discrete support on $\mathbb{R}$. To economize on notation, we use " $\int$ " to refer to both the integration over a continuum of signal realizations and the summation over discrete realizations.

ment corresponding to an effective strategy s. This is standard observation in the rational inattention literature; it is formally stated in this binary action context in Woodford (2008) and Yang (2015). Thus we will identify experiments with effective strategies, unless otherwise stated. We will write S for the set of effective strategies and $c : S \to [0,\ \overline{c}]$ for the cost functional restricted to effective strategies.

We equip the space of effective strategies with the L^1-metric, so that the distance between effective strategies s_1 and s_2 is given by

$$\|s_1, s_2\| = \int_{\mathbb{R}} |s_1(\theta) - s_2(\theta)|\, g(\theta)\, d\theta;$$

and write $B_\delta(s)$ for the set of effective strategies within δ of s under this metric.

We will collect together in Section 7.2 a discussion of how the maintained assumptions in this section and the body of the text can be relaxed.

2.2 Equilibrium

Now writing $[0,1]$ for the continuum of players, a player's ex ante payoff - if she chooses effective strategy s_i and the profile of others' effective strategies is $\{s_j\}_{j\in[0,1]}$ - is given by

$$u\left(s_i, \{s_j\}_{j\in[0,1]}\right) = \int_{\theta} s_i(\theta)\left(1_{\left\{\int s_j(\theta)dj \geq \beta(\theta)\right\}} - t\right) g(\theta)\, d\theta.$$

Definition 1 (Nash Equilibrium) *$\{s_j\}_{j\in[0,1]}$ is a Nash equilibrium of the λ-regime change game if*

$$s_i \in \arg\max_{s\in S} u\left(s, \{s_j\}_{j\in[0,1]}\right) - \lambda \cdot c(s)$$

for each $i \in [0,1]$.

We will be restricting attention to *monotonic* (non-decreasing) effective strategies. This is with loss of generality but is consistent with many applications (e.g., the effective strategy is always monotone in global game models) and allows us to highlight key insights. We write S_M for the set of monotone effective strategies.

Definition 2 (Monotone Nash Equilibrium) *$\{s_j\}_{j\in[0,1]}$ is a monotone Nash equilibrium if it is a Nash equilibrium and each s_j is monotone.*

2.3 The Threshold Decision Problem

A key ingredient of the analysis will be a simple class of "threshold decision problems". Suppose a player must choose an effective strategy s when (i) the cost to investing is t; (ii) there is a payoff 1 if she invests only if the state is at least ψ; and (iii) the information cost of effective strategy s is $\lambda \cdot c(s)$. Thus the payoff from choosing effective strategy s is

$$U(s|\psi,\lambda) = \left[\int s(\theta)\left(1_{\{\theta\geq\psi\}} - t\right) g(\theta)\, d\theta\right] - \lambda \cdot c(s) \ .$$

This decision problem is parameterized by ψ and λ, and we will refer to it as the (ψ,λ)-threshold decision problem. We write $S(\psi,\lambda)$ for the set of optimal monotone effective strategies in the (ψ,λ)-threshold decision problem, i.e.,

$$S(\psi,\lambda) = \underset{s\in S_M}{\arg\max}\; U(s|\psi,\lambda) \ .$$

3 Leading Example

Our leading example will illustrate the key ideas of the paper. We assume that

1. the threshold function β is given by

$$\beta(\theta) = 1-\theta;$$

2. the players' common prior g is the uniform distribution over $\left[\underline{\theta},\overline{\theta}\right]$ where $\underline{\theta} < 0$ and $\overline{\theta} > 1$;

3. the cost functional is given by

$$c(s) = \max\left(0, 1 - \left(\sup_{\theta}\ |s'(\theta)|\right)^{-\gamma}\right) \tag{2}$$

with $\gamma > 0$.

If s is discontinuous, then $s'(\theta)$ is understood to be infinity, and so the cost of any discontinuous s is 1.[3]

Under this cost functional, the cost depends only on the *slope* of the effective strategy and is increasing in the (maximum) slope: a high slope of the effective strategy requires a high ability to distinguish nearby states. The cost of a discontinuous effective strategy is set equal to 1. The cost is reduced if the slope is bounded with maximum $k > 0$. That cost saving is $k^{-\gamma}$, and thus the cost saving approaches 0 as the maximum slope k approaches ∞. A higher γ implies that the cost saving approaches zero faster. Thus - heuristically - a higher γ makes it cheaper to distinguish nearby states.

[3] If $s(\theta)$ is continuous but not differentiable at θ, we can take it to equal the maximum of the left and right derivatives.

3.1 The Threshold Decision Problem

We first solve the (ψ, λ)-threshold decision problem for arbitrary threshold $\psi \in [0,1]$; this analysis will then be an input into our analysis of the game. A player must choose an effective strategy s to maximize

$$U(s|\psi, \lambda) = \frac{1}{\overline{\theta} - \underline{\theta}} \left[\int s(\theta) \cdot \left(1_{\{\theta \geq \psi\}} - t\right) \cdot d\theta \right] - \lambda \cdot c(s). \tag{3}$$

If there was no cost of information (i.e., $\lambda = 0$), a player's optimal effective strategy in the threshold decision problem is the step function $1_{\{\theta \geq \psi\}}$, which perfectly distinguishes the threshold event $\left[\psi, \overline{\theta}\right]$ from its complement, $[\underline{\theta}, \psi)$. For small but positive λ, the optimal effective strategy will always take the form

$$s_{\xi,k}(\theta) = \begin{cases} 0, \text{ if } \theta \leq \xi - \frac{1}{2k} \\ \frac{1}{2} + k(\theta - \xi), \text{ if } \xi - \frac{1}{2k} \leq \theta \leq \xi + \frac{1}{2k} \\ 1, \text{ if } \theta \geq \xi + \frac{1}{2k} \end{cases} \tag{4}$$

for some $\xi \in \mathbb{R}$ and $k \in \mathbb{R}_+ \cup \{\infty\}$. A typical payoff gain function $1_{\{\theta \geq \psi\}} - t$ and effective strategy $s_{\xi,k}$ in this form are illustrated in the Figure 3.1 below.

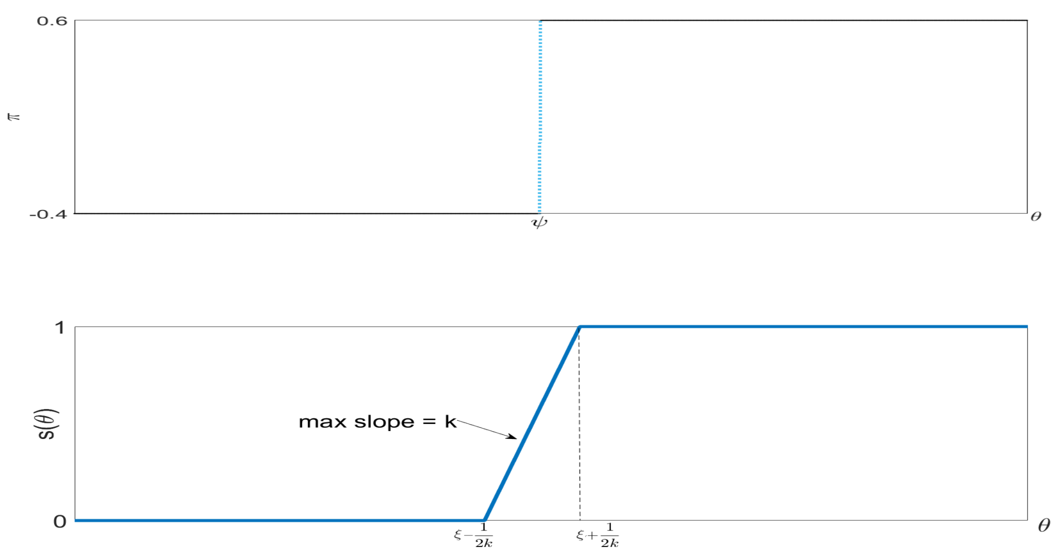

Figure 3.1: optimal effective strategies

Thus the effective strategy will take values 0 or 1 except in a region centered around a "cutoff" ξ where it will be linearly increasing with slope k. The step function $1_{\{\theta \geq \psi\}}$ is within this class, since it corresponds to $s_{\psi,\infty}$. If the optimal effective strategy had a finite maximum slope k, and the optimal strategy s did not take this form, then the effective strategy $s_{\xi,k}$ with

$$\xi = \psi - \frac{1}{k}\left(s(\psi) - \frac{1}{2}\right)$$

would be at least as good, since it is lower than s to the left of ψ, higher than s to the right of ψ and has the same cost.[4]

The cutoff ξ determines the "position" of the effective strategy. Because the cost functional has the property that $c(s_{\xi,k})$ is independent of ξ, we can first solve for ξ, taking k as given. Now

$$\frac{1}{\overline{\theta} - \underline{\theta}} \int s_{\xi,k}(\theta) \cdot \left(1_{\{\theta \geq \psi\}} - t\right) \cdot d\theta$$

is maximized - for any given k - when the cutoff ξ is set as

$$\xi = \psi + \frac{1}{k}\left(t - \frac{1}{2}\right). \tag{5}$$

By substituting (4) and (5) into (3), we obtain the following expression for the total payoff as a function of k:

$$\frac{1}{\overline{\theta} - \underline{\theta}}\left[\left(\overline{\theta} - \psi\right)(1-t) - \frac{1}{2}\left(t - t^2\right) \cdot k^{-1}\right] - \lambda \cdot \max\left(0, 1 - k^{-\gamma}\right) .$$

Simple calculations show that this expression is maximized by setting the maximum slope k equal to

$$k^* = \begin{cases} \left[\frac{t(1-t)}{2\lambda\gamma\left(\overline{\theta} - \underline{\theta}\right)}\right]^{\frac{1}{1-\gamma}}, \text{ if } \gamma < 1 \\ \infty, \text{ if } \gamma \geq 1 \end{cases} . \tag{6}$$

Thus the optimal effective strategy in the (ψ, λ)-threshold decision problem is s_{ξ^*,k^*} where k^* is given by (6) and ξ^* is given by

$$\xi^* = b(\psi) \triangleq \psi + \frac{1}{k^*}\left(t - \frac{1}{2}\right) . \tag{7}$$

Since k^* will not be varying in our analysis going forward, we will refer to s_{ξ,k^*} as the "ξ-cutoff strategy" because this strategy is centered on ξ.

We now discuss how the optimal effective strategy depends on γ. When $\gamma \geq 1$, the optimal slope is $k^* = \infty$ and the optimal cutoff is ψ, and the optimal effective strategy is then the step function $1_{\{\theta \geq \psi\}}$, which is discontinuous at the threshold ψ. The cost saving $k^{-\gamma}$ from replacing the step function by a continuous effective strategy with slope $k < \infty$

[4] This is true if λ is sufficiently small. If λ were sufficiently large, optimal effective strategies might have $s(\theta) \in (0,1)$ when θ is close to $\overline{\theta}$ or $\underline{\theta}$. A sufficient condition to rule this out is that $\lambda < \frac{t(1-t)}{2(\overline{\theta} - \underline{\theta})} \min\left(1, \left[\frac{\overline{\theta}-1}{t}\right]^{\max(0,1-\gamma)}, \left[\frac{-\underline{\theta}}{1-t}\right]^{\max(0,1-\gamma)}\right)$.

is too small to compensate the sacrificed benefit, and the player chooses the step function that sharply distinguishes states below ψ from those above it. In this case, it is relatively cheap to distinguish nearby states. But when $\gamma < 1$, the cost saving is large enough and the optimal effective strategy becomes continuous no matter how small $\lambda > 0$ is. The player chooses not to sharply distinguish any event $\left[\psi, \overline{\theta}\right]$ from its complement. It is worth noting that for sufficiently small λ whether the optimal effective strategy is continuous or discontinuous does not depend on λ, which controls the overall difficulty of information acquisition. The (dis)continuity is purely determined by γ and hence is a property of the cost functional.

This derivation of the optimal effective strategy links continuous choice and the relative cost of distinguishing nearby states. In Section 5, we characterize continuous choice within a rich parametric generalization of this cost functional and also provide further foundations for continuous choice.

3.2 Solving for Equilibrium

We now study the equilibrium in the regime change game for fixed λ. First observe that because players are choosing monotone effective strategies, there will be a regime change threshold ψ such that the regime will survive only if the state is above that threshold ψ. But now the analysis in the previous section implies that each player will then follow the ξ-cutoff strategy (s_{ξ,k^*}) with $\xi = b\left(\psi\right)$ given by (7). It remains to identify when the regime will survive if all players follow a ξ-cutoff strategy.

