Sequential Sampling and Equilibrium [*]

Duarte Gonçalves[†]

October 16, 2020

Job Market Paper

Click here for the latest version

Abstract

I propose an equilibrium solution concept in which players sequentially sample to resolve strategic uncertainty over their opponents' distribution of actions. Bayesian players sample from their opponents' distribution of actions at a cost and make optimal choices given their posterior beliefs. The solution concept makes predictions on the joint distribution of players' choices, beliefs, and decision times, and generates stochastic choice through the randomness inherent to sampling, without relying on indifference or choice mistakes. It rationalizes well-known deviations from Nash equilibrium such as the own-payoff effect and I show its novel predictions relating choices, beliefs, and decision times are supported by existing data.

Keywords: Belief Formation; Game Theory; Information Acquisition; Sequential Sampling; Bayesian Learning; Statistical Decision Theory.

JEL Classifications: C70, D83, D84, C41.

[*]I am very grateful to Yeon-Koo Che, Mark Dean, Navin Kartik and Kfir Eliaz for the continued encouragement and advice. I also thank Elliot Lipnowski and Evan Sadler, as well as Teresa Esteban-Casanelles, Laura Doval, Prajit Dutta, Evan Friedman, Bruno Furtado, Qingmin Liu, Jacopo Perego, Sara Shahanaghi, Yu Fu Wong and the participants at Columbia University's Micro Theory Colloquium and NYU's Student Micro Theory Lunch for valuable feedback.

[†] Department of Economics, Columbia University; duarte.goncalves@columbia.edu.

Table of Contents

In strategic settings, individuals' payoffs to alternative actions depend on others' *gameplay*, the probabilities with which others take their respective actions. Strategic uncertainty — uncertainty concerning others' gameplay — is therefore a crucial element of such settings, as others' gameplay is often not known in advance. In order to make a choice, individuals form beliefs about others' gameplay, be it by a process of internal deliberation or by acquiring new evidence. Moreover, the effort individuals exert to resolve their strategic uncertainty is likely to respond to incentives in the environment, rendering belief formation endogenous to the strategic setting. However, most solution concepts preclude strategic uncertainty. Although this is a useful simplification, in reality, individuals are typically unsure about others' gameplay when choosing their own actions

Strategic uncertainty can also help explain several otherwise puzzling patterns in individuals' beliefs documented by experimental evidence. First, experimental evidence shows that beliefs are typically biased, diverging both from Nash equilibrium as well as from others' actual behavior (e.g. Costa-Gomes and Weizsäcker 2008). Second, individuals report different beliefs when faced with the same strategic setting, suggesting randomness in belief formation (Friedman and Ward 2019). Finally, individuals' beliefs depend on their own incentives, even when holding others' behavior fixed (Esteban-Casanelles and Gonçalves 2020). In turn, these patterns in beliefs may bridge the gap between the fact that individuals do tend to best-respond to their beliefs and the abundance of evidence regarding gameplay deviations from Nash equilibrium. Therefore, to better model gameplay, a logical next step is to explicitly account for belief formation.

In this paper, I develop an equilibrium framework based on sequential sampling in which players form beliefs about their opponents' gameplay through a costly evidence accumulation process. Individuals have a belief about others' gameplay and, prior to making a choice, they can refine their belief by exerting costly effort to gather and process evidence, which I model as sampling from their opponents' distribution of actions. This can be interpreted as parsing historical data on past play or as an internal deliberation procedure in which, in line with recent neuroscience research (e.g. Shadlen and Shohamy 2016), individuals sample from their own past experiences. As a trade-off emerges between acquiring additional information to potentially make better choices, and the costs that obtaining such information entails, players have to decide when to optimally stop sampling and make a choice. However, the randomness in a player's samples, which depends on others' distribution of actions, induces randomness in that player's choices, as choices depend on the samples observed before stopping. I then define a *sequential sampling equilibrium* as a consistent distribution of actions of all players, closing the model with a fixed-point condition. Furthermore, I show that sequential sampling equilibria have a steady-state foundation in which short-lived players sample from past evidence, sidestepping the apparent circularity of the solution concept.

The solution concept builds on an individual decision-making foundation of sequential sampling in a rich environment of choice under uncertainty. Although sequential evidence accumulation has fared well in explaining behavior in non-strategic settings by relating belief formation and choices,[1] it has not been explored in the context of equilibrium solution concepts. In my model, players effectively act as decision makers as they take others' uncertain behavior as given. Players can sample at a cost from their opponents' choice distribution, which can be interpreted as sampling from past data. Thus, players face an optimal stopping problem, deciding when to stop sampling. As discussed below, in laying the ground for the analysis of this solution concept, I also contribute to the literature on optimal stopping by providing novel comparative statics results.

I provide a general existence result for sequential sampling equilibrium when players' Bayesian updating is consistent. For specific cases, such as when players have degenerate priors, no sampling occurs and existence follows immediately. Moreover, if such priors are correct, corresponding to others' gameplay, the set of sequential sampling equilibria matches the set of Nash equilibria of the game. When priors are non-degenerate, players may want to acquire samples in order to learn their opponents' distribution. However, I show that although players always believe they will stop sampling in finite time, they may actually never stop sampling and thus never take an action, preventing the existence of an equilibrium. I establish that whenever players' priors have full support, a sequential sampling equilibrium exists. The proof builds on a novel technical result relating Bayesian consistency and optimal stopping. I show that with a full support prior, which

[1] Pioneered in cognitive science (Ratcliff 1978; Forstmann et al. 2016), this approach linking choice and a specific process of belief formation has been used to explain individual decisions in many contexts of economic interest, from simple purchasing decisions (Krajbich et al. 2012) to response to advertising (Chiong et al. 2019). Further, it has led to new theoretical models that rationalize known patterns of how choices relate to decision times (e.g. Fudenberg et al. 2018). See Clithero (2018) for a review.

necessary and sufficient for Bayesian updating to be consistent for any true distribution of the samples, optimal stopping time is bounded which, in particular, precludes players from sampling forever. Although the assumption that priors have full support is retained throughout the paper, an extension of this argument to potentially misspecified priors with convex support is shown to hold, which may be relevant for related research.

As stated earlier, sequential sampling equilibrium can be thought of as a steady state of a simple dynamic process. I consider a sequence of populations of short-lived players who sample from data on past play according to the sequential sampling procedure that underlies my solution concept. I prove that if the distribution of data on actions converges as the data accumulates, the limit distribution is a sequential sampling equilibrium. Moreover, I show that this distribution always converges in 2×2 games with a unique Nash equilibrium.[2]

Sequential sampling also provides a foundation for Nash equilibrium. I establish that, as sampling costs vanish, sequential sampling equilibria converge to Nash equilibria. This follows from showing that optimal stopping implies that players, individually, asymptotically have no regret when priors have full support. It is important to point out that convergence to a Nash equilibrium requires optimal sampling behavior. In contrast, this result no longer holds when information acquisition is myopic, that is, when players keep sampling only if the next observation improves the expected payoff net of the sampling cost, neglecting the continuation value of further samples.

Not all Nash equilibria can be reached as sampling costs vanish. In particular, a Nash equilibrium involving weakly dominated actions cannot be reached as sequential sampling players holding non-degenerate beliefs would never choose these actions. In fact, I find that this is both necessary and sufficient for pure-strategy equilibria: a pure-strategy Nash equilibrium can be reached with some full support priors if and only if it does not involve weakly dominated actions. Further, I provide a sufficient condition for a more robust selection result, whereby a pure-strategy Nash equilibrium is a limit point *any* full support priors. This condition is more permissive than strict Nash equilibria, for which it is intuitive that this robust selection holds.

[2] As is well-known (e.g. Shapley 1964), under fictitious play — whereby players best-respond to the empirical distribution of past play — this empirical distribution of play can cycle and fail to converge to Nash equilibrium when players have three or more actions.

Away from the zero cost limit, sequential sampling equilibrium rationalizes deviations from Nash equilibrium. A pervasive pattern of choice distributions that emerges in 2×2 games in experimental settings is that increasing the payoffs associated with an action of a player leads that player to choose that action more often. This has been termed the "own-payoff effect" and has been widely documented in contexts where the Nash equilibrium predictions go against it, as in the case of generalized matching pennies games (e.g. Goeree and Holt 2001). I prove that sequential sampling equilibrium not only predicts the own-payoff effect but, in contrast to existing models, also matches specific patterns in the data regarding how the joint distribution of choices and decision time changes with payoffs. In particular, when increasing a player's payoffs to a given action, my model also predicts an "opponent-payoff time effect", with the opponent choosing the best response to such an action more often and *faster*. This provides a novel prediction regarding how choices relate to stopping times in this class of games, which is borne out by existing experimental evidence.

These results follow from two monotone comparative statics for the underlying individual choice process that may be of independent interest. First, I establish that increasing the payoffs to a given action increases the probability not only that it is chosen but also that it is chosen earlier, a result that generalizes beyond two-action settings. That is, even though increasing payoffs could lead to a greater value to sampling further and potentially discovering that such an action is suboptimal, the individual requires less information to be convinced to take such an action. Second, I show that the probability that an action is chosen and the time it takes to choose it is also monotone in the true probability that the action is optimal, which, in a strategic setting, is given by the opponent's gameplay. Moreover, in both cases, the opposite is true for the other action.

Sequential sampling equilibrium can also account for the experimental findings regarding reported beliefs mentioned earlier: *randomness*, *payoff-dependence* of the players' beliefs, and *bias*. Players' equilibrium beliefs, that is, the beliefs held at the time when they make their choices, depend on the sample path of signals they have observed. This implies that players' equilibrium beliefs are random. Moreover, as payoffs influence the benefits and extent of information acquisition, not only individuals' choices but also their beliefs and stopping time depend on their payoffs. This implies that, in arbitrary games, equilibrium beliefs may be biased on average. I explore this feature in 2×2 games and uncover a systematic relation between beliefs and decision times. I show that, under specific conditions, beliefs held at very fast or very slow decisions are more biased in mean. This result builds upon the fact that beliefs closer in mean to the distribution over opponent's actions that makes the player indifferent between the two actions, a prediction supported by experimental data.

Finally, I discuss how to extend sequential sampling equilibrium to games of incomplete information and more general information structures. It is straightforward to adjust my solution concept to Bayesian games by having samples include information on the realized actions as well as the state. An analogous result to that of convergence to Nash equilibrium is then obtained: limit points of Bayesian sequential sampling equilibria as sampling costs vanish are Bayesian Nash equilibria. Furthermore, I consider the case where players cannot perfectly distinguish between states in their samples. In this case, I prove that limit points of a sequence of equilibria with vanishing costs are analogy-based expectations equilibria (Jehiel 2005; Jehiel and Koessler 2008).

To summarize, sequential sampling equilibrium constitutes a flexible equilibrium framework for analyzing strategic interaction. As my results show, it provides a rationale for standard solution concepts, accounts for several behavioral patterns that have been documented in experiments, and makes novel predictions not just on choices that individuals make in strategic settings but also on timed stochastic choice data, the joint distribution of choices, beliefs and decision times.

Related Literature

This paper contributes to three different literatures: belief formation in games, costly information acquisition and cognitive limitations and, more broadly, sequential sampling.

Within the literature focusing on belief formation in games, the papers closest to mine use sampling as a mechanism to account for players' choices. Osborne and Rubinstein (2003) focus on the case where each player observes a fixed number of samples from their opponents' equilibrium distribution of actions with the mapping from samples to actions being exogenously specified. Salant and Cherry (2020) study a special case of this solution concept in mean-field games with binary actions, while keeping the sampling procedure exogeneous. In particular, this paper explicitly considers that players form beliefs based on the observed sample and best-respond to their beliefs. Osborne and Rubinstein (1998) examine a similar notion of equilibrium, where players receive a fixed number of samples from the payoffs of each of their actions and choose the action with the highest average payoff in the sample. Rubinstein and Wolinsky (1994) consider the case where players form beliefs by observing a signal about others' gameplay.

In contrast to these papers, my solution concept endogenizes the signals players obtain by making information acquisition the object of choice of the player. The endogeneity of the sampling process results, for instance, in players sampling more when payoffs are scaled up, thereby affecting equilibrium beliefs and thus gameplay, a phenomenon that cannot be captured with exogenous sampling. One example of how this is the fact that, in a class of dominance-solvable games, scaling up payoffs has no effect on sampling equilibrium predictions, but sequential sampling equilibrium predicts that it leads results in to choices consistent with more steps of iterated deletion of dominated actions.

Another reason why players may form mistaken beliefs is because they conflate behavior of players of different types. Jehiel and Koessler (2008) examine this phenomenon by adapting Jehiel's (2005) analogy-based expectations equilibrium to look at equilibrium gameplay in static Bayesian games, where players form beliefs by averaging gameplay of different types. Eyster and Rabin (2005) look at a similar concept, cursed equilibrium, where beliefs are formed by combining each type's equilibrium behavior and the average behavior of opponents of all types. As mentioned earlier, this paper provides a rationale for analogy-based expectations equilibrium, and thus, for fully cursed equilibrium, as a limit case of sequential sampling equilibria in static Bayesian games.

Beliefs may also be misspecified, whereby players assign zero probability to some situations. Esponda and Pouzo's (2016) Berk-Nash equilibrium allows for general forms of misspecification of the players' prior beliefs and is not restricted to either normal-form or complete information games. There, players best-respond to their equilibrium beliefs, those in the support of players' priors that minimize the Kullback–Leibler divergence to equilibrium gameplay. While their framework can be seen as the limit case of Bayesian learning from sampling with potentially misspecified priors, I focus on characterizing belief formation in games where such information acquisition is costly. In contrast to sequential sampling equilibria, whenever players' priors have full support, Berk-Nash equilibria coincide with Nash equilibria. Furthermore, as in Nash equilibrium, equilibrium beliefs in Berk-Nash equilibrium will in general preclude any strategic uncertainty as equilibrium beliefs will be degenerate except on knife-edge cases.

This paper is also related to existing models of costly information acquisition in games. Yang (2015) studies equilibrium in a coordination game where players can acquire unrestricted but costly information on an exogenous payoff-relevant parameter. As in much of the rational inattention literature (Sims 2003; Matějka and McKay 2015), the cost of information is given by the decrease of the priors' entropy. Denti (2018) allows for players to obtain correlated information and for more general information cost functions (as in Caplin and Dean 2015). While ex-ante costly information acquisition has been recently connected to costly sequential sampling (Morris and Strack 2019), the main difference with respect to my solution concept is that these papers examine the case where strategic uncertainty is fully driven by the uncertainty about the exogenous parameter. As in these models players' beliefs are correct, whenever there is no uncertainty about this exogenous parameter, equilibria in these papers correspond to Nash equilibria. Instead, sequential sampling equilibrium aims to capturing strategic uncertainty and belief formation even in complete information games.

A different approach to costly information acquisition is taken by Alaoui and Penta (2016; 2018). These authors provide an axiomatic basis for a model of choice deriving from a cost-benefit analysis of reasoning to form beliefs about their opponents, endogenizing types in the level-k model (Stahl and Wilson 1994; Nagel 1995). While their model allows for non-equilibrium gameplay, one of their representation results (Alaoui and Penta 2018, Theorem 4) corresponds to an analogue of sequential sampling equilibrium where players follow myopic information acquisition strategies. Differently from these papers, I focus on deriving comparative statics within an equilibrium framework where players are forward-looking in their information acquisition. As mentioned earlier, this distinction matters since in the forward-looking case gameplay converges to Nash equilibrium when sampling costs vanish, whereas in the myopic case this is no longer the case.

It is natural to compare sequential sampling equilibria with fictitious play and learning in games more generally. Following the original interpretation of equilibrium beliefs as given by a scenario where players "accumulate empirical evidence" (Nash 1950), fictitious play (Brown 1951) has players myopically best-responding to observed or simulated frequency of past play. Such an approach does not necessarily lead to gameplay converging to Nash equilibrium as shown by Shapley (1964). However, it does provide a rationale for Nash equilibria as it characterizes steady states of such a process (Fudenberg and Kreps 1993). Thus, the steady-state characterization of sequential sampling equilibria provides a clear analogue to the characterization of Nash equilibria as steady-states of fictitious play. The crucial difference between fictitious play — or the more general learning processes considered in Fudenberg and Kreps (1993) — and the dynamic process I analyze is that whereas data is freely observable in fictitious play, sequential sampling players face information acquisition costs.

A myopic learning model with sampling is given by Oyama et al. (2015). The authors show how a dynamic version of Osborne and Rubinstein's (2003) sampling equilibrium satisfying a condition on the distribution of samples leads to almost global asymptotic stability of p-dominant actions — actions that are strict best-responses to having at least p fraction of the population playing them. Their results fail to apply to the dynamic process in this paper as not only is the sampling distribution an endogenous object, sampling is not independent of sample sizes. In my model, as information acquisition optimally depends on payoffs and beliefs, the sample size is not independent from the sample path, as players will stop sampling earlier or later depending which observations realize.

A less related paper on learning in strategic settings is that by Kalai and Lehrer (1993), who focus instead on infinitely repeated finite games played by forward-looking Bayesian agents who maximize their discounted expected utility. The authors show that when players have subjective beliefs about their opponents' strategies for the repeated game, then gameplay converges to a Nash equilibrium of the repeated game. Differently, this paper looks not at repeated games but instead at games that are played infrequently and where learning occurs offline.

Lastly, this paper also contributes to the literature on sequential sampling. The optimal stopping problem that players as individuals solve in this paper differ from the standard problem in Wald (1947) and Arrow et al. (1949) in that the support of their priors is not finite. It is also different from the problem in Fudenberg et al. (2018), as there individuals have correct priors and can take binary actions. I contribute with novel results to this literature: comparative statics results on how choices, stopping times, and beliefs depend on payoffs and the true probability distributions, as well as more specific results in the special case of Beta priors. Additionally, I provide a general condition for optimal stopping time to be bounded regardless of the true probability distribution by proving an extension of a result on uniform consistency of Bayesian updating by Diaconis and Freedman (1990) that makes the problem of solving for optimal stopping computationally tractable for arbitrary finite actions.

Outline

An outline of the remainder of the paper is as follows: Section 1 introduces the players' information acquisition problem. In Section 2, I define sequential sampling equilibrium and discuss existence, its interpretation and how equilibria are steady states of a dynamic process related to fictitious play. Section 3 examines its relation to Nash equilibrium when sampling costs vanish. Section 4 derives several implications of optimal stopping in binary problems and uses these to obtain comparative statics results in 2×2 games. Extensions to Bayesian games and more general information structures are the focus of Section 5. Finally, I conclude with a discussion of specific avenues for further work in Section 6. The proofs can be found in the Appendix.

1. Sequential Sampling

1.1. Setup

Let $\Gamma = \langle I, A, u \rangle$ denote a normal-form game, where I denotes a finite set of players or roles, $A := \times_{i \in I} A_i$ where A_i is i's finite set of feasible actions and $u := (u_i)_{i \in I}$ where $u_i : A \to \mathbb{R}$ is i's payoff function. I will write $-i$ to denote $I \setminus \{i\}$, $\sigma_i \in \Sigma_i := \Delta(A_i)$, $\sigma_{-i} \in \Sigma_{-i} := \Delta(A_{-i})$, where $\sigma_{-i}(a_{-i}) = \prod_{j \in -i} \sigma_j(a_j)$. I also extend u_i to the space of probability distributions over actions with $u_i(\sigma_i, \sigma_{-i}) = \mathbb{E}_\sigma[u_i(a_i, a_{-i})]$.

In contrast to standard solution concepts, each player $i \in I$ is uncertain about others' gameplay and has a belief $\mu_i \in \Delta(\Sigma_{-i})$ about σ_{-i}. In the case where players are restricted to believing each of the opponents' gameplay is independent, I instead define their beliefs to be given by a product measure $\mu_i = \times_{j\in -i}\mu_{ij}$, where each μ_{ij} is a probability measure on Σ_j. When a prior belief μ_i can be written in such way as a product measure, I say it does not allow for correlation and that it does if otherwise.

Player i's problem is to choose σ_i in order to maximize their expected utility given their beliefs, $\mathbb{E}_{\mu_i}[u(\sigma_i, \sigma_{-i})]$, and their value function is then given by

$$v_i(\mu_i) := \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i, \sigma_{-i})].$$

Prior to making a choice, player i can sample from the unknown probability distribution σ_{-i} at a cost. That is, player i can observe realizations of a stochastic process $\mathbf{X}_i = \{X_{i,t}\}_{t\in\mathbb{N}}$ defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with $\mathbb{F}_i$ denoting the natural filtration of $\mathbf{X}_i$. I will write $X_i^t = (X_{i,\ell})_{\ell=1}^t$ to stand for the sample path up to time t, where each realization $X_{i,\ell}$ is distributed according to σ_{-i}, $\mathbb{P}(\omega : X_{i,\ell}(\omega) = a_{-i}) = \sigma_{-i}(a_{-i})$ or, equivalently, $X_{i,\ell} \sim \sigma_{-i}$, with the understanding that $X_i^0 = \emptyset$. The player's beliefs $\mu_i \in \Delta(\Sigma_{-i})$ then induce a joint distribution on Σ_{-i} and the realizations of the stochastic process, given by $\mathbb{P}_{\mu_i}(S_{-i} \times B) = \int_{S_{-i}} Q^{\infty}_{s_{-i}}(B)\mu_i(ds_{-i})$ where $Q^{\infty}_{s_{-i}}$ is the infinite product measure on the space of sample paths given $s_{-i} \in \Sigma_{-i}$, $S_{-i} \subseteq \Sigma_{-i}$ and measurable set B of sample paths.

I denote the set of all finite sample path realizations by $\mathscr{X}_i := \bigcup_{t\in\mathbb{N}} A_{-i}^t$. Upon observing a given sample path up to time t, X_i^t, player i updates beliefs on Σ_{-i} according to Bayes' rule, denoted by $\mu_i|X_i^t$ and which I assume is well defined — which is the case whenever, for instance, the prior puts strictly positive probability on fully mixed gameplay, $\mu_i(\text{ int }(\Sigma_{-i})) > 0$. I will write $\mu_i \mid x_i^t$ to denote $\mu_i \mid X_i^t = x_i^t$.

For the sake of simplicity, sampling costs are taken to be linearly increasing in the number of samples with $c_i \in \mathbb{R}_{++}$ denoting the constant flow cost. General results extend to the case where they are given by a non-negative, non-decreasing function on $\mathbb{N}$ that diverges to $+\infty$ and has non-decreasing successive differences, that is, $c_i(t+2) - c_i(t+1) \geq c_i(t+1) - c_i(t)$, $t \in \mathbb{N}$.

An **extended game** G is then a tuple comprising an underlying normal-form game Γ, a vector $c = (c_i)_{i\in I} \in \mathbb{R}_{++}^{|I|}$ where c_i is i's sampling cost, and $\mu = (\mu_i)_{i\in I}$ where $\mu_i \in \Delta(\Sigma_{-i})$ is i's prior about their opponents' gameplay.

1.2. Optimal Stopping

Before formally defining the solution concept, let us first consider the players' sampling problem in isolation.

In order to maximize their expected payoffs, each player i faces an optimal stopping problem: based on accumulated evidence, they can decide whether to stop and make a choice or to acquire a new sample. Player i decides on a stopping time t_i in the set $\mathbb{T}_i$ of all stopping times adapted with respect to $\mathbb{F}_i$ with values in $\mathbb{N}_0 \cup \{\infty\}$, taking into account the sampling cost c_i. Let V_i denote the resulting value function, where

$$V_i(\mu_i) := \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i} \left[v_i \left(\mu_i \mid X_i^{t_i} \right) - c_i \cdot t_i \right] \tag{SP$_i$}$$

with a slight abuse of notation by letting $\mathbb{E}_{\mu_i}(\cdot)$ correspond to the expectation taken with respect to $\mathbb{P}_{\mu_i}$. Note that V_i is bounded above by $\max_{a \in A} u_i(a_i, a_{-i})$, which ensures V_i is well-defined and finite-valued.

I recast i's optimal stopping problem as a dynamic programming problem. Let B_i be a function mapping from the set of real-valued functions on $\Delta(\Sigma_{-i})$ to itself where

$$B_i(V)[\mu_i] := \max \left\{ v_i(\mu_i) \, , \, \mathbb{E}_{\mu_i} \left[V(\mu_i \mid X_i) \right] - c_i \right\}$$

The following is a well-known result, given that V_i is bounded above:

Fact 1. Let V be a solution to the functional equation $B_i(V) = V$. Then, $V = V_i$; i.e., it is a value function satisfying (SP$_i$).

As is standard in optimal stopping problems, I focus on the earliest optimal stopping time τ_i, where

$$\tau_i = \inf \left\{ t \in \mathbb{N}_0 \,\middle|\, V_i(\mu_i \mid X_i^t) = v_i(\mu_i \mid X_i^t) \right\} \tag{1}$$

whenever the right-hand side is non-empty, and $+\infty$ when otherwise. As a consequence of the above fact, we have:

Fact 2. The stopping time τ_i is a solution to the optimal stopping problem (SP$_i$).

The proof for these facts follows standard arguments (see e.g. Ferguson 2008, ch. 3).

Some general properties of the solution to this optimal stopping problem are given in the next two propositions.

Proposition 1. The optimal stopping time τ_i is finite with probability 1 with respect to μ_i, that is, $\mathbb{P}_{\mu_i}(\tau_i < \infty) = 1$. Moreover, for any $t \in \mathbb{N}_0$ and for any true probability distribution $\sigma_{-i} \in \Sigma_{-i}$, both $\mathbb{P}_{\mu_i}(\tau_i \leq t)$ and $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq t)$ decrease with the sampling cost c_i.

Proposition 1 states not only that players expect their stopping time to be finite, but also that stopping time increases in a first-order stochastic dominance sense as the sampling cost decreases. The first claim follows from the fact that, as expected utility is bounded from above, if the player expects to sample for an infinite amount of time with positive probability, they will have an infinitely negative expected utility and is therefore better-off stopping immediately. The second is due simply to the fact that increasing the sampling cost reduces the continuation value, leading the player to stop earlier. As we will see, although the latter statement holds for both the distribution of players' stopping time under their prior and the true distribution of their samples, the former holds only with respect to their prior and, without further assumptions, may fail to hold under the true distribution of their samples.

To state the next results, let us introduce some further notation. For $a_i \in A_i$, let $\geq_{a_i}$ denote the partial order on utility functions $u_i : A \to \mathbb{R}$ such that $u'_i \geq_{a_i} u_i$ if, for every $a' \in A$, $u'_i(a'_i, a'_{-i}) \geq u_i(a'_i, a'_{-i})$, with equality for all $a'_i \neq a_i$. I will write $M_i(a_i)$ to denote the set of beliefs $\mu_i \in \Delta(\Sigma_{-i})$ at which i stops sampling and $a_i \in A_i$ is optimal, that is,

$$M_i(a_i) := \left\{ \mu_i \in \Delta(\Sigma_{-i}) \,\middle|\, V_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a_i, s_{-i})] \right\},$$

where $M_i(a_i)$ depends on the sampling cost c_i and the utility function, as these affect the value function V_i.

Proposition 2.

(i) For any $a_i \in A_i$, $M_i(a_i)$ is convex and increases (decreases) with respect to set inclusion in the utility function u_i with respect to $\geq_{a_i}$ ($\geq_{a'_i}$, $a'_i \in A_i \setminus \{a_i\}$).

(ii) For any $a_i \in A_i$, $\sigma_{-i} \in \Sigma_{-i}$ and $t \in \mathbb{N}_0$, $\mathbb{P}_{\sigma_{-i}}\left(\tau_i \leq t \cap (\mu_i \mid X_i^{\tau_i}) \in M_i(a_i)\right)$ is increasing (decreasing) in the utility function u_i with respect to $\geq_{a_i}$ ($\geq_{a'_i}$, $a'_i \in A_i \setminus \{a_i\}$).

This result provides comparative statics that are not only of interest to the study of decision time in non-strategic settings, but, importantly, it will also enable us to later on characterize how changes in payoffs affect sequential sampling equilibrium predictions. First, the proposition states that the set of beliefs at which decision-makers stop and take a given action is convex and characterizes how it changes with respect to the payoffs associated with that same action and to those of other actions. Second, it shows that, for any true distribution of the samples, the probability that a given action is optimal upon stopping increases with that action's payoffs and, furthermore, that the probability of taking such action earlier also increases. Note that this does not mean that the probability of stopping earlier increases. In short, as action a_i is now more attractive payoff-wise, the decision-maker requires less evidence to be sufficiently convinced to stop sampling and take that action. The opposite is true if, instead, $a_i' \neq a_i$ becomes more attractive.

The intuition for the proof is easily summarized. Convexity of this set follows from convexity of the value function. To see that it increases in the set inclusion order with higher payoffs to that action, note that the continuation payoff is only affected if the decision-maker stops with the same action. Therefore, if it is optimal to stop at a given belief under lower payoffs to that action, it will still be optimal to stop at those same beliefs under higher payoffs. An opposite reasoning holds when considering the effects of increasing payoffs to action a_i on the beliefs at which it is optimal to stop and take another action $a_i' \neq a_i$: If the decision-maker did not stop and take action a_i' under beliefs μ_i under lower payoffs to action a_i, they will certainly not be more convinced to stop and take action a_i' when payoffs to a_i increase. These two observations together imply the second claim in the proposition.

A corollary to Proposition 1 relates the sampling cost to changes in $M_i(a_i)$:

Corollary 1. For any $a_i \in A_i$, $M_i(a_i)$ increases in the set inclusion partial order in the sampling cost c_i.

The corollary allows us to observe that, while scaling up the payoffs to action a_i enlarges $M_i(a_i)$, scaling up the payoffs to all actions — which is equivalent to decreasing the sampling cost — shrinks $M_i(a_i)$. This implies that, as the sampling cost decreases, the player collects more information before stopping and taking action a_i.

2. Equilibrium

By observing different sample paths, players may end up making different choices, that is, their information acquisition will induce a distribution of choices. My equilibrium notion imposes a consistency requirement between the distribution of the observations they sample and the choices that their optimal sequential sampling procedures induce.

Let $\mathscr{X}_i(u_i,\mu_i,c_i)$ denote the set of sample paths at which player i stops according to the optimal stopping time τ_i when endowed with utility function u_i, prior μ_i and sampling cost c_i, that is,

$$\mathscr{X}_i(u_i,\mu_i,c_i) := \left\{x_i^t \in \mathscr{X} \;\middle|\; V_i(\mu_i \mid x_i^t) = v_i(\mu_i \mid x_i^t) \cap V_i(\mu_i \mid x_i^{t-h}) > v_i(\mu_i \mid x_i^{t-h}),\ \forall h = 1,\dots,t\right\}.$$

In other words, $\mathscr{X}_i(u_i,\mu_i,c_i)$ characterizes the set of *stopping sample paths*.

Upon stopping, player i chooses some $s_i \in \arg\max_{\sigma_i\in\Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i,\sigma_{-i})]$. Let $b_i : \Delta(\Sigma_{-i}) \to \Sigma_i$ denote a *selection of optimal choices* at a given belief $\nu_i \in \Delta(\Sigma_{-i})$,

$$b_i(\nu_i) \in \arg\max_{\sigma_i\in\Sigma_i} \mathbb{E}_{\nu_i}[u(\sigma_i,\sigma_{-i})].$$

As $X_{i,\ell} \sim \sigma_{-i}$, let $f_i : \sigma_{-i} \mapsto f_i(\sigma_{-i})$ denote player i's *expected gameplay* induced by optimal sequential sampling from their opponents' true distribution of choices, given a selection rule, that is,

$$\begin{aligned} f_i(\sigma_{-i}) &:= \mathbb{E}_{\sigma_{-i}}\left[b_i\left(\mu_i|X_i^{\tau_i}\right)\right] \qquad (2)\\ &= \underbrace{\sum_{x_i^t\in\mathscr{X}_i(u_i,\mu_i,c_i)}}_{\substack{\text{Sum over all}\\ \text{stopping sample paths}}} \underbrace{\prod_{\ell=1}^{t} \sigma_{-i}(x_{i,\ell})}_{\substack{\text{Prob. reaching}\\ \text{stopping sample path}\\ x_i^t \text{ given } \sigma_{-i}}} \underbrace{b_i(\mu_i \mid x_i^t)}_{\substack{\text{Best-response}\\ \text{to beliefs } \mu_i\mid x_i^t}}, \end{aligned}$$

assuming player i stops in finite time almost surely. That is, the probability of player i taking action a_i, $f_i(\sigma_{-i})[a_i]$, is given by the probability of taking such an action once player i stops after having observed sample path x_i^t, $b_i(\mu_i \mid x_i^t)[a_i]$, and the probability that such sample path is observed. The probability that this sample path is observed is given by $\prod_{\ell=1}^{t} \sigma_{-i}(x_{i,\ell})$, as each observation corresponds to an action profile $x_{i,\ell} \in A_{-i}$, sampled independently from i's opponents' gameplay, $\sigma_{-i} \in \Sigma_{-i}$.

My solution concept corresponds to a consistency condition on overall gameplay given the players' sequential information acquisition, whereby the distribution of a player's samples matches the distribution of their opponents' actions. Formally,

Definition 1. A strategy profile σ is a **sequential sampling equilibrium** of the extended game G if, $\forall i \in I$, there is a selection of optimal choices $b_i : \Delta(\Sigma_{-i}) \to \Sigma_i$ such that $\sigma_i = \mathbb{E}_{\sigma_{-i}}\left[b_i\left(\mu_i | X_i^{\tau_i}\right)\right]$, where $X_{i,\ell} \sim \sigma_{-i}$, and $\mathbb{P}_{\sigma_{-i}}(\tau_i < \infty) = 1$.

Given its fixed-point definition, one can view a sequential sampling equilibrium as a self-enforcing distribution of action data. When players sample from accumulated past data distributed according to a sequential sampling equilibrium, under the same selection of best-responses to their beliefs, their expected gameplay will match the distribution of accumulated past data. As I show in a later section, this interpretation is well-grounded in a steady-state foundation for the solution concept.

Many interpretations of the sampling process are possible. For instance, players can be thought of as parsing existing data on past actions or asking friends, incurring a cost per acquired data point or friend inquired. Another interpretation is that players are drawing from their own memory past realizations or more general information that helps them reason about how others would act in such situation,[3] echoing recent developments in the neuroscience literature on decision-making (see, e.g. Bornstein and Norman 2017; Bakkour et al. 2018).

Additionally, sequential sampling equilibrium notion provides a way to relax other solution concepts, which it nests as a special case. When players' priors assign probability one to a single probability distribution of their opponents associated with the same Nash equilibrium of the underlying game, that Nash equilibrium will coincide with a sequential

[3] I discuss more general signal structures in Section 5.

sampling equilibrium of the game.[4] Thus, sequential sampling equilibrium relaxes in a particular manner the implicit epistemic assumption in Nash equilibrium that, in equilibrium, players come to know their opponents' gameplay. Further, players in my solution concept need not know others' payoff functions, as neither their information acquisition nor the choices it entails rely on such knowledge. Hence, my model also dispenses with the assumption of mutual knowledge of the game and of others' rationality. Osborne and Rubinstein's (2003) sampling equilibrium also corresponds to a variant of my model where players can sample at no cost the first t observations and face an arbitrarily large cost to any further sampling.

2.1. Existence

Contrary to what one might think, existence is not immediate. For some priors and sampling costs, a sequential sampling equilibrium need not exist in the associated extended game. If the priors are degenerate or sampling costs are so high that players optimally choose not to sample, then sequential sampling equilibria are just the players' best-responses to their priors. However, existence is not ensured for general priors as players may want to sample indefinitely. The next example illustrates this fact.

Example 1.

		Player 2		
		L	M	R
Player 1	U	1,0	0,1	0,0
	D	0,0	0,1	1,0

Figure 1. **Non-existence**

[4]As it is implicit in this statement, even though beliefs are degenerate and coincide on the same Nash equilibrium, not all best-responses will coincide with that same Nash equilibrium, which explains why there may be multiple sequential sampling equilibria instead of there being a unique equilibrium coinciding with the Nash equilibrium players believe to occur. Moreover, such non-uniqueness can occur even when the game has a unique Nash equilibrium. This echoes Aumann and Brandenburger's (1995) results on the epistemic characterization of Nash equilibrium, whereby conjectures — and not choices — are found to coincide with Nash equilibrium.

Let Γ be a two-player normal-form game as given by Figure 1. As M is a strictly dominant action for player 2, for any extended game $G = \langle \Gamma, \mu, c \rangle$, any sequential sampling equilibrium $\overline{\sigma}$ must have that $\sigma_2(M) = 1$. Suppose that player 1's prior μ_1 assigns equal probability to $\sigma_2' = (1-3\epsilon, 2\epsilon, \epsilon)$ and to $\sigma_2'' = (\epsilon, 2\epsilon, 1-3\epsilon)$, where the first, second and third elements of σ_2 correspond to the probability with which L, M and R are played, respectively. Then, were player 1 to observe M, the posterior is the same as the prior, given the symmetry in σ_2' and σ_2'', that is, $\mu_1 \mid M = \mu_1$. When instead L (M) is observed, the posterior then places a larger probability on σ_2' (σ_2'') than σ_2'' (σ_2').

Then, given the symmetry in the problem, the value player 1 assigns to taking just one additional sample — also called the expected value of sample information (Raiffa and Schlaifer 1961, ch. 5A) — given by $\mathbb{E}_{\mu_1}\left[v_1(\mu_1|X_1)\right] - v_1(\mu_1)$, is strictly positive. As the value obtained from optimal stopping is always weakly greater than the value of myopic sampling policy, $V_1(\mu_1) \geq \mathbb{E}_{\mu_1}\left[v_1(\mu_1|X_1)\right]$, then for any sampling cost that is low enough, player 1 will deem it worthwhile to sample, as $V_1(\mu_1) - c_1 \geq \mathbb{E}_{\mu_1}\left[v_1(\mu_1|X_1)\right] - c_1 > v_1(\mu_1)$. Given that under any equilibrium $\sigma_2(M) = 1$, player 1's posterior is always identical to the prior for any sampling path where samples are distributed according to σ_2, i.e. $\mu_{1,t} = \mu_1 \mid X_1^t = \mu_1 \mid (M, \ldots, M) = \mu_1$. Consequently, player 1 will always consider it worthwhile to continue sampling as $V_1(\mu_{1,t}) - c_1 > v_1(\mu_{1,t})$ $\forall t \in \mathbb{N}$. As player 1 will then never stop sampling and, consequently, will never make a choice, a sequential sampling equilibrium does not exist in that extended game. □

Example 1 shows that, even if player i's optimal stopping time is finite with probability one with respect to i's own prior, it may actually be finite with probability zero under the true distribution of the sampling process when i's prior is misspecified. In fact, this is the sole impediment for existence of a sequential sampling equilibrium. That is, existence is ensured insofar as all players stop sampling in finite time almost surely for any gameplay of their opponents.

In order to address potential issues of non-existence, I will impose that players' priors have full support. Player i's prior μ_i is said to have **full support** if it assigns positive probability to every open neighborhood of every $\sigma_{-i} \in \Sigma_{-i} \subset \mathbb{R}^{|-i|-1}$, where Σ_{-i} is endowed with the Euclidean topology.

Under such condition, a strong result on the optimal stopping time ensues.

Proposition 3. Suppose that μ_i is has full support. Then, $\exists T(u_i, \mu_i, c_i) = T_i \in \mathbb{N}_0$ such that $\forall \sigma_{-i} \in \Sigma_{-i}$, $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq T_i) = 1$.

As in finite dimensional spaces, the Bayesian learning is consistent for any distribution if and only if the prior has full support (Freedman 1963), Proposition 3 uncovers an important consequence of Bayesian learning for optimal stopping: Not only is the decision-makers' optimal stopping time finite with probability one, for any true distribution of their samples, it is also bounded uniformly across all distributions of samples. This effectively transforms the optimal stopping problem from infinite to finite horizon, allowing for a solution to be obtained by backward recursion, simplifying the problem significantly.[5]

The intuition underlying the result is that if the prior has full support, the posterior accumulates around the empirical mean. Then, one can guarantee a bound on the rate at which the posterior accumulates around the empirical mean, depending on the number of samples but not on the sample path itself (Diaconis and Freedman 1990). With this, it is possible to bound the gains in expected payoff of sampling further regardless of the realized sample path and show that there is a number of observations after which the cost of an additional observation dwarfs the expected gain, regardless of realizations. Hence, one concludes that the decision-maker necessarily stops after such number of samples and we can find an explicit upper bound for the stopping time that depends only on the prior μ_i, payoffs u_i, and sampling cost c_i. This stands in contrast to the canonical problem in Arrow et al. (1949) where the prior has finite support, of which Example 1 is an illustration, and

[5]In Appendix C.1, I provide a refinement of this result. I show that for a specific class of priors that are commonly used in applications, Dirichlet priors, the optimal stopping time is bounded uniformly over both the true distribution of the samples and all priors in this family, characterizing the exact upper bound. This suggests a tractable manner to numerically solve for optimal stopping and, thus, for sequential sampling equilibria, which may be of use to practitioners in applying the solution concept to the data.

optimal stopping time is not bounded.[6] The argument in the proof also extends to more general sampling costs, provided these are non-decreasing, with non-decreasing successive differences and unbounded from above.

An immediate implication of Proposition 3 is that it ensures the induced expected gameplay given σ_{-i} is well-defined, with $f_i(\sigma_{-i}) \in \Sigma_i$. As a consequence, it is a sufficient that players have full-support priors for there to be a sequential sampling equilibrium.

Theorem 1. Let G be an extended game where players' priors have full support. Then, a sequential sampling equilibrium exists.

The result follows from standard arguments. Since f_i is not only well-defined but also continuous, Brouwer's fixed-point theorem applies.

As the discussion at the start of this subsection reflects, full-support priors are sufficient but not necessary for an equilibrium to exist. In fact, I provide sufficient conditions for non-degenerate but misspecified priors such that the result in Proposition 3 holds and, consequently, a sequential sampling equilibrium exists. An extension of Diaconis and Freedman's (1990) provided in the Appendix shows that Bayesian updating with a general class of misspecified priors leads to beliefs uniformly accumulating around the points in the support that minimize the Kullback-Leibler divergence with respect to the empirical mean, which contributes to the classical convergence result by Berk (1966). To ensure uniqueness of such a minimizer and preclude situations as that illustrated in Example 1, I require the prior's support to be convex.

2.2. Sampling from Past Data

For any equilibrium model, an important question is how players may come to behave according to the model's predictions. In the case of sequential sampling equilibrium, a valid concern is that it exhibits an inherent circularity, as players sample from their opponents' equilibrium distribution of choices. I previously stated that sequential sampling equilibria

[6] Similarly, optimal stopping time is also not bounded in the continuous-time version of the canonical problem, with Gaussian noise, be it with (Moscarini and Smith 1963) or without experimentation concerns (Chernoff 1961). In some cases with finite support prior, however, stopping time can be bounded, as in the case with Poisson arrival of conclusive information, but not when the decision-maker can choose from different information sources (Che and Konrad 2019).

can be thought of as a steady state of a process where a sequence of populations of players sampling from accumulated past data. This section formalizes that argument.

The dynamic process is as follows. Fix an extended game G. At each period, $n = 1, \ldots,$ a unit measure of agents $j \in J$ plays the game, evenly divided across the different roles I of the finite normal-form game underlying G and randomly matched. For every period n, agents from the previous period are replaced by a new population of agents as is standard in evolutionary models of learning in strategic settings.

Each agent in role i can sample from past accumulated data from opponents' realized actions, that is, each observation is sampled according to its empirical frequency in all past periods, with $\sigma_0 \in \Sigma$ given. Sampling is sequential and optimal with respect to the agent's sampling cost c_i, their prior μ_i and utility function u_i and I assume that agents cannot observe their calendar time or that of the observations. When the player stops sampling, they best-respond to their beliefs.[7] I call this process **dynamic sequential sampling**.

The process induced by dynamic sequential sampling is akin to fictitious play (Brown 1951). For instance, when beliefs are Dirichlet and agents sample a fixed number of times, this corresponds to a variant of fictitious play with sampling (Kaniovski and Young 1995). Under dynamic sequential sampling, however, the number of samples a given agent draws is an endogenous object and determined in a sequentially optimal manner.

It is now shown that indeed any steady state of the dynamic sequential sampling process is a sequential sampling equilibrium. Take the operator $f : \Sigma \to \Sigma$ from the definition of sequential sampling equilibrium, where $f(\sigma) = (f_i(\sigma_{-i}))_{i \in I}$ and $f_i(\sigma_{-i}) = \mathbb{E}_{\sigma_{-i}}\left[b_i\left(\mu_i \mid X_i^{\tau_i}\right)\right]$ denotes the expected gameplay of agents in role $i \in I$, given a selection of best-responses b_i to the agent's beliefs held when they stop sampling. Let $\sigma_{n-1} \in \Sigma$ denote the distribution of actions in the data accumulated up to time $n-1$, $n \in \mathbb{N}$. Then, dynamic sequential sampling gameplay at time t is distributed according to $f(\sigma_{n-1})$ and the resulting distribution of accumulated data at the end of the period is then given by

$$\sigma_n = \frac{n}{n+1}\sigma_{n-1} + \frac{1}{n+1} f(\sigma_{n-1}).$$

[7] The results in this section go through if both sampling costs and priors are idiosyncratic to each agent and drawn from a fixed distribution every period.

I will denote $\{\sigma_n\}_{n\in\mathbb{N}_0}$ as the **dynamic sequential sampling gameplay process** induced by the extended game G.

Our first result in this section shows that, whenever the distribution of accumulated data converges, it converges to a sequential sampling equilibrium, as does, of course, gameplay.

Theorem 2. For any extended game G and any $\sigma_0 \in \Sigma$, if the induced dynamic sequential sampling gameplay process converges, its limit is a sequential sampling equilibrium of G. Moreover, every sequential sampling equilibrium of G is a limit of a convergent dynamic sequential sampling gameplay process.

Theorem 2 establishes for sequential sampling equilibrium and the dynamic process I defined above an analogue to the seminal result in Fudenberg and Kreps (1993) relating Nash equilibria and fictitious play, in the sense that sequential sampling equilibria coincide with the limits of convergent dynamic processes. Moreover, I show that the dynamic process as I defined above can be generalized. In the process described above, each observation has the same weight. One could think of a setting where the observations from the near-past are more easily accessible. This can be modelled as a giving a different weight to each period, for instance, exponential weighting: $\sigma_n = \alpha\sigma_{n-1} + (1-\alpha)f(\sigma_{n-1})$, $\alpha \in (0,1)$. The claim in Theorem 2 also holds under this alternative definition. The assumption that there is a continuum of agents for each role is also not essential: A similar result holds when the populations are finite.

While cycling may occur and preclude convergence of gameplay — similarly to what occurs with fictitious play[8] — in specific classes of games, convergence and asymptotic stability are guaranteed.[9] This next proposition provides one such condition, albeit a very stringent one.

Proposition 4. Let $G = \langle\Gamma,\mu,c\rangle$ be an extended game such that Γ is a 2×2 game with a unique Nash equilibrium and players' priors are absolutely continuous. Then, for any selection of best-responses, the dynamic sequential sampling gameplay process $\{\sigma_n\}_{n\in\mathbb{N}_0}$

[8]The classical reference is Shapley (1964). Additionally, cycling can also occur with stochastic fictitious play: see Hommes and Ochea (2012).

[9]An equilibrium $\overline{\sigma}$ is asymptotically stable if for all $\epsilon > 0$, there is a $\delta > 0$ such that for any $\sigma_0 : \|\sigma_0 - \overline{\sigma}\|_\infty < \delta$, $\|\sigma_n - \overline{\sigma}\|_\infty < \epsilon$. That is, if the dynamic sequential sampling gameplay process, starting at σ_0 close enough to the equilibrium, remains close thereafter.

converges to a sequential sampling equilibrium and this sequential sampling equilibrium is globally asymptotically stable.

The proof for the first claim exploits results specific to 2×2 games that will be discussed in Section 4. In particular, uniqueness of a Nash equilibrium imposes specific conditions on the payoff structure. These conditions, as we will see later on, have a direct translation into properties of the operator f, which I use to show local stability based on the eigenvalues of the Jacobian matrix of the dynamic system. Then, local stability is extended to global stability on the simplex by relying on a proof of the Jacobian conjecture for $\mathbb{R}^2$.

3. Relation to Nash Equilibrium

Our solution concept has players forming equilibrium beliefs according to an initial interpretation of Nash equilibrium beliefs, whereby these are approximately reached by players "accumulat[ing] empirical information" (Nash 1950). However, differently from what is considered there, here players face a cost to acquire such information. Hence, a natural question is whether, as these costs vanish, sequential sampling equilibria converges to a Nash equilibrium. In this section I show that such intuition is indeed correct when priors have full support.

As the sampling costs decrease, players that do not have a dominant action will acquire more and more samples. Provided that the optimal stopping time grows unboundedly as costs go to zero, one would expect that, by the law of large numbers, players learn the true distribution of actions of their opponents and sequential sampling equilibrium converges to a Nash equilibrium. Unfortunately, once we condition on stopping, the observations in the stopping sample path cease to be independent or identically distributed, precluding such a proof strategy. Therefore, I take a different approach.

I rely on the fact that optimal stopping time minimizes expected regret gross of sampling costs and then show that a particular suboptimal stopping rule achieves zero expected regret gross of sampling costs. This particular stopping rule corresponds to stopping after a fixed number of samples, unconditionally on the realized sample path and depending only on the sampling cost. Then, optimal stopping will also attain null regret as costs vanish, which is then used to show that sequential sampling equilibria converge to Nash equilibria.

Let us formally define regret. The regret player i expects under stopping time $t_i \in \mathbb{T}_i$, given their utility function, sampling cost and prior and their choices upon stopping, is

$$R_i(t_i; u_i, \mu_i, c_i) := \mathbb{E}_{\mu_i}\left[\max_{\sigma_i' \in \Sigma_i} u_i(\sigma_i', s_{-i})\right] - \mathbb{E}_{\mu_i}\left[u_i\left(b_i(\mu_i \mid X_i^{t_i}), s_{-i}\right)\right].$$

Player i's regret is then the difference between the expectation over the maximum utility achievable, where the player knows the opponents' gameplay, and the expected utility attained from following stopping time t_i.

A first result is that the *optimal stopping time* τ_i for player i minimizes the sum of expected regret and total sampling costs.

Lemma 1. For any selection of best-responses b_i,

$$\tau_i \in \arg\min_{t_i \in \mathbb{T}_i} R_i(t_i; u_i, \mu_i, c_i) + \mathbb{E}_{\mu_i}[c_i \cdot t_i].$$

The proof of Lemma 1 extends Fudenberg et al.'s (2018) Proposition 2, which addresses the case of two actions, continuous time and correct priors, with payoffs following independent Gaussian distributions; I show that this holds in this environment as well. It states that the optimal stopping problem that player i faces is equivalent to a regret minimizing problem. This equivalence underlies the next lemma which is key for the main result in this section:

Lemma 2. Let μ_i have full support. Then, for any sequence of sampling costs $\{c_{i,n}\}_{n\in\mathbb{N}} \subset \mathbb{R}_{++}$ such that $c_{i,n} \to 0$ and for any selection of best-responses b_i, $R_i(\tau_{i,n}; u_i, \mu_i, c_{i,n}) + \mathbb{E}_{\mu_i}[c_{i,n} \cdot \tau_{i,n}] \to 0$, where $\tau_{i,n}$ is an optimal stopping time under sampling cost $c_{i,n}$.

As mentioned at the beginning of this section, the proof involves constructing a potentially suboptimal stopping time involving sampling a fixed number of samples that depends only on the sampling cost. Then, we can use the law of large numbers for the observations in stopping sample paths under this particular stopping rule and show that it attains zero regret, gross of total sampling costs. The result then follows from Lemma 1.

The main result of this section is that, as sampling costs vanish, sequential sampling equilibrium converges to a Nash equilibrium of the underlying game. Let $\Sigma^{SSE}(G)$ denote the set of sequential sampling equilibria of the extended game G. The formal statement is as follows:

Theorem 3. Let $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^{|I|}_{++}$ such that $c_n \to 0$ and $G_n = \langle \Gamma, \mu, c_n \rangle$ be an extended game with full-support priors. For any sequence $\{\sigma_n\}_{n\in\mathbb{N}} \in \times_{n\in\mathbb{N}} \Sigma^{SSE}(G_n)$, its limit points are Nash equilibria of the underlying game Γ.

The crucial argument in the proof of Theorem 3 is simple. If a limit point of such a sequence is not a Nash equilibrium, then some player must experience strictly positive ex-post regret. However, as sampling costs vanish, holding opponents' gameplay fixed, the player face no expected regret as shown in Lemma 2. The result follows by showing that the former conclusion implies a violation of the latter.

It is noteworthy that convergence of sequential sampling equilibria to Nash equilibrium crucially hinges on information acquisition to be sequentially optimal. In the Appendix, I introduce a version of sequential sampling equilibrium where, instead of optimal stopping, players are myopic with respect to information acquisition, considering the informational value of just the next sample and neglecting the continuation value.[10] Given such sampling rule, I then illustrate how it fails to converge to a Nash equilibrium. The main reason for this failure is that, in the limit, myopic sequential sampling players generically acquire a bounded number of samples as sampling costs vanish.

3.1. Reachability of Nash Equilibria

While equilibrium gameplay converges to Nash equilibria when priors have full support, this does not imply that *all* Nash equilibria are reachable as a limit. As such, reachability by sequential sampling equilibria with full-support priors leads to some selection principle. The purpose of this section is to discuss which Nash equilibria can be selected in this manner.

[10]Such model has a direct correspondence to a representation of the value of reasoning in by Alaoui and Penta (2018, Theorem 4).

Let us start by formally defining reachability.

Definition 2. Given game Γ, gameplay $\sigma \in \Sigma$ is **reachable** with priors μ if there is a sequence of sampling costs $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^{|I|}_{++}$ with $c_n \to 0$ and a sequence of sequential sampling equilibria $\{\sigma_n\}_{n\in\mathbb{N}} \in \times_{n\in\mathbb{N}}\Sigma^{SSE}(G_n)$, where $G_n = \langle \Gamma, \mu, c_n \rangle$ is an extended game, such that $\sigma_n \to \sigma$.

From Theorem 3, it is immediate that only Nash equilibria of the underlying game Γ are reachable with full-support priors. Three questions arise: which Nash equilibria cannot be reached, which can be reached with some collection of full-support priors, and which can be reached with all full-support priors. In the next proposition I provide a sharp characterization of the sets of Nash equilibria that can and cannot be reached.

Proposition 5.

(i) A Nash equilibrium is reachable with some full-support priors only if no weakly dominated action is chosen with positive probability.

(ii) A pure-strategy Nash equilibrium is reachable with some full-support priors μ (potentially allowing for correlation) if and only if it does not involve weakly dominated actions.

Proposition 5 uncovers a meaningful relation of reachability by sequential sampling equilibrium and dominance. First, it states that no Nash equilibrium involving weakly dominated strategies with strictly positive probability is reachable. This follows from the fact that full-support priors will, by definition, confer some positive mass to the event that opponents choose actions under which a weakly dominated action yields a strictly lower payoff than some other choice, making it unappealing for the player to ever choose it. Second, I provide a partial converse: that any pure-strategy Nash equilibrium not involving weakly dominated actions is reachable with some full-support priors. This results from the fact that an action which is not weakly dominated must be a best-response to some interior action distribution of the opponents (potentially correlated) (Pearce 1984; Weinstein 2020). Thus, for part (ii) of the proposition, priors may need to allow for correlated opponent gameplay. An immediate corollary to Proposition 5 is that any trembling-hand perfect Nash equilibrium in pure strategies is reachable with some full-support prior.

Another interesting question is to characterize the set of Nash equilibria which one can reach with *all* full-support priors. This stands as a robust selection criterion, as it requires the action distribution to be the limit of a sequence of sequential sampling equilibria regardless of the full-support priors players have. For this purpose, let us consider the following condition:

Definition 3. A Nash equilibrium σ^* of a game Γ is **strongly robust** if there is an $\epsilon > 0$ such that for any player $i \in I$, $\forall \sigma_{-i} \in B_\epsilon(\{\sigma^*_{-i}\})$, $b_i \in \arg\max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \sigma_{-i})$.

This condition simply requires that σ^*_i be a best-response to any distribution of actions of the opponents in an ϵ-neighborhood around σ^*_{-i}. This is satisfied, for instance, whenever σ^*_i is a strict best-response to σ^*_{-i}, but the requirement is weaker, as it admits other actions to be a best-response to σ^*_{-i}. On the other hand, it is stronger than trembling-hand perfection, as it requires that every players' Nash equilibrium strategy be a best response to any tremble of the opponents' strategies.[11] Under strong robustness, the following result holds:

Proposition 6. Let σ be a pure-strategy Nash equilibrium of game Γ. For any full-support priors μ, $\exists c \in \mathbb{R}^{|I|}_{++}$ such that σ is a sequential sampling equilibrium of the extended game $G = \langle \Gamma, \mu, c \rangle$ if an only if σ is a strongly robust Nash equilibrium,

This result directly implies that whenever a Nash equilibrium is strongly robust, it is reachable with any full-support priors. A simple corollary of Proposition 6 again relates dominance and reachability:

Corollary 2. Any strict Nash equilibrium is reachable with any full-support priors.

[11] Strong robustness strictly strengthens the requirements imposed by trembling-hand perfect as given by the equivalence laid out in Mas-Colell et al. (1995, Proposition 8.F.1). In fact, there are trembling-hand perfect equilibria that cannot be reached with some full-support priors. Moreover, it can be shown that any Nash equilibrium that satisfies strong robustness constitutes a singleton stable set (Kohlberg and Mertens 1986), but that the converse is not true. A discussion and proof can be found in the Appendix C.

4. 2×2 Games

In this section I will focus on 2×2 games, that is, two-player games where each player has two actions, which is a canonical class of games used both for simple theory as well as for experiments. I will label $A_i = A_{-i} = \{0,1\}$ and, consequently, identify Σ_i and Σ_{-i} with $[0,1]$, with σ_i denoting the probability that player i chooses action 1. For this section, I will consider only priors $\mu_i \in \Delta([0,1])$ that are absolutely continuous on $[0,1]$, the set of which I will denote by $\mathscr{M}$. I denote the density of prior μ_i by $d\mu_i$.

In order to derive predictions of my solution concept for this class of games, I will first provide some useful general implications of optimal stopping for a given player i, taking the opponent's behavior as exogenous. In other words, I characterize properties of the individual optimal stopping problem in binary settings upon which the analysis of equilibrium comparative statics will be developed.

4.1. Optimal Stopping and Comparative Statics

Let us first consider the optimal stopping problem for player i, with σ_{-i} as an exogenous distribution. There are three possible cases to consider: (i) player i has a weakly dominated action, (ii) player i is indifferent between the two actions no matter which action the opponent takes, and (iii) player i has no weakly dominated action nor is indifferent between the two actions. In cases (i) and (ii), it is immediate that player i will never benefit from acquiring any information and then $\tau_i = 0$. The main difference between (i) and (ii) is that in (i), under any full-support prior, player i will choose the weakly dominant action with probability 1 and in (ii) any mixing between the two actions is optimal. Focusing now on case (iii) and given that I will be focusing on player i's standpoint as a decision-maker, assume without loss that $u_i(1,1) \geq u_i(0,1)$,[12] for otherwise actions of player $-i$ can be relabeled. In this case, $u_i(1,1) - u_i(0,1) = \delta_1 > 0$ and $u_i(0,0) - u_i(1,0) = \delta_0 > 0$.

The following lemma shows that, in this class of games, a player will never stop sampling when indifferent between the two actions, provided the player samples at least once.

[12] Where $u_i(0,1)$ corresponds to the payoff player i obtains when player i chooses 0 and player $-i$ chooses 1.

Lemma 3. Let $\mu_i \in \mathcal{M}$. If $\tau_i > 0$, then for any selection b_i of optimal choices, $b_i(\mu_i \mid X_i^{\tau_i}) \in \{0,1\}$.

Lemma 3 simplifies the discussion of comparative statics by allowing us to focus on pure actions, ignoring randomization. The reasoning behind the proof is simple. Suppose that player i stops sampling after observing a 0-valued sample leaving player i indifferent between the two actions (the argument is symmetric if the last sample is 1-valued). Then, before observing the last observation, action 1 is optimal under player i's prior, as observing a 0-valued observation induces a lower belief mean. Moreover, if the last observation had instead realized to be 1-valued, player i would still want to choose action 1. This implies that if player i stops sampling when indifferent between the two actions, whichever action was optimal before taking the last sample is still optimal regardless of the realization of the sample. Therefore, given that the player will not sample any further, the last sample bears no informational value to the player. As the sample is costly, then it is suboptimal to take it.

In order to state the main result in this section, let us introduce some notation. I will say that a prior $\mu_i \in \mathcal{M}$ MLR-dominates another prior $\mu_i' \in \mathcal{M}$ — writing $\mu_i \geq_{MLR} \mu_i'$ —, if $d\mu_i(\sigma_{-i}) \cdot d\mu_i'(\sigma_{-i}') \geq d\mu_i(\sigma_{-i}') \cdot d\mu_i'(\sigma_{-i})$, whenever $\sigma_{-i} \geq \sigma_{-i}'$, for $\sigma_{-i}, \sigma_{-i}' \in [0,1]$. This simply means that the probability density functions associated with μ_i and μ_i' have the monotone likelihood ratio property, which is preserved under Bayesian updating.

The main theorem in this section characterizes comparative statics of the joint distribution of choices and stopping time with respect to payoffs, the opponents' gameplay and the player's prior.

Theorem 4. Let $u_i(1,1) - u_i(0,1) = \delta_1 > 0$, $u_i(0,0) - u_i(1,0) = \delta_0 > 0$ and $\mu_i \in \mathcal{M}$. Then, for any time t and any selection of optimal choices b_i,

$$\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \leq t\right)$$

(i) is increasing in the prior μ_i according to $\geq_{MLR}$;

(ii) is strictly increasing and infinitely continuously differentiable in the opponent's gameplay σ_{-i}; and

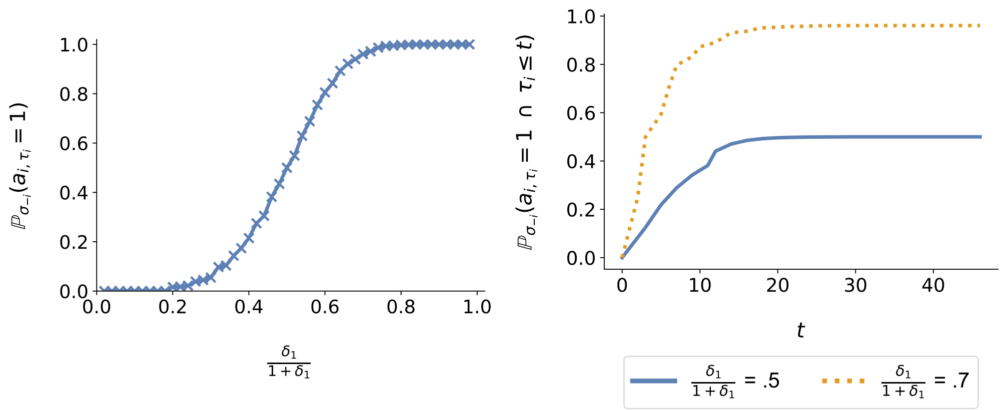
(iii) is increasing in payoff difference δ_1 and decreases in δ_0.