To answer this, observe that - assuming a continuum law of large numbers - if all players follow the ξ-cutoff strategy, the ξ-cutoff strategy will also describe the proportion of players investing as a function of the state, so that the proportion investing in state θ will be $s_{\xi,k^*}\left(\theta\right)$. Now the regime will survive only if the state is greater than a regime change threshold ψ, where ψ solves

$$s_{\xi,k^*}\left(\psi\right) = \frac{1}{2} + k^*\left(\psi - \xi\right) = 1 - \psi \ ;$$

re-arranging gives

$$\psi = \widehat{\psi}\left(\xi\right) \triangleq \frac{\frac{1}{2} + k^*\xi}{1 + k^*} \ . \tag{8}$$

Now an equilibrium will correspond to any (ψ, ξ) solving (7) and (8), where players follow the ξ-cutoff strategy and the regime threshold is ψ. Then an equilibrium regime change threshold must be a solution to

$$\begin{aligned}\psi &= \widehat{\psi}\left(b\left(\psi\right)\right) \\ &= \frac{\frac{1}{2}+k^*\left(\psi+\frac{1}{k^*}\left(t-\frac{1}{2}\right)\right)}{1+k^*} \\ &= \frac{k^*\psi+t}{1+k^*}\ .\end{aligned} \tag{9}$$

If $\gamma \geq 1$ and thus $k^* = \infty$, any $\psi \in [0,1]$ is a fixed point, and thus there is a continuum of equilibria corresponding to regime change thresholds $\theta^* \in [0,1]$: for any $\theta^* \in [0,1]$, given that $k^* = \infty$, the equilibrium strategy will be $1_{\{\theta \geq \theta^*\}}$.

But if $\gamma < 1$ and thus $k^* < \infty$, then (9) is a contraction with a unique solution

$$\theta^* = t\ , \tag{10}$$

and so the equilibrium strategy is centered on

$$\xi^* = t + \frac{1}{k^*}\left(t - \frac{1}{2}\right)\ . \tag{11}$$

To summarize, we have shown that when λ is sufficiently small, there are two cases to consider in solving equilibria. When $\gamma \geq 1$, for any $\theta^* \in [0,1]$, there exists an equilibrium where each player takes the effective strategy $1_{\{\theta \geq \theta^*\}}$ and there is regime change if $\theta \geq \theta^*$. When $\gamma < 1$, there exists a unique equilibrium in which each player takes the effective strategy s_{ξ^*,k^*} and the regime change threshold is θ^*, where ξ^*, k^* and θ^* are given by (11), (6) and (10), respectively. As $\lambda \to 0$, (6) implies that k^* tends to infinity, and the equilibrium threshold ξ^* converges to t.

This analysis of equilibrium in the example illustrates the property that there is limit uniqueness (a unique equilibrium for small enough λ) if a continuous choice rule is optimal in the threshold decision problems. This observation is generalized in Section 4, where it is shown that any cost functional giving rise to continuous choice will give rise to a unique equilibrium with the same equilibrium threshold t. The significance of this particular threshold (dubbed "Laplacian") will be explained there.

3.3 Interpretation of the Result

The straightforward interpretation of this result is that each player's effective strategy arises from a binary-signal experiment, where a player acquires an experiment with two possible values, say 0 and 1, and

$$q(x|\theta) = \begin{cases} s_{\xi^*,k^*}(\theta) & \text{if } x = 1 \\ 1 - s_{\xi^*,k^*}(\theta) & \text{if } x = 0 \end{cases};$$

and a decision rule specifying investment only if the signal is 1.

However, an alternative interpretation of the example gives a tight connection to the global games literature. Suppose that each player can observe a noisy signal of the state, where the signal is equal to the true state plus noise. In particular, suppose that the signal is distributed uniformly on the interval $\left[\theta - \frac{1}{2k}, \theta + \frac{1}{2k}\right]$. A higher value of k corresponds to more accurate information and in the limit - as $k \to \infty$ - players have perfect information about the state. If a player invests if and only if her signal is greater than cutoff ξ, her effective strategy is $s_{\xi,k}$ as defined above in equation (4). But now suppose that a player can choose the accuracy of her information: in particular, suppose that the cost of acquiring information with accuracy k is $\lambda \cdot \max(0, 1 - k^{-\gamma})$. Note that this implies that players are able to acquire perfect information at finite cost λ. Now the player faces exactly the (ψ, λ)-threshold decision problem, except that she is constrained to choose effective strategies of the form $s_{\xi,k}$. This discussion shows that the global game noisy signal model pins down the shape of optimal effective strategies but that if we fix the shape, the analysis is the same.

However, there is an important difference in interpretation. In the uniform noise interpretation of this example, there would have been a unique equilibrium even if players had made chosen sub-optimal choices of accuracy, as long as they did not acquire perfect information. In this sense, the uniqueness arises due to the assumed noisy information structure. On the other hand, with the binary signal interpretation, it is feasible for the players to choose information that does not lead to unique equilibrium, at a finite cost (approaching 0 as $\lambda \to 0$). But they choose not to do so. In this sense, limit uniqueness arises for general information structures via optimality whereas in the global games literature, it arises by assumption. This point is formalized in Section 6.

4 Continuous Choice and Laplacian Selection in Coordination Games

Optimal play in a game depends on what players think that other players are doing. In a symmetric binary action game, Morris and Shin (2003) defined the *Laplacian action* to be the best response to a uniform belief over the proportion of other players choosing each action. In the regime change game, if the state is θ and a player has a uniform belief over the proportion of other players investing, then she assigns probability $1 - \beta(\theta)$ to enough players investing to give rise to regime change. Thus invest is the Laplacian action if

$$1 - \beta(\theta) - t \geq 0$$

or

$$\theta \geq \beta^{-1}(1-t).$$

We want to identify conditions on the cost functional under which the Laplacian action is always played, in the limit as $\lambda \to 0$.

Definition 3 (Laplacian Selection) *Cost functional $c(\cdot)$ satisfies Laplacian selection if, for any $\delta > 0$, there exists $\overline{\lambda} > 0$ such that $\left\| s, 1_{\{\theta \geq \beta^{-1}(1-t)\}} \right\| \leq \delta$ whenever s is a monotone equilibrium strategy in the λ-regime change game and $\lambda \leq \overline{\lambda}$.*

As in the leading example, we first analyze behavior in the (ψ, λ)-threshold decision problem, showing that it is optimal for players to choose strategies that are close to a step function at ψ when the cost of information is small.

Lemma 4 (Optimal Effective Strategies in the Threshold Decision Problems) *The essentially unique monotone optimal effective strategy if $\lambda = 0$ is a step function at ψ, i.e.,*

$$S(\psi, 0) = \left\{1_{\{\theta \geq \psi\}}\right\}.$$

For any $\delta > 0$, there exists $\overline{\lambda} > 0$ such that $S(\psi, \lambda) \subset B_\delta\left(1_{\{\theta \geq \psi\}}\right)$ for all $\psi \in [0,1]$ and $\lambda \leq \overline{\lambda}$.

Proof. When $\lambda = 0$, it is straightforward to see that the player chooses $s(\theta) = 1$ if $1_{\{\theta \geq \psi\}} - t > 0$ and $s(\theta) = 0$ if $1_{\{\theta \geq \psi\}} - t < 0$. Hence, $S(\psi, 0) = \left\{1_{\{\theta \geq \psi\}}\right\}$.

Now consider the case of $\lambda > 0$. Recall that $\overline{c}$ is an upper bound on the cost of an experiment. For any $s \in S(\psi, \lambda)$ and $s \neq 1_{\{\theta \geq \psi\}}$, the optimality of s implies

$$\begin{aligned}
\int_{-\infty}^{\infty} \left[1_{\{\theta \geq \psi\}} - t\right] \cdot \left[1_{\{\theta \geq \psi\}} - s(\theta)\right] g(\theta)\, d\theta &< \lambda \cdot \left[c\left(1_{\{\theta \geq \psi\}}\right) - c(s)\right] \\
&\leq \lambda \cdot \overline{c}\,.
\end{aligned}$$

But

$$\int_{-\infty}^{\infty} \left[1_{\{\theta\geq\psi\}} - t\right] \cdot \left[1_{\{\theta\geq\psi\}} - s(\theta)\right] g(\theta)\, d\theta$$

$$\begin{aligned}
&= t \cdot \int_{-\infty}^{\psi} s(\theta)\, g(\theta)\, d\theta + (1-t) \cdot \int_{\psi}^{\infty} [1 - s(\theta)]\, g(\theta)\, d\theta \\
&\geq \min(t, 1-t) \cdot \left[\int_{-\infty}^{\psi} s(\theta)\, g(\theta)\, d\theta + \int_{\psi}^{\infty} [1 - s(\theta)]\, g(\theta)\, d\theta\right] \\
&= \min(t, 1-t) \cdot \left\| 1_{\{\theta\geq\psi\}}, s \right\| \ .
\end{aligned}$$

Now the above two inequalities imply

$$\left\| 1_{\{\theta\geq\psi\}}, s \right\| < \frac{\lambda \cdot \ \bar{c}}{\min(t, 1-t)} \ . \tag{12}$$

Hence for any $\delta > 0$, $\left\| 1_{\{\theta\geq\psi\}}, s \right\| < \delta$ if $\lambda < \overline{\lambda} = \frac{\delta \cdot \min(t,1-t)}{\bar{c}}$. ■

The fact that the decision maker's optimal effective strategies approximates $1_{\{\theta\geq\psi\}}$ as $\lambda \to 0$ reflects her motive to sharply identify event $\{\theta \geq \psi\}$ from its complement. In a decision problem, whether this is achieved by a continuous or discontinuous $s \in S(\psi, \lambda)$ is not important, since the loss caused by deviating from $1_{\{\theta\geq\psi\}}$ is of the order of magnitude of $\left\| 1_{\{\theta\geq\psi\}}, s \right\|$. In contrast, in the game considered here, the continuity of s is crucial in determining the equilibrium outcomes. Our first key property of the cost functional is continuous choice: under that cost functional - optimal effective strategies are always absolutely continuous whenever $\lambda > 0$.

Definition 5 (Continuous Choice) *Cost functional $c(\cdot)$ satisfies continuous choice if all optimal strategies are absolutely continuous, i.e., $S(\psi, \lambda)$ consists only of absolutely continuous functions, for all $\psi \in [0, 1]$ and $\lambda \in \mathbb{R}_{++}$.*

In Section 5, we will identify sufficient conditions for continuous choice as well as sufficient conditions for its failure, and argue that the sufficient conditions for continuous choice reflect a high relative cost of distinguishing nearby states. Here we study the implications of this assumption.

A profile of effective strategies $\{s_i\}_{i\in[0,1]}$ will induce an *aggregate effective strategy*

$$\widehat{s}(\theta) = \int_{i\in[0,1]} s_i(\theta)\, di \tag{13}$$

which can be interpreted, assuming a continuum law of large numbers, as the proportion of players that invest conditional on the state being θ. If all individual effective strategies are monotone, then so is the aggregate effective strategy.

Now a profile of monotone effective strategies $\{s_i\}_{i\in[0,1]}$ will induce a unique threshold $\psi \in [0,1]$ such that $\widehat{s}(\theta) > \beta(\theta)$ for $\theta > \psi$ and $\widehat{s}(\theta) < \beta(\theta)$ for $\theta < \psi$. Thus an aggregate effective strategy gives rise to an event in the payoff state space

$$F_\psi = \{\theta \in \mathbb{R} : \theta \geq \psi\} \ .$$

We will call this a regime change event since it characterizes the set of states where there is regime change (i.e., the proportion investing exceeds $\beta(\theta)$). Now any player's opponents' strategies are summarized by a threshold ψ. Hence, her optimal best response is equivalent to maximizing $U(s|\psi,\lambda)$.

Lemma 6 *A strategy profile $\{s_i\}_{i\in[0,1]}$ of monotone strategies is an equilibrium of the λ-regime change game if they induce a threshold θ^* such that each strategy is optimal in the (θ^*,λ)-decision problem, i.e., each $s_i \in S(\theta^*,\lambda)$.*

The proof is straightforward and hence omitted.

A second important condition for our main result concerns how costs vary as we translate the effective strategy. Let $T_\Delta : S \to S$ be a translation operator: that is, for any $\Delta \in \mathbb{R}$ and $s \in S$,

$$(T_\Delta s)(\theta) = s(\theta + \Delta) \ .$$

Definition 7 (Translation Insensitivity) *Cost functional $c(\cdot)$ satisfies translation insensitivity if there exists $K > 0$ such that, for all s, $|c(T_\Delta s) - c(s)| < K \cdot |\Delta|$.*

This property requires that the information cost responds at most linearly to translations of the effective strategies. Translation insensitivity captures the idea that the cost of information acquisition reflects the cost of paying attention to some neighborhood of the state space, but is not too sensitive to where attention is paid. Now we have:

Proposition 8 (Laplacian Selection) *If $c(\cdot)$ satisfies continuous choice and translation insensitivity, then $c(\cdot)$ satisfies Laplacian selection.*

Thus when $c(\cdot)$ satisfies continuous choice and translation insensitivity, and information costs are low, all equilibria are close to the Laplacian switching strategy.[5] The property of continuous choice is essential to the limit uniqueness result. Recall that a player's ideal strategy is to sharply identify the event of regime change whenever it occurs. This requires

[5] The continuous choice property and the translation insensitivity can be relaxed to local versions: they only need to hold in a small neighborbood of the step functions. See Subsection 7.2 for the formal definitions and results.

perfectly distinguishing the states above the threshold of regime change from those below it, calling for an effective strategy discontinuous at the threshold. The property of continuous choice means that the players will never choose such sharp strategies for any relevant decision problems (i.e., (ψ, λ)-threshold decision problem with $\psi \in [0,1]$). This property limits the players' ability to coordinate in acquiring information except on the regime change event $F_{\beta^{-1}(1-t)}$.

5 Foundations for Continuous Choice and the Local Cost of Distinguishing Nearby States

Continuous choice captures the idea that it is relatively hard to distinguish states that are close together: it will not be optimal to pay for discontinuities in the effective strategies if those discontinuities are expensive. But continuous choice is a property that is endogenous to the threshold decision problem. In this section, we discuss alternative ways of interpreting continuous choice in terms of the underlying properties of the cost functional and identify conditions on cost functionals that are sufficient for (or necessary for) the continuous choice property.

5.1 Continuous Approximation and the Relative Cost of Distinguishing Nearby States

We first provide a characterization of continuous choice in terms of the cost of approximating a discontinuous effective strategy. A *continuous approximation* of an effective strategy s that is not absolutely continuous is a sequence of absolutely continuous effective strategies $\{s^n\}_{n=1}^{\infty}$ with

$$\lim_{n\to\infty} \|s, s^n\| = 0 \ .$$

The approximation is *cheap* if, first, $c(s^n) < c(s)$ for all n; and, second,

$$\lim_{n\to\infty} \frac{c(s) - c(s^n)}{\|s, s^n\|} = \infty \ .$$

That is, choosing s^n instead of s, gives a non-negative cost saving $c(s) - c(s^n)$ and this saving relative to the degree of approximation becomes arbitrarily large. Note that this is a property of cost functional alone.

Lemma 9 (continuous choice) *If every effective strategy that is not absolutely continuous has a cheap continuous approximation under $c(\cdot)$, then $c(\cdot)$ satisfies continuous choice.*

Proof. Suppose $s \in S(\psi, \lambda)$ is not absolutely continuous but has a cheap continuous approximation. Then we can find an absolutely continuous $\widetilde{s}$ such that

$$\|\widetilde{s}, s\| < \lambda \cdot [c(s) - c(\widetilde{s})].$$

Then, the gain from replacing s by $\widetilde{s}$ is

$$\begin{aligned} & U(\widetilde{s}|\psi,\lambda) - U(s|\psi,\lambda) \\ = & \int [\widetilde{s}(\theta) - s(\theta)] \cdot \left[1_{\{\theta \geq \psi\}} - t\right] g(\theta)\, d\theta + \lambda \cdot [c(s) - c(\widetilde{s})] \\ > & -\int |\widetilde{s}(\theta) - s(\theta)| \cdot 1 \cdot g(\theta)\, d\theta + \|\widetilde{s}, s\| \\ = & \ 0, \end{aligned}$$

which contradicts the optimality of s. ■

This characterization of continuous choice is essentially a reinterpretation of the property. We now use this characterization in a more substantive characterization.

5.2 Averaging Slopes

As suggested by our leading example, the slope $k(\theta)$ at state θ of an effective strategy is a natural measure of how much attention is devoted to distinguishing states in the neighborhood of that point. We now consider a rich class of cost functionals based on this natural measure. Suppose that there is a local cost which is increasing in the slope k: let $h(\cdot)$ be a strictly increasing function with $h(0) = 0$; in order to characterize behavior for large k, we assume $\lim_{k\to\infty} k^{-r} \cdot h(k)$ exists and is strictly positive for some $r > 0$. The interpretation is that $h(k)$ is a cost of the local attention when the derivative is k, and r measures how fast h increases for large k. Suppose that local cost can also depend on the location, because it is easier to distinguish states on different regions of the state space: let the location cost be linear in $w : \mathbb{R} \to \mathbb{R}_{++}$. Together, h and w give rise to a local cost of effective strategy s at state θ:

$$\kappa_{w.h}[s](\theta) = w(\theta)\, h(s'(\theta)) \ .$$

To aggregate the local cost at different states, we take an L^p-norm of $\kappa_{w.h}[s]$ and further identify a cost functional

$$c_{w,h,H,p}(s) = H\left(\|\kappa_{w.h}[s]\|_p\right) \ ,$$

where $p > \frac{1}{r}$ and $H(\cdot)$ is an increasing function.[6] Now we have a cost functional characterized by the number p and the functions w, h and H. If H and thus $c_{w,h,H,p}$ is not bounded

[6] Note that $\left\|\kappa_{w.h}\left[s_{\xi,k}\right]\right\|_p$ increases in slope k if and only if $p > \frac{1}{r}$. We thus assume $p > \frac{1}{r}$ to make sure that sharper effective strategies cost more.

above, it is trivial that it will satisfy cheap continuous approximation. The interesting case is if H and thus $c_{w,h,H,p}$ have an upper bound. Recall that this was a maintained assumption in our earlier analysis. For this characterization, we write

$$\overline{c} = \lim_{x \to \infty} H(x)$$

for the least upper bound. To characterize behavior for large x, we assume that $\lim_{x \to \infty} x^{\gamma} \cdot [\overline{c} - H(x)]$ exists and is strictly positive for some $\gamma > 0$.

Proposition 10 *Cost functional $c_{w,h,H,p}$ satisfies continuous choice if*

$$\frac{1}{\gamma} + \frac{1}{p} > r \ . \tag{14}$$

Thus slow growth of costs as they approach the upper bound favors continuous choice. This is directly generated by slow growth of h (i.e., small r) and slow growth of H (i.e., small γ). If $p = \frac{1}{r}$, the p-norm of $\kappa_{w.h}[s]$ is roughly the simple average of s' and thus independent of slopes. Thus as p approaches $\frac{1}{r}$ from above, the growth of costs becomes slower. Thus smaller values of γ, r and p all favor continuous choice. We can establish a partial converse when $p \to \infty$; that is, if $c_{w,h,H,p=\infty}$ satisfies continuous choice then $\frac{1}{\gamma} > r$. The proof strategy is similar to that of Proposition 10 and omitted here. Following the same intuition, we expect this converse to hold for all p, although some additional technical assumptions might be required.

This proposition generalizes the analysis in the leading example. Recall that we studied the cost functional

$$\max\left(0, 1 - \left(\sup_{\theta} \ s'(\theta)\right)^{-\gamma}\right) \ .$$

This is a special case of the cost functional above where h is the identity function (so that $r = 1$), w is a constant function, $p \to \infty$, and

$$H(x) = \max\left(0, 1 - x^{-\gamma}\right) .$$

Note that the cost functionals of this section will fail translation insensitivity if the function w is not constant. However, they will satisfy a local translation insensitivity property (Definition 18) in Subsection 7.2); we show there that this version can be substituted for translation insensitivity in our main result.

5.3 Perception and Attention

This subsection relates the continuous choice to perceptual constraints and information processing costs. The noisy signal information structure in global games can be motivated by perceptual constraints. The entropy cost function of information can be motivated by information processing constraints. It is natural to incorporate both of these constraints into the cost function.[7] As long as there are perceptual constraints (i.e., the cost for perfect perception is unbounded), it is immediate that the cost functional will satisfy the continuous choice property. But by including information processing costs, we capture the idea that players will have no more information than they use, and so binary signals in our application. In this section, we formally describe a cost functional of this form.

Let $z = \theta + \sigma \cdot \varepsilon$ be an information source that a player can pay attention to, where ε is a continuous random variable with a density f. We assume f' exists and is L^1-integrable. Without loss of generality, let the mean of ε be zero. The player can choose the standard deviation σ of her information source z (investing in the precision of perception) and then acquire information about θ only through paying attention to z. The cost of acquiring an information source with standard deviation σ and acquiring I bits of information is $h(\sigma, I) \geq 0$; we assume that (i) h is continuous in σ and I; (ii) $h(\sigma, I) = 0$ if and only if $\sigma = \infty$ and $I = 0$; (iii) for any fixed I, $h(\sigma, I)$ is strictly decreasing in σ with $\lim_{\sigma \to 0} h(\sigma, I) = \infty$; and (iv) $h(\sigma, I)$ is strictly increasing in I . All these assumptions are standard for information costs. In point (iii), we drop the maintained assumption on the upper bound of information costs motivated by the view that perception error is unavoidable.

A player's strategy now consists of a decision rule σ, the amount of perception error, and the choice of attention allocation. As noted earlier, we can without loss of generality restrict attention allocation choice to be represented by a function $m : \mathbb{R} \to [0, 1]$, with $m(z) = \Pr(\text{invest}|\ z)$. The amount of information, measured by the reduction of entropy, is given by

$$I = \mathbf{E}\rho(m(z)) - \rho(\mathbf{E}(m(z))) \ , \tag{15}$$

where

$$\rho(m) = m \ln m + (1 - m) \ln(1 - m) \, .$$

Then the effective strategy induced by σ and m is

$$s(\theta) = \sigma^{-1} \int m(z) f\left(\frac{z - \theta}{\sigma}\right) dz \ .$$

Hence, the set of feasible effective strategies is

$$S_f = \left\{ s : \exists\ \sigma \text{ and } m \text{ s.t. } s(\theta) = \sigma^{-1} \int m(z) f\left(\frac{z - \theta}{\sigma}\right) dz \right\} \ .$$

[7] We are grateful to Chris Sims for discussion about this combination.

and the information cost functional is given by

$$c_{h,f}(s) = \begin{cases} h(\sigma, I), & \text{if } s \in S_f \\ \infty, & \text{otherwise} \end{cases},$$

where I is given by (15).

Proposition 11 *For any f and h satisfying the above properties, the cost functional $c_{h,f}$ satisfies continuous choice.*

If we allowed the cost function to be bounded (i.e., relaxed assumption (iii)), we would have continuous choice depend on how quickly the upper bound was approached as continuous approximations, as in the previous sub-section.

5.4 Drift Diffusion[8]

One specific model of information acquisition is that players learn by observing a drift diffusion process that is informative about the state. Any stopping rule will then give rise to an experiment. It is natural to identify the (ex ante) cost of the experiment with the expected stopping time. We do not in general know which experiments could be derived this way, nor - in this continuum state case - if there going to be simple expression for the cost functional.

Nonetheless, one can provide direct arguments that there will be continuous choice in some settings, as shown by Strack (2016). Suppose that a diffusion has a drift given by the state. A player will stop and make a binary choice whenever the process hits a time dependent barrier. Suppose that the stopping time is bounded whenever the barrier is uniformly bounded. Then - for any action - the ex ante probability that that action will be played is continuous in the state. This is true without any additional assumptions about the stopping rule. In particular, it will occur if stopping times are chosen optimally (as long as the uniform stopping time property holds).

5.5 General Sequential Information Acquisition

Drift diffusion with stopping times is one particular model of sequential sampling of information. Hebert and Woodford (2016) have considered a sequential model where there is a general model of well-behaved learning at each time. They then study what happens as the time period and information increment both get shorter. They show that in the continuous time limit there is a cost functional of information acquisition which is pinned down by a matrix specifying the costs of pairwise distinguishability of state. In a continuous state limit of their finite state model, we would have a natural parameterized class of cost

[8] We are grateful to Philipp Strack for discussion of the material in this section.

functionals. This would provide a foundation for a class of cost functionals that could be tested for continuous choice.

6 Feasibility and Optimality of Information Acquisition

When there is complete information - or common certainty of payoffs - in a coordination game like the regime change game, there are often multiple equilibria. The global games literature (Carlsson and Damme (1993) and Morris and Shin (1998)) has shown that if players observe noisy signals about payoffs - and thus common certainty about payoffs is relaxed - then there is equilibrium uniqueness. In the context of the continuum player symmetric binary actions games studied in this paper, there is also a particular action played in the limit: the Laplacian action discussed at the beginning of Section 4. This is true no matter how small the variance of the noise (as long as it is positive); and it remains true if the variance of the noise is a choice variable for the players (Szkup and Trevino (2015) and Yang (2015)). These observations have sometimes been crudely summarized by the claim that relaxing common certainty of payoffs gives equilibrium uniqueness and Laplacian selection.

But it matters how common certainty of payoffs is relaxed. Weinstein and Yildiz (2007) show that for any (interim correlated) rationalizable action, there is a small perturbation of beliefs and higher-order beliefs (formally, close in the product topology on the universal space of belief types of Mertens and Zamir (1985)) such that that action is uniquely rationalizable and thus played in any equilibrium. Unifying the two literatures, Morris, Shin, and Yildiz (2016) have identified the restriction on higher order beliefs under which the Laplacian selection occurs, highlighting not only what drives the usual global game selection in terms of universal types but also why restrictions stronger than closeness in the product topology are required.

These results taken together beg two questions. If players chose their experiments, would the resulting information structure imply Laplacian selection? And, if yes, are there feasible information structures that would not have led to Laplacian selection? The purpose of this section is to answer these questions. We will argue that in the global game literature, with or without endogenous choice of precision of signals, players are constrained to choose information structures giving rise to Laplacian selection. In contrast, in our analysis, there exist feasible and finite-cost information structures that are consistent with any selection, but the endogenous choice of equilibrium information structures gives rise to Laplacian selection. Thus we show that the usual global game selection of Laplacian play in our analysis is a consequence of endogenous information acquisition but not a consequence of the special information structures allowed in the global game analysis. We will illustrate this using our leading example.

In order to establish these results, we must first give a characterization of equilibria in regime change games with arbitrary information structures. Thus fix an arbitrary information structure, i.e., a profile of players' experiments $\{q_i\}_{i\in[0,1]}$. It is convenient to characterize equilibria in terms of *regime change events*, i.e., sets of states where the regime changes. An event $F \subseteq \mathbb{R}$ will be an equilibrium regime change event if F is the set of states with the property that at least proportion $\beta(\theta)$ of players assign probability at least t to that event.