Theorem 4 indicates how the probability that a player choosing an action up to time t varies with (i) the player's prior, (ii) the gameplay of the opponent, and (iii) the player's payoffs. As one might suspect, these results will serve as key inputs in providing comparative statics for sequential sampling equilibrium in 2×2 games.

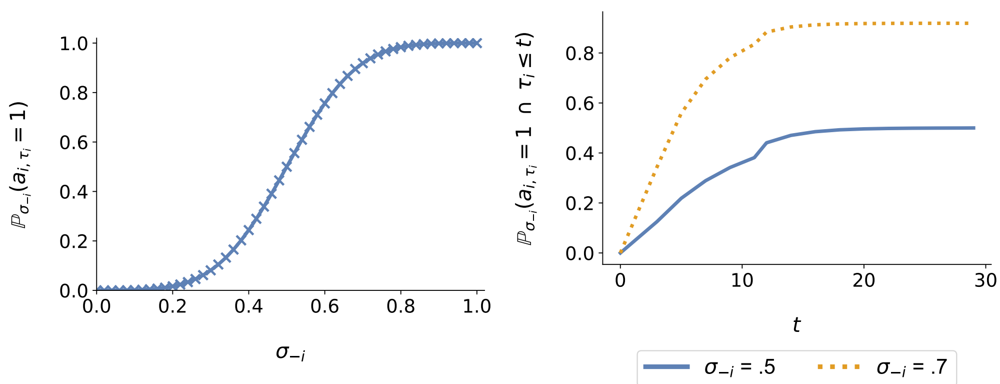
Let us go over the intuition for behind this result, keeping in mind that the expected value to action 1 is increasing in σ_{-i}. The proof for claim (i) follows from showing that the value function is increasing in $\geq_{MLR}$. From there, I show that given two priors, $\mu_i, \mu_i' \in \mathcal{M}$ where $\mu_i \geq_{MLR} \mu_i'$, if the player stops at μ_i' and takes action 1, then stopping at μ_i and taking action 1 is also optimal.[13] The argument for why claim (ii) should hold is straightforward: higher σ_{-i} will increase the probability any given sample takes value 1 which makes player i more likely to observe a sample path conducing to taking action 1. The fact that the probability of choosing action 1 before time t is a polynomial with respect to σ_{-i} implies the claim on differentiability. Importantly, note that claim (ii) does not depend on whether player i's beliefs are correct or not. Instead, it relates both the choices and the stopping time to the unknown true distribution of opponent's actions that player i is uncertain about, regardless of the prior beliefs that player i holds. Finally, claim (iii) follows immediately from the fact that the set of stopping beliefs at which a given action is optimal increases in set inclusion with respect to that action's payoffs (Proposition 2) and the observation that the player will never stop sampling when indifferent (Lemma 3).

Figure 2 illustrates the findings regarding the effects of varying payoffs, opponent's gameplay, and the prior on the joint probability of taking a given action up to time t. As shown in Theorem 4, the comparative statics are monotone. On the left-hand-side panels, I show how the probability that action 1 is chosen is increasing in the payoffs associated to action 1 (panel 2a); the probability with which the opponent chooses action 1 (σ_{-i}), for which player i's action 1 is optimal (panel 2b); and player i's prior, in the MLR partial order (panel 2c). To easily obtain MLR-ranked priors, I use the fact that, for Beta priors with parameters $\hat{t}_i \cdot (\hat{\sigma}_{-i}\,,\, 1 - \hat{\sigma}_{-i}) \in \mathbb{R}^2_{++}$, the priors with higher $\hat{\sigma}_{-i} \in (0,1)$ dominate those with lower $\hat{\sigma}_{-i}$ in the MLR partial order, for any fixed $\hat{t}_i > 0$. On the right-hand-side panels, I

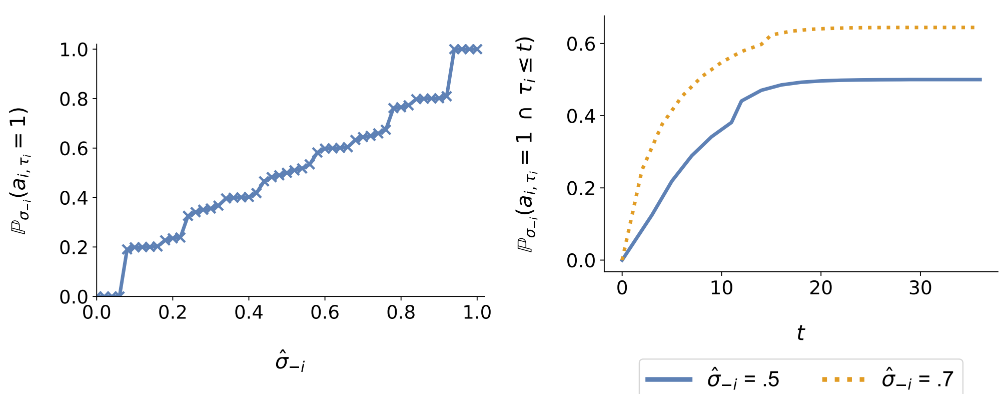
[13] This result can be extended beyond absolutely continuous priors, provided they are ranked according to strong stochastic dominance (Lehrer and Wang 2020) which is necessary and sufficient for the priors to be ranked according to first-order stochastic dominance and, for any realized sample path, for the posteriors will retain the ranking. When the priors are absolutely continuous, strong stochastic dominance corresponds to the monotone likelihood ratio ranking, $\geq_{MLR}$.



(a) **Monotonicity in Payoffs**



(b) **Monotonicity in Opponent's Gameplay**



(c) **Monotonicity in the Prior**

Figure 2. **Comparative Statics of Timed Stochastic Choice Data**

Note: The figures illustrate the comparative statics results. Figures on the left illustrate the changes in the probability that player i takes action 1; on the right the probability of taking such an action up to time t is contrasted for two different parameter values. Payoffs are given by $u_i(1,1) - u_i(0,1) = \delta_1$ and $u_i(0,0) - u_i(1,0) = 1$ and the prior is given by a Beta distribution with parameters $2 \cdot (\hat{\sigma}_{-i}, 1 - \hat{\sigma}_{-i})$. Panel (a) varies payoffs $\delta_1 \in \mathbb{R}_+$. Panel (b) shows how different opponent's gameplay $\sigma_{-i} \in [0,1]$ affect optimal stopping. Panel (c) compares optimal stopping under different priors, varying $\hat{\sigma}_{-i}$, where priors with higher $\hat{\sigma}_{-i}$ MLR-dominate priors with lower values of $\hat{\sigma}_{-i}$. When unspecified, the prior is the uniform distribution ($\hat{\sigma}_{-i} = 1/2$), payoffs are symmetric ($\delta_1 = 1$) and opponent's gameplay is $\sigma_{-i} = 1/2$.

illustrate the fact that these comparative statics also extend to the probability with which player i takes action 1 *earlier*, that is, up to any given time t. It is worthy of mention that, while opponent's gameplay affects player i's choices in a continuous manner, the same does not happen with changes in either payoffs or the prior, owing to the discrete nature of the signal structure that player i has available.

A useful property that enables sharper predictions from optimal stopping is linearity of the prior in new information. This provides structure that mimics the behavior of Bayesian updating for Gaussian priors.

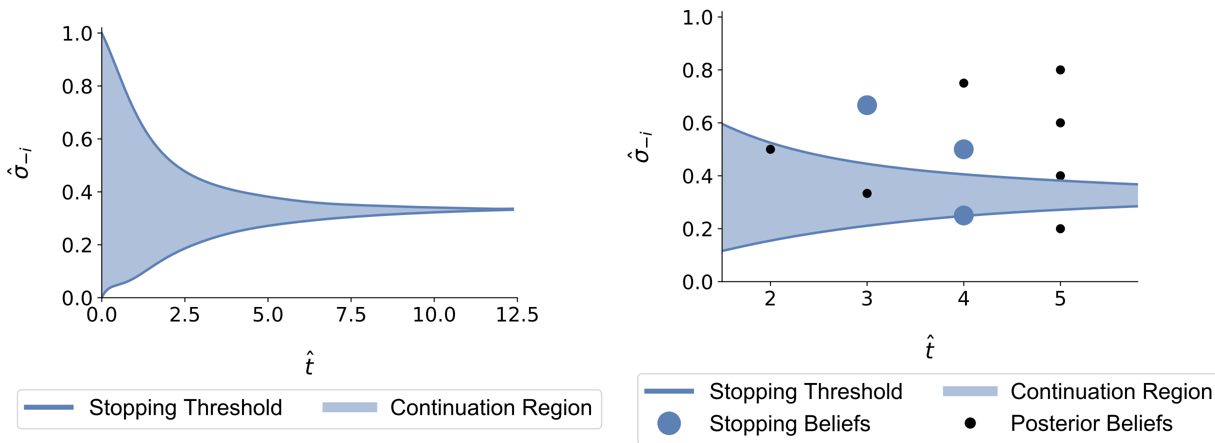
Definition 4. A prior μ_i is said to be **linear in the accumulated information** if it is non-degenerate and there are constants $a_t, b_t \in \mathbb{R}$ such that for any sample path $x_i^t \in \mathscr{X}_i$ of t observations the posterior mean satisfies $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i^t] = a_t \sum_{\ell=1}^t x_{i,\ell}^t + b_t$.

This property, together with the fact that beliefs are a martingale and some algebraic manipulation, allows us to write the posterior mean as a convex combination of the prior mean and the empirical mean of the accumulated information, $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i^t] = \alpha_t/t \cdot \sum_{\ell=1}^t x_{i,\ell}^t + (1-\alpha_t)\cdot \mathbb{E}_{\mu_i}[s_{-i}]$, where $\alpha_t/t = 1/((1-\alpha_1)/\alpha_1 + t) \in (0,1)$. This is extremely convenient as, by linearity of expected utility, one can then analyze optimal stopping just relying on the belief mean and the number of samples. In fact, as shown by Diaconis and Ylvisaker (1979, Theorem 5), identifies a specific parametric class of priors: a prior μ_i is linear in the accumulated information if and only if it is a Beta distribution.

Under such parametric restriction, we have the following characterization of the set of beliefs at which player i optimally stops:

Proposition 7. Let $u_i(1,1) - u_i(0,1) = \delta_1 > 0$, $u_i(0,0) - u_i(1,0) = \delta_0 > 0$ and μ_i be linear in the accumulated information. Then, there are functions $\overline{\sigma}_{-i}, \underline{\sigma}_{-i} : \mathbb{R}_{++} \to [0,1] \cup \{\emptyset\}$ such that player i keeps sampling at time t if and only if $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i^t] \in (\underline{\sigma}_{-i}(t), \overline{\sigma}_{-i}(t))$, where $\overline{\sigma}_{-i}$ is decreasing and $\underline{\sigma}_{-i}$ is increasing in t. Moreover, whenever $\overline{\sigma}_{-i}(t), \underline{\sigma}_{-i}(t) \in [0,1]$, $\overline{\sigma}_{-i}(t) \geq \delta_0/(\delta_0 + \delta_1) \geq \underline{\sigma}_{-i}(t)$.

Proposition 7 shows that, under the condition that beliefs follow a Beta distribution, it is sufficient to consider the posterior mean to characterize the beliefs at which player i continues sampling at any given moment as is illustrated in Figure 3. In particular, player



(a) **Stopping Regions for Beta Priors**

(b) **Optimal Stopping with Uniform Prior**

Figure 3. **Stopping Regions for Beta Priors**

Note: The panel on the left shows the stopping threshold for posterior means at which player i with a Beta prior stops. The upper and lower bound collapse to the indifference point, in this case 1/3. The panel on the right illustrates the possible posterior means that a player starting with a uniform prior (Beta(1,1)) can attain depending on the sample path observed. When this sample path exists the continuation region, the player stops sampling and takes the action which is optimal at that posterior belief; in this case, action 1, if the belief is above 1/3 and action 0 if otherwise. The parameters of the Beta distribution are defined for convenience as $\hat{t}\cdot(\hat{\sigma}_{-i}\,,\,1-\hat{\sigma}_{-i})$, where $\hat{t}>0$ and $\hat{\sigma}_{-i}\in[0,1]$ coincides with the posterior mean. Payoffs are such that $u_i(1,1)-u_i(0,1)=2(u_i(0,0)-u_i(1,0))>0$.

i samples if and only if the player is sufficiently close to being indifferent between the two alternatives. Moreover, the more information is accumulated, the closer to indifferent player i needs to be to continue sampling. This translates to this setting what is commonly known in the neuroscience literature as "collapsing boundaries" (e.g. Hawkins et al. 2015; Bhui 2019).

It is interesting to note that when the absolute difference in the expected payoffs is known — the case where the prior's support is a doubleton — the stopping region is characterized by fixed bounds in terms of the posterior means as shown by Arrow et al. (1949). In contrast, when there is richer uncertainty about the difference in expected payoffs, as when the prior is given by a Beta distribution, the stopping region is characterized by bounds that collapse to the posterior mean that makes the individual indifferent between the two alternatives. Thus, a clear parallel between the setup in this paper and that in Fudenberg et al. (2018) emerges, where the individual infers the difference in payoffs of two alternatives from the drift of a Brownian motion and a similar contrast between known and unknown payoff differences gives rise to, respectively, fixed and collapsing stopping bounds. An important difference is that in Fudenberg et al. (2018), collapsing boundaries hold on average and when individuals have correct priors, while in my model they hold even without these qualifications.

Exploring the structure provided by Beta priors, I uncover two further predictions for a special case of our environment. In order to state these results, let us define the probability that the player i chooses the best response to the true distribution of actions of player $-i$ (σ_{-i}) before time t as

$$p_i(t;\sigma_{-i}) := \mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) \in \arg\max_{\sigma_i \in [0,1]} u_i(\sigma_i, \sigma_{-i}) \,\middle|\, \tau_i \leq t\right).$$

Let us also define the expected bias in beliefs at time t as

$$\beta_i(t;\sigma_{-i}) := \left|\mathbb{E}_{\mu_i}[s_{-i} \,|\, X_i^{\tau_i}, \tau_i = t] - \sigma_{-i}\right|.$$

Finally, let μ_i be said to be symmetric whenever $d\mu_i(\sigma_{-i}) = d\mu_i(1-\sigma_{-i})$. Then, we have that:

Proposition 8. Let $u_i(1,1) - u_i(0,1) = u_i(0,0) - u_i(1,0) > 0$ and μ_i be linear in the accumulated information and symmetric. Then,

(i) $p_i(t;\sigma_{-i})$ is decreasing in t; and

(ii) $\beta_i(t;\sigma_{-i})$ is quasiconvex in t.

Proposition 8 is delivering two sharp predictions. First, it states that in symmetric environments, where the payoff differences to the two actions are the same, later actions are more likely to be worse. That is, there is a speed-accuracy complementarity. It is interesting to note that while collapsing bounds for the continuation region are necessary, are not sufficient to guarantee the result. Instead, the result follows from the exact functional form that these bounds take and that I characterize (for arbitrary payoffs).[14] However, in contrast to the findings in Fudenberg et al. (2018), this relation between speed and accuracy may be non-monotone once we step away from symmetric environments, even when

[14]This characterization may prove useful for practitioners and fitting the model to data and can be found in the Appendix B.

prior beliefs are correct on average. That is, even if $\mathbb{E}_{\mu_i}[s_{-i}] = \sigma_{-i}$, $p_i(t;\sigma_{-i})$ need not be monotone in t.[15]

The second claim in Proposition 8 regards the bias in beliefs that player i has and how it relates to decision time. I show that bias exhibits a clear pattern: it decreasing up to a given time t and increasing for later decisions.

4.2. Equilibrium and Comparative Statics

The analysis for equilibrium comparative statics in 2×2 games stems directly from the results on optimal stopping time.

A first observation is that uniqueness of a Nash equilibrium in a 2×2 game generically implies uniqueness of a sequential sampling equilibrium whenever players' priors are absolutely continuous.

Proposition 9. Let $G = \langle \Gamma, \mu, c \rangle$ be an extended game such that Γ is a 2×2 game and priors are absolutely continuous. Suppose each player $i \in I$ samples at least once, $\tau_i > 0$, or is not indifferent between the two actions under their prior, $\mathbb{E}_{\mu_i}[u_i(0, \sigma_{-i})] \neq \mathbb{E}_{\mu_i}[u_i(1, \sigma_{-i})]$. Then if Γ has a unique Nash equilibrium, there is a unique sequential sampling equilibrium. Furthermore, a converse holds if both players sample at least once.

The requirement in Proposition 9 for uniqueness of a sequential sampling equilibrium rules out cases where a player faces a sampling cost c_i so high that it is optimal to not sample at all *and*, at the same time, is indifferent between the two alternatives according to their prior. In such knife-edge cases, there will be multiple sequential sampling equilibria as any distribution of choices of that player is optimal.

The result follows from the comparative statics results derived in Theorem 4. For the case where the unique Nash equilibrium is in pure strategies, one of the players has a dominant strategy and will play one of the actions with probability one. Then, the other player will either does not sample and has a unique best response to their prior, as they are not indifferent between the two actions; or they do sample and, as the players never stop sampling when indifferent and as they will always observe the same sample path (the opponent plays the dominant action), the will also always play the best-response to

[15] Examples are provided in Appendix C.

		Player $-i$ 0	Player $-i$ 1
Player i	0	δ_0 , 0	0 , γ_0
	1	0 , γ_1	δ_1 , 0

Figure 4. **Generalized Matching Pennies**

Note: $\delta_k, \gamma_k > 0$, $k = 0, 1$.

their unique posterior mean upon stopping. In particular, when there is a unique Nash equilibrium in pure strategies and both players sample, the unique sequential sampling equilibrium coincides with the unique Nash equilibrium.

When, instead, the unique Nash equilibrium is in mixed strategies, the result is obtained by noting that, whenever a player i samples at least once, the probability of choosing any given action, $f_i(\sigma_{-i}) = \mathbb{P}_{\sigma_{-i}}(b_i(\mu_i \mid X_i^{\tau_i}) = 1)$ is continuous and monotone, being either strictly increasing or strictly decreasing for $\sigma_{-i} \in (0,1)$. Monotonicity in the opponent's gameplay together with the fixed point condition then implies uniqueness of a sequential sampling equilibrium. This unique sequential sampling equilibrium corresponds to the unique and asymptotically globally stable steady state of the dynamic sequential sampling gameplay process analyzed in Section 2.2. Moreover, it should be easy to see that, following the same arguments, a similar result holds when discussing symmetric equilibria: if there is a unique symmetric Nash equilibrium then, when both players hold the same prior and face the same sampling cost, there will be a unique symmetric sequential sampling equilibrium under the conditions of Proposition 9.

While uniqueness of a Nash equilibrium implies uniqueness of a sequential sampling equilibrium, it is not the case that the two coincide. As I mentioned before, they coincide whenever the Nash equilibrium is in pure strategies and one of the players samples at least once.[16] When the Nash equilibrium is in fully-mixed strategies — the only other case possible when the Nash equilibrium is unique in a 2×2 game —, the unique sequential sampling equilibrium will converge to the Nash equilibrium in the limit as sampling costs vanish (Proposition 5). However, for any fixed sampling costs, in general, not only can the Nash equilibrium and the sequential sampling equilibrium differ, the two will also imply different comparative statics, as we shall see next.

[16] This conclusion is immediate from the proof of Proposition 9.

A well-known and counter-intuitive prediction of Nash equilibrium pertains to generalized matching pennies, that is, 2×2 games with a unique Nash equilibrium in fully mixed strategies, whose structure is illustrated in Figure 4. When the payoffs to action 1 of player i increase, Nash equilibrium predicts that the probability with which action 1 is chosen remains the same and it is, instead, the opponent's mixed strategy that changes to make player i indifferent between choosing any of the two actions. However, experimental evidence shows that increasing player i's payoffs to an action leads that player to choose that action more often, what has been since termed the **own-payoff effect**.[17]

Sequential sampling equilibrium not only implies the own-payoff effect, it also provides additional predictions regarding beliefs and decision times. While quantal response equilibrium also rationalizes the own-payoff effect, it does so by directly embedding monotonicity of choices with respect to payoffs in the assumptions for players' behavior, being monotonicity one of its defining assumptions (Goeree et al. 2005). In contrast, in my solution concept this monotonicity in payoffs follows from the effects that payoffs have on how players acquire information and, consequently, on their stopping time, as this next proposition highlights.

Proposition 10. Let $G = \langle \Gamma, \mu, c \rangle$ be an extended game such that priors are absolutely continuous and Γ is a generalized matching pennies game. Suppose that both players sample at least once. Then,

(i) the probability of player i choosing action 1 increases in action 1's payoffs; and

(ii) the probability of player $-i$ choosing action 0 up to time t increases in player i's payoffs to action 1 for any $t \geq 0$.

The reasoning for the first claim in Proposition 10 is straightforward and illustrated in Figure 5. The higher payoffs to action 1 for player i, the more likely that player is to choose it for any distribution of player $-i$'s choices. As player i's probability of choosing action 1 is increasing in the probability that the opponent also chooses it and as the inverse is true for

[17] This finding has been replicated several times, namely by Ochs (1995), McKelvey et al. (2000) and Goeree and Holt (2001).

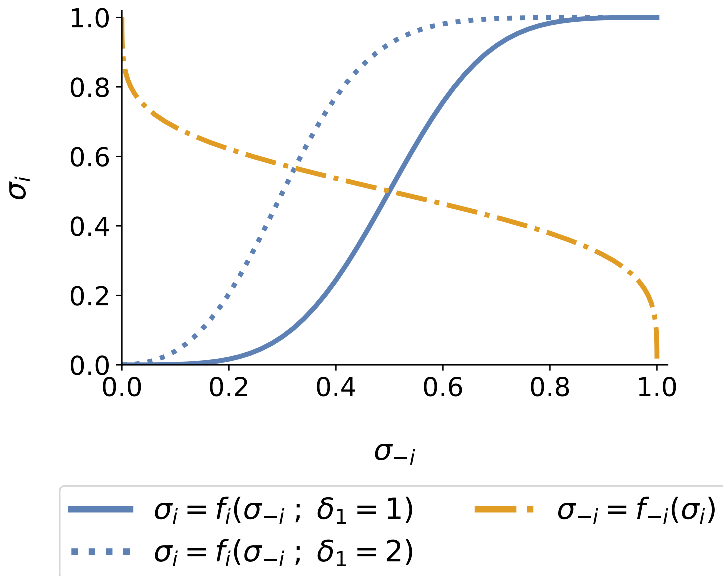


Figure 5. **Comparative Statics in Generalized Matching Pennies**

Note: The figure illustrates how sequential sampling equilibrium implies the own-payoff effect in the context of generalized matching pennies. The unique sequential sampling equilibrium is given by intersection of the blue and the orange curves. Priors correspond to the uniform distribution and payoffs for player i are given by $u_i(1,1) - u_i(0,1) = \delta_1$ and $u_i(0,0) - u_i(1,0) = 1$, and for player $-i$ $u_{-i}(0,1) - u_{-i}(1,1) = u_{-i}(1,0) - u_{-i}(0,0) = 1$.

player $-i$, the unique sequential sampling equilibrium shifts according to the own-payoff effect.[18]

The second comparative statics result regards stopping time and, therefore, has no parallel in the literature: player $-i$ chooses action 0 more often and faster. That is, for any time t, the probability with which player $-i$ stops sampling before t and takes action 0 increases. This result holds regardless of whether player i is choosing action 1 with high or low probability to start with. For player i, the increase in the payoffs to action 1 and the decrease in the probability their opponent chooses action 1 lead to opposing forces. In contrast to their net effect on choices alone, the net effect on the joint distribution of choices and stopping time is ambiguous. I call this the **opponent-payoff time effect**.

[18] A similar result holds in my model with respect to symmetric anti-coordination (extended) games. In such case, the unique symmetric sequential sampling equilibrium exhibits the own-payoff effect under the same conditions as in generalized matching pennies. This matches gameplay patterns documented in experimental settings by Chierchia et al. (2018) in the context of symmetric two-player anti-coordination games. As for 2×2 games with multiple Nash equilibria, all pure strategy Nash equilibria in undominated strategies can be sequential sampling equilibria of a given extended game with full-support priors, provided the sampling costs are sufficiently small — a proof of this claim follows the same arguments to that of Proposition 6. Consequently, no meaningful comparative statics results are possible.

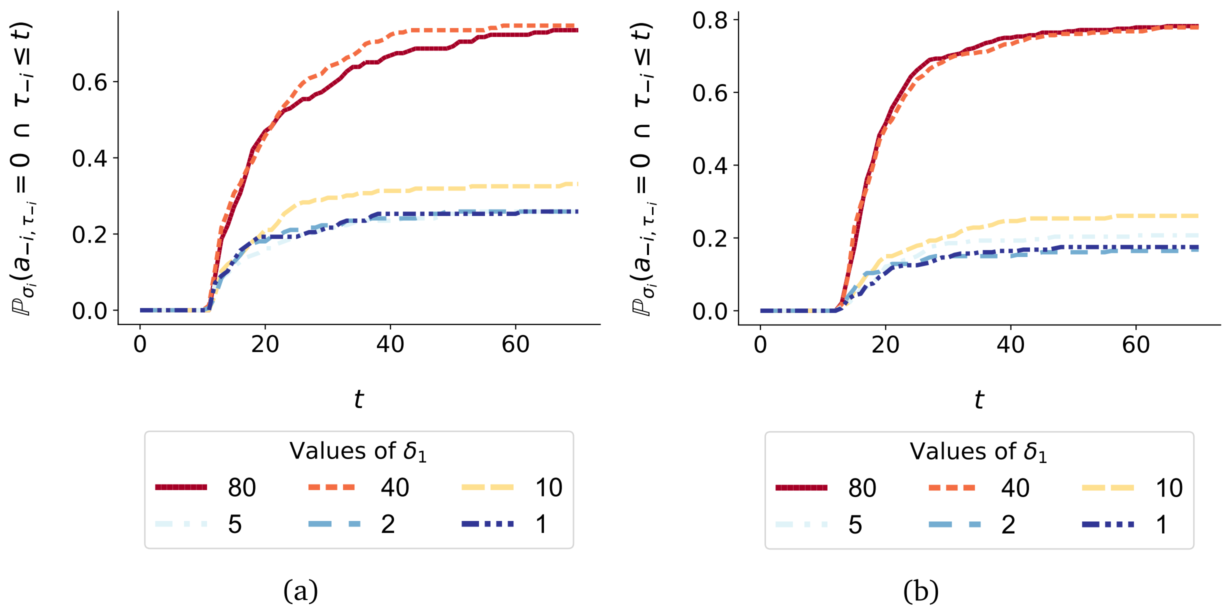


Figure 6. **Player 2's Joint Distribution of Choices and Decision Time**

Note: The figure uses the data from Friedman and Ward (2019) and shows how player $-i$'s joint probability of taking action 1 by time t shifts monotonically with player i's payoffs for taking action 1. The left (right) panel corresponds to treatments where beliefs are not (are) elicited. Payoffs are $u_i(1,1) = \delta_1$, $u_i(0,0) = u_{-i}(0,1) = u_{-i}(1,0) = 20$ and zero elsewhere.

Figure 6 provides suggestive evidence that the novel prediction in Proposition 10 is supported by existing experimental data if one is to interpret stopping time as a proxy for decision times. The data are from Friedman and Ward (2019)[19] who collect experimental data on choices and beliefs and record decision times for different generalized matching pennies games where only the payoffs to action 1 of player i vary. The figure suggests the exact pattern predicted by sequential sampling equilibrium: as the payoffs to player i's action 1 increase, the probability that player $-i$ subjects take action 0 up to any given time t also increases.

An important feature of sequential sampling equilibrium is that players' equilibrium beliefs — that is, beliefs held at the time where players make their choices — are themselves random. This follows from the fact that players' equilibrium beliefs depend on the sample path observed up to the moment when they stop sampling. Thus, the randomness in the observed samples translates into randomness of the player's beliefs.

[19] I thank the authors for generously making the dataset available.

Table 1. **Decision Time and Reported Beliefs**

Dep. Variable:	Distance between Reported Beliefs and Indifference Point		
	Player i	Player $-i$	Both
Decision Time (secs)	– 0.10**	– 0.07**	– 0.09***
	(0.04)	(0.03)	(0.03)
Player i	–	–	–7.55***
			(0.60)
Intercept	33.31***	40.77***	40.90***
	(1.89)	(0.92)	(0.97)
Controls	Yes	Yes	Yes
N	1620	1680	3300
R^2	0.08	0.27	0.17

Note: The table presents regression results on the relation between decision times and the distance between reported beliefs to indifference points with data from Friedman and Ward (2019), made available by the authors. Reported beliefs refer to the elicited beliefs — in an incentive-compatible manner and without feedback — about the probability the opponents would play action 1. Indifference point refers to the posterior mean that would make the player indifferent between taking either action and distance refers to the Euclidean metric. The left-most and middle columns use data for subjects in the role of player 1 and player 2 only, respectively; the right-most column uses both. The subjects played multiple generalized matching pennies with payoffs $u_i(1,1) = \delta_1 \in \{1,2,5,10,40,80\}$, $u_i(0,0) = u_{-i}(0,1) = u_{-i}(1,0) = 20$ and zero elsewhere. Controls for player role (i or $-i$) and for each game played were included. Heteroskedasticity-robust standard errors in parentheses. $*$: p-value $<.1$, $**$: p-value $<.05$, $***$: p-value $<.001$.

The results on optimal stopping also have a direct implication for equilibrium beliefs. If players have Beta-distributed priors, beliefs held at later stopping times are closer to their indifference point, that is, the opponent action distribution that makes the player indifferent between the two actions. Indeed, as shown in Proposition 7, under the assumption of Beta priors, the continuation region in binary settings is characterized by an upper and a lower bound on posterior means that monotonically collapse to this indifference point. Consistent with this analytical result, I find that, in Friedman and Ward's (2019) data, decision time is negatively correlated with the distance between reported beliefs and players indifference point, as shown in Table 1.

Furthermore, Propositions 8 and 10 together have direct implications for how the players' bias evolves over time as well as for how decision time relates to the probability of best-responding to the true distribution of actions of the opponent. This next corollary explicitly states that the predictions found earlier carry over to strategic settings.[20]

Corollary 3. Let $G = \langle \Gamma, \mu, c \rangle$ be an extended game such that priors are absolutely continuous and Γ is a generalized matching pennies game. If player i's prior is linear in accumulated information and symmetric, and if $u_i(1,1) - u_i(1,0) = u_i(0,0) - u_i(0,1) > 0$, then, at any sequential sampling equilibrium σ, player i's bias $\beta_i(t;\sigma_{-i})$ is quasiconvex in t and the probability that player i best-responds to the true distribution of the opponent, σ_{-i}, is decreasing in t.

5. Extensions to Games of Incomplete Information

In this section, I focus on games of incomplete information. I assume that players have access to data on gameplay and types as is, for example, the case with bids in past auctions. In such cases, however, it may be sensible to assume that players cannot perfectly distinguish types in the data. In keeping with the auction example, while the bidders' actual willingness-to-pay may not be observable, there could be data available on socio-demographic covariates providing some coarse signal on willingness-to-pay.

In what follows, I extend my solution concept to model these more general environments and discuss how it relates to both Bayesian Nash equilibrium and analogy-based expecta-

[20] As one would need several observations per individual — given that whether bias is increasing or decreasing depends on the idiosyncratic sampling cost —, it was not possible to test the implications on bias.

tion equilibrium (Jehiel and Koessler 2008). In brief, when players can observe types in their samples, as sampling costs vanish, the solution concept converges to a Bayesian Nash equilibrium, but when players cannot fully differentiate between types, limiting sequential sampling equilibria correspond to analogy-based expectations equilibria.

Let $\Gamma = \langle I, A, \Theta, u, \rho \rangle$ denote a game of incomplete information, where I is a finite set of players, $A = \times_{i\in I} A_i$, with A_i being i's finite set of actions, $\Theta = \times_{i\in I}\Theta_i$ with Θ_i denoting the finite set of player i's possible types, $u = (u_i)_{i\in I}$ a vector of payoff functions where $u_i : A \times \Theta \to \mathbb{R}$ denotes i's payoff function and $\rho \in \Delta(\Theta)$ a probability distribution over type profiles.

For each player i, let $\Xi_i := A_{-i} \times \Theta$. Each player is endowed with a prior $\mu_i \in \Delta(\Delta(\Xi_i))$, which allows for player i to know ρ or to be uncertain about the exact distribution of types. As before, each player is able to observe samples from a stochastic process $\mathbf{X}_i$ at a cost per observation of $c_i > 0$.

In contrast to previous sections, I am going to allow for more general information structures in that players need not perfectly observe the realized types and actions of their opponents. Let $\mathscr{E}_i$ be a partition of Ξ_i with generic element ϵ_i. The process $\mathbf{X}_i = \{X_{i,t}\}_{t\in\mathbb{N}}$ is now $\mathscr{E}_i$-valued, being defined on a general probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and $\mathbb{P}(\omega : X_{i,t}(\omega) = \epsilon_i) = \sum_{(a_{-i},\theta)\in e_i} \rho(\theta)\sigma_{-i,\theta_{-i}}(a_{-i})$, where $\sigma_{-i,\theta_{-i}}(a_{-i}) = \prod_{j\in -i}\sigma_{j,\theta_j}(a_j)$ and $\sigma_{j,\theta_j}(a_j)$ will denote the probability with which player j with type θ_j chooses action a_j.

Player i's problem is then analogous to that in complete information games, with $v_{i,\theta_i}(\mu_i) := \max_{\sigma_i\in\Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i, \sigma_{-i}, \theta) \mid \theta_i]$ and

$$V_{i,\theta_i}(\mu_i) := \sup_{t_i \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[v_{i,\theta_i}(\mu_i \mid X_i^{t_i}) - c_i \cdot t_i \mid \theta_i\right]$$

where $\mathbb{T}_i$ denotes the set of all stopping times adapted with respect to the natural filtration of $\mathbf{X}_i$ and $X_i^t = (X_{i,\ell})_{\ell=1}^t$. By the same arguments, the earliest optimal stopping time for a player i with type θ_i is given by $\tau_{i,\theta_i} := \inf\{t \in \mathbb{N}_0 \mid V_{i,\theta_i}(\mu_i \mid X_i^t) = v_{i,\theta_i}(\mu_i \mid X_i^t)\}$.

I will call an **analogy-based sequential sampling equilibrium** of the extended game $\langle \Gamma, \mathscr{E}, \mu, c\rangle$ a strategy profile $\overline{\sigma} = (\sigma_{i,\theta_i}\theta_i \in \Theta_i, i \in I)$ such that $\forall i \in I$ and $\forall \theta_i \in \Theta_i$, there is an optimal selection $\sigma^*_{i,\theta_i} : \Delta(\Delta(\Xi_i)) \to \Sigma_i$ where $\sigma^*_{i,\theta_i}(\mu_i) \in \arg\max_{\sigma_i\in\Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i, a_{-i}, \theta) \mid \theta_i]$

such that $\overline{\sigma}_{i,\theta} = \mathbb{E}_{\rho,\overline{\sigma}_{-i,\theta_{-i}}}\left[\sigma^*_{i,\theta_i}\left(\mu_i | X_i^{\tau_{i,\theta_i}}\right)\right]$, where with probability $\sum_{(a_{-i},\theta)\in\epsilon_i}\rho(\theta)\overline{\sigma}_{-i,\theta_{-i}}(a_{-i})$ $X_{i,t}=\epsilon_i$, with τ_{i,θ_i} finite almost surely.

The definition is analogous to sequential sampling equilibrium as defined earlier in the paper except that now players have types that refine their beliefs and a general signal structure that provides information on both gameplay and types. In particular, the partitions $\mathscr{E}_i$ allow for a player to not be able to distinguish between some actions that their opponents take at a specific state or type profile θ, which is precluded from the type of analogy partitions considered by Jehiel and Koessler (2008).

I will consider two specific kinds of analogy partitions:

Condition 1. For every $i \in I$, all elements of $\mathscr{E}_i$ are singletons.

Condition 2. For every player $i \in I$, there is a partition E_i of Θ_i such that for every $\epsilon_i \in \mathscr{E}_i$, there $e_i \in E_i$ and $a_{-i} \in A_{-i}$ where $\epsilon_i = \{a_{-i}\} \times e_i$. Moreover, for every $\theta_i, \theta'_i \in \Theta_i$ and $\theta_{-i}, \theta'_{-i} \in \Theta_{-i}$, if $(\theta_i, \theta_{-i}), (\theta'_i, \theta'_{-i}) \in e_i$, then $\theta_i = \theta'_i$.

Condition 1 states that the analogy partitions are the finest possible, implying that each player i will be able to perfectly observe the actions and types that are sampled. That is, the player could potentially fully learn the joint distribution of actions and types with enough observations being sampled and an adequate prior. Condition 2 instead states that the analogy partitions are finer than the partition induced by information the player possesses by knowing their type.[21] In particular, Condition 2 implies that actions of the opponents are perfectly observed but their types may not be. It provides an analogue to the kinds of analogy partitions in Proposition 2 of Jehiel and Koessler (2008). For every $e_i \in E_i$, I will denote by $e_i(\theta)$ the set of $\theta' \in \Theta$ such that $\theta, \theta' \in e_i(\theta)$.

The next proposition relates the limit of a sequence of analogy-based sequential sampling equilibria to Bayesian Nash equilibria of the underlying game of incomplete information Γ and to the analogy-based expectation equilibria of the strategic environment $\langle \Gamma, \mathscr{E} \rangle$. A Bayesian Nash equilibrium of Γ is an action distribution $\sigma = (\sigma_{i,\theta_i}, i \in I, \theta_i \in \Theta_i)$ such that

[21]Such a partition is given by $\{P_i(\theta_i), \theta_i \in \Theta_i\}$, where $P_i : \Theta_i \rightrightarrows \Xi_i$ be such that $P_i(\theta_i) = \{(a_{-i}, \theta_i, \theta_{-i}), a_{-i} \in A_{-i}, \theta_{-i} \in \Theta_{-i}\}$.

for every player $i \in I$ with type $\theta_i \in \Theta_i$,

$$\sigma_{i,\theta_i} \in \arg\max_{\sigma_i \in \Sigma_i} \sum_{\theta_{-i} \in \Theta_{-i}} \rho(\theta_{-i} \mid \theta_i) u_i(\sigma_i, \sigma_{-i,\theta_{-i}}, \theta_i, \theta_{-i}),$$

where $\rho(\theta_{-i} \mid \theta_i) = \left(\rho(\theta_i, \theta_{-i})\right) / \left(\sum_{\theta_{-i} \in \Theta_{-i}} \rho(\theta_i, \theta_{-i})\right)$.

Adjusting Jehiel and Koessler's (2008) notation to this setting, an analogy-based expectation equilibrium of the strategic environment $\langle \Gamma, \mathscr{E} \rangle$[22] where the analogy partitions satisfy Condition 2 is an action distribution $\sigma = (\sigma_{i,\theta_i}, i \in I, \theta_i \in \Theta_i)$ such that for every player $i \in I$ with type $\theta_i \in \Theta_i$,

$$\sigma_{i,\theta_i} \in \arg\max_{\sigma_i \in \Sigma_i} \sum_{\theta_{-i} \in \Theta_{-i}} \rho(\theta_{-i} \mid \theta_i) \sum_{a_{-i} \in A_{-i}} \overline{\sigma}_{-i}(a_{-i} \mid \theta_i, \theta_{-i}) u_i(\sigma_i, a_{-i}, \theta_i, \theta_{-i}),$$

where $\overline{\sigma}_{-i}(a_{-i}, \theta)$ denotes the probability with which average gameplay of i's opponents within the element of the partition $e_i(\theta)$, that is,

$$\overline{\sigma}_{-i}(a_{-i} \mid \theta) = \left(\sum_{\theta' \in e_i(\theta)} \rho(\theta') \cdot \sigma_{-i,\theta'_{-i}}(a_{-i}) \right) \Bigg/ \left(\sum_{\theta' \in e_i(\theta)} \rho(\theta') \right).$$

When the analogy partitions satisfy instead Condition 1, analogy-based expectation equilibrium coincides with Bayesian Nash equilibrium.

We then have the following result:

Proposition 11. Let $\{c_n\}_{n \in \mathbb{N}} \subseteq \mathbb{R}^{|I|}_{++}$ such that $c_n \to 0$ and $G_n = \langle \Gamma, \mathscr{E}, \mu, c_n \rangle$ where priors are absolutely continuous. Fix any sequence $\{\sigma_n\}_{n \in \mathbb{N}}$ where for every $n \in \mathbb{N}$ σ_n is an analogy-based sequential sampling equilibrium of G_n.

(i) If analogy partitions satisfy Condition 1, then the limit points of $\{\sigma_n\}_{n \in \mathbb{N}}$ are Bayesian Nash equilibria of the underlying game Γ.

(ii) If analogy partitions satisfy Condition 2 and priors are uniform, then the limit points of $\{\sigma_n\}_{n \in \mathbb{N}}$ are analogy-based expectation equilibria of the strategic environment $\langle \Gamma, \mathscr{E} \rangle$.

[22] The definition of analogy-based expectation equilibria in Jehiel and Koessler (2008) corresponds to static games of incomplete information counterpart of the original formulation by Jehiel (2005) which concerns multi-stage games with observable actions.

When players are able to perfectly distinguish their opponents' actions and types in their samples, in the limit when sampling costs vanish, my solution concept extended to Bayesian games selects a Bayesian Nash equilibrium of the underlying game. This is quite natural, as they are then able to learn the joint distribution of types and opponents' actions with enough observations and thus a no-regret argument akin to the one in Lemma 2 delivers the result for the limiting gameplay.

When players can observe their opponents' actions but not necessarily their types, in the limit, it selects an analogy-based expectation equilibrium of the associated strategic environment when priors are uniform. Hence, analogy-based sequential sampling provides a foundation for the learning interpretation posited by Jehiel and Koessler (2008), where the analogy partitions correspond to a constraint on the observable past data by the players. Finally, as noted by Jehiel and Koessler (2008), when the players analogy partitions pool together all the opponents' types, analogy-based expectation equilibrium that the limiting analogy-based sequential sampling equilibrium selects will also be a fully cursed equilibrium (Eyster and Rabin 2005).

There are two caveats to this result. First, I restrict to analogy partitions that are finer than the players' private information. Second, I require priors to be uniform. When this is not the case, in general, priors will still exert some influence over the posteriors, as the data that players observe is coarse, and thus, in the limit, the beliefs about the opponents' gameplay may not coincide with the beliefs given by analogy-based expectations.[23]

6. Conclusion

This paper proposes an equilibrium framework for strategic settings where players face strategic uncertainty regarding their opponents' gameplay and form their beliefs by accumulating evidence based on a sequential sampling procedure. Equilibrium conditions impose a consistency requirement between the gameplay and the distribution of the evidence that is shown to have a natural interpretation as a steady state of a dynamic process akin to fictitious play. Importantly, and differently from other models of information ac-

[23] For instance, if player i cannot distinguish between samples that regard opponents with types θ_{-i} and θ'_{-i}, then the posterior probability that $(a_{-i}, \theta_i, \theta_{-i})$ relative to $(a_{-i}, \theta_i, \theta'_{-i})$ remains unchanged regardless of the sample path observed, as the observations will not allow player i to distinguish between the two.

quisition in strategic settings, there is no exogenous uncertainty: players seek evidence to inform their beliefs about how others behave.

My solution concept is closely related to other equilibrium models. It provides a foundation for Nash equilibrium, given that, as sampling costs vanish, sequential sampling equilibrium approaches Nash equilibria. Moreover, the model extends to games of incomplete information where it is similarly related to Bayesian Nash equilibrium and analogy-based expectation equilibrium. In drawing such connections, it highlights the importance of strategic uncertainty in rationalizing deviations from these benchmark models.

The sequential sampling equilibria provides a disciplined framework to study the joint distribution of choices, gameplay and decision time in strategic settings. Underlying my solution concept, this paper develops and analyzes a model of individual decision-making with costly information acquisition based on sequential sampling which delivers stochastic choice without relying on indifference or mistakes in a rich environment. Several monotone comparative statics results are established and shown to match well-known patterns in experimental data that elude Nash equilibrium such as the own-payoff effect. Moreover, as we have seen, the model makes novel predictions regarding how choices relate to decision times that are supported by existing data.

While not emphasized in this paper, the model is able to account for play of non-rationalizable actions. Extending results in Gonçalves (2020) — which regard a new class of dominance-solvable games — one can show[24] that sequential sampling equilibrium is not only able to account for play of iteratedly dominated actions, it also supports an experimental finding in Esteban-Casanelles and Gonçalves (2020): as payoffs are scaled up, actions taken correspond to more sophisticated level-k play. Thus, sequential sampling equilibrium can also be seen as bridging the standard equilibrium framework and models of endogenous and costly reasoning as Alaoui and Penta (2016).

One key feature of my solution concept is the exogenous nature of the priors. Imposing priors to be correct eliminates any strategic uncertainty players may face and makes my solution concept coincide with Nash equilibrium. One may wonder whether, if priors are required to be *consistent* with gameplay — that is, correct on average — but non-degenerate,

[24]Relying on Propositions 3 and 7 in that paper and on Proposition 1 in this.

retaining some element of strategic uncertainty, one still obtains a coincidence with Nash equilibrium. The answer is negative: even if priors are restricted to be consistent, equilibrium gameplay and comparative statics would not coincide with those predicted by Nash equilibrium.

Two avenues for future work are logical next steps and, indeed, work in progress. The first is to test the solution concept's predictions. Despite supporting evidence for the model's predictions having been presented, its explanatory power remains unclear. Its sharp predictions in 2×2 games merit testing in an experimental context. The second is to utilize this novel framework to provide new insights on canonical economic problems such as voting, pricing, bargaining, and investment decisions and speculative trade. In these economic settings, actors surely face strategic uncertainty, and yet can access past data from analogous previous interactions. I view this as an exciting avenue for future work.

References

Alaoui, L. and A. Penta (2016). Endogenous Depth of Reasoning. *The Review of Economic Studies 83*(4), 1297–1333.

Alaoui, L. and A. Penta (2018). Cost-Benefit Analysis in Reasoning. *Mimeo*, 1–42.

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica 17*(3/4), 213–244.

Aumann, R. and A. Brandenburger (1995). Epistemic conditions for nash equilibrium. *Econometrica 63*(5), 1161.

Bakkour, A., A. Zylberberg, M. N. Shadlen, and D. Shohamy (2018). Value-based decisions involve sequential sampling from memory.

Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Annals of Mathematical Statistics 37*(1), 51–58.

Bhui, R. (2019). Testing optimal timing in value-linked decision making. *Computational Brain & Behavior 2*(2), 85–94.

Bornstein, A. M. and K. A. Norman (2017). Reinstated episodic context guides sampling-based decisions for reward. *Nature Neuroscience 20*(7), 997–1003.

Brown, G. W. (1951). *Activity Analysis of Production and Allocation*, Chapter XXIV. Iterative Solution of Games by Fictitious Play, pp. 374–376. John Wiley and Sons.

Caplin, A. and M. Dean (2015). Revealed Preference, Rational Inattention, and Costly Information Acquisition. *American Economic Review 105*(7), 2183–2203.

Che, Y.-K. and M. Konrad (2019). Optimal dynamic allocation of attention. *American Economic Review forthcoming*.

Chen, P. N., J. X. He, and H. S. Qin (2001). A proof of the jacobian conjecture on global asymptotic stability. *Acta Mathematica Sinica 17*(1), 119–132.

Chernoff, H. (1961). Sequential tests for the mean of a normal distribution. In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Volume 1, pp. 79–91.

Chierchia, G., R. Nagel, and G. Coricelli (2018). "betting on nature" or "betting on others": anti-coordination induces uniquely high levels of entropy. *Scientific Reports 8*(1).

Chiong, K., M. Shum, R. Webb, and R. Chen (2019). Combining choices and response times in the field: A drift-diffusion model of mobile advertisements. *Mimeo*.

Clithero, J. A. (2018). Response times in economics: Looking through the lens of sequential sampling models. *Journal of Economic Psychology 69*, 61–86.

Costa-Gomes, M. A. and G. Weizsäcker (2008). Stated Beliefs and Play in Normal Form Games. *Review of Economic Studies 75*(3), 729–762.

Denti, T. (2018). Unrestricted Information Acquisition. *Mimeo*, 1–44.

Diaconis, P. and D. Freedman (1990). On the uniform consistency of bayes estimates for multinomial probabilities. *Annals of Statistics 18*(3), 1317–1327.

Diaconis, P. and D. Ylvisaker (1979). Conjugate priors for exponential families. *The Annals of Statistics 7*(2), 269–281.

Esponda, I. and D. Pouzo (2016). Berk-nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica 84*(3), 1093–1130.

Esteban-Casanelles, T. and D. Gonçalves (2020). The effect of incentives on choices and beliefs in games: An experiment. *Mimeo*.

Eyster, E. and M. Rabin (2005). Cursed equilibrium. *Econometrica 73*(5), 1623–1672.

Ferguson, T. S. (2008). *Optimal Stopping and Applications*. UCLA.

Forstmann, B. U., R. Ratcliff, and E.-J. Wagenmakers (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology 67*(1), 641–666.

Freedman, D. (1963). On the asymptotic behavior of bayes' estimates in the discrete case. *Annals of Mathematical Statistics 34*(4), 1386–1403.

Friedman, E. and J. Ward (2019). Stochastic choice and noisy beliefs in games: an experiment. *Mimeo*.

Fudenberg, D. and D. M. Kreps (1993). Learning mixed equilibria. *Games and Economic Behavior 5*(3), 320–367.

Fudenberg, D. and D. K. Levine (1998). *The Theory of Learning in Games*. MIT Press.

Fudenberg, D., P. Strack, and T. Strzalecki (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review 108*(12), 3651–3684.

Goeree, J. K. and C. A. Holt (2001). Ten Little Treasures of Game Theory and Ten Intuitive Contradictions. *American Economic Review 91*(5), 1402–1422.

Goeree, J. K., C. A. Holt, and T. R. Palfrey (2005). Regular Quantal Response Equilibrium. *Experimental Economics 8*(4), 347–367.

Gonçalves, D. (2020). Diagonal games: A tool for experiments and theory. *Mimeo*.

Hawkins, G. E., B. U. Forstmann, E.-J. Wagenmakers, R. Ratcliff, and S. D. Brown (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience 35*(6), 2476–2484.

Hommes, C. H. and M. I. Ochea (2012). Multiple equilibria and limit cycles in evolutionary games with logit dynamics. *Games and Economic Behavior 74*(1), 434–441.

Jehiel, P. (2005). Analogy-based expectation equilibrium. *Journal of Economic Theory 123*(2), 81–104.

Jehiel, P. and F. Koessler (2008). Revisiting games of incomplete information with analogy-based expectations. *Games and Economic Behavior 62*(2), 533–557.

Kalai, E. and E. Lehrer (1993). Rational learning leads to nash equilibrium. *Econometrica 61*(5), 1019.

Kaniovski, Y. M. and H. Young (1995). Learning dynamics in games with stochastic perturbations. *Games and Economic Behavior 11*(2), 330–363.

Kohlberg, E. and J.-F. Mertens (1986). On the strategic stability of equilibria. *Econometrica 54*(5), 1003–1037.

Krajbich, I., D. Lu, C. Camerer, and A. Rangel (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology 3*, 193.

Lehrer, E. and T. Wang (2020). Strong stochastic dominance. *Mimeo*, 1–31.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford University Press.

Matějka, F. and A. McKay (2015). Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model. *American Economic Review 105*(1), 272–298.

McKelvey, R. D., T. R. Palfrey, and R. A. Weber (2000). The effects of payoff magnitude and heterogeneity on behavior in 2×2 games with unique mixed strategy equilibria. *Journal of Economic Behavior and Organization 42*(4), 523–548.

Morris, S. and P. Strack (2019). The wald problem and the equivalence of sequential sampling and static information costs. *Mimeo*.

Moscarini, G. and L. Smith (1963). The optimal level of experimentation. *Econometrica 69*(6), 1629–1644.

Nagel, R. (1995). Unraveling in Guessing Games: An Experimental Study. *American Economic Review 85*(5), 1313–1326.

Nash, J. (1950). *Non-Cooperative Games*. Ph. D. thesis, Princeton University.

Ochs, J. (1995). Games with unique, mixed strategy equilibria: An experimental study. *Games and Economic Behavior 10*(1), 202–217.

Osborne, M. J. and A. Rubinstein (1998). Games with procedurally rational players. *American Economic Review 88*(4), 834–847.

Osborne, M. J. and A. Rubinstein (2003). Sampling equilibrium, with an application to strategic voting. *Games and Economic Behavior 45*(2), 434–441.

Oyama, D., W. H. Sandholm, and O. Tercieux (2015). Sampling best response dynamics and deterministic equilibrium selection. *Theoretical Economics 10*(1), 243–281.

Pearce, D. G. (1984). Rationalizable strategic behavior and the problem of perfection. *Econometrica 52*(4), 1029.

Raiffa, H. and R. Schlaifer (1961). *Applied Statistical Decision Theory*. MIT Press.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review 85*(2), 59–108.

Rubinstein, A. and A. Wolinsky (1994). Rationalizable conjectural equilibrium: Between nash and rationalizability. *Games and Economic Behavior 6*(2), 299–311.

Salant, Y. and J. Cherry (2020). Statistical inference in games. *Econometrica 88*(4), 1725–1752.

Selten, R. (1975). Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory 4*(1), 25–55.

Shadlen, M. N. and D. Shohamy (2016, jun). Decision Making and Sequential Sampling from Memory. *Neuron 90*(5), 927–939.

Shapley, L. S. (1964). Some topics in two-person games. In M. Dresher, L. S. Shapley, and A. W. Tucker (Eds.), *Advances in Game Theory. (AM-52)*, pp. 1–28. Princeton University Press.

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

Stahl, D. O. and P. W. Wilson (1994). Experimental evidence on players' models of other players. *Journal of Economic Behavior & Organization 25*(3), 309–327.

Wald, A. (1947). Foundations of a general theory of sequential decision functions. *Econometrica 15*(4), 279–313.

Weinstein, J. (2020). Best-reply sets. *Economic Theory Bulletin 8*(1), 105–112.

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

Appendices

A. Table of Results

B. Omitted Proofs

Proof of Proposition 1

Let us first argue that any optimal stopping time is finite with probability 1, with respect to the prior μ_i. Suppose, for the purpose of contradiction, that, optimally, player i does not stop in finite time with probability 1, that is, $\mathbb{P}_{\mu_i}(\tau_i < \infty) \leq 1-k$, for $k > 0$. Then, as payoffs are finite, the player's expected payoff is negative infinity and the player is strictly better off by stopping immediately, leading to a contradiction regarding the optimality of τ_i.

Let $V_i(\mu_i; c_i)$ be the value function V_i given sampling cost c_i.

Lemma 4. Let $c_i' > c_i$. Then, $V_i(\mu_i; c_i) = v_i(\mu_i) \implies V_i(\mu_i; c_i') = v_i(\mu_i)$.

Proof. Let τ_i and τ_i' denote the optimal stopping times under sampling costs c_i and c_i', respectively. Then,

$$\begin{aligned} V_i(\mu_i; c_i') = \mathbb{E}_{\mu_i}\left[v_i(\mu_i \mid X_i^{\tau_i'}) - c_i' \cdot \tau_i'\right] &\leq \mathbb{E}_{\mu_i}\left[v_i(\mu_i \mid X_i^{\tau_i'}) - c_i \cdot \tau_i'\right] \\ &\leq \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[v_i(\mu_i \mid X_i^{t_i}) - c_i \cdot t_i\right] \\ &= V_i(\mu_i; c_i). \end{aligned}$$

□

Finally, we obtain a distributional implication on optimal stopping time as sampling costs vary:

Lemma 5. The optimal stopping time decreases in a first-order stochastic dominance sense in sampling costs, with respect to both μ_i and σ_{-i}.

Proof. Let $c_i' > c_i$ and let τ_i and τ_i' denote the optimal stopping times under sampling costs c_i and c_i', respectively. Then, by Lemma 4, as $V_i(\mu_i; c_i) = v_i(\mu_i) \implies V_i(\mu_i; c_i') = v_i(\mu_i)$, we have that, for any $t \in \mathbb{N}_0$,

$$\begin{aligned} \{\omega \in \Omega : \tau_i(\omega) \leq t\} &= \left\{\omega \in \Omega : V_i(\mu_i \mid X_i^{t'}(\omega); c_i) = v_i(\mu_i \mid X_i^{t'}(\omega)),\ t' \leq t\right\} \\ &\subseteq \left\{\omega \in \Omega : V_i(\mu_i \mid X_i^{t'}(\omega); c_i') = v_i(\mu_i \mid X_i^{t'}(\omega)),\ t' \leq t\right\} \\ &= \left\{\omega \in \Omega : \tau_i'(\omega) \leq t\right\}. \end{aligned}$$

The claim follows immediately from the above. □

Proof of Proposition 2

To show that $M_i(a_i)$ is convex, let us prove first a useful property of V_i.

Lemma 6. V_i is convex.

Proof. Let $\sigma_i^d : \mathscr{X}_i \cup \{\emptyset\} \to \Sigma_i$ and let $\tilde{V}_i(t_i, \sigma_i^d, \mu_i) := \mathbb{E}_{\mu_i}\left[\mathbb{E}_{\mu_i}\left[u_i(\sigma_i^d(X_i^{t_i}), \sigma_{-i}) \mid X_i^{t_i}\right] - c_i \cdot t_i\right]$. Note that

$$\begin{aligned}\tilde{V}_i(t_i, \sigma_i^d, \mu_i) &= \mathbb{E}_{\mu_i}\left[\mathbb{E}_{\mu_i}\left[u_i(\sigma_i^d(X_i^{t_i}), \sigma_{-i}) \mid X_i^{t_i}\right] - c_i \cdot t_i\right] \\ &= \mathbb{E}_{\mu_i}\left[u_i(\sigma_i^d(X_i^{t_i}), \sigma_{-i}) - c_i \cdot t_i\right].\end{aligned}$$

Furthermore, note that $V_i(\mu_i) = \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[v_i(\mu_i \mid X_i^{t_i}) - c_i \cdot t_i\right] = \sup_{t_i \in \mathbb{T}_i, \sigma_i^d} \tilde{V}_i(t_i, \sigma_i^d, \mu_i)$. Consequently, $\forall \mu_i, \mu_i' \in \Delta(\Sigma_{-i})$, $\lambda \in [0,1]$,

$$\begin{aligned}V_i(\lambda\mu_i + (1-\lambda)\mu_i') &= \sup_{t_i \in \mathbb{T}_i, \sigma_i^d} \tilde{V}_i(t_i, \sigma_i^d, \lambda\mu_i + (1-\lambda)\mu_i') \\ &= \sup_{t_i \in \mathbb{T}_i, \sigma_i^d} \lambda\tilde{V}_i(t_i, \sigma_i^d, \mu_i) + (1-\lambda)\tilde{V}_i(t_i, \sigma_i^d, \mu_i') \\ &\leq \lambda \sup_{t_i \in \mathbb{T}_i, \sigma_i^d} \tilde{V}_i(t_i, \sigma_i^d, \mu_i) + (1-\lambda) \sup_{t_i \in \mathbb{T}_i, \sigma_i^d} \tilde{V}_i(t_i, \sigma_i^d, \mu_i') \\ &= \lambda V_i(\mu_i) + (1-\lambda)V_i(\mu_i').\end{aligned}$$

□

Lemma 7. For any $a_i \in A_i$, $M_i(a_i)$ is convex.

Proof. That, for $a_i \in A_i$, $M_i(a_i)$ is convex is a direct consequence of Lemma 6. Take $\mu_i, \mu_i' \in M_i(a_i)$, then $V_i(\mu_i) = v_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})]$ and $V_i(\mu_i') = v_i(\mu_i') = \mathbb{E}_{\mu_i'}[u_i(a_i, \sigma_{-i})]$, then $\forall \lambda \in [0,1]$, by Lemma 6, we have that $V_i(\lambda\mu_i + (1-\lambda)\mu_i') \leq \lambda\mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] + (1-\lambda) = \mathbb{E}_{\mu_i'}[u_i(a_i, \sigma_{-i})] = \mathbb{E}_{\lambda\mu_i + (1-\lambda)\mu_i'}[u_i(a_i, \sigma_{-i})]$. Moreover, as, $\forall a_i' \in A_i$, $\mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] \geq \mathbb{E}_{\mu_i}[u_i(a_i', \sigma_{-i})]$ and $\mathbb{E}_{\mu_i'}[u_i(a_i, \sigma_{-i})] \geq \mathbb{E}_{\mu_i'}[u_i(a_i', \sigma_{-i})]$, by linearity, $v_i(\lambda\mu_i + (1-\lambda)\mu_i') = \mathbb{E}_{\lambda\mu_i + (1-\lambda)\mu_i'}[u_i(a_i, \sigma_{-i})] \geq \mathbb{E}_{\mu_i}[u_i(a_i', \sigma_{-i})]$. Hence, $V_i(\lambda\mu_i + (1-\lambda)\mu_i') \leq v_i(\lambda\mu_i + (1-\lambda)\mu_i') = \mathbb{E}_{\lambda\mu_i + (1-\lambda)\mu_i'}[u_i(a_i, \sigma_{-i})]$, which implies that $M_i(a_i)$ is convex. □

Now denote $V_i(\mu_i; u_i)$ the value function V_i given utility function $u_i : A \to \mathbb{R}$.

The first claim in Proposition 2 follows from the next couple of lemmata.

Lemma 8. Let $g : \Sigma_{-i} \to \mathbb{R}_+$ and let $u_i'(\tilde{a}_i, \sigma_{-i}) = u_i(\tilde{a}_i, \sigma_{-i}) + \mathbf{1}_{\tilde{a}_i = a_i} \cdot g(\sigma_{-i})$. Then,

(i) If $V_i(\mu_i; u_i) = \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})]$, then $V_i(\mu_i; u_i') = \mathbb{E}_{\mu_i}[u_i'(a_i, \sigma_{-i})]$;

(ii) If $V_i(\mu_i; u_i') = \mathbb{E}_{\mu_i}[u_i'(\tilde{a}_i, \sigma_{-i})]$ and $\tilde{a}_i \neq a_i$, then $V_i(\mu_i; u_i) = \mathbb{E}_{\mu_i}[u_i(\tilde{a}_i, \sigma_{-i})]$.

Proof.

(i) Suppose not. Then $\exists \tau_i' \in \mathbb{T}_i$ such that

$$\begin{aligned}
&\mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{\tau_i'}\right] - \tau_i' \cdot c_i\right] \\
&> \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] + \mathbb{E}_{\mu_i}[g(\sigma_{-i})].
\end{aligned}$$

Meanwhile, we have

$$\begin{aligned}
V_i(\mu_i; u_i) &= \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] \\
&\geq \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{t_i}\right] - t_i \cdot c_i\right] \\
&\geq \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{\tau_i'}\right] - \tau_i' \cdot c_i\right].
\end{aligned}$$

But a contradiction then arises:

$$\begin{aligned}
&\mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{\tau_i'}\right] - \tau_i' \cdot c_i\right] \\
&> \mathbb{E}_{\mu_i}[u_i(a_k, \sigma_{-i})] + \mathbb{E}_{\mu_i}[g(\sigma_{-i})] \\
&\geq \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{\tau_i'}\right] - \tau_i' \cdot c_i,\right] + \mathbb{E}_{\mu_i}[g(\sigma_{-i})] \\
\implies \quad \mathbb{E}_{\mu_i}[g(\sigma_{-i})] &< \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{\tau_i'}\right] - \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{\tau_i'}\right]\right] \\
&\leq \mathbb{E}_{\mu_i}\left[\mathbb{E}_{\mu_i}\left[g(\sigma_{-i}) \mid X_i^{\tau_i'}\right]\right] \leq \mathbb{E}_{\mu_i}[g(\sigma_{-i})],
\end{aligned}$$

a contradiction.