More formally, for a fixed information structure $\{q_i\}_{i\in[0,1]}$, let $\widetilde{q}_i(\cdot|x_i) \in \Delta(\mathbb{R})$ denote player i's posterior belief upon observing signal x_i. Suppose that F was an equilibrium regime change event. Player i would invest if and only if she observed a signal x_i with

$$\widetilde{q}_i(F|x_i) \geq t.$$

Thus the probability that i would invest conditional on θ would be

$$q_i(\{x_i : \widetilde{q}_i(F|x_i) \geq t\} |\theta).$$

By a continuum law of large numbers assumption, the proportion of players investing at θ would be

$$\int_{i\in[0,1]} q_i(\{x_i : \widetilde{q}_i(F|x_i) \geq t\} |\theta)\, di \ .$$

So the regime would change if this expression were greater than $\beta(\theta)$. Writing

$$B^{t,\beta}(F) = \left\{ \theta \in \mathbb{R} \left| \int_{i\in[0,1]} q_i(\{x_i : \widetilde{q}_i(F|x_i) \geq t\} |\theta)\, di \geq \beta(\theta) \right. \right\},$$

we have the following:

Proposition 12 *For a given information structure $\{q_i\}_{i\in[0,1]}$, there is an equilibrium strategy profile where the regime changes on event F if and only if $B^{t,\beta}(F) = F$. In this case, we say that F is an equilibrium regime change event under information structure $\{q_i\}_{i\in[0,1]}$.*

The proof is straightforward and hence omitted here. In a previous version of this paper Morris and Yang (2016), we gave a characterization of equilibrium regime change events in terms of players' beliefs and higher order beliefs about the states θ. This exercise is analogous to the common belief foundations of global game result reported in Morris, Shin, and Yildiz (2016).[9]

[9] This exercise is less general as it relies on the one dimensional state space and continuum player assumption to get a common belief characterization, although it gives a common belief characterization for a different (regime change) game.

Now we introduce information acquisition. As before, write F_ψ for the event

$$F_\psi = \{\theta \,|\theta \geq \psi\}.$$

Say that a regime change event F_ψ is *feasible* if it is technologically feasible to choose an information structure under which F_ψ is an equilibrium regime change event. It is *incentive compatible* if there is a Nash equilibrium of the information acquisition game giving rise to an information structure under which F_ψ is an equilibrium regime change event. By definition, an F_ψ must be feasible if it is incentive compatible, and could be feasible but not incentive compatible.

Proposition 13 (Feasilibility Under Binary Information Structures) *Any threshold event F_ψ with threshold $\psi \in [0,1]$ is feasible if all players choose a binary-signal experiment*

$$q(x|\theta) = \begin{cases} s_{\xi,k}(\theta) & \text{if } x = 1 \\ 1 - s_{\xi,k}(\theta) & \text{if } x = 0 \end{cases},$$

where $s_{\xi,k}$ is defined by (4) with

$$\xi = \left(1 + \frac{1}{k}\right)\psi - \frac{1}{2k}$$

and k sufficiently large.

Proposition 13 states that by acquiring information appropriately, it is feasible for the players to let any threshold event F_ψ be an equilibrium regime change event. The players can choose a binary-signal experiment (while the signals are still conditionally independent across players) that focuses around the threshold ψ to distinguish F_ψ from its complement sharply (i.e., k large enough). This seems to be a very natural result for binary-signal experiments.

This result about the feasibility of attaining equilibrium regime change events can be contrasted with our earlier results about the incentive compatibility of information choices. Recall the leading example and Proposition 13. When $\gamma \geq 1$, any threshold event F_ψ with $\psi \in [0,1]$ is feasible and also incentive compatible. In contrast, when $\gamma < 1$, by Proposition 13 it remains feasible for the players to coordinate on equilibrium regime change event F_ψ with $\psi \in [0,1]$, but our analysis in the leading example shows that it is not incentive compatible for them to do so in the limit unless $\psi = t$.

We can contrast the results for binary-signal experiments with those in a global game model. The setting follows Subsection 3.3. Each player i's experiment consists of a signal $z_i = \theta + k^{-1} \cdot \varepsilon_i$, where $\varepsilon_i \sim U_{[-1/2,1/2]}$ is independent from θ and across the players, and k^{-1} measures the magnitude of noise. We will refer to the collection of these experiments as the global game information structure. Fixing $k < \infty$, consider the corresponding operator

$B^{t,\beta}(\cdot)$ with $\beta(\theta) = 1 - \theta$. For any ψ, we can verify, using arguments from the leading example, that

$$B^{t,\beta}(F_\psi) = F_{\frac{k\psi+t}{1+k}}.$$

Then for any $k < \infty$, it is clear that F_t is the only fixed point of $B^{t,\beta}(\cdot)$. Thus we obtain the following:

Proposition 14 (Feasibility Under Global Game Information Structures) *A threshold event F_ψ is an equilibrium regime change event when players choose a global game information structure for any $k < \infty$, only if $\psi = t$.*

The proof is straightforward and hence omitted here. In order to compare to the case of information acquisition with binary signal experiments, we let the players acquire information about θ by increasing k as in Subsection 3.3. The players choose a precision $k^* < \infty$ if and only if $\gamma < 1$, which leads to exactly the same effective strategies as with binary information structures. Hence if $\gamma < 1$, it is only feasible for the players to "choose" equilibrium regime change event F_t through acquiring global game information structures. Since the players have no other choice, it is also incentive compatible to choose information structures giving rise to F_t, which results in the same event of regime change in equilibrium as with binary information structures. Thus there are conceptually very different mechanisms behind the same equilibrium outcome of the two technologies. While under binary information structures any threshold event F_ψ with $\psi \in [0,1]$ is feasible but only F_t is incentive compatible, the global game information structure setting directly shrinks the players' "choice" of equilibrium regime change events to a singleton, which can only be F_t.

7 Extensions

In this section, we discuss extensions to our main result. We report a partial converse. And we present a version of our main result, where various assumptions are relaxed, most importantly allowing general binary action symmetric coordination games with a continuum of players but also noting how various technical assumptions can be relaxed.

7.1 A Converse Result: Limit Multiplicity

In order to appreciate the importance of the conditions for limit uniqueness, this section contrasts these conditions to a sufficient condition for limit multiplicity.

Definition 15 (Lipschitz) *Cost functional $c(\cdot)$ is Lipschitz, if there exists a $K > 0$ such that $|c(s_1) - c(s_2)| < K \cdot \|s_1, s_2\|$ for all $s_1, s_2 \in S$.*

The Lipschitz property requires that the information cost responds at most linearly to any change of the effective strategies. This property is sufficient for limit multiplicity.

Proposition 16 *If cost functional $c(\cdot)$ is Lipschitz, then there exists $\overline{\lambda} > 0$ such that the game has multiple equilibria for all $\lambda \in (0, \overline{\lambda})$. In particular, for every $\theta^* \in [0,1]$, $s^*_{i,\ \lambda} = 1_{\{\theta \geq \theta^*\}}$ for each $i \in [0,1]$ is an equilibrium.*

Yang (2015) also showed this sufficient condition for multiplicity (in a closely related setting), as well as showing that there was multiplicity when the cost functional corresponded to entropy reduction. Entropy reduction is not a special case of Lipschitz, but the argument takes a very similar form (see Subsection 8.1 for more discussion of this point).

Proof. Let $\overline{\lambda} = \min\left(\frac{t}{K}, \frac{1-t}{K}\right)$. It suffices to show that $1_{\{\theta \geq \theta^*\}} \in S(\theta^*, \lambda)$ for any $\theta^* \in [0,1]$ and $\lambda \in (0, \overline{\lambda})$. This is true because for any $s \neq 1_{\{\theta \geq \theta^*\}}$,

$$\begin{aligned}
& U\left(1_{\{\theta \geq \theta^*\}}|\theta^*, \lambda\right) - U\left(s|\theta^*, \lambda\right) \\
= & \int_{-\infty}^{\infty} \left[1_{\{\theta \geq \theta^*\}} - t\right] \cdot \left[1_{\{\theta \geq \theta^*\}} - s(\theta)\right]\ g(\theta)\, d\theta - \ \lambda \cdot \left[c\left(1_{\{\theta \geq \theta^*\}}\right) - c(s)\right] \\
> & \ t \cdot \int_{-\infty}^{\theta^*} s(\theta)\ g(\theta)\, d\theta + (1-t) \cdot \int_{\theta^*}^{\infty} [1 - s(\theta)]\ g(\theta)\, d\theta - \ \lambda K \cdot \left\|1_{\{\theta \geq \theta^*\}}, s\right\| \\
\geq & \ \min(t, 1-t) \cdot \left[\int_{-\infty}^{\theta^*} s(\theta)\ g(\theta)\, d\theta + \int_{\theta^*}^{\infty} [1 - s(\theta)]\ g(\theta)\, d\theta\right] - \ \lambda K \cdot \left\|1_{\{\theta \geq \theta^*\}}, s\right\| \\
= & \ [\min(t, 1-t) - \ \lambda K] \cdot \left\|1_{\{\theta \geq \theta^*\}}, s\right\| > 0,
\end{aligned}$$

where the first inequality follows the Lipschitz property. ■

It is easy to see that the Lipschitz property implies the translation insensitivity.[10] The proof also establishes that the Lipschitz property implies the failure of continuous choice. Thus the Lipschitz property preserves translation insensitivity but fails continuous choice, highlighting the importance of the latter condition for our main result.

In our leading example, the Lipschitz property corresponds to the case of $\gamma \geq 1$. To see this, choose s_1 and s_2 and let k_1 and k_2 be their maximum slopes, respectively. Since the optimal effective strategies converge to step function $1_{\{\theta \geq \theta^*\}}$ as $\lambda \to 0$, let $k_2 > k_1 \geq 1$ and s_1 and s_2 belong to $B_\delta\left(1_{\{\theta \geq \theta^*\}}\right)$ for some small $\delta > 0$. Then

$$\|s_1, s_2\| \geq \left[4 \cdot \left(\overline{\theta} - \underline{\theta}\right)\right]^{-1} \cdot \left[k_1^{-1} - k_2^{-1}\right] + O(\delta).$$

[10] This is because

$$\begin{aligned}
\|T_\Delta s, s\| &= \int |s(\theta + \Delta) - s(\theta)| \cdot g(\theta)\, d\theta \leq \overline{g} \cdot \int [s(\theta + \Delta) - s(\theta)]\, d\theta \\
&= \overline{g} \cdot \int s'(\theta)\, d\theta \cdot \Delta + o(\Delta) \leq \overline{g} \cdot 1 \cdot \Delta + o(\Delta) < K \cdot \Delta\ ,
\end{aligned}$$

where $K > \overline{g} = \sup_{\theta \in \mathbb{R}} g(\theta)$.

The first term of the right hand side is what the distance would be if $s_1 = s_{\xi,k_1}$ and $s_2 = s_{\xi,k_2}$ for some center point ξ, which provides a lower bound. So

$$\frac{c(s_2) - c(s_1)}{\|s_1, s_2\|} \leq 4 \cdot \left(\overline{\theta} - \underline{\theta}\right) \cdot \frac{k_1^{-\gamma} - k_2^{-\gamma}}{k_1^{-1} - k_2^{-1}} + O(\delta).$$

Since k_1^{-1} and k_2^{-1} belong to $[0, 1]$ and the derivative of $f(y) = y^\gamma$ is bounded if and only if $\gamma \geq 1$, the information cost in our leading example has the Lipschitz property if and only if $\gamma \geq 1$.

7.2 Relaxing Assumptions: General Payoffs and Local Assumptions

This subsection extends our main results in two directions: generalizing payoffs of the game and relaxing key assumptions. Our generalization of payoffs admits all binary action, continuum player, symmetric games with strategic complementarities. Our relaxation of key assumptions includes replacing global properties of information cost functionals with local properties.

We earlier defined $\pi(l, \theta)$ to be a player's payoff when $l \in [0, 1]$ is the proportion of players that invest and θ is the state of the world. However, our regime change game assumption corresponded to a particular functional form for $\pi(l, \theta)$ specified in equation (1). We now relax the functional form and replace it with the following more general restrictions.