(ii) Suppose not. Then, $\exists \tau_i \in \mathbb{T}_i$ such that

$$\begin{aligned}
&\mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{\tau_i}\right] - \tau_i \cdot c_i\right] \\
&\leq \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{t_i}\right] - t_i \cdot c_i\right] \\
&= V_i(\mu_i; u_i') = \mathbb{E}_{\mu_i}\left[u_i'(\tilde{a}_i, \sigma_{-i})\right] = \mathbb{E}_{\mu_i}\left[u_i(\tilde{a}_i, \sigma_{-i})\right] \\
&< \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{\tau_i}\right] - \tau_i \cdot c_i\right] \\
\implies 0 &< \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{\tau_i}\right] - \tau_i \cdot c_i\right] \\
&\quad - \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{\tau_i}\right] - \tau_i \cdot c_i\right] \\
&= \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) \mid X_i^{\tau_i}\right] - \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, \sigma_{-i}) + \mathbf{1}_{\sigma_i(a_i)=1} g(\sigma_{-i}) \mid X_i^{\tau_i}\right]\right] \\
&\leq 0,
\end{aligned}$$

a contradiction.

□

Hence, we have that, $\forall a_i \in A_i$, if $u_i \geq_{a_i} u_i'$, then $M_i(a_i; u_i) \subseteq M_i(a_i; u_i')$.

The second claim in part (i) of the proposition is that $M_i(a_i)$ decreases with respect to set inclusion in the utility function u_i with respect to $\geq_{a_i'}$, for $a_i' \neq a_i$. This follows from the fact that $V_i(\mu_i; u_i') \geq V_i(\mu_i; u_i)$ and that, for $a_i' \in A_i \setminus \{a_i\}$,

$$\begin{aligned}
M_i(a_i'; u_i) &= \left\{\mu_i \in \Delta(\Sigma_{-i}) \mid V_i(\mu_i; u_i) = \mathbb{E}_{\mu_i}\left[u_i(a_i', \sigma_{-i})\right]\right\} \\
&\supseteq \left\{\mu_i \in \Delta(\Sigma_{-i}) \mid V_i(\mu_i; u_i') = \mathbb{E}_{\mu_i}\left[u_i(a_i', \sigma_{-i})\right]\right\} \\
&= M_i(a_i'; u_i').
\end{aligned}$$

Finally, the second part of the proposition is an immediate consequence of the results above. To see this, let $u_i' \geq_{a_i} u_i$ and denote τ_i' and τ_i the earliest optimal stopping time associated with u'_i and u_i respectively. Note that, $\forall t \in \mathbb{N}_0$,

$$\begin{aligned}
&\{\tau_i \leq t \cap (\mu_i \mid X_i^{\tau_i}) \in M_i(a_i; u_i)\} \\
&= \left\{ \omega \in \Omega : \exists \ell \leq t, X_i^\ell(\omega) = x_i^\ell \text{ such that } \begin{array}{l} V_i(\mu_i \mid x_i^r; u_i) = \mathbb{E}_{\mu_i}\left[u_i(a_i, s_{-i}) \mid x_i^r\right] \text{ and} \\ V_i(\mu_i \mid x_i^r; u_i) > \max_{a'_i \in A_i} \mathbb{E}_{\mu_i}\left[u_i(a'_i, s_{-i}) \mid x_i^r\right], r \leq \ell - 1 \end{array} \right\} \\
&\subseteq \left\{ \omega \in \Omega : \exists \ell \leq t, X_i^\ell(\omega) = x_i^\ell \text{ such that } \begin{array}{l} V_i(\mu_i \mid x_i^r; u'_i) = \mathbb{E}_{\mu_i}\left[u'_i(a_i, s_{-i}) \mid x_i^r\right] \text{ and} \\ V_i(\mu_i \mid x_i^r; u'_i) > \max_{a'_i \in A_i} \mathbb{E}_{\mu_i}\left[u'_i(a'_i, s_{-i}) \mid x_i^r\right], r \leq \ell - 1 \end{array} \right\} \\
&= \left\{\tau'_i \leq t \cap (\mu_i \mid X_i^{\tau'_i}) \in M_i(a_i; u'_i)\right\},
\end{aligned}$$

where subset inclusion is a consequence of Lemma 8 given that stopping the earliest at x_i^ℓ under u_i and choosing a_i implies that under u'_i, i either also stops at x_i^r and chooses a_i or has stopped earlier and has chosen a_i — in which case the probability of choosing a_i cannot be smaller — or has stopped earlier, say at $x_i^{\ell-h}$, and chosen $a'_i \in A_i \setminus \{a_i\}$. But if the latter holds, then by Lemma 8, i would also have stopped earlier at $x_i^{\ell-h}$ and have chosen a'_i, a contradiction.

It then follows that

$$\mathbb{P}_{\mu_i}\left(\tau'_i \leq t \cap (\mu_i \mid X_i^{\tau'_i}) \in M_i(a_i; u'_i)\right) \geq \mathbb{P}_{\mu_i}\left(\tau_i \leq t \cap (\mu_i \mid X_i^{\tau_i}) \in M_i(a_i; u_i)\right),$$

and $\forall \sigma_{-i} \in \Sigma_{-i}$,

$$\mathbb{P}_{\sigma_{-i}}\left(\tau'_i \leq t \cap (\mu_i \mid X_i^{\tau'_i}) \in M_i(a_i; u'_i)\right) \geq \mathbb{P}_{\sigma_{-i}}\left(\tau_i \leq t \cap (\mu_i \mid X_i^{\tau_i}) \in M_i(a_i; u_i)\right).$$

The argument is analogous for action $a'_i \in A_i \setminus \{a_i\}$.

Proof of Corollary 1

Let $V_i(\mu_i; c_i)$ be the value function V_i given sampling cost c_i and $M_i(a_i; c_i)$ the corresponding set of stopping beliefs at which a_i is optimal. Take $c'_i > c_i$. As, from Lemma 4, $V_i(\mu_i; c_i) = v_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] \implies V_i(\mu_i; c'_i) = v_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})]$, the claim follows immediately.

Proof of Proposition 3

I will start by proving the result for the case where μ_i allows for correlation.

Let $\overline{X}_i^t \in \Sigma_{-i}$ denote the empirical mean of sample path X_i^t, that is, $\frac{1}{t}\sum_{\ell=1}^t X_{i,\ell}$. I will use $B_\epsilon(s_{-i})$ to denote the ϵ-neighborhood around $s_{-i} \in \Sigma_{-i}$ in the sup-norm, that is $B_\epsilon(s_{-i}) := \{s'_{-i} \in \Sigma_{-i} \mid \|s_{-i} - s'_{-i}\|_\infty \leq \epsilon\}$.

Noting that, as Σ_{-i} is compact, a prior having full support is equivalent to Diaconis and Freedman's (1990) condition of ϕ-positivity, we have, from their Theorem (4.2) and Corollary (2.6) that for all full-support $\mu_i \in \Delta(\Sigma_{-i})$, for any $\epsilon < 1/(2 \cdot |A_{-i}|)$, any $t \in \mathbb{N}$ and any $\overline{X}_i^t \in \Sigma_{-i}$,

$$\frac{\mu_i\left(B_\epsilon(\overline{X}_i^t) \mid X_i^t\right)}{1 - \mu\left(B_\epsilon(\overline{X}_i^t) \mid X_i^t\right)} \geq \psi(\epsilon) \cdot \exp\left(t \cdot 2\epsilon^2 \cdot \lambda\right),$$

where $\psi(\epsilon) > 0$, $\forall \epsilon > 0$, and λ is a fixed and strictly positive scalar. For small $\epsilon > 0$, one then has that

$$\begin{aligned}
& \mu_i\left(B_\epsilon(\overline{X}_i^t) \mid X_i^t\right) \geq \frac{\psi(\epsilon) \cdot \exp\left(t \cdot 2\epsilon^2 \cdot \lambda\right)}{1 + \psi(\epsilon) \cdot \exp\left(t \cdot 2\epsilon^2 \cdot \lambda\right)} =: h(\epsilon, t) \\
\Longrightarrow \quad & h(\epsilon,t) \cdot (\overline{X}_i^t + \epsilon \cdot \mathbb{1}) + (1 - h(\epsilon,t)) \cdot \mathbb{1} \;\geq\; \mathbb{E}_{\mu_i | X_i^t}[\sigma_{-i}] \;\geq\; h(\epsilon,t) \cdot (\overline{X}_i^t - \epsilon \cdot \mathbb{1}),
\end{aligned}$$

where $\mathbb{1}$ denotes a vector of 1s.

Then, for any $t \in \mathbb{N}$, any sample paths X_i^t and any $\epsilon > 0$ sufficiently small,

$$\begin{aligned}
\|\mathbb{E}_{\mu_i | X_i^t}[\sigma_{-i}] - \overline{X}_i^t\|_\infty &= \max_{a_{-i} \in A_{-i}} \left|\mathbb{E}_{\mu_i | X_i^t}[\sigma_{-i}(a_{-i})] - \overline{X}_i^t(a_{-i})\right| \\
&\leq \max_{a_{-i} \in A_{-i}} \max\left\{h(\epsilon,t)(\overline{X}_i^t(a_{-i}) + \epsilon) + (1 - h(\epsilon,t)) - \overline{X}_i^t(a_{-i})\,,\; \overline{X}_i^t(a_{-i}) - h(\epsilon,t) \cdot (\overline{X}_i^t - \epsilon)\right\} \\
&\leq \max_{a_{-i} \in A_{-i}} \max\left\{h(\epsilon,t) \cdot \epsilon + (1 - h(\epsilon,t)) \cdot \overline{X}_i^t(a_{-i})\,,\; h(\epsilon,t) \cdot \epsilon + (1 - h(\epsilon,t)) \cdot \overline{X}_i^t(a_{-i})\right\} \\
&\leq (h(\epsilon,t) \cdot \epsilon + (1 - h(\epsilon,t))).
\end{aligned}$$

Moreover, for any sample paths such that X_i^t, X_i^{t+1} such that $X_{i,\ell}^t = X_{i,\ell}^{t+1}$ for $\ell = 1, \ldots, t$, it is immediate that $\|\overline{X}_i^{t+1} - \overline{X}_i^t\|_\infty \le 2/(t+1)$ and, therefore, we have that

$$\begin{aligned}
0 &\le \mathbb{E}_{\mu_i|X_i^t}\left[v_i(\mu_i \mid X_i^{t+1}) - v_i(\mu_i \mid X_i^t)\right] \\
&= \mathbb{E}_{\mu_i|X_i^t}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i|X_i^{t+1}}[u_i(\sigma_i, \sigma_{-i})] - \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i|X_i^t}[u_i(\sigma_i, \sigma_{-i})]\right] \\
&= \mathbb{E}_{\mu_i|X_i^t}\Bigg[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i|X_i^{t+1}}[u_i(\sigma_i, \sigma_{-i})] - \mathbb{E}_{\mu_i|X_i^t}[u_i(b_i(\mu_i \mid X_i^{t+1}), \sigma_{-i})] \\
&\qquad + \mathbb{E}_{\mu_i|X_i^t}[u_i(b_i(\mu_i \mid X_i^{t+1}), \sigma_{-i})] - \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i|X_i^t}[u_i(\sigma_i, \sigma_{-i})]\Bigg] \\
&\le \mathbb{E}_{\mu_i|X_i^t}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i|X_i^{t+1}}[u_i(\sigma_i, \sigma_{-i})] - \mathbb{E}_{\mu_i|X_i^t}[u_i(b_i(\mu_i \mid X_i^{t+1}), \sigma_{-i})]\right] \\
&= \mathbb{E}_{\mu_i|X_i^t}\left[u_i\left(b_i(\mu_i \mid X_i^{t+1}), \mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}]\right) - u_i\left(b_i(\mu_i \mid X_i^{t+1}), \mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}]\right)\right] \\
&= \mathbb{E}_{\mu_i|X_i^t}\left[\sum_{a_{-i} \in A_{-i}} \left(\mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}](a_{-i}) - \mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}](a_{-i})\right) \cdot u_i(b_i(\mu_i \mid X_i^{t+1}), a_{-i})\right] \\
&\le \mathbb{E}_{\mu_i|X_i^t}\left[\sum_{a_{-i} \in A_{-i}} \left|\mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}](a_{-i}) - \mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}](a_{-i})\right| \cdot \max_{a \in A} |u_i(a)|\right] \\
&= \max_{a \in A} |u_i(a)| \cdot \mathbb{E}_{\mu_i|X_i^t}\left[\left\|\mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}] - \mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}]\right\|_1\right] \\
&\le \max_{a \in A} |u_i(a)| \cdot \mathbb{E}_{\mu_i|X_i^t}\left[\left\|\mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}] - \mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}]\right\|_\infty\right] \\
&\le \max_{a \in A} |u_i(a)| \cdot \mathbb{E}_{\mu_i|X_i^t}\left[\left\|\mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}] - \overline{X}_i^{t+1}\right\|_\infty + \left\|\mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}] - \overline{X}_i^t\right\|_\infty + \left\|\overline{X}_i^{t+1} - \overline{X}_i^t\right\|_\infty\right] \\
&\le \max_{a \in A} |u_i(a)| \cdot \left[(h(\epsilon, t) + h(\epsilon, t+1)) \cdot \epsilon + (1 - h(\epsilon, t)) + (1 - h(\epsilon, t+1)) + \frac{2}{t+1}\right] \\
&\le 2 \max_{a \in A} |u_i(a)| \cdot \left[h(\epsilon, t+1) \cdot \epsilon + (1 - h(\epsilon, t)) + \frac{1}{t+1}\right],
\end{aligned}$$

where the first inequality follows from the fact that the value of additional information is always weakly positive, the second inequality from optimality and linearity of u_i is used in the ensuing equality and the last inequality from the fact that $h(\epsilon, t)$ is strictly increasing in t. As $\lim_{t \to \infty} h(\epsilon, t) = 1$, we have that $2 \max_{a \in A} |u_i(a)| \cdot \left[h(\epsilon, t+1) \cdot \epsilon + (1 - h(\epsilon, t)) + \frac{1}{t+1}\right] \to 2 \max_{a \in A} |u_i(a)| \cdot \epsilon$, monotonically, as $t \to \infty$. As this holds for any $\epsilon > 0$ and for any sample path, it holds for $\epsilon \le c_i/(2 \max_{a \in A} |u_i(a)|)$. Thus, $\exists T_i \in \mathbb{N}$ such that $\forall t \ge T_i$, $\mathbb{E}_{\mu_i|X_i^t}\left[v_i(\mu_i \mid X_i^{t+1})\right] - c_i < v_i(\mu_i \mid X_i^t)$, $\forall X_i^t, X_i^{t+1} \in \mathscr{X}_i$ such that $X_{i,\ell}^t = X_{i,\ell}^{t+1}$ for $\ell = 1, \ldots, t$.

I will now show that this implies that $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq T_i)$, $\forall \sigma_{-i} \in \Sigma_{-i}$. First, suppose that there is a sample path x_i^t at which i stops, i.e. $\exists x_i^t \in \mathscr{X}_i(u_i, \mu_i, c_i)$, $t \geq T_i + 1$. Then,

$$v_i(\mu_i \mid x_i^{t-1}) < V_i(\mu_i \mid x_i^{t-1}) - c_i = \mathbb{E}_{\mu_i \mid x_i^{t-1}}[v_i(\mu_i \mid x_i^{t-1}, X_{i,t}] - c_i,$$

a contradiction. Hence, $\forall t \geq T_i + 1$, $\nexists x_i^t \in \mathscr{X}_i(u_i, \mu_i, c_i)$.

Suppose that for some $\sigma_{-i} \in \Sigma_{-i}$, $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq T_i) < 1$. Then, $\exists x_i^t \in \mathscr{X}_i$ such that $t > T_i$ and $\mathbb{P}_{\sigma_{-i}}(X_i^t = x_i^t) > 0$. As $\mu_i(\text{ int }(\Sigma_{-i})) > 0$, then $\mathbb{P}_{\mu_i}(X_i^t = x_i^t) > 0$, which then implies that $0 < \mathbb{P}_{\mu_i}(X_i \notin \mathscr{X}_i(u_i, \mu_i, c_i) \leq \mathbb{P}_{\mu_i}(\neg \tau_i < \infty)$. As, by Proposition 1, $\mathbb{P}_{\mu_i}(\tau_i < \infty) = 1$, we reach a contradiction.

To see that the same argument holds when the prior μ_i does not allow for correlation, note that, for any $a, b, c, d \in [0,1]$, $|ab - cd| = |ab - cb + cb - cd| \leq |a - c|b + |b - d|c \leq |a - c| + |b - d|$ and thus,

$$\begin{aligned}
&\mathbb{E}_{\mu_i \mid X_i^t}\left[\sum_{a_{-i} \in A_{-i}} \left|\mathbb{E}_{\mu_i \mid X_i^{t+1}}[\sigma_{-i}](a_{-i}) - \mathbb{E}_{\mu_i \mid X_i^{t+1}}[\sigma_{-i}](a_{-i})\right| \cdot \max_{a \in A} |u_i(a)|\right] \\
=&\mathbb{E}_{\mu_i \mid X_i^t}\left[\sum_{a_{-i} \in A_{-i}} \left|\prod_{j \in -i} \mathbb{E}_{\mu_{ij} \mid X_i^{t+1}}[\sigma_j](a_j) - \prod_{j \in -i} \mathbb{E}_{\mu_{ij} \mid X_i^t}[\sigma_j](a_j)\right| \cdot \max_{a \in A} |u_i(a)|\right] \\
\leq&\mathbb{E}_{\mu_i \mid X_i^t}\left[\sum_{a_{-i} \in A_{-i}} \sum_{j \in -i} \left|\mathbb{E}_{\mu_{ij} \mid X_i^{t+1}}[\sigma_j](a_j) - \mathbb{E}_{\mu_{ij} \mid X_i^t}[\sigma_j](a_j)\right| \cdot \max_{a \in A} |u_i(a)|\right] \\
\leq&|A_{-i}| \cdot (|I| - 1)\mathbb{E}_{\mu_i \mid X_i^t}\left[\max_{j \in -i} \left\|\mathbb{E}_{\mu_{ij} \mid X_i^{t+1}}[\sigma_j] - \mathbb{E}_{\mu_{ij} \mid X_i^t}[\sigma_j]\right\|_\infty \cdot \max_{a \in A} |u_i(a)|\right] \\
\leq&\max_{a \in A} |u_i(a)| \cdot |A_{-i}| \cdot (|I| - 1) \cdot \\
&\quad \mathbb{E}_{\mu_i \mid X_i^t}\left[\max_{j \in -i}\left(\left\|\mathbb{E}_{\mu_{ij} \mid X_i^t}[\sigma_j] - \overline{X}_i^{j,t}\right\|_\infty + \left\|\mathbb{E}_{\mu_{ij} \mid X_i^{t+1}}[\sigma_j] - \overline{X}_i^{j,t+1}\right\|_\infty + \left\|\overline{X}_i^{j,t+1} - \overline{X}_i^{j,t}\right\|_\infty\right)\right],
\end{aligned}$$

where $\overline{X}_i^{j,t} = \frac{1}{t}\sum_{\ell=1}^t X_{i,\ell}^{j,t}$ denoting the empirical mean of the observations pertaining to player j, with $X_{i,\ell}^{j,t}$ denoting the j-th element — corresponding to player j —, of the ℓ-th observation of the sample path X_i^t. As μ_{ij} has full support on Σ_j, we have, by the same arguments, that μ_{ij} accumulates around the empirical mean of observations pertaining to player j, uniformly across all sample realizations of length t. The remaining steps to complete the proof for when μ_i does not allow for correlation are therefore similar to before.

Proof of Theorem 1

For every player $i \in I$, fix a selection of optimal choices $b_i : \Delta(\Sigma_{-i}) \to \Sigma_i$. By Proposition 3, $\mathscr{X}_i(u_i, \mu_i, c_i)$ is finite and, by definition (2), f_i — the induced expected gameplay of player i as a function of σ_{-i} and given b_i — is continuous in $(\sigma_j)_{j \in -i}$ with respect to the L^p-norm, $p \geq 1$. Let $f : \Sigma \to \Sigma$, where $f(\sigma) = (f_i(\sigma_{-i}))_{i \in I}$. Hence, by Brouwer's fixed-point theorem, there is $\overline{\sigma} \in \Sigma$ such that $\overline{\sigma} = f(\overline{\sigma})$.

Proof of Theorem 2

Let $\bar{\sigma}$ denote the limit of σ_t. Then,

$$0 = \lim_{n\to\infty} \|\sigma_n - \bar{\sigma}\|_\infty = \left\|\lim_{n\to\infty} \sigma_n - \bar{\sigma}\right\|_\infty = \left\|\lim_{n\to\infty} \frac{1}{n+1}\left(\sigma_0 + \sum_{\ell=0}^{n-1} f(\sigma_\ell)\right) - \bar{\sigma}\right\|_\infty$$
$$= \left\|\lim_{n\to\infty} \frac{1}{n+1} \sum_{\ell=0}^{n-1} f(\sigma_\ell) - \bar{\sigma}\right\|_\infty .$$

As for any $i \in I$, f_i is a polynomial, it is continuous and then so it f. As $\sigma_t \to \bar{\sigma}$ and f is continuous, then $f(\sigma_t) \to f(\bar{\sigma})$. Consequently, the Cesàro mean $C_t := \frac{1}{t+1}\sum_{\ell=0}^{t-1} f(\sigma_\ell)$ also converges to $f(\bar{\sigma})$ and therefore $0 = \|f(\bar{\sigma}) - \bar{\sigma}\|_\infty \implies f(\bar{\sigma}) = \bar{\sigma}$.

Suppose now that $\sigma_n = \alpha \cdot \sigma_{n-1} + (1-\alpha) \cdot f(\sigma_{n-1})$, $\alpha \in (0,1)$. Then, as $\sigma_n \to \bar{\sigma} \implies f(\sigma_n) \to f(\bar{\sigma})$ and as for any fixed $\ell \in \mathbb{N}_0$, $\alpha^{n-1-\ell} \cdot f(\sigma_\ell) \to 0$, we have $\sigma_n = \alpha^n \cdot \sigma_0 + (1-\alpha) \cdot \sum_{\ell=0}^{n-1} \alpha^{n-1-\ell} \cdot f(\sigma_\ell) \to f(\bar{\sigma}) \implies \bar{\sigma} = f(\bar{\sigma})$.

As mentioned, the result also holds for finite populations. Let that the maximum of samples any agent in any role takes is bounded above by T. Suppose that, change the sampling process such that players cannot sample more than the available observations — this can be accommodated by having a period-dependent sampling cost that, for every period $n < T$, makes the cost of the n-th sample infinity. Alternatively, start off with T observations, as in fictitious play. For each period n, let the realized actions be denoted by a_n. Let $\sigma_n = \frac{1}{n+1} \cdot a_0 + \frac{1}{n+1} \cdot \sum_{\ell=0}^{n-1} a_\ell$ and, for simplicity, assume that sampling is with replacement, so that $a_n \sim f(\sigma_{n-1})$. Then, $\sigma_n \to \bar{\sigma}$ implies that $f(\sigma_n) \to f(\bar{\sigma})$ and thus a_n converges in distribution. As players will sample at most T observations, sampling with and without replacement is asymptotically equivalent, although the dynamics will differ.

Proof of Proposition 4

To ease notation, I will denote σ_i as the probability that player $i = 1,2$ plays strategy s_1. By manner of a continuous-time approximation as in (Fudenberg and Levine 1998), the dynamic system can be written as $\left(\dot{\sigma}_i = f_i(\sigma_j) - \sigma_i\right)$, $i, j = 1,2$, $i \neq j$. The Jacobian of the dynamic system is given by

$$\begin{pmatrix} -1 & f_1'(\sigma_2) \\ f_2'(\sigma_1) & -1 \end{pmatrix}$$

and its eigenvalues are given by $\lambda = -1 \pm \sqrt{f_1'(\sigma_2) f_2'(\sigma_1)}$. As any game with a unique Nash equilibrium that is fully mixed implies that it must be a asymmetric matching pennies game. Thus, by claim (ii) in Theorem 4, $f_1' f_2' \leq 0$ and, by Proposition 9, there is a unique sequential sampling equilibrium equilibrium. Therefore, as the real parts of the eigenvalues of the Jacobian matrix are strictly negative, the unique sequential sampling equilibrium is locally stable. Moreover, by the Jacobian conjecture on global asymptotic stability — proved to hold on the plane (Chen et al. 2001) —, the unique equilibrium is globally asymptotically stable as the eigenvalues of the Jacobian are always strictly for any $(\sigma_1, \sigma_2) \in [0,1]^2$.

If the game has a unique Nash equilibrium and is not fully mixed, then one of the players must have a weakly dominant strategy. To see why the claim holds then, suppose, without loss, that for player 1 it is weakly dominant to play $\sigma_1 = 1$. Then, for any full-support prior, $f_1(\sigma_{2,n}) = 1$, $\forall n = 1,2,\ldots$, $\implies \sigma_{1,n} \to 1$ and, consequently, $f_2(\sigma_{1,n}) \to f_2(1)$, resulting in $\sigma_{2,n} \to f_2(1)$.

Proof of Lemma 1

I will start by showing an equivalence between the optimal stopping problem players face and the regret minimization problem which extends Fudenberg et al.'s (2018) proposition 2 holds to this environment.

Let $\kappa := \max_{a_i \in A_i} u_i(a_i, s_{-i})$ denote the utility that player i would experience were the player to know that the opponents' gameplay was s_{-i} and let

$$V_i(t_i, \mu_i; u_i, c_i) := \mathbb{E}_{\mu_i}\left[v_i(\mu_i | X_i^{t_i}) - c_i \cdot t_i\right].$$

We then have the following equivalence:

Lemma 9. $\forall t_i \in \mathbb{T}_i$, $-R_i(t_i; u_i, \mu_i, c_i) - \mathbb{E}_{\mu_i}[c_i \cdot t_i] = V_i(t_i, \mu_i; u_i, c_i) - \mathbb{E}_{\mu_i}[\kappa]$.

Proof. First, note that, for any selection of best-responses b_i,

$$\begin{aligned}
\mathbb{E}_{\mu_i}\left[v_i(\mu_i | X_i^{t_i})\right] &= \mathbb{E}_{\mu_i}\left[\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}\left[u_i(\sigma_i, s_{-i}) \mid X_i^{t_i}\right]\right] \\
&= \mathbb{E}_{\mu_i}\left[\mathbb{E}_{\mu_i}\left[u_i\left(b_i\left(\mu_i \mid X_i^{t_i}\right), s_{-i}\right) \mid X_i^{t_i}\right]\right] \\
&= \mathbb{E}_{\mu_i}\left[u_i\left(b_i\left(\mu_i \mid X_i^{t_i}\right), s_{-i}\right)\right].
\end{aligned}$$

Then, it immediately follows that

$$\begin{aligned}
&V_i(t_i, \mu_i; u_i, c_i) - \mathbb{E}_{\mu_i}[\kappa] \\
&= -R_i(t_i; u_i, \mu_i, c_i) - \mathbb{E}_{\mu_i}[c_i \cdot t_i].
\end{aligned}$$

□

The claim in Lemma 1 follows immediately as a corollary of Lemma 9, given that

$$\begin{aligned}
\tau_i(c_i) \in &\arg\max_{t_i \in \mathbb{T}_i} V_i(t_i, \mu_i; u_i, c_i) \\
&= \arg\max_{t_i \in \mathbb{T}_i} V_i(t_i, \mu_i; u_i, c_i) - \mathbb{E}_{\mu_i}[\kappa] \\
&= \arg\min_{t_i \in \mathbb{T}_i} R_i(t_i; u_i, \mu_i, c_i) + \mathbb{E}_{\mu_i}[c_i \cdot t_i].
\end{aligned}$$

Proof of Lemma 2

Let $t_{i,n} \in \mathbb{T}_i$ be a stopping time according to which i stops after sampling $\lfloor 1/\sqrt{c_{i,n}} \rfloor$ observations, regardless of their realization. Let $\overline{X}_i^t$ denote the empirical average, $\frac{1}{t}\sum_{\ell=1}^{t} X_{i,\ell}$.

Then, conditional on σ_{-i} being the true distribution, we have that $t_{i,n} \to \infty$ and $\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_{i,n}}] \overset{a.s.}{\to} \sigma_{-i}$ as $n \to \infty$ given that, $\|\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_{i,n}}] - \overline{X}_i^{t_{i,n}}\|_\infty \to 0$ as shown in Proposition 3 and $\|\overline{X}_i^{t_{i,n}} - \sigma_{-i}\|_\infty \overset{a.s.}{\to} 0$ by the strong law of large numbers. Then, letting $u_i^*(\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^t]) := \max_{\sigma_i \in \Sigma_i} u_i(\sigma_i, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^t]) = \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i, s_{-i}) \mid X_i^t]$, we have that u_i^* is continuous in $\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^t]$ by Berge's theorem of the maximum and, thus, by the continuous mapping theorem, $\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i, s_{-i}) \mid X_i^{t_{i,n}}] \overset{a.s.}{\to} u_i^*(\sigma_{-i})$. Consequently, we find that

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}_{\mu_i}\left[v_i(\mu_i \mid X_i^{t_{i,n}}) - c_{i,n} \cdot t_{i,n}\right] &= \lim_{n\to\infty} \mathbb{E}_{\mu_i}\left[u_i^*(\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_{i,n}}])\right] - \lim_{n\to\infty} c_{i,n} \cdot \left\lfloor \sqrt{1/c_{i,n}} \right\rfloor \\
&= \mathbb{E}_{\mu_i}\left[\mathbb{E}_{\mu_i}\left[\lim_{n\to\infty} u_i^*(\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_{i,n}}]) \mid \sigma_{-i}\right]\right] \\
&= 0,
\end{aligned}$$

where the before last equality follows from the Lebesgue dominated convergence theorem. Finally, we have that

$$0 \le \lim_{n\to\infty} R_i(\tau_{i,n}; u_i, \mu_i, c_{i,n}) = \lim_{n\to\infty} \inf_{t_i \in \mathbb{T}_i} R_i(t_i; u_i, \mu_i, c_{i,n}) \le \lim_{n\to\infty} R_i(t_{i,n}; u_i, \mu_i, c_{i,n}) = 0,$$

proving the claim.

Proof of Theorem 3

Take any limit point σ^* and let $\{\sigma_m\}_m$ be a subsequence of $\{\sigma_n\}_n$ that converges to σ^*. Define

$$r_i(\sigma_i, \sigma_{-i}) := \max_{\sigma_i' \in \Sigma_i} u_i(\sigma_i', \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}).$$

Suppose that σ^* is not a Nash equilibrium. Then, $\exists i \in I$ such that $r_i(\sigma_i^*, \sigma_{-i}^*) \ge k > 0$. As r_i is continuous in (σ_i, σ_{-i}), it is continuous in $\sigma \in \Sigma$ and, therefore, as $\sigma_m \to \sigma^*$, $\exists M \in \mathbb{N}$ such that $\forall m \ge M$, $r_i(\sigma_{i,m}, \sigma_{-i,m}) \ge k/2$. Recall that $\sigma_{i,m} = f_i(\sigma_{-i,m})$, by definition of a sequential sampling equilibrium and let $\bar{r}_i(\sigma_{-i}) := r_i(f_i(\sigma_{-i}), \sigma_{-i})$. As f_i is continuous, so is $\bar{r}_i$ and thus, $\exists \delta > 0$ such that $\forall s_{-i} \in B_\delta(\sigma_{-i}^*)$, $\bar{r}_i(s_{-i}) \ge k/4$. But then, as $\exists M' \in \mathbb{N}$ such that $\forall m \ge M'$, $\|\sigma_{-i,m} - \sigma_{-i}^*\|_\infty < \delta/2$, $\forall s_{-i} \in B_{\delta/2}(\sigma_{-i,m})$, $s_{-i} \in B_\delta(\sigma_{-i}^*)$ and $\bar{r}_i(s_{-i}) \ge k/4 > 0$. This implies that

$\forall m \geq M'$,

$$R_i(\tau_{i,m}; u_i, \mu_i, c_{i,m}) \geq \mu_i(B_{\delta/2}(\sigma_{-i,m})) \cdot k/2 \geq \phi_i(\delta/2) \cdot k/4 > 0,$$

where $\phi_i(\epsilon) := \inf_{\sigma_{-i} \in \Sigma_{-i}} \mu_i(B_\epsilon(\sigma_{-i}))$ and, as μ_i has full support, $\forall \epsilon > 0$, $\phi_i(\epsilon) > 0$. Consequently, $\lim_{m\to\infty} R_i(\tau_{i,m}; u_i, \mu_i, c_{i,m}) \geq k/4 > 0$, what contradicts Lemma 2.

Proof of Proposition 5

Claim (i) follows from the observation that if μ_i has full support and a_i is weakly dominated by $\sigma_i \in \Sigma_i$, $u_i(a_i, s_{-i}) \leq u_i(\sigma_i, s_{-i})$ $\forall s_{-i} \in \Sigma_{-i}$ and for some $s'_{-i} \in \Sigma_{-i}$ such inequality is strict and $u_i(a_i, s_{-i}) - u_i(\sigma_i, s_{-i}) < -\delta$, for some $\delta > 0$. Hence, it will remain strict in a neighborhood ϵ_i around $s'_{-i} \in \Sigma_{-i}$ and, therefore,

$$\mathbb{E}_{\mu_i}[u_i(\sigma_i, s_{-i}) \mid x_i^t] - \mathbb{E}_{\mu_i}[u_i(a_i, s_{-i}) \mid x_i^t] \geq \mu_i(B_\epsilon(s'_{-i}) \mid x_i^t)\delta.$$

As μ_i has full support, $\mu_i(B_\epsilon(s'_{-i})) \geq \inf_{s_{-i} \in \Sigma_{-i}} \mu_i(B_\epsilon(s'_{-i})) =: \phi_i(\epsilon) > 0$. Take $s''_{-i} \in B_\epsilon(s'_{-i})$ such that $s''_{-i}(a_{-i}) \geq \epsilon/(2 \cdot |A_{-i}|)$ $\forall a_{-i} \in A_{-i}$ and note that

$$\begin{aligned} \mu_i(B_\epsilon(s'_{-i}) \mid x_i^t) &\geq \mu_i(B_{\epsilon/(4\cdot|A_{-i}|)}(s''_{-i}) \mid x_i^t) \geq \int_{B_{\epsilon/(4\cdot|A_{-i}|)}(s''_{-i})} \prod_{\ell=1}^{t} s_{-i}(x_{i,\ell})\mu_i(ds_{-i}) \\ &\geq \left(\frac{\epsilon}{4 \cdot |A_{-i}|}\right)^t \mu_i(B_{\epsilon/4}(s''_{-i})) \geq \left(\frac{\epsilon}{4 \cdot |A_{-i}|}\right)^t \cdot \phi_i(\epsilon/(4 \cdot |A_{-i}|)) > 0. \end{aligned}$$

Thus, $\forall x_i^t \in \mathscr{X}_i$,

$$\mathbb{E}_{\mu_i}[u_i(\sigma_i, s_{-i}) \mid x_i^t] - \mathbb{E}_{\mu_i}[u_i(a_i, s_{-i}) \mid x_i^t] > 0,$$

and, consequently, $b_i(\mu_i \mid x_i)(a_i) = 0$ $\forall x_i \in \mathscr{X}_i$ which implies that for any sampling cost c_i and full-support prior μ_i and opponents' gameplay $\sigma_{-i} \in \Sigma_{-i}$, $f_i(\sigma_{-i})(a_i) = 0$. Hence, a_i will never be chosen with positive probability at a sequential sampling equilibrium and thus, no Nash equilibrium involving weakly dominated actions is reachable with full-support priors.

The only if part of claim (ii) follows directly from claim (i). For the if part, we will need this next lemma:

Lemma 10. (Pearce 1984, Lemma 4; Weinstein 2020, Proposition 2)
For any game Γ, $i \in I$ and $B \subseteq A_i$, exactly one of the following is true:

(i) There is a belief $s_{-i} \in \text{int}(\Sigma_{-i})$ such that $B \subseteq \arg\max_{a_i \in A_i} u_i(a_i, s_{-i})$;

(ii) There is a pair $\sigma_i \in \Sigma_i$, $\sigma_i' \in \Delta(B)$ such that σ_i weakly dominates σ_i', where σ_i, σ_i' can be selected such that $\text{supp}(\sigma_i) \cap \text{supp}(\sigma_i') = \emptyset$.

Let a^* be such a pure-strategy Nash equilibrium not involving any weakly dominated actions. For every $i \in I$, and $a_i \in A_i$, let $M_i^*(a_i)$ denote the set of opponents' gameplay for which it is optimal for player i to choose action a_i. Then, by Lemma 10, for any $i \in I$, there is $\tilde{\sigma}_{-i} \in \text{int}(\Sigma_{-i}) \cap M_i^*(a_i^*)$. For every i, let μ_i be a Dirichlet prior on Σ_{-i} with parameters $\tilde{\sigma}_{-i} \gg 0$. Note that, $\forall t \in \mathbb{N}$, $\mathbb{E}[s_{-i} \mid (a_{-i}^*)^t] = \frac{1}{t+1}\tilde{\sigma}_{-i} + \frac{t}{t+1}\delta_{a_{-i}^*} \in M_i^*(a_i^*)$, given that $\{\delta_{a_i^*}, \tilde{\sigma}_{-i}\} \subset M_i^*(a_i^*)$ and $M_i^*(a_i^*)$ is convex, where $\delta_{a_{-i}}$ is a Dirac measure on a_{-i}. Let $b_i(\mu_i \mid x_i)(a_i^*) = 1$ whenever $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i] \in M_i^*(a_i^*)$ for $x_i \in \mathscr{X}_i$. Then, if $\sigma_{-i} = \delta_{a_{-i}^*}$, $\mathbb{E}_{\sigma_{-i}}[b_i(\mu_i \mid X_i^{\tau_{i,n}})](a_i^*) = 1$, for the optimal stopping time $\tau_{i,n}$ given $c_{i,n}$ and, consequently, $\delta_{a^*} \in \Sigma^{SSE}(\langle \Gamma, \mu, c_n \rangle)\ \forall n \in \mathbb{N}$.

Proof of Proposition 6

I will start by showing the only if part of the claim in the proposition.

Let $M_i^*(a_i)$ denote the set of opponents' gameplay to which action a_i is a best-response. Suppose, for the purpose of contradiction, that $\exists i \in I$ such that, $\forall \epsilon > 0$, $B_\epsilon(\delta_{a_{-i}^*}) \backslash M_i^*(a_i^*) \neq 0$, where $\delta_{a_{-i}}$ is a Dirac measure on a_{-i}. For all $m \in \mathbb{N}$, let $\sigma_{-in} \in B_{1/m}(\delta_{a_{-i}^*}) \setminus M_i^*(a_i^*)$. Then, $\tilde{\Sigma}_i^*(\sigma_{-i,m}) := \arg\max_{\sigma_i \in \Delta(A_i \setminus \{a_i^*\})} u_i(\sigma_i, \sigma_{-i,m}) \subseteq \arg\max_{\sigma_i \in \Delta(A_i)} u_i(\sigma_i, \sigma_{-i,m})$ and, by Berge's theorem of the maximum, $\tilde{\Sigma}_i^*(s_{-i})$ is upper-hemicontinuous and compact-valued and let $\sigma_{i,m} \in \tilde{\Sigma}_i^*(\sigma_{-i,m})$. Thus, there is a converging subsequence $(\sigma_k)_k \subseteq (\sigma_{i,m})_m$ and let $\tilde{\sigma}_i := \lim_{k\to\infty} \sigma_{i,k}$. As $u_i(\sigma_{i,k}, \sigma_{-i,k}) \geq u_i(a_i^*, \sigma_{-i,k})$, $\sigma_{i,k} \to \tilde{\sigma}_i$, and $\sigma_{-i,k} \to \delta_{a_{-i}^*} \in M_i^*(a_i^*)$, then $u_i(\tilde{\sigma}_i, a_{-i}^*) = u_i(a_i^*, a_{-i}^*)$. Consequently, $\exists k \in \mathbb{N}$ such that $\text{supp}(\tilde{\sigma}_i) \cap \text{supp}(\tilde{\sigma}_{i,k}) \neq \emptyset$. Let $a_i' \in \text{supp}(\tilde{\sigma}_i) \cap \text{supp}(\tilde{\sigma}_{i,k})$. Then, $u_i(a_i', \sigma_{-i,k}) = u_i(\sigma_{i,k}, \sigma_{-i,k}) \geq u_i(a_i^*, \sigma_{-i,k}) + \delta$, for some $\delta > 0$. As such, $\exists \epsilon' > 0$ such that $\forall s_{-i} \in B_{\epsilon'}(\sigma_{-i,k})$, $u_i(a_i', s_{-i}) \geq u_i(a_i^*, s_{-i}) + \delta/2$ and $B_{\epsilon'}(\sigma_{-i,k}) \cap \text{int}(\Sigma_{-i}) \setminus M_i^*(a_i^*) \neq \emptyset$. Let μ_i be given by a Dirichlet distribution with parameters $s_{-i}' \in B_{\epsilon'}(\sigma_{-i,k}) \cap \text{int}(\Sigma_{-i}) \setminus M_i^*(a_i^*)$. Then, $\forall t \in \mathbb{N}$, $\mathbb{E}_{\mu_i}[u_i(a_i', s_{-i}) \mid (a_{-i}^*)^t] - \mathbb{E}_{\mu_i}[u_i(a_i^*, s_{-i}) \mid (a_{-i}^*)^t] =$

$\frac{1}{t+1}\delta/2$, which then implies that $E_{\mu_i}[s_{-i} \mid (a^*_{-i})^t] \notin M^*_i(a^*_i)$. Hence, $\forall c_i > 0$, if $\sigma_{-i} = \delta_{a^*_{-i}}$, $\mathbb{E}_{\sigma_{-i}}[b_i(\mu_i \mid X_i^{\tau_i})](a^*_i) = 0$.

To show the if part of the claim, we need the following lemma, which proves that, whenever there are two undominated actions that do not give identical payoffs, a player's earliest stopping time grows unboundedly as the sampling costs vanish.

Lemma 11. Suppose that player i has no weakly better action, that is, $\nexists a_i \in A_i$ such that $u_i(a_i, a_{-i}) \geq \max_{a'_i \in A_i} u_i(a'_i, a_{-i})$. Let $\{c_{i,n}\}_{n\in\mathbb{N}} \subset \mathbb{R}_{++}$ be a sequence of sampling costs such that $c_{i,n} \to 0$ and let $\tau_{i,n}$ denote the associated optimal stopping time. Then, if player i's prior has full support, for any distribution of samples, the earliest stopping time diverges σ_{-i}, that is, $\inf \operatorname{supp}(\tau_{i,n}) \to \infty$, $\forall \sigma_{-i} \in \Sigma_{-i}$.

Proof. If there is more than one undominated action, for every action $a_i \in A_i$, $\exists a'_i \in A_i$ and $\tilde{s}_{-i} \in \Sigma_{-i}$ such that $u_i(a'_i, \tilde{s}_{-i}) > u_i(a_i, \tilde{s}_{-i})$ and thus, for every action a_i there are $\epsilon_i, \delta_i > 0$ such that $\forall s''_{-i} \in B_{\epsilon_i}(\tilde{s}_{-i}) \subseteq \Sigma_{-i} \setminus M^*_i(a_i)$,

$$\max_{a''_i \in A_i} u_i(a''_i, s''_{-i}) - u_i(a_i, s''_{-i}) \geq u_i(a'_i, s''_{-i}) - u_i(a_i, s''_{-i}) \geq \delta_i > 0.$$

As the number of actions is finite, there is a $(\epsilon, \delta) \gg 0$ for which the above conditions are satisfied by every $a_i \in A_i$.

Suppose, for the purpose of contradiction, that the is $T \in \mathbb{N}$ such that $\inf \operatorname{supp}(\tau_{i,n}) \leq T$, $\forall n \in \mathbb{N}$, for any $\sigma_{-i} \in \Sigma_{-i}$.

Let us note that for any $\tilde{s}_{-i} \in \Sigma_{-i}$ and $\epsilon > 0$, $\exists s'_{-i} \in B_\epsilon(\tilde{s}_{-i})$ such that $|\tilde{s}'_{-i}(a_{-i}) - \tilde{s}_{-i}(a_{-i})| = \epsilon/(2 \cdot |A_{-i}|)$ $\forall a_{-i} \in A_{-i}$. As such, we have that $B_{\epsilon/4}(\tilde{s}'_{-i}) \subset B_\epsilon(\tilde{s}_{-i})$ and that $\forall s''_{-i} \in B_{\epsilon/4}(\tilde{s}'_{-i})$, $\min_{a_{-i} \in A_{-i}} s''_{-i}(a_{-i}) \geq \epsilon/(4 \cdot |A_{-i}|)$.

This then implies that $\forall x_i^t \in \mathscr{X}_i$,

$$\begin{aligned}
\mu_i \mid x_i^t(B_\epsilon(\tilde{s}_{-i}) &\geq \mu_i \mid x_i^t(B_{\epsilon/4}(\tilde{s}'_{-i})) \\
&= \int_{B_{\epsilon/4}(\tilde{s}'_{-i})} \frac{\prod_{\ell=1}^t s_{-i}(x_{i,\ell})\mu_i(ds_{-i})}{\int_{\Sigma_{-i}} \prod_{\ell=1}^t s_{-i}(x_{i,\ell})\mu_i(ds_{-i})} \\
&\geq \int_{B_{\epsilon/4}(\tilde{s}'_{-i})} \prod_{\ell=1}^t s_{-i}(x_{i,\ell})\mu_i(ds_{-i}) \\
&\geq \left(\frac{\epsilon}{4\cdot|A_{-i}|}\right)^t \phi_i(\epsilon/4) > 0,
\end{aligned}$$

where $\phi(\epsilon) := \inf_{s_{-i}\in\Sigma_{-i}} \mu_i(B_\epsilon(s_{-i}))$, which is strictly positive for any $\epsilon > 0$ as μ_i has full support.

Then, as inf supp $(\tau_{i,n}) \leq T$, for any $n \in \mathbb{N}$, $\exists x_i^t \in \mathscr{X}_i$ with $t \leq T$, at which player i stops. Take any $a_i \in$ supp $b_i(\mu_i \mid x_i^t) \geq 1/|A_i|$. We have that

$$\begin{aligned}
&R_i(\tau_{i,n}; u_i, \mu_i, c_{i,n}) \\
&\geq \mathbb{P}_{\mu_i}\left(X_i^t = x_i^t \;\&\; b_i(\mu_i \mid x_i^t) = a_i \;\&\; s''_{-i} \in B_\epsilon(\tilde{s}_{-i})\right) \cdot \mathbb{E}_{\mu_i}\left[\max_{a''_i \in A_i} u_i(a''_i, s_{-i}) - u_i(a_i, s_{-i}) \mid x_i^t, s_{-i} \in B_\epsilon(\tilde{s}_{-i})\right] \\
&\geq b_i(\mu_i \mid x_i^t)(a_i) \cdot \mu_i \mid x_i^t(B_\epsilon(\tilde{s}'_{-i})) \cdot \delta \\
&\geq \frac{1}{|A_i|} \cdot \left(\frac{\epsilon}{4\cdot|A_{-i}|}\right)^T \phi_i(\epsilon/4) \cdot \delta > 0.
\end{aligned}$$

As this holds, $\forall c_{i,n}$, then we have that $\lim_{n\to\infty} R_i(\tau_{i,n}; u_i, \mu_i, c_{i,n}) \geq k$ for some $k > 0$, which contradicts Lemma 2. □

Note that, for any full-support μ_i, $\exists T < \infty$ such that $\mathbb{E}_{\mu_i}[s_{-i} \mid (a^*_{-i})^t] \in B_\epsilon(a_-i^*)$, $\forall t \geq T$, $\forall i \in I$. Let I' denote the set of players for whom a_i^* is not weakly better than any other action, i.e. $\exists a_{-i} \in A_{-i}$ such that $\max_{a_i \in A_i} u_i(a_i, a_{-i}) > u_i(a_i^*, a_{-i})$. As, for any player i in I', there is $\epsilon > 0$ such that $B_\epsilon(\delta_{a^*_{-i}}) \subseteq M_i^*(a_i^*)$, then, by Lemma 11, $\exists \bar{c} \gg 0$ such that $\inf_{i\in I'} \inf \text{supp } (\tau_i(c)) \geq T$ $\forall c \leq \bar{c}$. Hence, letting $b_i(\mu_i \mid x_i)(a_i^*) = 1$ whenever $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i] \in M_i^*(a_i^*)$ for $x_i \in \mathscr{X}_i$, we have that $\forall c \leq \bar{c}$, $a^* \in \Sigma^{SSE}(\langle \Gamma, \mu, c \rangle)$.

Proof of Theorem 4 and Lemma 3

The proof of the claims in Theorem 4 and Lemma 3 requires us to derive several properties of the value function V_i and the optimal stopping time for the binary setting.

First, let us show a useful equivalence.

Lemma 12. Let $A_i = A_{-i} = \{0,1\}$ and let $u_i : A \to \mathbb{R}$ such that $u_i(1,1) - u_i(0,1) = \delta_1 > 0$, $u_i(0,0) - u_i(1,0) = \delta_0 > 0$. Then, $\tau_i = \tau_i'$, where τ_i denotes the optimal stopping time associated with u_i and τ_i' that associated with utility function $u_i' : A \to \mathbb{R}$ such that $u_i'(1,1) = \delta_1$, $u_i'(1,0) = -\delta_0$ and $u_i'(0,1) = u_i'(0,0) = 0$.

Proof.

$$
\begin{aligned}
&\arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[v_i(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i\right] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[\max\{u_i(1, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]), u_i(0, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}])\} - t_i \cdot c_i\right] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[\max\{u_i(1, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]) - u_i(0, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]), 0\} - t_i \cdot c_i\right] + \mathbb{E}_{\mu_i}\left[u_i(0, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}])\right] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[\max\{\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]\delta_1 + (1 - \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}])\delta_0, 0\} - t_i \cdot c_i\right] + u_i(0, \mathbb{E}_{\mu_i}[s_{-i}]) \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}\left[\max\{u_i'(1, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]), u_i'(0, \mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}])\} - t_i \cdot c_i\right].
\end{aligned}
$$

□

Thus, throughout the remainder of the proof, assume without loss that $u_i(a_i, a_{-i}) = a_i \cdot ((\delta_1 + \delta_0) \cdot a_{-i} - \delta_0)$.

I will now note the following facts:

Lemma 13.

(i) For any $\mu_i \in \mathcal{M}$, $\mu_i \mid x_i \in \mathcal{M}$ for any $x_i \in \{0,1\}$ and $\mu_i \mid x_i = 1 \geq_{MLR} \mu_i \geq_{MLR} \mu_i \mid x_i = 0$.

(ii) For $\mu_i, \mu_i' \in \mathcal{M}$, $\mu_i \geq_{MLR} \mu_i' \implies \mu_i \mid x_i \geq_{MLR} \mu_i' \mid x_i$, for any $x_i \in \{0,1\}$.

Proof. As, when $d\mu_i(\sigma) > 0$, for $x_i \in \{0,1\}$,

$$
\frac{d\mu_i|x_i(\sigma)}{d\mu_i(\sigma)} = \frac{\sigma^{x_i}(1-\sigma)^{1-x_i}}{\mathbb{E}_{\mu_i}[s^{x_i}(1-s)^{1-x_i}]},
$$

claim (i) follows immediately. For claim (ii), note that

$$\frac{d\mu_i|x_i(\sigma)}{d\mu_i'|x_i(\sigma)} = \frac{d\mu_i(\sigma)}{d\mu_i'(\sigma)} \frac{\mathbb{E}_{\mu_i'}[s^{x_i}(1-s)^{1-x_i}]}{\mathbb{E}_{\mu_i}[s^{x_i}(1-s)^{1-x_i}]}$$

is increasing in σ for $x_i \in \{0,1\}$ as, by assumption, $\mu_i \geq_{MLR} \mu_i'$. □

Recall the definition of B_i,

$$B_i(\tilde{V}_i)[\mu_i] := \max\{v_i(\mu), \mathbb{E}_{\mu_i}[\tilde{V}_i(\mu_i \mid X_i)] - c_i\}.$$

In this binary setting, we further have that

$$\mathbb{E}_{\mu_i}[\tilde{V}_i(\mu_i \mid X_i] = \mathbb{E}_{\mu_i}[s]\tilde{V}_i(\mu_i \mid 1) + \mathbb{E}_{\mu_i}[1-s]\tilde{V}_i(\mu_i \mid 0).$$

Let $B_i^{(k+1)}(\tilde{V}_i) := B_i(B_i^{(k)}(\tilde{V}_i))$, for $k \in \mathbb{N}$, with $B_i^{(1)} \equiv B_i$. Note that, for any full-support μ_i, given Proposition 3, it must be the case that $V_i = B_i^{(k)}(v_i)$, $\forall k \geq T_i + 1$. This next lemma characterizes a useful property of V_i:

Lemma 14. V_i is increasing in $\geq_{MLR}$.

Proof. Let $\mu_i, \mu_i' \in \mathscr{M}$ be such that $\mu_i \geq_{MLR} \mu_i'$. Note that $v_i(\mu_i) \geq v_i(\mu_i')$. Suppose that $\forall k \leq n$, $B_i^{(k)}(v_i)[\mu_i] \geq B_i^{(k)}(v_i)[\mu_i']$. Then,

$$\begin{aligned} B_i^{(n+1)}(v_i)[\mu_i] &= \max\{v_i(\mu_i), \mathbb{E}_{\mu_i}[s]B_i^{(n)}(v_i)[\mu_i \mid 1] + \mathbb{E}_{\mu_i}[1-s]B_i^{(n)}(v_i)[\mu_i \mid 0] - c_i\} \\ &\geq \max\{v_i(\mu_i), \mathbb{E}_{\mu_i'}[s]B_i^{(n)}(v_i)[\mu_i \mid 1] + \mathbb{E}_{\mu_i'}[1-s]B_i^{(n)}(v_i)[\mu_i \mid 0] - c_i\} \\ &\geq \max\{v_i(\mu_i), \mathbb{E}_{\mu_i'}[s]B_i^{(n)}(v_i)[\mu_i' \mid 1] + \mathbb{E}_{\mu_i'}[1-s]B_i^{(n)}(v_i)[\mu_i' \mid 0] - c_i\} \\ &= B_i^{(n+1)}(v_i)[\mu_i']. \end{aligned}$$

Hence, by induction, $V_i(\mu_i) \geq V_i(\mu_i')$. □

It will be useful to introduce some notation. Recall that $u_i(a_i, \sigma_{-i}) := \mathbf{1}_{a_i=1}((\delta_1 + \delta_0)\sigma_{-i} - \delta_0)$ and let $u_i'(a_i, \sigma) := \mathbf{1}_{a_i=0}((\delta_1 + \delta_0)(1 - \sigma_{-i}) - \delta_1)$ and $v_i'(\mu_i) = \max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}[u_i'(\sigma_i, s_{-i})]$, where $s_{-i} \sim \mu_i$. Define $W_i : \Delta([0,1]) \to \mathbb{R}$ be such that $W_i(\mu_i) = V_i(\mu_i) - \mathbb{E}_{\mu_i}[u_i(1, s_{-i})]$.

Lemma 15. $W_i(\mu_i) = \max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i'(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i]$. Moreover,

$$\arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i'(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i] = \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i].$$

Proof. By Facts 1 and 2,

$$\arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i] \neq \emptyset.$$

Then,

$$\begin{aligned}
\tau_i &\in \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[\max\{(\delta_1 + \delta_0)\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}] - \delta_0 \,,\, 0\} - t_i \cdot c_i] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[\max\{0 \,,\, (\delta_1 + \delta_0)\mathbb{E}_{\mu_i}[1 - s_{-i} \mid X_i^{t_i}] - \delta_1\} - t_i \cdot c_i] + (\delta_1 + \delta_0)\mathbb{E}_{\mu_i}[\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]] - \delta_0 \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i'(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i] + (\delta_1 + \delta_0)\mathbb{E}_{\mu_i}[\mathbb{E}_{\mu_i}[s_{-i} \mid X_i^{t_i}]] + \mathbb{E}_{\mu_i}[u_i(1, s_{-i})] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i'(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i] + (\delta_1 + \delta_0)\mathbb{E}_{\mu_i}[s_{-i}] \\
&= \arg\max_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i'(\mu_i \mid X_i^{t_i}) - t_i \cdot c_i].
\end{aligned}$$

The claim then follows immediately. □

It should be easy to see that, analogously to what occurs with V_i, W_i is decreasing in $\geq_{MLR}$.

Lemma 16. Let $\mu_i \in \mathcal{M}$ and let $x_i, x_i' \in \mathcal{X}_i$ such that $\sum_\ell \mathbf{1}_{x_{i,\ell}=1} \geq \sum_\ell \mathbf{1}_{x_{i,\ell}'=1}$ and $\sum_\ell \mathbf{1}_{x_{i,\ell}=0} \leq \sum_\ell \mathbf{1}_{x_{i,\ell}'=0}$. Then, $V_i(\mu_i \mid x_i) = 0 \implies V_i(\mu_i \mid x_i') = 0$ and $W_i(\mu_i \mid x_i') = 0 \implies W_i(\mu_i \mid x_i) = 0$.

Proof. By transitivity of $\geq_{MLR}$, $\mu_i \mid x_i \geq_{MLR} \mu_i \mid x_i'$. Thus, $0 = V_i(\mu_i \mid x_i) \geq V_i(\mu_i \mid x_i') \geq 0$. The proof for W_i is analogous. □

Let us now prove Lemma 3:

Suppose that player i stops after realization x_i, with beliefs μ_i prior to observing realization x_i, and that $u_i(1, \mathbb{E}_{\mu_i}[s_{-i} \mid a]) = u_i(0, \mathbb{E}_{\mu_i}[s_{-i} \mid a]) = 0$. Note that this implies that $V_i(\mu_i) > v_i(\mu_i)$ as otherwise player i would not deem it optimal to acquire this last observation.

Then, $V_i(\mu_i \mid a) = u_i(1, \mathbb{E}_{\mu_i}[s_{-i} \mid a]) = u_i(0, \mathbb{E}_{\mu_i}[s_{-i} \mid a]) = 0 = W_i(\mu_i \mid a)$. Then, by Lemma 14,

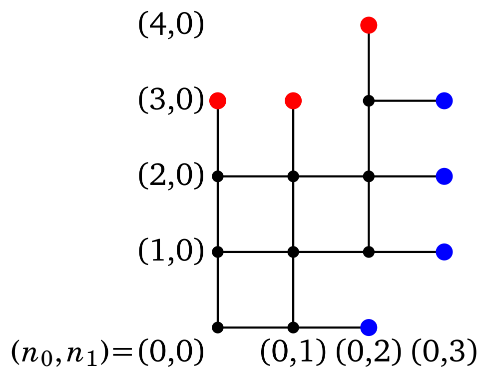


Figure 7. **Grid 1**

Note: Each node represents a possible information set when ignoring the order of the observations. Colored nodes (n_0,n_1) represent nodes at which the process $(Y^t)_t$ stops, that is, information sets such that the sample path x_i contains n_j j-valued observations and $V_i(\mu_i \mid x_i) = v_i(\mu_i \mid x_i)$; at red (blue) stopping nodes action 0 (1) is optimal.

if $a = 1$, $0 = V_i(\mu_i \mid 1) \geq V_i(\mu_i \mid 0) \geq 0$ and thus $V_i(\mu_i) = v_i(\mu_i) = 0 > \mathbb{E}_{\mu_i}[V_i(\mu_i \mid X_i)] - c = -c$, a contradiction.

If instead $a = 0$, again by monotonicity of W_i with respect to $\geq_{MLR}$, $0 = W_i(\mu_i \mid 0) \geq W_i(\mu_i \mid 1) \geq 0$, which implies $V_i(\mu_i \mid 1) = v_i(\mu_i \mid 1) = u_i(1, \mathbb{E}_{\mu_i}[s_{-i} \mid 1])$ and $V_i(\mu_i \mid 0) = u_i(1, \mathbb{E}_{\mu_i}[s_{-i} \mid 0]) = 0$. And thus, $V_i(\mu_i) = u_i(1, \mathbb{E}_{\mu_i}[s_{-i}]) > \mathbb{E}_{\mu_i} V_i(\mu_i \mid X_i) - c$, again a contradiction.

Given Lemma 3, claim in Theorem 4(i) immediately follows from Lemma 16.

In order to show monotonicity of $\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \leq t\right)$ in $\sigma_{-i} \in [0,1]$, I will express the information accumulation as a stochastic process on a "grid", i.e. $\mathbb{N}_0^2$. Each node, a point in $\mathbb{N}_0^2$, denotes the information acquired by the player, that is, the number of 0-valued and 1-valued observations, and, thus, given a prior μ_i, pins down player i's posterior. Stopping times characterize stopping sample paths, up to order, and, therefore, correspond to stopping posterior beliefs.

The structure of the proof is intuitive. First, I characterize properties that these "stopping nodes" have to satisfy. Then, I show that starting from an arbitrary "continuation node", the stopping nodes in the "continuation grid" will share these same properties. Finally, the proof follows from an induction argument. The algebra, however, is tedious.

To illustrate it, consider Figure 7. If player i has already sampled n_0 0-valued observations and n_1 1-valued observations, the player either stops if $V_i(\mu_i \mid x_i) = v_i(\mu_i \mid x_i$, where x_i is a sample path with n_j j-valued observations, $j = 0,1$, or takes another sample. In the lat-

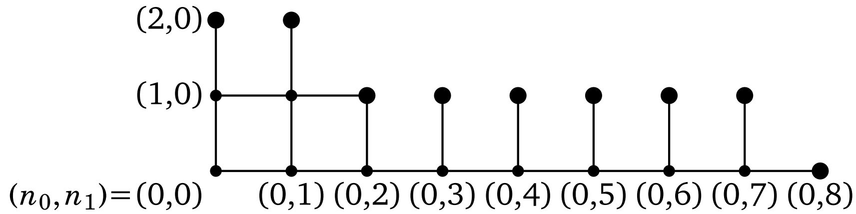


Figure 8. **Grid 2**

Note: In this grid, the probability that a process higher σ_{-i} may increase of decrease the probability of stopping conditional after moving to the right. The terminal nodes in this grid violate Condition 3(i) which is met when the terminal nodes are induced by optimal stopping.

ter case, player i will move to node (n_0, n_1+1) with probability σ_{-i} and to node (n_0+1, n_1) with complementary probability. As from Proposition 3 that the optimal stopping time is bounded, we know that this stochastic process will eventually get to a node where player i stops. Moreover, whether the last observation is 0-valued or 1-valued is enough to determine whether player i takes action 0 or 1 upon stopping.[25] Hence, the probability of taking action 1 corresponds to the probability of the stochastic process stopping after "moving to the right", an argument which I formal below and monotonicity just implies that higher σ_{-i} leads to a greater probability of the stochastic process stopping at this boundary on the right-hand-side of the grid. While it seems intuitive that increasing the probability that the process "moves right" at each step should increase the probility that it stops after moving to the right, this is not generally the case and depends on the structure of the stochastic process induced by player i's optimal stopping which must comply with. Figure 8 presents a counterexample, where the probability of stopping after moving to the right is given by $\sigma_{-i}^8 + 2\sigma_{-i}(1-\sigma_{-i})$, which is strictly decreasing in $\sigma_{-i} \in (.521, .737)$.