Assumption A1 (Monotonicity and Boundedness): *a) $\pi(l, \theta)$ is non-decreasing in l and θ; b) $|\pi(l, \theta)|$ is uniformly bounded.*

Assumption A2 (State Single Crossing): *For any $l \in [0, 1]$, there exists a $\theta_l \in \mathbb{R}$ such that $\pi(l, \theta) > 0$ if $\theta > \theta_l$ and $\pi(l, \theta) < 0$ if $\theta < \theta_l$.*

Assumption A3 (Strict Laplacian State Monotonicity and Continuity): *Let $\Pi(\theta) = \int_0^1 \pi(l, \theta)\, dl$. Then, a) there exists a unique $\theta^{**} \in \mathbb{R}$, such that $\Pi(\theta^{**}) = 0$; b) Π is continuous, and Π^{-1} exists on an open neighborhood of $\Pi(\theta^{**})$.*

These assumptions are standard in the global game literature.[11] In particular, Assumptions A1 and A2 imply that if all players follow effective strategy $s \in S_M$, there exists a threshold $\theta_s \in \mathbb{R}$ such that $\pi(s(\theta), \theta) > 0$ if $\theta > \theta_s$ and $\pi(s(\theta), \theta) < 0$ if $\theta < \theta_s$. Assumption A1 further implies that $\theta_s \in [\theta_{\min}, \theta_{\max}]$ for all $s \in S_M$, where $\theta_{\min}$ and $\theta_{\max}$ are defined by choosing $l = 1$ and $l = 0$ in Assumption A2, respectively. Consequently, we obtain the limit dominance condition often assumed in the global game literature. That is, $\pi(l, \theta) > 0$ for all $l \in [0, 1]$ and $\theta > \theta_{\max}$, and $\pi(l, \theta) < 0$ for all $l \in [0, 1]$ and $\theta < \theta_{\min}$.[12] Assumption A3 ensures the existence of a unique Laplacian action almost everywhere. If the state is θ and a player has a uniform belief over the proportion of other players investing,

[11] See the general assumptions surveyed in Subsection 2.2 of Morris and Shin (2003).
[12] Note that we have $\theta_{\min} = 0$ and $\theta_{\max} = 1$ in the regime change game.

then she enjoys $\Pi(\theta)$ from investing. Thus invest is the Laplacian action if

$$\Pi(\theta) \geq 0 \ ,$$

or, in other words, if $\theta \geq \theta^{**}$.

Slightly abusing earlier notation, let

$$U(\widetilde{s}|s,\lambda) = \int \pi(s(\theta),\theta) \cdot \widetilde{s}(\theta) g(\theta) d\theta - \lambda \cdot c(\widetilde{s})$$

denote a player's expected payoff from playing effective strategy $\widetilde{s}$ if all other players choose strategy s.[13] Following the convention of the main model, we will refer to this game as the λ-general coordination game. We will again focus on the monotone equilibria. Then a player's decision problem is

$$\max_{\widetilde{s} \in S_M} U(\widetilde{s}|s,\lambda) \ . \tag{16}$$

Again, slightly abusing earlier notations, call this problem the (s,λ)-decision problem and write $S(s,\lambda)$ for the set of optimal monotonic effective strategies, i.e.,

$$S(s,\lambda) = \arg\max_{\widetilde{s} \in S_M} U(\widetilde{s}|s,\lambda) \ .$$

Definition 17 (generalized limit uniqueness) *Cost functional $c(\cdot)$ satisfies Laplacian selection if, for any $\delta > 0$, there exists $\overline{\lambda} > 0$ such that $\left\| s, 1_{\{\theta \geq \theta^{**}\}} \right\| \leq \delta$ whenever s is a monotone equilibrium strategy of the λ-general coordination game and $\lambda \leq \overline{\lambda}$.*

We want to identify conditions on the cost functional under which there is a unique equilibrium where the Laplacian action is always played, in the limit as $\lambda \to 0$. We next relax the assumptions on the information cost functionals.

First, we define the local translation insensitivity instead of the global translation insensitivity.

Definition 18 (local translation insensitivity) *Cost functional $c(\cdot)$ is said to be locally translation insensitive at $s \in S$, if there exists a $\delta > 0$ and $K > 0$ such that $|c(T_\Delta \widetilde{s}) - c(\widetilde{s})| < K \cdot |\Delta|$ holds for all $\widetilde{s} \in B_\delta(s)$ and $\Delta \in \mathbb{R}$ providing that $T_\Delta \widetilde{s} \in B_\delta(s)$.*

Second, we define a local version of the continuous choice property.

Definition 19 (locally continuous choice) *Cost functional $c(\cdot)$ satisfies locally continuous choice at $s \in S$, if there exists a $\delta > 0$ such that $S(\widetilde{s},\lambda)$ consists only of absolutely continuous functions for all $\widetilde{s} \in B_\delta(s)$ and $\lambda \in \mathbb{R}_{++}$.*

[13] Equivalently, $s(\theta)$ can be interpreted as the aggregate effective strategy, which is the proportion of the players that invest when the state is θ.

These local properties are weaker than their counterparts in Section 4. We gave an example where translation insensitivity failed but local translation insensitivity held in Section 5.2. Now our main result continues to hold with general payoffs and local restrictions.

Proposition 20 *If $c(\cdot)$ satisfies locally continuous choice and is locally translation insensitive at $1_{\{\theta\geq\theta_s\}}$ for all $\theta_s \in [\theta_{\min}, \theta_{\max}]$, then $c(\cdot)$ satisfies Laplacian selection.*

This proposition generalizes the results of Proposition 8 and shares the same intuition. In particular, as shown in the proof, the maintained assumption that the information cost is uniformly bounded for all effective strategies can be further relaxed. Indeed, the proof goes through when the information cost is bounded on a subset $\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$ instead of all effective strategies. As shown by Lemma 23 in the appendix, this condition guarantees that the optimal strategies in $S(s, \lambda)$ uniformly converge to $1_{\{\theta\geq\theta_s\}}$ for all $s \in S_M$. Hence, even in the applications where the information cost is unbounded on $\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$, our results are still valid as long as there is uniform convergence of the optimal effective strategies, a property satisfied in most models of information acquisition. The locally continuous choice property can be justified by a local version of cheap continuous approximation following the logic of its counterpart in Subsection 5.1; this argument is omitted here.

The analogous converse result holds.

Definition 21 (locally Lipschitz) *Cost functional $c(\cdot)$ is locally Lipschitz at $s \in S$, if there exists a $\delta > 0$ and $K > 0$ such that $|c(s_2) - c(s_1)| < K \cdot \|s_1, s_2\|$ for all $s_1, s_2 \in B_\delta(s)$.*

The local Lipschitz property implies the local translation insensitivity; local Lipschitz implies the failure of local continuous choice; thus local Lipschitz preserves local translation insensitivity but fails local continuous choice; so we have:

Proposition 22 *If the information cost $c(\cdot)$ is locally Lipschitz at $1_{\{\theta\geq\theta_s\}}$ for some $\theta_s \in [\theta_{\min}, \theta_{\max}]$, then there exists a $\overline{\lambda} > 0$ such that the game has multiple equilibria for all $\lambda \in (0, \overline{\lambda})$. In particular, for every $\theta'_s \in (\theta_{\min}, \theta_{\max})$ in a neighborhood of θ_s, $s^*_{i,\lambda} = 1_{\{\theta\geq\theta'_s\}}$ for each $i \in [0, 1]$ is an equilibrium.*

This proposition is a generalization of Proposition 16 and the two propositions share the same intuition.

8 Discussion

8.1 Entropy Reduction

Shannon's entropy measures uncertainty and its reduction measures the amount of information. In a closely related setting, Yang (2015) obtains limit multiplicity when the information cost is given by entropy reduction. This is consistent with our intuition because entropy reduction has the feature that it is equally easy to distinguish nearby and distant states. Hence, continuous choice fails for sufficiently small values of λ. This result suggests that the Lipschitz property is sufficient but not necessary for limit multiplicity, since it is not satisfied by entropy reduction. In particular, the Lipschitz property rules out the possibility that the marginal cost of changing $s(\theta)$ going to infinity, but the marginal entropy reduction of pushing $s(\theta)$ to 0 or 1 tends to infinity. Consequently, in any threshold decision problem, it would be optimal to choose a discontinuous effective strategy but not a step function $1_{\{\theta \geq \psi\}}$. While this distinction is not important for our analysis, it is important in other contexts (e.g., Denti (2016)).

8.2 Learning about Others' Actions

A maintained assumption in our analysis is that players acquire information about the state only. Hence, their signals are conditionally independent given the state. Denti (2016) considers the problem when (finitely many) players can acquire information about others' information, which essentially allows the players' signals to be correlated even if conditioned on the state. The information cost takes the form of entropy reduction, which prevents the effective strategies from attaining 0 or 1 as the marginal entropy reduction of doing so is infinity. Consequently, the players' actions contain residual uncertainty other than that originated from the uncertain state, which allows the players to correlate their actions through acquiring others' information. This gives rise to smoother best responses and a different answer for us - limit uniqueness - in this case. Note that if we were assuming an information cost with the Lipschitz property, we too would get limit multiplicity. This is because the players would choose step functions like $1_{\{\theta \geq \psi\}}$ so that their actions are deterministic functions of the state. Hence, acquiring information about others' information amounts to acquiring information about the state. Therefore, the game reduces to the one studied in Section 7.1 and limit multiplicity follows.

8.3 Evidence on Informational Costs

The property that nearby states are harder to distinguish than distant states seems natural in any setting where states have a natural metric and correspond to physical outcomes. Jazayeri and Movshon (2007) examine decision makers' ability to discriminate the direction of dots on the screen when they face a threshold decision problem. There is evidence that subjects are better at discriminating states on either side of the threshold, consistent with optimal allocation of scarce resources to discriminate. However, the ability to discriminate between states on either side of the threshold disappears as we approach the threshold, giving rise to continuous choice in our sense in this setting.[14] The allocation of resources in

[14] We are grateful to Michael Woodford for providing this reference.

this case is at the unconscious neuro level. Subjects in Caplin and Dean (2015) are asked to discriminate between the number of balls on the screen, where allocation of resources is presumably a conscious choice (e.g., how much time to devote to the task). Ongoing work in Dean, Morris, and Trevino (2016) confirms that, given a threshold decision problem, an inability to distinguish nearby states arises as expected.

References

Blackwell, D. (1953). Equivalent comparisons of experiments. *Annals of Mathematical Statistics 24*(2), 265–272.

Caplin, A. and M. Dean (2015, July). Revealed preference, rational inattention, and costly information acquisition. *American Economic Review 105*(7), 2183–2203.

Carlsson, H. and E. v. Damme (1993). Global games and equilibrium selection. *Econometrica 61*(5), 989–1018.

Dean, M., S. Morris, and I. Trevino (2016). Endogenous information structures and play in global games. Technical report.

Denti, T. (2016). Games with unrestricted information acquisition. Technical report, MIT.

Frankel, D. M., S. Morris, and A. Pauzner (2003). Equilibrium selection in global games with strategic complementarities. *Journal of Economic Theory 108*(1), 1–44.

Hebert, B. and M. Woodford (2016). Rational inattention with sequential information sampling. Technical report.

Jazayeri, M. and J. A. Movshon (2007). A new perceptual illusion reveals mechanisms of sensory decoding. *Nature 446*(7138), 912–915.

Mathevet, L. (2008). A contraction principle for finite global games. *Economic Theory 42*(3), 539.

Mathevet, L. and J. Steiner (2013). Tractable dynamic global games and applications. *Journal of Economic Theory 148*(6), 2583 – 2619.

Mertens, J. and S. Zamir (1985). Formalization of bayesian analysis for games with incomplete information. *International Journal of Game Theory 14*, 1–29.

Morris, S. and H. Shin (2003). Global games: Theory and applications. In M. Dewatripont, L. Hansen, and S. Turnovsky (Eds.), *Advances in Economics and Econometrics: Proceedins of the Eight World Congress of the Econometric Society*, pp. 56–114. Cambridge: Cambridge University Press.

Morris, S. and H. S. Shin (1998). Unique equilibrium in a model of self-fulfilling currency attacks. *The American Economic Review 88*(3), pp. 587–597.

Morris, S., H. S. Shin, and M. Yildiz (2016). Common belief foundations of global games. *Journal of Economic Theory 163*, 826 – 848.

Morris, S. and M. Yang (2016). Coordination and the relative cost of distinguishing nearby states. techreport, Duke University and Princeton University.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics 50*(3), 665–690.

Strack, P. (2016). Smoothness of choice probabilities. Technical report.

Szkup, M. and I. Trevino (2015). Information acquisition in global games of regime change. *Journal of Economic Theory 160*, 387 – 428.

Weinstein, J. and M. Yildiz (2007). A structure theorem for rationalizability with application to robust predictions of refinements. *Econometrica 75*(2), 365–400.

Woodford, M. (2008). Inattention as a source of randomized discrete adjustment. Working paper, Columbia University.

Yang, M. (2015). Coordination with flexible information acquisition. *Journal of Economic Theory 158, Part B*, 721 – 738.

9 Appendix

Proof of Proposition 8.

Proof. Lemma 4 implies

$$\lim_{\lambda\to 0} \sup_{s\in\ S(\psi,\lambda) \text{ and } \psi\in[0,1]} \left\| s, 1_{\{\theta\geq\psi\}} \right\| = 0. \tag{17}$$

Let $\left\{ s^*_{i,\ \lambda} \right\}_{i\in[0,1]}$ denote an equilibrium strategy profile of the λ-regime change game. Then the aggregate effective strategy is given by

$$\widehat{s}^*{}_{\lambda}(\theta) = \int_{i\in[0,1]} s^*_{i,\ \lambda}(\theta)\, di\ .$$

Assuming the continuum law of large numbers, the proportion of players that take action 1 conditional on θ is $\widehat{s}^*{}_{\lambda}(\theta)$. Since all $\left\{ s^*_{i,\ \lambda} \right\}_{i\in[0,1]}$ are absolutely continuous, $\widehat{s}^*{}_{\lambda}$ is also absolutely continuous. Hence, there exists a unique $\theta^*{}_{\lambda}$ such that

$$\widehat{s}^*{}_{\lambda}(\theta^*{}_{\lambda}) = \beta(\theta^*{}_{\lambda})\,. \tag{18}$$

Note that $\theta^*{}_{\lambda}$ is the threshold of regime change in this equilibrium. Let $\theta^{**} = \beta^{-1}(1-t)$. Since

$$\left\| s^*_{i,\ \lambda}, 1_{\{\theta\geq\theta^{**}\}} \right\| \leq \left\| s^*_{i,\ \lambda}, 1_{\left\{\theta\geq\theta^*{}_{\lambda}\right\}} \right\| + \left\| 1_{\left\{\theta\geq\theta^*{}_{\lambda}\right\}}, 1_{\{\theta\geq\theta^{**}\}} \right\|$$

and (17) implies

$$\lim_{\lambda\to 0} \left\| s^*_{i,\ \lambda}, 1_{\left\{\theta\geq\theta^*{}_{\lambda}\right\}} \right\| = 0\ ,$$

it suffices to show that $\theta^*{}_{\lambda} \to \theta^{**}$ as $\lambda \to 0$.