To proceed, let us define the process $\left(Y^t\right)_{t\in\mathbb{N}_0}$ such that $Y^t = (\sum_{\ell=1}^t X_{i,\ell}^{\tau_i \wedge t}, \tau_i \wedge t - \sum_{\ell=1}^t X_{i,\ell}^{\tau_i\wedge t})$. Then, $\exists B_0, B_1 : \mathbb{N}_0 \to \mathbb{N}$ such that $V_i(\mu_i \mid x_i) = v_i(\mu_i \mid x_i) = \mathbb{E}_{\mu_i}[u(j, s_{-i})]$ if $n_j \geq B_j(n_{1-j})$, where n_j is the number of j-valued observations in the sample path x_i, that is, $n_j = \sum_\ell \mathbf{1}_{x_{i,\ell}=j}$, $j = 0, 1$. To see this, note that as the support of the optimal stopping time τ_i is bounded above for any distribution σ_{-i} and as, for any n_{1-j} and any full-support prior, there is a finite number of j-valued observations such that makes the posterior arbitrarily concentrated around the degenerate distribution $s_{-i} = j$, then there is a $x_i \in \mathcal{X}_i$ such that $\sum_\ell \mathbf{1}_{x_{i,\ell}=1-j} =$

[25]To see this, suppose that player i with beliefs μ_i stops after observing a 1 and takes action 0. Then, by MLR monotonicity of V_i, it must be that player i would also stop when observing a 0-valued sample and would also take action 0. But then this last sample has no value for player i as regardless of the realization, player i will take the same action and, thus, it is strictly better, for any sampling cost $c_i > 0$, to stop at μ_i.

n_{1-j} and $V_i(\mu_i \mid x_i) = \mathbb{E}_{\mu_i}[u_i(j, s_{-i}) \mid x_i]$. Moreover, by MLR monotonicity of W_i, $\forall x_i'$ such that $\sum_\ell \mathbf{1}_{x'_{i,\ell}=0} = n_0$ and $\sum_\ell \mathbf{1}_{x'_{i,\ell}=1} \geq \sum_\ell \mathbf{1}_{x_{i,\ell}=1}$, we have that $V_i(\mu_i \mid x_i) - \mathbb{E}_{\mu_i}[u_i(1, s_{-i}) \mid x_i] = W_i(\mu_i \mid x_i) = 0 \geq W_i(\mu_i \mid x_i') \geq 0$ and $W_i(\mu_i \mid x_i') = 0 = V_i(\mu_i \mid x_i') - \mathbb{E}_{\mu_i}[u_i(1, s_{-i}) \mid x_i']$, and so player i stops at x_i' as well and takes action 1. Thus, there is a smallest n_1 such that $\sum_\ell \mathbf{1}_{x_{i,\ell}=1} = n_1$ and $\sum_\ell \mathbf{1}_{x_{i,\ell}=0} = n_0$ and which player i stops and takes action 1 upon observing n_0 0-valued samples. Then, $B_1(n_0) = n_1$. The argument is symmetric for B_0. Hence, B_0, B_1 are well-defined and $B_1(n)$ ($B_0(n)$) corresponds to the blue (red) nodes in Figure 7.

I will now introduce a set of conditions on B_0, B_1 which are satisfied by stopping points of the process $(Y^t)_t$.

Condition 3. $B_0, B_1 : \mathbb{N}_0 \to \mathbb{N}$ satisfy condition 3 if (i) B_0, B_1 are non-decreasing; (ii) $\exists N_0, N_1 \in \mathbb{N} : B_j(N_{1-j} - 1) = N_j$, $j \in \{0,1\}$ and $\forall (n_0, n_1) \ll (N_0, N_1)$, if $n_j \geq B_j(n_{1-j})$, then $n_{1-j} < B_{1-j}(n_j)$, $j \in \{0,1\}$

Condition 3(i) is satisfied due to MLR monotonicity of the value function V_i. To see this, suppose that $B_1(n) > B_1(n+1)$ for some $n \in \mathbb{N}_0$. Then, let $x_i, x_i' \in \mathscr{X}_i$ such that $\sum_\ell \mathbf{1}_{x_{i,\ell}=0} = 1 + \sum_\ell \mathbf{1}_{x'_{i,\ell}=0} = n+1$ and $\sum_\ell \mathbf{1}_{x_{i,\ell}=1} = \sum_\ell \mathbf{1}_{x'_{i,\ell}=1} = B_1(n+1) < B_1(n)$. Then, as player i stops at $\mu_i \mid x_i$, by MLR monotonicity, the player also stops at $\mu_i \mid x_i'$. But then, $B_1(n)$ cannot correspond to the smallest number of 1-valued observations at which player i stops after having also drawn n 0-valued observations, which contradicts the definition of B_1. The argument for B_0 is symmetric. Condition 3(ii) follows from the fact that optimal stopping time is bounded from above and thus there is a latest stopping time in the support of τ_i. As I have argued above, the last observation drawn has to determine the action of player i and thus existence of such N_0, N_1 is assured. Moreover, as the player never stops at an posterior at which the player is indifference between the two actions (unless i does not sample at all), this fact and condition (i) imply that condition (ii) is also met.

Prior to stopping, the process defined above is a Markov chain on $\mathbb{N}_0^2$, with the $Y^{t+1} = Y^t + (1 - z_t, z_t)$ where z_t is a random variable drawn from a Bernoulli distribution with the same parameter for all t. As Lemma 17 shows that whenever B_0, B_1 determine the stopping points of such stochastic process satisfy Condition 3, the probability of stopping at $(n, B_1(n))$ increases in the parameter of the Bernoulli distribution.

Lemma 17. Suppose B_0, B_1 satisfy condition 3 and let N_0, N_1 be such as in condition 3(ii). Let z_t be a random variable distributed according to a Bernoulli distribution with parameter $\theta \in [0,1]$ and $\left(Y^t\right)_{t\in\mathbb{N}}$ be a stochastic process such that, given $Y^0 \in \mathbb{N}_0^2$,

$$Y^{t+1} = \begin{cases} Y^t = (Y_0^t, Y_1^t) & \text{if } Y_0^t \geq B_0(Y_1^t) \text{ or } Y_1^t \geq B_1(Y_0^t) \\ Y^t + (z_t, 1 - z_t) & \text{if otherwise.} \end{cases}$$

Then, for any $Y^0 \ll (B_0(Y_1^0), B_1(Y_0^0))$,

(i) $\mathbb{P}(Y_1^t \geq B_1(Y_0^t)) + \mathbb{P}(Y_0^t \geq B_0(Y_1^t)) = 1$;

(ii) $\mathbb{P}(Y_0^t \geq B_0(Y_1^t))$ is non-decreasing in $\theta \in [0,1]$;

(iii) $\frac{d}{d\theta}\mathbb{P}(Y_0^t \geq B_0(Y_1^t)) > 0$, $\forall \theta \in (0,1)$.

Proof. For claim (i), note that by condition 3(iii), if $Y_j^t \geq B_j(Y_{1-j}^t)$, then $Y_{1-j}^t \geq B_{1-j}(Y_j^t)$, which implies that $\mathbb{P}(Y_1^t \geq B_1(Y_0^t)) + \mathbb{P}(Y_0^t \geq B_0(Y_1^t)) \leq 1$. Suppose that the inequality is strict. Then, $\mathbb{P}(Y^t \ll (B_0(Y_1^t), B_1(Y_0^t)) \,\forall t \in \mathbb{N}_0) > 0$. Let $T = N_0 + N_1 - 1 - Y_0^0 - Y_1^0$. Then, it must be that $Y_0^T + Y_1^T = T + Y_0^0 + Y_1^0 = N_0 + N_1 - 1$ which implies that (a) $Y_0^T \geq N_0$ and $Y_1^T \leq N_1 - 1$ or (b) $Y_1^T \geq N_1$ and $Y_0^T \leq N_0 - 1$. If (a) holds, then $Y_0^T \geq N_0 = B_0(N_1 - 1) \geq B_0(Y_1^T)$ by monotonicity of B_j and, therefore, $\forall t \geq T$, $Y^t = Y^T$, a contradiction. If (b) holds, the argument is symmetric.

For claim (ii), for $(n_0, n_1) \in \mathbb{N}_0^2$ such that $(n_0, n_1) \ll (N_0, N_1)$, let $p(n_0, n_1) = \mathbb{P}(Y_0^t \geq B_0(Y_1^t) \mid Y^0 = (n_0, n_1))$. First, note that $p(N_0 - 1, N_1 - 1) = \theta$. Then, $\forall n_1 = 0, \ldots, N_1 - 1$, and $n_0 = 0, \ldots, N_0 - 1$,

$$p(N_0 - n_0, N_1 - 1) = \begin{cases} 0 & \text{if } N_1 - 1 \geq B_1(N_0 - n_0) \\ \theta^{n_0} & \text{if otherwise} \end{cases}$$

$$p(N_0 - 1, N_1 - n_1) = \begin{cases} 1 & \text{if } N_0 - 1 \geq B_0(N_1 - n_1) \\ 1 - (1-\theta)^{n_1} & \text{if otherwise.} \end{cases}$$

Now, I will show that for every $(n_0, n_1) \in \mathbb{N}_0^2$ such that $(n_0, n_1) \leq (N_0 - 1, N_1 - 1)$, $p(n_0 + 1, n_1) \geq p(n_0, n_1) \geq p(n_0, n_1 + 1)$. First, note that $p(N_0 - 2, N_1 - 2) = 0 \implies p(N_0 - 2, N_1 - 1) = 0$

and $p(N_0-2,N_1-2)=1 \implies p(N_0-1,N_1-2)=1$. Thus, we have that $\{1,\theta+(1-\theta)\theta\} \ni p(N_0-1,N_1-2) \geq p(N_0-2,N_1-2) \geq p(N_0-2,N_1-1) \in \{0,\theta^2\}$. Suppose that $\forall n_0,n_1$ such that $N_0-2 \geq n_0 \geq M_0$ and $N_1-2 \geq n_1 \geq M_1$, $p(n_0,N_1-1) \leq p(n_0+1,N_1-1)$ and $p(N_0-1,n_1) \geq p(N_0-1,n_1+1)$. Then,

$$p(M_0-1,N_1-1)=\begin{cases} 0 & \text{if } N_1-1 \geq B_1(M_0-1) \\ \theta p(M_0,N_1-1) & \text{if otherwise} \end{cases}$$

$$p(N_0-1,M_1-1)=\begin{cases} 1 & \text{if } N_0-1 \geq B_0(M_1-1) \\ \theta+(1-\theta)p(N_0-1,M_1) & \text{if otherwise.} \end{cases}$$

and, by induction, $\forall(n_0,n_1) \leq (N_0-1,N_1-1)$, $p(n_0,N_1-1) \leq p(n_0+1,N_1-1)$ and $p(N_0-1,n_1) \geq p(N_0-1,n_1+1)$.

Now suppose that (i) $\forall n_1 : N_1-1 \geq n_1 \geq M_1 \geq 1$, $\forall n_0 \leq N_0-1$, $p(n_0,n_1) \leq p(n_0+1,n_1)$ and (ii) $\forall n_0 : N_0-1 \geq n_0 \geq M_0 \geq 1$, $\forall n_1 \leq N_1-1$, $p(n_0,n_1) \geq p(n_0,n_1+1)$ and observe that it holds for $M_j = N_j-2$, $j=0,1$. I will show that (a) $\forall n_0 \leq N_0-1$, $\forall n_1 \leq \underline{M_1-1}$, $p(n_0,n_1) \leq p(n_0+1,n_1)$ and (b) $\forall n_0 : N_0-1 \geq n_0 \geq \underline{M_0-1} \geq 1$, $\forall n_1 \leq N_1-1$, $p(n_0,n_1) \geq p(n_0,n_1+1)$.

For all $N_0-1 \geq n_0 \geq M_0-1$ and $N_1-1 \geq n_1 \geq M_1$,

$$\begin{aligned} p(n_0,n_1) &= \theta p(n_0+1,n_1)+(1-\theta)p(n_0,n_1+1) \\ &\leq \theta p(n_0+2,n_1)+(1-\theta)p(n_0+1,n_1+1) \\ &= p(n_0+1,n_1) \end{aligned}$$

and for all $N_0-1 \geq n_0 \geq M_0$ and $N_1-1 \geq n_1 \geq M_1-1$, $j=0,1$,

$$\begin{aligned} p(n_0,n_1) &= \theta p(n_0+1,n_1)+(1-\theta)p(n_0,n_1+1) \\ &\geq \theta p(n_0+1,n_1+1)+(1-\theta)p(n_0,n_1+2) \\ &\geq p(n_0,n_1+1). \end{aligned}$$

In particular, we have that $p(M_0-1,M_1) \leq p(M_0,M_1) \leq p(M_0,M_1-1)$. Thus,

$$p(M_0-1,M_1) \leq p(M_0-1,M_1-1) = \theta p(M_0,M_1-1)+(1-\theta)p(M_0-1,M_1) \leq p(M_0-1,M_1),$$

and therefore,

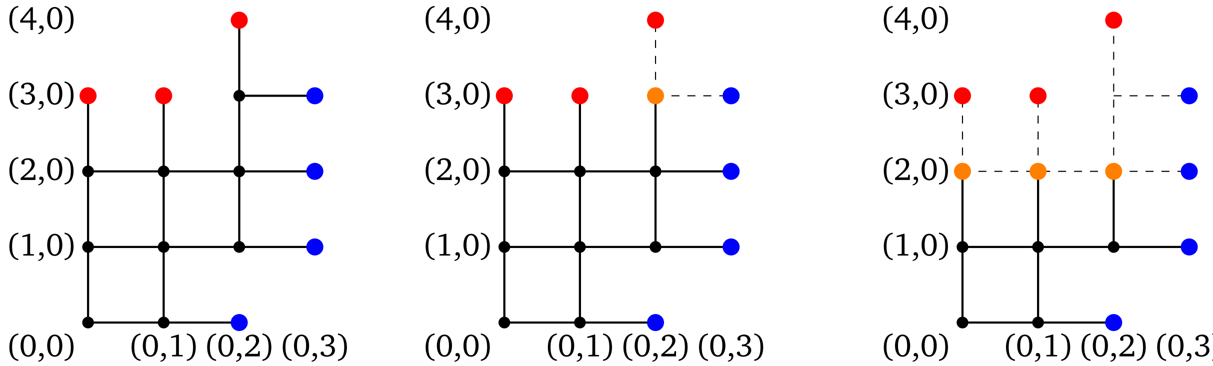
$$p(M_0-2,M_1-1)=\theta p(M_0-1,M_1-1)+(1-\theta)p(M_0-2,M_1)\le p(M_0-1,M_1-1)$$
$$p(M_0-1,M_1-2)=\theta p(M_0,M_1-2)+(1-\theta)p(M_0-1,M_1-1)\ge p(M_0-1,M_1-1)$$

and, by the same argument, $p(M_0-2,M_1-1)\le p(M_0-2,M_1-2)\ge p(M_0-1,M_1-2)$. Then, by induction, (a) and (b) hold and we further have that for all $N_0-1\ge n_0$ and $N_1-1\ge n_1$, $p(n_0,n_1+1)\le p(n_0,n_1)\le p(n_0+1,n_1)$.
It remains to show that $p(n_0,n_1)$ is increasing in θ and that $\frac{d}{d\theta}p(0,0)>0$ for $\theta\in(0,1)$.
It is straightforward to check that $\frac{d}{d\theta}p(N_0-1,N_1-1)>0$ and $\forall(n_0,n_1)\le(N_0-1,N_1-1)$, $\frac{d}{d\theta}p(N_0-1,n_1)\ge 0$ and $\frac{d}{d\theta}p(n_0,N_1-1)\ge 0$. Suppose that $\forall(m_0,m_1):(M_0,M_1)\le(m_0,m_1)\le(N_0-1,N_1-1)$ and $\forall(n_0,n_1)\le(N_0-1,N_1-1)$, $\frac{d}{d\theta}p(m_0,n_1)\ge 0$ and $\frac{d}{d\theta}p(n_0,m_1)\ge 0$. I will show that $\frac{d}{d\theta}p(m_0-1,n_1)\ge 0$ and $\frac{d}{d\theta}p(n_0,m_1-1)\ge 0$.

$$\begin{aligned}\frac{d}{d\theta}p(m_0-1,n_1)&=\frac{d}{d\theta}(\theta p(m_0,n_1)+(1-\theta)p(m_0-1,n_1+1))\\&=p(m_0,n_1)-p(m_0-1,n_1+1)+\theta\frac{d}{d\theta}p(m_0,n_1)+(1-\theta)\frac{d}{d\theta}p(m_0-1,n_1+1)\\&=p(m_0,n_1)-p(m_0-1,n_1+1)+\theta\frac{d}{d\theta}p(m_0,n_1)+\\&\quad+(1-\theta)(p(m_0,n_1+1)-p(m_0-1,n_1+2))+(1-\theta)\theta\frac{d}{d\theta}p(m_0,n_1+1)+\\&\quad+(1-\theta)^2\frac{d}{d\theta}p(m_0-1,n_1+2).\end{aligned}$$

As for any $n_1\le N_1-1$, (1) $\frac{d}{d\theta}p(m_0,n_1)\ge 0$, and (2) $p(m_0,n_1)-p(m_0-1,n_1+1)\ge 0$, and as by iterating the last term of the expression $h=m_1-n_1$ times we get $(1-\theta)^h\frac{d}{d\theta}p(m_0-1,n_1+h)\ge 0$, we have that $\frac{d}{d\theta}p(m_0-1,n_1)\ge 0$.
To see that $\frac{d}{d\theta}p(0,0)>0$, note that if this is not the case, then $\frac{d}{d\theta}p(0,0)=0$ which implies $\frac{d}{d\theta}p(1,0)=\frac{d}{d\theta}p(0,1)=p(1,0)-p(0,1)=0$ and, by induction, $0=p(B_0(0)-1,0)-p(0,0)=p(0,B_1(0)-1)-p(0,0)$. Then, if $\theta\in(0,1]$, $0=p(B_0(0)-1,0)-p(0,0)=\theta+(1-\theta)p(B_0(0)-1,1)-p(0,0)\implies p(0,0)=1$ and if $\theta\in[0,1)$, $0=p(0,B_1(0)-1)-p(0,0)=\theta p(1,B_1(0)-1)-p(0,0)\implies p(0,0)=0$, and we reach a contradiction. □

Finally, let us note that the probability of stopping before time t and taking action 1, $\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i\mid X_i^{\tau_i})=1\cap\tau_i\le t\right)$ corresponds to the probability that $Y^t=(n,B_1(n))$ for some



(a) (b) (c)

Figure 9.

$n \in \mathbb{N}_0$. In order to show that $\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \le t\right)$ is non-decreasing in σ_{-i} for all t, I will make use of Lemma 17 by redefining the bounds B_0, B_1.

To provide some intuition, let us consider Figure 9. In panel 9a, I have represented the terminal nodes — blue and red — corresponding to the stopping time τ_i, where upon stopping on a red (blue) node, the player takes action 0 (1). In panel 9b, in order to examine how the probability that event $\{b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \le 5\}$ realizes depends on σ_{-i}, we change the number of 0-valued observations that are required to stop and take action 0 having observed two 1-valued observations from $B_0(2) = 4$ to $\tilde{B}_0(2) = 3$, with the red terminal node at (4,2) being replaced by the orange terminal node at (3,2). This way, we keep the process from ever reaching the blue node at (3,3) while keeping unchanged the probability that it reaches any other blue node. In panel 9c, I illustrate the modifications to the terminal nodes that allow us to study how the probability that the event $\{b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \le 4\}$ realizes varies with σ_{-i}, by replacing the red terminal nodes with the orange terminal nodes, that is, by requiring fewer 0-valued observations in order to stop. Note that, under both transformations, B_1 is unchanged and the new implied $\tilde{B}_0$, determining the number of 0-valued observations required to stop, also complies with Condition 3 and so, Lemma 17 applies. The last part of the proof, below, formalizes this argument.

Let T denote the latest stopping time according to τ_i for any $\sigma_{-i} \in (0,1)$ and N_j^T denote the corresponding number of j-valued samples that player i observed upon stopping at T and taking action j, $j = 0,1$. Then, $B_j(N_{1-j}^T) = N_j$. For any $t < B_1(0)$, $\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \le t\right) =$

0, $\forall\sigma_{-i} \in [0,1]$. Let $B_j^T = B_j$, $j = 0,1$ and, for any $t = B_1(0), B_1(0)+1, ..., T-1$, let (a)

$$N(B_0^t, B_1^t) := \max\left\{ (n_0, n_1) \in \mathbb{N}_0^2 \;\middle|\; \begin{array}{l} B_j^t(n_{1-j}) = B_j^t(n_{1-j}-1)+1,\ j = 0,1; \\ \mathbb{P}_{\sigma_{-i}}(Y^t = (n_0-1, n_1-1)) > 0,\ \sigma_{-i} \in (0,1) \end{array} \right\};$$

(b) $B_0^t(n) = \min\{B_0^{t+1}(n), B_0^{t+1}(N_1(B_0^{t+1}, B_1^{t+1})-1)-1\}$ $B_1^t(n) = B_1(n)$. (a) defines the number of j-valued samples at the latest stopping node according to bounds B^t; (b) redefines the stopping nodes on the grid. To see that for $z_n \sim Bernoulli(\sigma_{-i})$ and, given $Y^0 = (0,0)$,

$$Y^{n+1} = \begin{cases} Y^n = (Y_0^n, Y_1^n) & \text{if } Y_0^n \geq B_0^t(Y_1^n) \text{ or } Y_1^n \geq B_1^t(Y_0^n) \\ Y^n + (1-z_n, z_n) & \text{if otherwise} \end{cases}$$

we have $\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i \leq t\right) = \mathbb{P}\left(Y^n = (n_0, B_1^t(n_0)),\ n, n_0 \in \mathbb{N}_0\right)$, note that when $t = T-1$, we have that $B_0^{T-1}(N_1(B_0^T, B_1^T)-1) = B_0^{T-1}(B_1^{T-1}(N_0-1)-1) = B_0^T(N_1-1)-1 = N_0-1$. Consequently,

$$\begin{aligned}
&\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) = 1 \cap \tau_i > T-1\right) \\
&= \mathbb{P}\left(Y^n = (N_0-1, B_1^{T-1}(N_0-1)),\ n \in \mathbb{N}_0\right) \\
&= \mathbb{P}\left(Y^{n-1} = (N_0-2, B_1^{T-1}(N_0-1))\right) \cdot \mathbb{P}\left(Y^n = (N_0-1, N_1-1) \mid Y^{n-1} = (N_0-1, B_1^{T-1}(N_0-1))\right) + \\
&\quad + \mathbb{P}\left(Y^{n-1} = (N_0-1, N_1-1)\right) \cdot \mathbb{P}\left(Y^n = (N_0-1, N_1) \mid Y^{n-1} = (N_0-1, N_1-1)\right) \\
&= \mathbb{P}\left(Y^{n-1} = (N_0-2, B_1^{T-1}(N_0-1))\right) \cdot 0 + \mathbb{P}\left(Y^{n-1} = (N_0-1, N_1-1)\right) \cdot 0 \\
&= 0,
\end{aligned}$$

given that at $Y^{n-1} = (N_0-2, B_1^{T-1}(N_0-1))$, $Y_1^{n-1} \geq B_1^{T-1}(N_0-1) = B_1(N_0-1) \geq B_1(N_0-2)$ and thus Y^{n-1} stops yielding $\mathbb{P}\left(Y^n = (N_0-1, N_1-1) \mid Y^{n-1} = (N_0-1, B_1^{T-1}(N_0-1))\right) = 0$ and at $Y^{n-1} = (N_0-1, N_1-1)$, $Y_0^{n-1} \geq B_0^{T-1}(N_1-1) = \min\{B_0^T(N_1-1), B_0^T(N_1-1)-1\} = N_0-1$ and therefore the process stops as well. The argument is analogous for other t.

As it is straightforward to check that B_0^t, B_1^t satisfy Condition 3 by construction, we have that Lemma 17 applies, which concludes the proof for claim (ii) in Theorem 4.

Finally, claim (iii) in Theorem 4, as mentioned, follows immediately from Proposition 2 and Lemma 3.

Proof of Proposition 7

Given that the family of Beta distributions is closed under Bayesian updating, for any measure μ_i corresponding to a Beta distribution with parameters $(\hat{t}\cdot\hat{\sigma}_{-i}, \hat{t}\cdot(1-\hat{\sigma}_{-i}))$ I will denote it by $(\hat{\sigma}_{-i}, \hat{t})$. Let us first establish some properties of V_i when μ_i is restricted to the class of Beta distributions that we will need.

Lemma 18. For any $\hat{t} > 0$ and any $\hat{\sigma}'_{-i}, \hat{\sigma}_{-i} \in (0,1)$ such that $\hat{\sigma}'_{-i} > \hat{\sigma}_{-i}$, $V_i(\hat{\sigma}'_{-i}, \hat{t}) \geq V_i(\hat{\sigma}_{-i}, \hat{t})$.

Proof. The result follows from the observation that $(\hat{\sigma}'_{-i}, \hat{t}) \geq_{MLR} (\hat{\sigma}_{-i}, \hat{t})$ and the fact that, by Lemma 14, V_i is monotone in $\geq_{MLR}$. □

I will now show that $V_i(\hat{\sigma}_{-i}, \hat{t})$ is convex in $\hat{\sigma}_{-i} \in (0,1)$. Note that this does not follow from convexity of $V_i(\mu_i)$ in μ_i as the space of Beta distributions is not convex, that is, the convex combination of parameters of two Beta distributions is not in general equivalent to the mixture of two Beta distributions with those parameters.

Lemma 19. For any $\hat{t} > 0$, $V_i(\hat{\sigma}_{-i}, \hat{t})$ is convex in $\hat{\sigma}_{-i} \in (0,1)$.

Proof. Let $\tilde{V}_i(\hat{\sigma}_{-i}, \hat{t})$ be convex and increasing in $\hat{\sigma}_{-i} \in (0,1)$ for any $\hat{t} > 0$. Then, it has increasing differences in $\hat{\sigma}_{-i}$. Consequently, for any $1 > \hat{\sigma}_{-i} > \hat{\sigma}'_{-i} > 0$, $\forall \lambda \in (0,1)$, letting $\hat{\sigma}''_{-i} = \lambda\hat{\sigma}_{-i} + (1-\lambda)\hat{\sigma}'_{-i}$, we have that

$$\begin{aligned}
&\lambda\left[\hat{\sigma}''_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right) + (1-\hat{\sigma}''_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}\,,\,\hat{t}+1\right)\right] + \\
&\quad + (1-\lambda)\left[\hat{\sigma}''_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right) + (1-\hat{\sigma}''_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}\,,\,\hat{t}+1\right)\right] \\
&\leq \lambda\left[\hat{\sigma}_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right) + (1-\hat{\sigma}_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}\,,\,\hat{t}+1\right)\right] + \\
&\quad + (1-\lambda)\left[\hat{\sigma}'_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right) + (1-\hat{\sigma}'_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}\,,\,\hat{t}+1\right)\right] \\
&\Leftrightarrow 0 \geq \lambda(1-\lambda)\Big[\left(\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right) - \tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right)\right) - \\
&\qquad - \left(\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}\,,\,\hat{t}+1\right) - \tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}\,,\,\hat{t}+1\right)\right)\Big].
\end{aligned}$$

Let

$$B_i(\tilde{V}_i)[(\hat{\sigma}_{-i}, \hat{t})] := \max\{v_i(\hat{\sigma}_{-i}, \hat{t})\,,\,\hat{\sigma}_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}_{-i} + 1/(\hat{t}+1)\,,\,\hat{t}+1\right) + (1-\hat{\sigma}_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}_{-i}\,,\,\hat{t}+1\right) - c_i\}$$

which corresponds to the operator B_i from Section 1 when beliefs are given by Beta distribution. Then,

$$
\begin{aligned}
& B_i(\tilde{V}_i)[(\hat{\sigma}''_{-i},\hat{t})] \\
&= \max\left\{v_i(\hat{\sigma}''_{-i},\hat{t})\,,\,\hat{\sigma}''_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}''_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}''_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}''_{-i}\,,\,\hat{t}+1\right)-c_i\right\} \\
&\leq \max\left\{\begin{array}{ll} \lambda\cdot v_i(\hat{\sigma}_{-i},\hat{t}) & \lambda\left[\hat{\sigma}''_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}''_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}\,,\,\hat{t}+1\right)\right]+ \\ +(1-\lambda)\cdot v_i(\hat{\sigma}'_{-i},\hat{t}) & , \quad +(1-\lambda)\left[\hat{\sigma}''_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}''_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}\,,\,\hat{t}+1\right)\right]-c_i \end{array}\right\} \\
&\leq \max\left\{\begin{array}{ll} \lambda\cdot v_i(\hat{\sigma}_{-i},\hat{t}) & \lambda\left[\hat{\sigma}_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}\,,\,\hat{t}+1\right)\right]+ \\ +(1-\lambda)\cdot v_i(\hat{\sigma}'_{-i},\hat{t}) & , \quad +(1-\lambda)\left[\hat{\sigma}'_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}'_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}'_{-i}\,,\,\hat{t}+1\right)\right]-c_i \end{array}\right\} \\
&\leq \lambda\cdot\max\left\{v_i(\hat{\sigma}_{-i},\hat{t})\,,\,\hat{\sigma}_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}_{-i}\,,\,\hat{t}+1\right)-c_i\right\}+ \\
&\quad +(1-\lambda)\cdot\max\left\{v_i(\hat{\sigma}'_{-i},\hat{t})\,,\,\hat{\sigma}'_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}'_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}'_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\cdot\hat{\sigma}'_{-i}\,,\,\hat{t}+1\right)-c_i\right\} \\
&= \lambda B_i(\tilde{V}_i)[(\hat{\sigma}_{-i},\hat{t})]+(1-\lambda)B_i(\tilde{V}_i)[(\hat{\sigma}'_{-i},\hat{t})].
\end{aligned}
$$

Let $B_i^{(k+1)}(\tilde{V}_i) := B_i(B_i^{(k)}(\tilde{V}_i)$ for $k \in \mathbb{N}$ with $B_i^{(1)} \equiv B_i$. As the optimal stopping time is uniformly bounded (Proposition 3), $V_i(\hat{\sigma}_{-i},\hat{t}) = B_i^{(k)}(v_i)[(\hat{\sigma}_{-i},\hat{t})]$ for some $k \in \mathbb{N}$. As $v_i(\hat{\sigma}_{-i},\hat{t})$ is clearly convex and increasing in $\hat{\sigma}_{-i}$ for any $\hat{t}$, then so is $V_i(\hat{\sigma}_{-i},\hat{t})$. □

A third needed property to establish the claims in Proposition 7 is the following:

Lemma 20. For any $\hat{\sigma}_{-i} \in (0,1)$ and $\hat{t} > 0$, $V_i(\hat{\sigma}_{-i},\hat{t})$ is non-increasing in $\hat{t}$.