We first show that $\int_{-\infty}^{\infty} \left[1_{\left\{\theta\geq\theta^*{}_{\lambda}\right\}} - t \right] \cdot g(\theta)\, d\widehat{s}^*{}_{\lambda}(\theta)$ is arbitrarily close to zero when λ is small enough. Consider player i's expected payoff from slightly shifting her equilibrium strategy $s^*_{i,\ \lambda}$ to $T_{\Delta} s^*_{i,\ \lambda}$, which is given by

$$W(\Delta) = \int_{-\infty}^{\infty} \left[1_{\left\{\theta\geq\theta^*{}_{\lambda}\right\}} - t \right] \cdot s^*_{i,\ \lambda}(\theta+\Delta) \cdot g(\theta)\, d\theta - \ \lambda \cdot c\left(T_{\Delta} s^*_{i,\ \lambda} \right).$$

The player should not benefit from this deviation, which implies $W'(0) = 0$, i.e.,

$$\int_{-\infty}^{\infty} \left[1_{\{\theta \geq \theta^*{}_\lambda\}} - t\right] \cdot \frac{ds^*_{i,\ \lambda}(\theta)}{d\theta} \cdot g(\theta)\, d\theta - \ \lambda \cdot \left.\frac{dc\left(T_\Delta s^*_{i,\ \lambda}\right)}{d\Delta}\right|_{\Delta=0}$$

$$= \int_{-\infty}^{\infty} \left[1_{\{\theta \geq \theta^*{}_\lambda\}} - t\right] \cdot g(\theta)\, ds^*_{i,\ \lambda}(\theta) - \ \lambda \cdot \left.\frac{dc\left(T_\Delta s^*_{i,\ \lambda}\right)}{d\Delta}\right|_{\Delta=0} = 0.$$

Here $W'(0)$ exists because $s^*_{i,\ \lambda}$ is absolutely continuous. In addition, the translation insensitivity implies $-K < \left.\frac{dc\left(T_\Delta s^*_{i,\ \lambda}\right)}{d\Delta}\right|_{\Delta=0} < K$. Hence, for any small $\varepsilon > 0$, by choosing $\lambda \in (0, \varepsilon)$ we obtain

$$-K\varepsilon < \int_{-\infty}^{\infty} \left[1_{\{\theta \geq \theta^*{}_\lambda\}} - t\right] \cdot g(\theta)\, ds^*_{i,\ \lambda}(\theta) < K\varepsilon\ .$$

The above inequality holds for all $i \in [0, 1]$, and thus implies

$$-K\varepsilon < \int_{-\infty}^{\infty} \left[1_{\{\theta \geq \theta^*{}_\lambda\}} - t\right] \cdot g(\theta)\, d\widehat{s}^*{}_\lambda(\theta) < K\varepsilon\ ,$$

i.e.,

$$\left|\int_{-\infty}^{\infty} \left[1_{\{\theta \geq \theta^*{}_\lambda\}} - t\right] \cdot g(\theta)\, d\widehat{s}^*{}_\lambda(\theta)\right| < K\varepsilon\ . \tag{19}$$

Since the density function $g(\theta)$ is continuous on $[0, 1]$, it is also uniformly continuous on $[0, 1]$. Hence, for any $\varepsilon > 0$, we can find an $\eta > 0$ such that $\left|g(\theta) - g\left(\theta'\right)\right| < \varepsilon$ for all $\theta, \theta' \in [0, 1]$ and $\left|\theta - \theta'\right| < \eta$. By (17), for all i, the effective strategy $s^*_{i,\ \lambda}$ converges to $1_{\{\theta \geq \theta^*{}_\lambda\}}$ in $L^1 - norm$, so does the aggregate effective strategy $\widehat{s}^*{}_\lambda$. Together with the monotonicity of $\widehat{s}^*{}_\lambda$, this implies the existence of a $\lambda_1 > 0$ such that for all $\lambda \in (0,\ \lambda_1)$, $\left|\widehat{s}^*{}_\lambda(\theta) - 1_{\{\theta \geq \theta^*{}_\lambda\}}\right| < \varepsilon$ for all $\theta \in (-\infty, \theta^*{}_\lambda - \eta) \cup (\theta^*{}_\lambda + \eta, \infty)$. Choosing $\lambda \in$

$(0, \min(\lambda_1, \varepsilon))$, by (19), we obtain

$$\begin{aligned}
& \left| \int_{\theta^*_\lambda - \eta}^{\theta^*_\lambda + \eta} \left[1_{\{\theta \geq \theta^*_\lambda\}} - t \right] \cdot g(\theta) \, d\widehat{s}^*_\lambda(\theta) \right| \\
< & \left| \int_{-\infty}^{\theta^*_\lambda - \eta} \left[1_{\{\theta \geq \theta^*_\lambda\}} - t \right] \cdot g(\theta) \, d\widehat{s}^*_\lambda(\theta) + \int_{\theta^*_\lambda + \eta}^{\infty} \left[1_{\{\theta \geq \theta^*_\lambda\}} - t \right] \cdot g(\theta) \, d\widehat{s}^*_\lambda(\theta) \right| + K\varepsilon \\
\leq & \int_{-\infty}^{\theta^*_\lambda - \eta} \left| 1_{\{\theta \geq \theta^*_\lambda\}} - t \right| \cdot g(\theta) \, d\widehat{s}^*_\lambda(\theta) + \int_{\theta^*_\lambda + \eta}^{\infty} \left| 1_{\{\theta \geq \theta^*_\lambda\}} - t \right| \cdot g(\theta) \, d\widehat{s}^*_\lambda(\theta) + K\varepsilon \\
\leq & \; 2\overline{g}\varepsilon + K\varepsilon \; , \qquad (20)
\end{aligned}$$

where $\overline{g} = \sup_{\theta \in \mathbb{R}} g(\theta) < \infty$. By the definition of η, $|g(\theta) - g(\theta^*_\lambda)| < \varepsilon$ for all $\theta \in [\theta^*_\lambda - \eta, \theta^*_\lambda + \eta]$. Hence,

$$\left| g(\theta^*_\lambda) \cdot \int_{\theta^*_\lambda - \eta}^{\theta^*_\lambda + \eta} \left[1_{\{\theta \geq \theta^*_\lambda\}} - t \right] d\widehat{s}^*_\lambda(\theta) - \int_{\theta^*_\lambda - \eta}^{\theta^*_\lambda + \eta} \left[1_{\{\theta \geq \theta^*_\lambda\}} - t \right] \cdot g(\theta) \, d\widehat{s}^*_\lambda(\theta) \right| < \varepsilon. \qquad (21)$$

Further note that

$$\begin{aligned}
& \left| 1 - \beta(\theta^*_\lambda) - t - \int_{\theta^*_\lambda - \eta}^{\theta^*_\lambda + \eta} \left[1_{\{\theta \geq \theta^*_\lambda\}} - t \right] d\widehat{s}^*_\lambda(\theta) \right| \\
= & \left| 1 - \beta(\theta^*_\lambda) - t - \widehat{s}^*_\lambda(\theta^*_\lambda + \eta) + \widehat{s}^*_\lambda(\theta^*_\lambda) + t \cdot [\widehat{s}^*_\lambda(\theta^*_\lambda + \eta) - \widehat{s}^*_\lambda(\theta^*_\lambda - \eta)] \right| \\
= & \left| (1 - t) \cdot [1 - \widehat{s}^*_\lambda(\theta^*_\lambda + \eta)] - t \cdot \widehat{s}^*_\lambda(\theta^*_\lambda - \eta) \right| \\
\leq & \; (1 - t) \cdot |1 - \widehat{s}^*_\lambda(\theta^*_\lambda + \eta)| + t \cdot |\widehat{s}^*_\lambda(\theta^*_\lambda - \eta)| \\
\leq & \; \varepsilon \; , \qquad (22)
\end{aligned}$$

where the second equality follows (18), the last inequality follows the facts that $\widehat{s}^*_\lambda(\theta^*_\lambda - \eta) \leq \varepsilon$ and $1 - \widehat{s}^*_\lambda(\theta^*_\lambda + \eta) \leq \varepsilon$ when $\lambda \in (0, \lambda_1)$.

Inequalities (20), (21) and (22) together imply that

$$\begin{aligned}
|1 - \beta(\theta^*_\lambda) - t| & < \varepsilon + \frac{2\overline{g} + K + 1}{g(\theta^*_\lambda)} \varepsilon \\
& \leq \varepsilon + \frac{2\overline{g} + K + 1}{\underline{g}} \varepsilon \; ,
\end{aligned}$$

where $\underline{g} = \inf_{\theta \in [0,1]} g(\theta) > 0$ since g is assumed to be continuous and strictly positive on $[0, 1]$. Hence, $\beta(\theta^*_\lambda)$ is arbitrarily close to $1 - t$ as $\lambda \to 0$. Therefore, the continuity of

$\beta^{-1}(\cdot)$ implies $\lim_{\lambda\to 0} \theta^*{}_\lambda = \beta^{-1}(1-t)$. ■

Proof of Proposition 10.

Proof. Suppose $s \in S(\psi, \lambda)$ is not absolutely continuous. Then the derivative of s has a Dirac Delta component, which together with $rp > 1$ implies $\|\kappa_{w.h}[s]\|_p = \infty$ and thus $c_{w,h,H,p}(s) = \overline{H}$. Since $1_{\{\theta\geq\psi\}}$ also costs $\overline{H}$, $1_{\{\theta\geq\psi\}}$ dominates s if $s \neq 1_{\{\theta\geq\psi\}}$. Hence we obtain $s = 1_{\{\theta\geq\psi\}}$. Now consider replacing $1_{\{\theta\geq\psi\}}$ with $s_{\psi,k}$ (defined by (4)). The cost saving will be

$$\begin{aligned} & \lambda \cdot \left[\overline{H} - H\left(\left[\int_{\psi-\frac{1}{2k}}^{\psi+\frac{1}{2k}} [w(\theta)\cdot h(k)]^p g(\theta)\, d\theta\right]^{1/p}\right)\right] \\ \geq\ & \lambda \cdot \left[\overline{H} - H\left(h(k)\cdot k^{-1/p}\cdot \overline{w}\cdot \overline{g}^{1/p}\right)\right] , \end{aligned}$$

where $\overline{w} = \sup_{\theta\in[\psi-\epsilon,\psi+\epsilon]} w(\theta)$ and $\overline{g} = \sup_{\theta\in[\psi-\epsilon,\psi+\epsilon]} g(\theta)$ for some fixed $\epsilon > 1/k$. Note that since both $w(\theta)$ and $g(\theta)$ are strictly positive on the closed interval $[\psi-\epsilon, \psi+\epsilon]$, so are $\overline{w}$ and $\overline{g}$. The utility loss from using $s_{\psi,k}$ instead of $1_{\{\theta\geq\psi\}}$ is

$$\begin{aligned} & \int_{-\infty}^{\infty} \left[1_{\{\theta\geq\psi\}} - t\right]\cdot\left[1_{\{\theta\geq\psi\}} - s_{\psi,k}(\theta)\right] g(\theta)\, d\theta \\ =\ & t\cdot \int_{-\infty}^{\psi} s_{\psi,k}(\theta)\, g(\theta)\, d\theta + (1-t)\cdot \int_{\psi}^{\infty} [1 - s_{\psi,k}(\theta)]\, g(\theta)\, d\theta \\ \leq\ & \max(t, 1-t)\cdot \left\|1_{\{\theta\geq\psi\}}, s_{\psi,k}\right\| \\ \leq\ & 4\cdot \overline{g}\cdot \max(t,1-t)\cdot k^{-1}. \end{aligned}$$

Hence,

$$\begin{aligned} \lim_{k\to\infty} \frac{\text{cost saving}}{\text{utility loss}} &= \lim_{k\to\infty} \frac{\lambda\cdot\left[\overline{H} - H\left(h(k)\cdot k^{-1/p}\cdot\overline{w}\cdot\overline{g}^{1/p}\right)\right]}{4\cdot\overline{g}\cdot\max(t,1-t)\cdot k^{-1}} \\ &\geq \frac{\lambda\cdot\overline{w}^{-\gamma}\cdot\overline{g}^{-\gamma/p}\cdot A}{4\cdot\overline{g}\cdot\max(t,1-t)}\cdot\lim_{k\to\infty}\frac{k^{-\gamma(r-1/p)}}{k^{-1}} \\ &= \infty , \end{aligned}$$

where A is a positive scaler and the last equality follows $\frac{1}{\gamma} + \frac{1}{p} > r$. Therefore, by choosing $s_{\psi,k}$ with $k > 0$ sufficiently large instead of s, the cost saving dominates the utility loss and the player will be strictly better off. This contradicts the optimality of s. ■

Proof of Proposition 11.

Proof. Note that since $\lim_{\sigma\to 0} h(\sigma, I) = \infty$, the decision maker will always choose a $\sigma > 0$.

Then

$$
\begin{aligned}
s'(\theta) &= \sigma^{-1} \cdot \frac{d}{d\theta} \int m(z) f\left(\frac{z-\theta}{\sigma}\right) dz \\
&= \sigma^{-1} \int m(z) \left[-\sigma^{-1} f'\left(\frac{z-\theta}{\sigma}\right)\right] dz \\
&< \infty ,
\end{aligned}
$$

where the second equality follows the dominated convergence theorem and the continuous differentiability of f, and the inequality follows the boundedness of m and that f' is L^1-integrable. Therefore, any optimal effective strategy is differentiable and thus absolutely continuous. ■

Proof of Proposition 13.

Proof. Let $s(\theta)$ be the proportion of players whose signal realization is 1, conditional on the true state being θ. Then

$$
\xi = \left(1 + \frac{1}{k}\right)\psi - \frac{1}{2k} \tag{23}
$$

implies that $s(\theta) \geq 1-\theta$ if and only if F_ψ is true. Together with (23),

$$
k \geq \frac{t \cdot (1-\psi)^2 + (1-t) \cdot \psi^2}{2 \cdot \min\left[t \cdot (\psi - \underline{\theta}), (1-t) \cdot (\overline{\theta} - \psi)\right]}
$$

implies that each player t-believes F_ψ if and only if her signal realization is 1. Hence, $F_\psi = B^{t,\beta}(F_\psi)$. Proposition **??** then leads to the desired result. ■

Lemma 23 *If $c\left(1_{\{\theta \geq \theta_s\}}\right)$ as a function of θ_s is bounded for $\theta_s \in [\theta_{\min}, \theta_{\max}]$, then for any $\rho > 0$, there exists a $\lambda_1 > 0$ such that $S(s,\lambda) \subset B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ for all $s \in S_M$ and $\lambda \in (0, \lambda_1)$.*

Proof. For any $\delta > 0$, define

$$
z(\delta) = \inf_{l \in [0,1]} \min\left(\pi(l, \theta_l + \delta), -\pi(l, \theta_l - \delta)\right) .
$$

Note that given $\delta > 0$, $\min\left(\pi(l, \theta_l + \delta), -\pi(l, \theta_l - \delta)\right)$ is a function of l on a compact set $[0,1]$. By Assumption A2, this function is always strictly positive. Hence, its infimum on $[0,1]$ exists and is strictly positive. That is, $z(\delta) > 0$ for all $\delta > 0$. In addition, for any $s \in S_M$ and $\theta \notin [\theta_s - \delta, \theta_s + \delta]$, we have

$$
|\pi(s(\theta), \theta)| \geq |\pi(s(\theta_s), \theta)| \geq z(\delta), \tag{24}
$$

where the first inequality follows Assumptions A1 and A2, and the second inequality follows the definition of $z(\delta)$.

If $S(s,\lambda) = \left\{1_{\{\theta \geq \theta_s\}}\right\}$, we are done. Now for any $\widetilde{s} \in S(s,\lambda)$ such that $\widetilde{s} \neq 1_{\{\theta \geq \theta_s\}}$, the optimality of $\widetilde{s}$ implies

$$\int_{-\infty}^{\infty} \pi\left(s\left(\theta\right),\theta\right)\cdot\left[1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right]g\left(\theta\right)d\theta \;<\; \lambda\cdot\left[c\left(1_{\{\theta\geq\theta_s\}}\right)-c\left(\widetilde{s}\right)\right] \;\leq\; \lambda\cdot c\left(1_{\{\theta\geq\theta_s\}}\right). \quad (25)$$

Note that

$$\begin{aligned}
&\int_{-\infty}^{\infty} \pi\left(s\left(\theta\right),\theta\right)\cdot\left[1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right]g\left(\theta\right)d\theta \\
\geq\; &\int_{\theta\notin[\theta_s-\delta,\theta_s+\delta]} \pi\left(s\left(\theta\right),\theta\right)\cdot\left[1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right]g\left(\theta\right)d\theta \\
\geq\; &\int_{-\infty}^{\infty} z\left(\delta\right)\cdot\left|1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right|g\left(\theta\right)d\theta-\int_{\theta_s-\delta}^{\theta_s+\delta} z\left(\delta\right)\cdot\left|1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right|g\left(\theta\right)d\theta \\
\geq\; & z\left(\delta\right)\cdot\left\|1_{\{\theta\geq\theta_s\}},\widetilde{s}\right\|-2\cdot z\left(\delta\right)\cdot\overline{g}\cdot\delta, \qquad (26)
\end{aligned}$$

where $\overline{g}=\sup_{\theta\in\mathbb{R}} g\left(\theta\right)<\infty$, the first inequality holds since $\pi\left(s\left(\theta\right),\theta\right)$ and $\left[1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right]$ always have the same sign and thus

$$\int_{\theta_s-\delta}^{\theta_s+\delta} \pi\left(s\left(\theta\right),\theta\right)\cdot\left[1_{\{\theta\geq\theta_s\}}-\widetilde{s}\left(\theta\right)\right]g\left(\theta\right)d\theta>0\ ,$$

and the second inequality follows (24). Inequalities (25) and (26) imply

$$\left\|1_{\{\theta\geq\theta_s\}},\widetilde{s}\right\|<\frac{\lambda\cdot c\left(1_{\{\theta\geq\theta_s\}}\right)}{z\left(\delta\right)}+2\cdot\overline{g}\cdot\delta\ . \qquad (27)$$

Hence, for any $\rho>0$, choose $\delta<\frac{\rho}{4\cdot\overline{g}}$ and $\lambda_1<\frac{z(\delta)\cdot\rho}{2\cdot c\left(1_{\{\theta\geq\theta_s\}}\right)}$, we obtain $\left\|1_{\{\theta\geq\theta_s\}},\widetilde{s}\right\|<\rho$ for all $\lambda\in(0,\lambda_1)$. (Note that $c\left(1_{\{\theta\geq\theta_s\}}\right)>0$, otherwise we return to the case $S\left(s,\lambda\right)=\left\{1_{\{\theta\geq\theta_s\}}\right\}$.)

Let $c_1=\sup_{\theta_s\in[\theta_{\min},\theta_{\max}]} c\left(1_{\{\theta\geq\theta_s\}}\right)$. For any $\rho>0$, choose $\delta<\frac{\rho}{4\cdot\overline{g}}$ and $\lambda_1<\frac{z(\delta)\cdot\rho}{2\cdot c_1}$. Then inequality (27) implies $\left\|1_{\{\theta\geq\theta_s\}},\widetilde{s}\right\|<\rho$ for all $s\in S_M$ and $\lambda\in(0,\lambda_1)$. ■

Proof of Proposition 20.

Proof. The idea of the proof is similar to that of Proposition 8.

We have assumed that the information cost functional is uniformly bounded for all effective strategies. Here we will prove our results under a weaker condition that $c\left(1_{\{\theta\geq\theta_s\}}\right)$ as a function of θ_s is bounded for $\theta_s\in[\theta_{\min},\theta_{\max}]$.

The local translation insensitivity implies that $c\left(1_{\{\theta\geq\theta_s\}}\right)$ is a continuous function of θ_s for $\theta_s \in [\theta_{\min}, \theta_{\max}]$, which further implies $\sup_{\theta_s\in[\theta_{\min},\theta_{\max}]} c\left(1_{\{\theta\geq\theta_s\}}\right) < \infty$. Then by Lemma 23, we have

$$\lim_{\lambda\to 0} \sup_{\widetilde{s}\in S(s,\lambda) \text{ and } s\in S_M} \left\|\widetilde{s}, 1_{\{\theta\geq\theta_s\}}\right\| = 0. \tag{28}$$

Let $\left\{s^*_{i,\lambda}\right\}_{i\in[0,1]}$ denote a monotone equilibrium of the λ-general coordination game. Then the aggregate effective strategy is given by

$$\widehat{s}^*_\lambda(\theta) = \int_{i\in[0,1]} s^*_{i,\lambda}(\theta)\, di\ ,$$

which by Assumptions A1 and A2 induces a threshold θ^*_λ such that $\pi\left(\widehat{s}^*_\lambda(\theta), \theta\right) > 0$ if $\theta > \theta^*_\lambda$ and $\pi\left(\widehat{s}^*_\lambda(\theta), \theta\right) < 0$ if $\theta < \theta^*_\lambda$. By (28),

$$\lim_{\lambda\to 0} \left\|s^*_{i,\lambda}, 1_{\left\{\theta\geq\theta^*_\lambda\right\}}\right\| = 0.$$

Since

$$\left\|s^*_{i,\lambda}, 1_{\{\theta\geq\theta^{**}\}}\right\| \leq \left\|s^*_{i,\lambda}, 1_{\left\{\theta\geq\theta^*_\lambda\right\}}\right\| + \left\|1_{\left\{\theta\geq\theta^*_\lambda\right\}}, 1_{\{\theta\geq\theta^{**}\}}\right\|,$$

it suffices to show that θ^*_λ becomes arbitrarily close to θ^{**} as $\lambda \to 0$.

We first show that the local translation insensitivity and local continuous choice property can be extended to a neighborhood of $\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$. For any $\theta_s \in [\theta_{\min}, \theta_{\max}]$, since $c(\cdot)$ is locally translation insensitive at $1_{\{\theta\geq\theta_s\}}$, there exists $\delta(\theta_s) > 0$ and $K(\theta_s) > 0$ such that $|c(T_\Delta\widetilde{s}) - c(\widetilde{s})| < K(\theta_s)\cdot|\Delta|$ holds for all $\widetilde{s} \in B_{\delta(\theta_s)}\left(1_{\{\theta\geq\theta_s\}}\right)$ and $\Delta \in \mathbb{R}$, providing that $c(\widetilde{s}) < \infty$ and $T_\Delta\widetilde{s} \in B_{\delta(\theta_s)}\left(1_{\{\theta\geq\theta_s\}}\right)$. It is straightforward to see that $\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$ is a sequentially compact subset of the metric space S and thus it is also compact. Since $\left\{B_{\delta(\theta_s)}\left(1_{\{\theta\geq\theta_s\}}\right) : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$ is an open cover of $\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$, it has a finite subcover, denoted by

$$\left\{B_{\delta(\theta^1_s)}\left(1_{\{\theta\geq\theta^1_s\}}\right), B_{\delta(\theta^2_s)}\left(1_{\{\theta\geq\theta^2_s\}}\right), \cdots, B_{\delta(\theta^n_s)}\left(1_{\{\theta\geq\theta^n_s\}}\right)\right\}\ .$$

Then

$$B_{\delta(\theta^1_s)}\left(1_{\{\theta\geq\theta^1_s\}}\right) \cup B_{\delta(\theta^2_s)}\left(1_{\{\theta\geq\theta^2_s\}}\right) \cup \cdots \cup B_{\delta(\theta^n_s)}\left(1_{\{\theta\geq\theta^n_s\}}\right)$$

is a finite open cover of $\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$. Hence, there exists a $\rho > 0$ such that

$$\begin{aligned}\left\{1_{\{\theta\geq\theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\} &\subset \cup_{\theta_s\in[\theta_{\min},\theta_{\max}]} B_\rho\left(1_{\{\theta\geq\theta_s\}}\right) \\ &\subset B_{\delta(\theta^1_s)}\left(1_{\{\theta\geq\theta^1_s\}}\right) \cup B_{\delta(\theta^2_s)}\left(1_{\{\theta\geq\theta^2_s\}}\right) \cup \cdots \cup B_{\delta(\theta^n_s)}\left(1_{\{\theta\geq\theta^n_s\}}\right)\ .\end{aligned}$$

Let $K = \max\left\{K\left(\theta^1_s\right), K\left(\theta^2_s\right), \cdots, K\left(\theta^n_s\right)\right\}$. Therefore, $|c(T_\Delta\widetilde{s}) - c(\widetilde{s})| < K\cdot|\Delta|$ holds for all $\widetilde{s} \in \cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ and $\Delta \in \mathbb{R}$, providing that $c(\widetilde{s}) < \infty$ and $T_\Delta \widetilde{s} \in \cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$. The same argument shows that $c(\cdot)$ satisfies the locally continuous choice in an open neighborhood of $\left\{1_{\{\theta \geq \theta_s\}} : \theta_s \in [\theta_{\min}, \theta_{\max}]\right\}$. That is, for all s in such neighborhood, $S(s, \lambda)$ consists of only absolutely continuous effective strategies. Slightly abusing the notation but without loss of generality, we denote the neighborhood in which both locally continuous choice property and local translation insensitivity hold by $\cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$.

We next show that $\int_{-\infty}^{\infty} \pi\left(\widehat{s}_\lambda^*(\theta), \theta\right) \cdot g(\theta)\, d\widehat{s}_\lambda^*(\theta)$ is arbitrarily close to zero when λ is small enough. By (28), there exists a $\lambda_1 > 0$ such that $S(s, \lambda) \subset \cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ for all $s \in S_M$ and $\lambda \in (0, \lambda_1)$. Hence, by choosing $\lambda < \lambda_1$, we have $s_{i,\lambda}^* \in S\left(\widehat{s}_\lambda^*, \lambda\right) \subset \cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ for all $i \in [0,1]$. This also implies that the aggregate effective strategy $\widehat{s}_\lambda^* \in \cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ and thus $s_{i,\lambda}^*$ is absolutely continuous for all $i \in [0,1]$. Now consider player i's expected payoff from slightly shifting her equilibrium strategy $s_{i,\lambda}^*$ to $T_\Delta s_{i,\lambda}^* \in \cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$, which is given by

$$W(\Delta) = \int_{-\infty}^{\infty} \pi\left(\widehat{s}_\lambda^*(\theta), \theta\right) \cdot s_{i,\lambda}^*(\theta + \Delta) \cdot g(\theta)\, d\theta - \lambda \cdot c\left(T_\Delta s_{i,\lambda}^*\right).$$

The player should not benefit from this deviation, which implies $W'(0) = 0$, i.e.,

$$\begin{aligned} & \int_{-\infty}^{\infty} \pi\left(\widehat{s}_\lambda^*(\theta), \theta\right) \cdot \frac{ds_{i,\lambda}^*(\theta)}{d\theta} \cdot g(\theta)\, d\theta - \lambda \cdot \left.\frac{dc\left(T_\Delta s_{i,\lambda}^*\right)}{d\Delta}\right|_{\Delta=0} \\ = & \int_{-\infty}^{\infty} \pi\left(\widehat{s}_\lambda^*(\theta), \theta\right) \cdot g(\theta)\, ds_{i,\lambda}^*(\theta) - \lambda \cdot \left.\frac{dc\left(T_\Delta s_{i,\lambda}^*\right)}{d\Delta}\right|_{\Delta=0} = 0. \end{aligned}$$

Here $W'(0)$ exists because $s_{i,\lambda}^*$ is absolutely continuous. Since the local translation insensitivity has been extended to $\cup_{\theta_s \in [\theta_{\min}, \theta_{\max}]} B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$, we have $-K < \left.\frac{dc\left(T_\Delta s_{i,\lambda}^*\right)}{d\Delta}\right|_{\Delta=0} < K$. Hence, for any small $\varepsilon > 0$, by choosing $\lambda \in (0, \min(\lambda_1, \varepsilon))$ we obtain

$$-K\varepsilon < \int_{-\infty}^{\infty} \pi\left(\widehat{s}_\lambda^*(\theta), \theta\right) \cdot g(\theta)\, ds_{i,\lambda}^*(\theta) < K\varepsilon \ .$$

The above inequality holds for all $i \in [0,1]$, and thus implies

$$-K\varepsilon < \int_{-\infty}^{\infty} \pi\left(\widehat{s}_\lambda^*(\theta), \theta\right) \cdot g(\theta)\, d\widehat{s}_\lambda^*(\theta) < K\varepsilon \ ,$$

i.e.,

$$\left| \int_{-\infty}^{\infty} \pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right) \cdot g(\theta)\, d\widehat{s}_{\lambda}^{*}(\theta) \right| < K\varepsilon\ . \tag{29}$$

Since the density function $g(\theta)$ is continuous on $[\theta_{\min}, \theta_{\max}]$, it is also uniformly continuous on $[\theta_{\min}, \theta_{\max}]$. For the same reason, $\Pi(\theta)$ is also uniformly continuous on $[\theta_{\min}, \theta_{\max}]$. Hence, for any $\varepsilon > 0$, we can find an $\eta > 0$ such that $\left|g(\theta) - g(\theta')\right| < \varepsilon$ and $\left|\Pi(\theta) - \Pi(\theta')\right| < \varepsilon$ for all $\theta, \theta' \in [\theta_{\min}, \theta_{\max}]$ and $\left|\theta - \theta'\right| < 2\eta$. Without loss of generality, we can choose $\eta < \varepsilon$. By (28), for all i, the effective strategy $s_{i,\lambda}^{*}$ converges to $1_{\{\theta \geq \theta_{\lambda}^{*}\}}$ in $L^1 - norm$, so does the aggregate effective strategy $\widehat{s}_{\lambda}^{*}$. Together with the monotonicity of $\widehat{s}_{\lambda}^{*}$, this implies the existence of a $\lambda_2 > 0$ such that for all $\lambda \in (0, \lambda_2)$, $\left|\widehat{s}_{\lambda}^{*}(\theta) - 1_{\{\theta \geq \theta_{\lambda}^{*}\}}\right| < \varepsilon$ for all $\theta \in (-\infty, \theta_{\lambda}^{*} - \eta) \cup (\theta_{\lambda}^{*} + \eta, \infty)$. Choosing $\lambda \in (0, \min(\lambda_1, \lambda_2, \varepsilon))$, by (29), we obtain

$$\begin{aligned} & \left| \int_{\theta_{\lambda}^{*}-\eta}^{\theta_{\lambda}^{*}+\eta} \pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right) \cdot g(\theta)\, d\widehat{s}_{\lambda}^{*}(\theta) \right| \\ < & \int_{-\infty}^{\theta_{\lambda}^{*}-\eta} \left|\pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right)\right| \cdot g(\theta)\, d\widehat{s}_{\lambda}^{*}(\theta) + \int_{\theta_{\lambda}^{*}+\eta}^{\infty} \left|\pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right)\right| \cdot g(\theta)\, d\widehat{s}_{\lambda}^{*}(\theta) + K\varepsilon \\ \leq & \ 2L\overline{g}\varepsilon + K\varepsilon\ , \end{aligned} \tag{30}$$

where $L > 0$ is the uniform bound for $|\pi(l, \theta)|$ and $\overline{g} = \sup_{\theta \in \mathbb{R}} g(\theta) < \infty$. By the definition of η, $|g(\theta) - g(\theta_{\lambda}^{*})| < \varepsilon$ for all $\theta \in [\theta_{\lambda}^{*} - \eta, \theta_{\lambda}^{*} + \eta]$. Hence,

$$\left| g(\theta_{\lambda}^{*}) \cdot \int_{\theta_{\lambda}^{*}-\eta}^{\theta_{\lambda}^{*}+\eta} \pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right) d\widehat{s}_{\lambda}^{*}(\theta) - \int_{\theta_{\lambda}^{*}-\eta}^{\theta_{\lambda}^{*}+\eta} \pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right) \cdot g(\theta)\, d\widehat{s}_{\lambda}^{*}(\theta) \right| < L\varepsilon. \tag{31}$$

Inequalities (30) and (31) imply

$$\left| \int_{\theta_{\lambda}^{*}-\eta}^{\theta_{\lambda}^{*}+\eta} \pi\left(\widehat{s}_{\lambda}^{*}(\theta), \theta\right) d\widehat{s}_{\lambda}^{*}(\theta) \right| < \frac{2L\overline{g} + K + L}{\underline{g}} \varepsilon\ , \tag{32}$$

where $\underline{g} = \inf_{\theta \in [\theta_{\min}, \theta_{\max}]} g(\theta) > 0$ since g is assumed to be continuous and strictly positive on $[\theta_{\min}, \theta_{\max}]$.

Next note that

$$\left|\int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*+\eta\right)ds - \int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*-\eta\right)ds\right| \leq \left|\Pi\left(\theta_\lambda^*+\eta\right)-\Pi\left(\theta_\lambda^*-\eta\right)\right| + 4L\varepsilon$$

$$< \varepsilon + 4L\varepsilon, \tag{33}$$

where the first inequality follows the fact that $\left|\widehat{s}_\lambda^*(\theta) - 1_{\{\theta\geq\theta_\lambda^*\}}\right| < \varepsilon$ for all $\theta \in (-\infty,\theta_\lambda^*-\eta)\cup(\theta_\lambda^*+\eta,\infty)$, and the second inequality follows the uniform continuity of $\Pi(\theta)$ on $[\theta_{\min},\theta_{\max}]$.

Further note that Assumption A1 implies

$$\int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*-\eta\right)ds \leq \int_{\theta_\lambda^*-\eta}^{\theta_\lambda^*+\eta} \pi\left(\widehat{s}_\lambda^*(\theta),\theta\right)d\widehat{s}_\lambda^*(\theta) \leq \int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*+\eta\right)ds,$$

which together with (32) and (33) implies

$$-\left(\frac{2L\overline{g}+K+L}{\underline{g}}+4L+1\right)\varepsilon < \int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*-\eta\right)ds$$

$$\leq \int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*+\eta\right)ds < \left(\frac{2L\overline{g}+K+L}{\underline{g}}+4L+1\right)\varepsilon \tag{34}$$

By Assumption A1, the monotonicity of $\pi(s,\theta)$ in θ implies

$$\left|\int_{\widehat{s}_\lambda^*(\theta_\lambda^*-\eta)}^{\widehat{s}_\lambda^*(\theta_\lambda^*+\eta)} \pi\left(s,\theta_\lambda^*\right)ds\right| < \left(\frac{2L\overline{g}+K+L}{\underline{g}}+4L+1\right)\varepsilon\ .$$

Again, using the fact that $\left|\widehat{s}_\lambda^*(\theta) - 1_{\{\theta\geq\theta_\lambda^*\}}\right| < \varepsilon$ for all $\theta \in (-\infty,\theta_\lambda^*-\eta)\cup(\theta_\lambda^*+\eta,\infty)$, the above inequality implies

$$\left|\int_0^1 \pi\left(s,\theta_\lambda^*\right)ds\right| < \left(\frac{2L\overline{g}+K+L}{\underline{g}}+6L+1\right)\varepsilon\ .$$

Therefore, we have

$$\lim_{\lambda\to 0}\Pi\left(\theta_\lambda^*\right) = 0,$$

which implies

$$\lim_{\lambda \to 0} \theta_\lambda^* = \theta^{**}$$

according to Assumption A3. ■

Proof. Choose $\rho > 0$ and $K > 0$ such that $|c(s_2) - c(s_1)| < K \cdot \|s_1, s_2\|$ for all $s_1, s_2 \in B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$. Note that we can always let $\theta_s \in (\theta_{\min}, \theta_{\max})$. This is without loss of generality because by definition, for any $1_{\{\theta \geq \theta'_s\}} \in B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ with $\theta'_s \in (\theta_{\min}, \theta_{\max})$, the information cost is also locally Lipschitz at $1_{\{\theta \geq \theta'_s\}}$.

Let $s \in S_M$ denote the effective strategy that induces the cutoff θ_s. Assumption A1 then implies that θ_s is also the cutoff for $\pi\left(1_{\{\theta \geq \theta_s\}}, \theta\right)$. This is because $\pi\left(1_{\{\theta \geq \theta_s\}}, \theta\right) \geq \pi(s(\theta), \theta) > 0$ for $\theta > \theta_s$ and $\pi\left(1_{\{\theta \geq \theta_s\}}, \theta\right) \leq \pi(s(\theta), \theta) < 0$ for $\theta < \theta_s$. In addition, $\theta_s \in (\theta_{\min}, \theta_{\max})$ implies

$$\inf\left(\{\pi(1, \theta) : \theta > \theta_s\}\right) > 0$$

and

$$\sup\left(\{\pi(0, \theta) : \theta < \theta_s\}\right) < 0 \ .$$

Let

$$b = \min\left\{\inf\left(\{\pi(1, \theta) : \theta > \theta_s\}\right), -\sup\left(\{\pi(0, \theta) : \theta < \theta_s\}\right)\right\} \ .$$

We next show that $s_{i,\lambda}^* = 1_{\{\theta \geq \theta_s\}}$ for all $i \in [0,1]$ is an equilibrium. Since θ_s is the cutoff for $\pi\left(1_{\{\theta \geq \theta_s\}}, \theta\right)$, Lemma 23 implies the existence of a $\lambda_1 > 0$ such that $S\left(1_{\{\theta \geq \theta_s\}}, \lambda\right) \in B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ for all $\lambda \in (0, \lambda_1)$. Let $\overline{\lambda} = \min\left(\lambda_1, \frac{b}{K}\right)$. It thus suffices to show that $1_{\{\theta \geq \theta_s\}}$ dominates all $\widetilde{s} \in B_\rho\left(1_{\{\theta \geq \theta_s\}}\right)$ when $\lambda \in \left(0, \overline{\lambda}\right)$. This is true because

$$\begin{aligned} & \int_{-\infty}^{\infty} \pi\left(1_{\{\theta \geq \theta_s\}}, \theta\right) \cdot \left[1_{\{\theta \geq \theta_s\}} - \widetilde{s}(\theta)\right] g(\theta)\, d\theta - \lambda \cdot \left[c\left(1_{\{\theta \geq \theta_s\}}\right) - c(\widetilde{s})\right] \\ \geq \quad & b \cdot \left\|1_{\{\theta \geq \theta_s\}}, \widetilde{s}\right\| - \lambda \cdot \left[c\left(1_{\{\theta \geq \theta_s\}}\right) - c(\widetilde{s})\right] \\ > \quad & (b - \lambda K) \cdot \left\|1_{\{\theta \geq \theta_s\}}, s\right\| > 0, \end{aligned}$$

where the first inequality follows the definition of b and the second inequality follows the local Lipschitz property. ■