Proof. First, note that $v_i(\hat{\sigma}_{-i},\hat{t}) = \max\{u_i(1,\hat{\sigma}_{-i}), u_i(0,\hat{\sigma}_{-i})\}$, that is, $v_i(\hat{\sigma}_{-i},\hat{t})$ is invariant with respect to $\hat{t}$. Let $z(t)$ be a random variable such that $z(t) = t/(t+1)\hat{\sigma}_{-i} + 1/(t+1)$ with probability $\hat{\sigma}_{-i}$ and $z(t) = t/(t+1)\hat{\sigma}_{-i}$ with complementary probability and let $\tilde{V}_i(\hat{\sigma}_{-i},\hat{t})$ be increasing and convex in the first argument, non-increasing in the second. Then, $\mathbb{E}[z(t)] = \hat{\sigma}_{-i}$ and for $t' > t$, $z(t)$ is a mean-preserving spread of $z(t')$ and thus $\mathbb{E}[\tilde{V}_i(z(t'),t')] \leq \mathbb{E}[\tilde{V}_i(z(t),t')] \leq \mathbb{E}[\tilde{V}_i(z(t),t)]$. Hence,

$$
\begin{aligned}
B_i(\tilde{V}_i)[(\hat{\sigma}_{-i},\hat{t})] &= \max\{v_i(\hat{\sigma}_{-i},\hat{t})\,,\,\hat{\sigma}_{-i}\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}+1/(\hat{t}+1)\,,\,\hat{t}+1\right)+(1-\hat{\sigma}_{-i})\tilde{V}_i\left(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}\,,\,\hat{t}+1\right)-c_i\} \\
&= \max\{v_i(\hat{\sigma}_{-i},\hat{t})\,,\,\mathbb{E}[\tilde{V}_i(z(\hat{t}),\hat{t})]-c_i\} \\
&\geq \max\{v_i(\hat{\sigma}_{-i},\hat{t}')\,,\,\mathbb{E}[\tilde{V}_i(z(\hat{t}'),\hat{t}')]-c_i\} \\
&= B_i(\tilde{V}_i)[(\hat{\sigma}_{-i},\hat{t}')].
\end{aligned}
$$

By the same argument as in Lemma 19, we have that $V_i(\hat{\sigma}_{-i}, \hat{t})$ is decreasing in $\hat{t}$. □

Let $\overline{\sigma}_{-i}, \underline{\sigma}_{-i} : \mathbb{R}_{++} \to [0,1] \cup \{\emptyset\}$ be such that $\overline{\sigma}_{-i}(t) := \sup\{\hat{\sigma}_{-i} \in [0,1] \mid V_i(\hat{\sigma}_{-i}, t) > v_i(\hat{\sigma}_{-i}, t)\}$ and $\underline{\sigma}_{-i}(t) := \inf\{\hat{\sigma}_{-i} \in [0,1] \mid V_i(\hat{\sigma}_{-i}, t) > v_i(\hat{\sigma}_{-i}, t)\}$. As $V_i(1,t) = v_i(1,t) = u_i(1,1)$ and $V_i(0,t) = v_i(0,t) = u_i(0,0)$, $\overline{\sigma}_{-i}$ and $\underline{\sigma}_{-i}$ are well-defined for all $t \in (0, T_i)$, where T_i denotes the smallest upper bound on optimal stopping across all Beta priors as obtained in Proposition 12, recalling that a Beta distribution is just a Dirichlet distribution over two categories. In particular, it is easy to check that, when the prior is given by a Beta distribution, the value of taking one additional sample and then stopping, $\mathbb{E}_{\mu_i}[v_i(\mu_i \mid X_i)] - v_i(\mu_i)$, is highest when the prior mean is centered at player i's indifference point, that is, $\hat{\sigma}_{-i} = \tilde{\sigma}_{-i}$, where $\tilde{\sigma}_{-i} : u_i(1, \tilde{\sigma}_{-i}) = u_i(0, \tilde{\sigma}_{-i})$. For any $\hat{t} \in (T_i - 1, T_i)$,

$$V_i(\hat{\sigma}_{-i}, \hat{t}) = \max\{v_i(\hat{\sigma}_{-i}, \hat{t})\, ,\, \hat{\sigma}_{-i} v_i(\hat{t}/(\hat{t}+1)\sigma_{-i} + 1/(\hat{t}+1), \hat{t}+1) + (1-\hat{\sigma}_{-i}) v_i(\hat{t}/(\hat{t}+1)\sigma_{-i}, \hat{t}+1) - c_i\}$$

and from analyzing such an expression, one can check that $V_i(\hat{\sigma}_{-i}, \hat{t}) > v_i(\hat{\sigma}_{-i}, \hat{t}) \implies V_i(\tilde{\sigma}_{-i}, \hat{t}) > v_i(\tilde{\sigma}_{-i}, \hat{t})$, which implies claim (ii). Moreover, $T_i = \frac{(\delta_1+\delta_0)\cdot\tilde{\sigma}_{-i}\cdot(1-\tilde{\sigma}_{-i}) - c_i}{c_i}$ as we show in the following lemma:

Lemma 21.

$$\begin{aligned} T_i &= \sup\{t > 0 : \sup\{\hat{\sigma}_{-i} \in [0,1] : V_i(\hat{\sigma}_{-i}, t) - v_i(\hat{\sigma}_{-i}, t)\} > 0\} \\ &= \sup\{t > 0 : V_i(\tilde{\sigma}_{-i}, t) - v_i(\tilde{\sigma}_{-i}, t) > 0\} \\ &= \frac{(\delta_1 + \delta_0)\cdot\tilde{\sigma}_{-i}\cdot(1-\tilde{\sigma}_{-i}) - c_i}{c_i} \end{aligned}$$

for any $c_i < (\delta_1 + \delta_0)\cdot\tilde{\sigma}_{-i}\cdot(1-\tilde{\sigma}_{-i})$.

Proof. Note that, for any $\hat{\sigma}_{-i} \geq \tilde{\sigma}_{-i}$, $V_i(\hat{\sigma}_{-i}, \hat{t}) - v_i(\hat{\sigma}_{-i}, \hat{t}) = W_i(\hat{\sigma}_{-i}, \hat{t})$ and for $\hat{\sigma}_{-i} \leq \tilde{\sigma}_{-i}$, $V_i(\hat{\sigma}_{-i}, \hat{t}) - v_i(\hat{\sigma}_{-i}, \hat{t}) = V_i(\hat{\sigma}_{-i}, \hat{t})$. As $V_i(\hat{\sigma}_{-i}, \hat{t})$ is non-decreasing in $\hat{\sigma}_{-i}$ and $W_i(\hat{\sigma}_{-i}, \hat{t})$ is non-increasing in $\hat{\sigma}_{-i}$, we have that $V_i(\hat{\sigma}_{-i}, \hat{t}) - v_i(\hat{\sigma}_{-i}, \hat{t}) \leq V_i(\tilde{\sigma}_{-i}, \hat{t})$ for any $\hat{\sigma}_{-i}$. Consequently,

$$\sup\{t > 0 : \sup\{\hat{\sigma}_{-i} \in [0,1] : V_i(\hat{\sigma}_{-i}, t) - v_i(\hat{\sigma}_{-i}, t)\} > 0\} = \sup\{t > 0 : V_i(\tilde{\sigma}_{-i}, t) - v_i(\tilde{\sigma}_{-i}, t) > 0\}.$$

Let

$$Ev_i(\hat{\sigma}_{-i},\hat{t}) := \hat{\sigma}_{-i}v_i(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i}+1/(\hat{t}+1),\hat{t}+1)+(1-\hat{\sigma}_{-i})v_i(\hat{t}/(\hat{t}+1)\hat{\sigma}_{-i},\hat{t}+1)-v_i(\hat{\sigma}_{-i},\hat{t}),$$

denoting the value of taking one additional sample and stopping immediately after. It is clear that $V_i(\hat{\sigma}_{-i},\hat{t})-v_i(\hat{\sigma}_{-i},\hat{t}) \geq Ev_i(\hat{\sigma}_{-i},\hat{t})$ and Ev_i does not account for the value of possibly keep sampling depending on the sample realization. Straightforward algebra shows that $Ev_i(\hat{\sigma}_{-i},\hat{t}) \leq Ev_i(\tilde{\sigma}_{-i},\hat{t})$, for any $\hat{\sigma}_{-i} \in [0,1]$ and $\hat{t} > 0$ and $Ev_i(\hat{\sigma}_{-i},\hat{t})$ is strictly decreasing in $\hat{t}$. Then, if $Ev_i(\tilde{\sigma}_{-i},\hat{t}) \leq c_i \implies Ev_i(\hat{\sigma}'_{-i},\hat{t}') \leq c_i$ for any $\hat{\sigma}'_{-i} \in [0,1]$ and for any $\hat{t}' \geq \hat{t}$. As $V_i(\hat{\sigma}_{-i},\hat{t}) = B_i^{(k)}(v_i)[(\hat{\sigma}_{-i},\hat{t})]$ where $B_i^{(k)}$ is the k-th composition of B_i with itself, then $Ev_i(\tilde{\sigma}_{-i},\hat{t}) \leq c_i \implies V_i(\hat{\sigma}'_{-i},\hat{t}') \leq c_i$ for any $\hat{\sigma}'_{-i} \in [0,1]$ and for any $\hat{t}' \geq \hat{t}$. Consequently, $T_i : Ev_i(\tilde{\sigma}_{-i},T_i) = c_i$. Thus,

$$\begin{aligned} 0 = Ev_i(\tilde{\sigma}_{-i},T_i) - c_i &= \tilde{\sigma}_{-i}\left(\frac{T_i}{T_i+1}\tilde{\sigma}_{-i} + \frac{1}{T_i+1}\right)(\delta_1+\delta_0) - c_i \\ \Leftrightarrow T_i &= \frac{(\delta_1+\delta_0)\cdot\tilde{\sigma}_{-i}\cdot(1-\tilde{\sigma}_{-i}) - c_i}{c_i} \end{aligned}$$

□

As for any $t \in (0,T_i)$, $\overline{\sigma}_{-i}(t) > \tilde{\sigma}_{-i} > \underline{\sigma}_{-i}(t)$, $v_i(\overline{\sigma}_{-i}(t),t) = u_i(1,\overline{\sigma}_{-i}(t))$. To see that for $\hat{\sigma}_{-i} \in (\overline{\sigma}_{-i}(t),1)$, $V_i(\hat{\sigma}_{-i},t) = u_i(1,\hat{\sigma}_{-i})$ note that $\exists\lambda \in (0,1)$ such that $\hat{\sigma}_{-i} = \lambda\overline{\sigma}_{-i}(t)+(1-\lambda)1$ and thus $u_i(1,\hat{\sigma}_{-i}) \leq V_i(\hat{\sigma}_{-i},t) \leq \lambda V_i(\overline{\sigma}_{-i}(t),t)+(1-\lambda)V_i(1,t) = \lambda u_i(1,\overline{\sigma}_{-i}(t))+(1-\lambda)u_i(1,1) = u_i(1,\hat{\sigma}_{-i})$. The arguments are symmetric for $\underline{\sigma}_{-i}$. Consequently, the continuation region is given by $\{(\hat{\sigma}_{-i},\hat{t}) \in (0,1)\times(0,T_i) : \hat{\sigma}_{-i} \in (\underline{\sigma}_{-i}(\hat{t}),\overline{\sigma}_{-i}(\hat{t}))\}$. The fact that $\overline{\sigma}_{-i}$ ($\underline{\sigma}_{-i}$) is decreasing in $t \in (0,T_i)$ follows from the fact that $V_i(\hat{\sigma}_{-i},\hat{t})$ is decreasing in $\hat{t}$ and, thus, $\forall\hat{t} < \hat{t}'$, $V_i(\hat{\sigma}_{-i},\hat{t}) = v_i(\hat{\sigma}_{-i},\hat{t}) \implies V_i(\hat{\sigma}_{-i},\hat{t}') = v_i(\hat{\sigma}_{-i},\hat{t}')$, which then implies claim (i).

Proof of Proposition 8

Let us denote player i's indifference point as $\tilde{\sigma}_{-i} := \frac{\delta_0}{\delta_0+\delta_1}$. The proof follows from providing a characterization of the bounds defining the continuation region.

Lemma 22. For any $\hat{\sigma}_{-i} \geq \frac{t+1}{t}\tilde{\sigma}_{-i}$, for any $k \geq 0$, $V_i(\hat{\sigma}_{-i},\hat{t}) \geq \frac{\hat{t}}{\hat{t}+k}V_i(\frac{\hat{t}}{\hat{t}+k}\hat{\sigma}_{-i}+\frac{k}{\hat{t}+k}\tilde{\sigma}_{-i},\hat{t}+k)$.

Proof. Suppose $\tilde{V}$ satisfies such a property. Then,

$$\begin{aligned}
B_i(\tilde{V}_i)\left[(\hat{\sigma}_{-i},\hat{t})\right] &= \max\left\{v_i\left(\frac{\hat{t}}{\hat{t}+k}\hat{\sigma}_{-i}+\frac{k}{\hat{t}+k}\tilde{\sigma}_{-i}\right),\left(\frac{\hat{t}}{\hat{t}+k}\hat{\sigma}_{-i}+\frac{k}{\hat{t}+k}\tilde{\sigma}_{-i}\right)\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+k+1}\hat{\sigma}_{-i}+\frac{k}{\hat{t}+k+1}\tilde{\sigma}_{-i}+\frac{1}{\hat{t}+k+1},\hat{t}+k+1\right)+\right.\\
&\left.+\left(\frac{\hat{t}}{\hat{t}+k}(1-\hat{\sigma}_{-i})+\frac{k}{\hat{t}+k}(1-\tilde{\sigma}_{-i})\right)\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+k+1}\hat{\sigma}_{-i}+\frac{h}{\hat{t}+k+1}\tilde{\sigma}_{-i},\hat{t}+k+1\right)-c_i\right\}\\
&\leq \max\left\{\frac{\hat{t}}{\hat{t}+k}v_i(\hat{\sigma}_{-i}),\frac{\hat{t}}{\hat{t}+k}\left(\frac{\hat{t}}{\hat{t}+k}\hat{\sigma}_{-i}+\frac{k}{\hat{t}+k}\tilde{\sigma}_{-i}\right)\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+1}\hat{\sigma}_{-i}+\frac{1}{\hat{t}+1},\hat{t}+1\right)+\right.\\
&\left.+\frac{\hat{t}}{\hat{t}+k}\left(\frac{\hat{t}}{\hat{t}+k}(1-\hat{\sigma}_{-i})+\frac{h}{\hat{t}+k}(1-\tilde{\sigma}_{-i})\right)\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+1}\hat{\sigma}_{-i},\hat{t}+1\right)-c_i\right\}\\
&\leq \frac{\hat{t}}{\hat{t}+k}\max\left\{v_i(\hat{\sigma}_{-i}),\left(\frac{\hat{t}}{\hat{t}+k}\hat{\sigma}_{-i}+\frac{k}{\hat{t}+k}\tilde{\sigma}_{-i}\right)\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+1}\hat{\sigma}_{-i}+\frac{1}{\hat{t}+1},\hat{t}+1\right)+\right.\\
&\left.+\left(\frac{\hat{t}}{\hat{t}+k}(1-\hat{\sigma}_{-i})+\frac{h}{\hat{t}+k}(1-\tilde{\sigma}_{-i})\right)\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+1}\hat{\sigma}_{-i},\hat{t}+1\right)-c_i\right\}\\
&\leq \frac{\hat{t}}{\hat{t}+k}\max\left\{v_i(\hat{\sigma}_{-i}),\hat{\sigma}_{-i}\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+1}\hat{\sigma}_{-i}+\frac{1}{\hat{t}+1},\hat{t}+1\right)+(1-\hat{\sigma}_{-i})\cdot\tilde{V}_i\left(\frac{\hat{t}}{\hat{t}+1}\hat{\sigma}_{-i},\hat{t}+1\right)-c_i\right\}\\
&= B_i(\tilde{V}_i)[(\hat{\sigma}_{-i},\hat{t})]
\end{aligned}$$

By the same argument as in Lemma 19, the result follows. □

It is straightforward to check that a symmetric argument to that in Lemma 22 holds as well.

Let $T := \inf\left\{t\in\mathbb{R}_+ \mid \frac{t+1}{t}\tilde{\sigma}_{-i} < 1\right\}$.

Lemma 23. If T is well defined, then $\exists T' \geq T$ such that $\forall t\in[T,T']$, $\overline{\sigma}_{-i}(t) = \frac{T'}{t}\overline{\sigma}_{-i}(T') - \frac{T'-t}{t}\tilde{\sigma}_{-i} \geq \frac{t+1}{t}\tilde{\sigma}_{-i}$ and, $\forall t > T'$, $\overline{\sigma}_{-i}(t) < \frac{t+1}{t}\cdot\tilde{\sigma}_{-i}$.

Proof. If $\overline{\sigma}_{-i}(T') \geq \frac{T'+1}{T'}\cdot\tilde{\sigma}_{-i}$, then $\frac{T'}{t}\cdot\overline{\sigma}_{-i}(T') - \frac{T'-t}{t}\cdot\tilde{\sigma}_{-i} \geq \frac{t+1}{t}\tilde{\sigma}_{-i}$. Note that, by definition,

$$\overline{\sigma}_{-i}(t)\cdot v_i\left(\frac{t}{t+1}\overline{\sigma}_{-i}(t)+\frac{1}{t+1},t+1\right)+(1-\overline{\sigma}_{-i}(t))\cdot V_i\left(\frac{t}{t+1}\overline{\sigma}_{-i}(t),t+1\right)-v_i(\overline{\sigma}_{-i}(t)) = c_i.$$

Then, suppose that $\overline{\sigma}_{-i}(t) < \frac{T'}{t}\cdot\overline{\sigma}_{-i}(T') - \frac{T'-t}{t}\tilde{\sigma}_{-i}$.

$$\begin{aligned}
c_i &= \overline{\sigma}_{-i}(T')\cdot v_i\left(\frac{T'}{T'+1}\overline{\sigma}_{-i}(T')+\frac{1}{T'+1},T'+1\right)+(1-\overline{\sigma}_{-i}(T'))\cdot V_i\left(\frac{T'}{T'+1}\overline{\sigma}_{-i}(t),T'+1\right)-v_i(\overline{\sigma}_{-i}(T')) = c_i\\
&\leq \overline{\sigma}_{-i}(T')\cdot v_i\left(\frac{T'}{t+1}\overline{\sigma}_{-i}(T')-\frac{T'-t}{t+1}\tilde{\sigma}_{-i}+\frac{1}{t+1},t+1\right)+(1-\overline{\sigma}_{-i}(T'))\cdot V_i\left(\frac{T'}{t+1}\overline{\sigma}_{-i}(T')-\frac{T'-t}{t+1}\tilde{\sigma}_{-i},t+1\right)-\\
&-v_i\left(\frac{T'}{t}\overline{\sigma}_{-i}(T')-\frac{T'-t}{t}\tilde{\sigma}_{-i}\right)\\
&< \overline{\sigma}_{-i}(t)\cdot v_i\left(\frac{t}{t+1}\overline{\sigma}_{-i}(t)+\frac{1}{t+1},t+1\right)+(1-\overline{\sigma}_{-i}(t))\cdot V_i\left(\frac{t}{t+1}\overline{\sigma}_{-i}(t),t+1\right)-v_i(\overline{\sigma}_{-i}(t))
\end{aligned}$$

where the first inequality follows from Lemma 22 and the second from Lemma 18. Thus, we reach a contradiction. Hence, $\overline{\sigma}_{-i}(t+h) \geq \frac{t+h}{t}\tilde{\sigma}_{-i} \implies \overline{\sigma}_{-i}(t) \geq \frac{t+1}{t}$. Let $T' := \sup\left\{t \in \mathbb{R}_+ \mid \frac{t+1}{t}\tilde{\sigma}_{-i} \leq \overline{\sigma}_{-i}(t)\right\}$. The result then follows. □

Now suppose that μ_i is symmetric, in that $\mathbb{E}_{\mu_i}[s_{-i}] = \frac{1}{2}$; denote by $\hat{t}$ its other parameter (the sum of the usual parameters of the Beta distribution). Suppose that $\sigma_{-i} \geq 1/2$ (the argument is symmetric if $\sigma_{-i} \leq 1/2$). Let us further define $\overline{K}_i(t) := \overline{\sigma}_{-i}(t)/t$. With some manipulations, we then have that

$$\frac{p_i(t;\sigma_{-i})}{1-p_i(t;\sigma_{-i})} = \left(\frac{\sigma_{-i}}{1-\sigma_{-i}}\right)^{2\cdot\lceil\overline{K}_i(t+\hat{t})\rceil-(t+\hat{t})}$$

Moreover, note that $\sup\operatorname{supp}(\tau_i) \leq \lceil T'\rceil - \hat{t} + 1$, where again $T' := \sup\left\{t \in \mathbb{R}_+ \mid \frac{t+1}{t}\tilde{\sigma}_{-i} \leq \overline{\sigma}_{-i}(t)\right\}$. Then, either $\tau_i \leq 1$ — in which case the proposition holds vacuously —, or $\forall t, t+h \in \operatorname{supp}(\tau_i)$

$$\frac{p_i(t;\sigma_{-i})}{1-p_i(t;\sigma_{-i})} \Big/ \frac{p_i(t+h;\sigma_{-i})}{1-p_i(t+h;\sigma_{-i})} = \left(\frac{\sigma_{-i}}{1-\sigma_{-i}}\right)^{2\cdot\lceil\overline{K}_i(t+\hat{t})-\overline{K}_i(t+h+\hat{t})\rceil+h} \geq 1$$

where the inequality comes from Lemma 23. This proves claim (i) in Proposition 8.

Player i's posterior mean when stopping and taking action 1 at time t is given by $\hat{\sigma}_{-i}(t) = \frac{\hat{t}}{\hat{t}+t}\frac{1}{2} + \frac{t}{\hat{t}+t}\lceil(\hat{t}+t)\cdot\overline{\sigma}_{-i}(t+\hat{t})\rceil$. By symmetry, when taking action 0 at time t, the posterior mean is $1-\hat{\sigma}_{-i}(t)$. Then, the bias $\beta_i(t;\sigma_{-i})$ can be written as

$$\begin{aligned}\beta_i(t;\sigma_{-i}) &= |p_i(t;\sigma_{-i})\cdot\hat{\sigma}_{-i}(t) + (1-p_i(t;\sigma_{-i}))\cdot(1-\hat{\sigma}_{-i}(t)) - \sigma_{-i}| \\ &= |2(p_i(t;\sigma_{-i})-1/2)\cdot(\hat{\sigma}_{-i}(t)-1/2) - (\sigma_{-i}-1/2)|.\end{aligned}$$

Claim (ii) then follows from noting that both $p_i(t;\sigma_{-i})$ and $\hat{\sigma}_{-i}(t)$ are greater than 1/2 and decreasing in t, implying that $\beta_i(t;\sigma_{-i})$ is quasiconvex in t.

Proof of Proposition 9

Suppose that the unique Nash equilibrium is in pure strategies and, without loss of generality, suppose it is $(a_i, a_{-i}) = (1,1)$. Then, for one of the players, say player i, action 0 must be weakly dominated. Hence, at any sequential sampling equilibrium, player i will play action 1 with probability 1 regardless of the opponent's gameplay. As $(1,0)$ must not be an equilibrium, it must be that $u_{-i}(0,1) < u_{-i}(1,1)$, where the arguments of u_{-i} are (a_{-i}, a_i). If player $-i$ does not sample, then the player best-responds to the prior μ_{-i} and, by assumption, there is a unique best-response. If player $-i$ does sample and as, by Lemma 3, player $-i$ will never stop sampling when indifferent between the two actions, then, there is a unique selection of best-responses at each sample path that induces stopping. Further, as, at any equilibrium, all player $-i$'s sampled observations are 1-valued, we have that at an equilibrium $\sigma_{-i} = f_{-i}(1)$. By similar arguments as those made in the proof of Theorem 4, whenever player $-i$ samples at least once, we have that the value of the last sampled observation determines the choice and thus $f_{-i}(1) = 1$.

Now suppose that the unique Nash equilibrium involves mixed strategies. Then, without loss of generality, up to relabelling, we have that $\min\{u_i(1,1)-u_i(0,1), u_i(0,0)-u_i(1,0)\} > 0$ and $\max\{u_{-i}(1,1)-u_{-i}(1,0), u_{-i}(0,0)-u_{-i}(0,1)\} < 0$. Consequently, by Theorem 4(ii), $f_i(\sigma_{-i})$ is non-decreasing and continuous and $f_{-i}(\sigma_i)$ is non-increasing and continuous. If no player samples, by assumption, $\sigma_i = \arg\max_{\sigma_i' \in [0,1]} \mathbb{E}_{\mu_i}[u_i(\sigma_i', s_{-i})] \in \{0,1\}$ and similarly for $-i$, making the sequential sampling equilibrium unique. If one of the players samples, say player $-i$, and the other does not, then we still have that $\sigma_i = \arg\max_{\sigma_i' \in [0,1]} \mathbb{E}_{\mu_i}[u_i(\sigma_i', s_{-i})] \in \{0,1\}$ and $f_{-i}(\sigma_i) \in \{0,1\}$. If both players sample, then, by Theorem 4(ii), $f_i(\sigma_{-i})$ ($f_{-i}(\sigma_i)$) is continuous and strictly increasing (decreasing) for $\sigma_{-i} \in (0,1)$ ($\sigma_i \in (0,1)$). Furthermore, $f_i(0) = f_{-i}(1) = 0$ and $f_i(1) = f_{-i}(0) = 1$. Hence, there is a unique and interior sequential sampling equilibrium given by $(\sigma_i, \sigma_{-i}) = (f_i(\sigma_{-i}, f_{-i}(\sigma_i)) \in (0,1)^2$.

To show that a converse holds when both players sample at least once, note that by Lemma 3 implies that the player i $(-i)$ will choose the action that (mis)matches the last observation sampled. Then, if there are two or more Nash equilibria, it must be that both f_i and f_{-i} are increasing. Consequently, $f_i(0) = f_{-i}(0) = 0$ and $f_i(1) = f_{-i}(1) = 1$, which implies there are multiple sequential sampling equilibria.

Proof of Proposition 10

For claim (i), note that, whenever both players sample at least once and the underlying game Γ has a unique Nash equilibrium in fully mixed strategies, by Proposition 9, there is a unique sequential sampling equilibrium where $\sigma_i = f_i(\sigma_{-i}) \in (0,1)$, $i = 1,2$. Let

$\delta_1' > \delta_1 > 0$ and denote the probability with which player 1 chooses action 1 given the payoffs to action 1 — keeping all other payoffs constant — and given the opponent's gameplay $s_2 \in [0,1]$ by $f_1(s_2;\delta_1)$ and $f_1(s_2;\delta_1')$, respectively. Let $(\sigma_1,\sigma_2) = (f_1(\sigma_2;\delta_1), f_2(\sigma_1))$ and $(\sigma_1',\sigma_2') = (f_1(\sigma_2';\delta_1'), f_2(\sigma_1'))$ denote the sequential sampling equilibria under δ_1 and δ_1' and suppose, for the purpose of contradiction, that $\sigma_1 > \sigma_1'$. Then, $\sigma_2 = f_2(\sigma_1) < f_2(\sigma_1') = \sigma_2'$ as, by Theorem 4(ii), f_2 is decreasing in σ_1. Finally, $\sigma_1 = f_1(\sigma_2;\delta_1) \leq f_1(\sigma_2;\delta_1') < f_1(\sigma_2';\delta_1') = \sigma_1'$, a contradiction, where the first inequality follows from the fact that f_1 is non-decreasing in δ_1 (Proposition 2) and it is strictly increasing in $\sigma_2 \in (0,1)$. As $\sigma_1' \geq \sigma_1$, claim (ii) then follows directly from Theorem 4(ii).

Proof of Proposition 11

Claim (i) in the proposition follows the same arguments in the proof of Theorem 3. First, note that when taken individually, fixing their opponents' gameplay, player i's posterior μ_i upon observing sample path x_i^t, accumulates about the empirical frequency. Then, the posterior mean conditional on the player's type will too accumulate about the empirical frequency conditional on the player's type. By the strong law of large numbers, the empirical frequency converges almost surely to the true probability distribution and by the continuous mapping theorem, the conditional empirical frequency (which lies on a finite dimensional space) also converges almost surely to the true conditional probability. It is then straightforward to adjust the arguments in Lemmata 1 and 2 and Theorem 3 to obtain the result.

For claim (ii), fix i's opponents' gameplay. For $\theta_i \in \Theta_i$, let $\tilde{P}_i(\theta_i) := \{(\theta_i,\theta_{-i}), \theta_{-i} \in \Theta_{-i}\}$ and denote the generic element of $\Delta(\Xi_i)$ by p_i. Note that

$$\mathbb{P}_{\mu_i}\left(\{(a_{-i},\theta), \theta \in e_i\} \mid X_i^t\right) = \frac{|e_i| + \sum_{\ell=1}^t \mathbf{1}_{X_{i,\ell}^t = (a_{-i},e_i)}}{|A_{-i}| \times \left(\sum_{e_i' \in E_i} |e_i'|\right) + t} \xrightarrow{a.s.} \sum_{\theta' \in e_i} \rho(\theta') \cdot \sigma_{-i,\theta_{-i}'}(a_{-i}) \tag{i}$$

$$\begin{aligned} \mathbb{P}_{\mu_i}\left(\{(a_{-i},\theta), \theta \in e_i\} \mid X_i^t, \theta_i\right) &= \mathbf{1}_{e_i \in \tilde{P}_i(\theta_i)} \cdot \frac{\mathbb{P}_{\mu_i}\left(\{(a_{-i},\theta), \theta \in e_i\} \mid X_i^t\right)}{\sum_{a_{-i}' \in A_{-i}, e_i' \in \tilde{P}_i(\theta_i)} \mathbb{P}_{\mu_i}\left(\{(a_{-i}',\theta'), \theta' \in e_i'\} \mid X_i^t\right)} \\ &\xrightarrow{a.s.} \mathbf{1}_{e_i \in \tilde{P}_i(\theta_i)} \cdot \frac{\sum_{\theta' \in e_i(\theta_i)} \rho(\theta') \cdot \sigma_{-i,\theta_{-i}'}(a_{-i})}{\sum_{\theta' \in \tilde{P}_i(\theta_i)} \rho(\theta')} \end{aligned} \tag{ii}$$

$$\mathbb{P}_{\mu_i}\left((a_{-i},\theta) \mid X_i^t, \theta_i, e_i\right) = \mathbf{1}_{e_i \in \tilde{P}_i(\theta_i)} \cdot \frac{1}{|e_i|}, \tag{iii}$$

where convergence almost surely is with respect to the true probability distribution and (i) follows from the strong law of large numbers, (ii) from (i) and the continuous mapping theorem and (iii) from the fact that μ_i is correspond to a uniform distribution, that is, a Dirichlet distribution on $\Delta(\Xi_i)$ with $|\Xi_i|$ parameters, $(1,1,...,1)$, and observation that, as X_i is a $\mathscr{E}_i$-valued process, we have that by condition 2, $\epsilon_i = \{a_{-i}\} \times e_i$. Thus,

$$
\begin{aligned}
&\mathbb{E}_{\mu_i|X_i^t,\theta_i}[u_i(\sigma_i,s_{-i},\theta_i,\theta_{-i})] \\
&= \sum_{a_{-i}\in A_{-i},\theta\in\tilde{P}_i(\theta_i)} \mathbb{E}_{\mu_i}[p_i(a_{-i},\theta)\,|\,X_i^t,\theta_i]u_i(\sigma_i,a_{-i},\theta) \\
&= \sum_{a_{-i}\in A_{-i},\theta\in\tilde{P}_i(\theta_i)} \mathbb{P}_{\mu_i}\left((a_{-i},\theta)\,|\,X_i^t,\theta_i\right)u_i(\sigma_i,a_{-i},\theta) \\
&= \sum_{a_{-i}\in A_{-i},e_i\subseteq\tilde{P}_i(\theta_i)}\sum_{\theta\in e_i} \mathbb{P}_{\mu_i}\left(\{(a_{-i},\theta),\theta\in e_i\}\,|\,X_i^t,\theta_i\right)\mathbb{P}_{\mu_i}\left((a_{-i},\theta)\,|\,X_i^t,\theta_i,e_i\right)u_i(\sigma_i,a_{-i},\theta) \\
&= \sum_{a_{-i}\in A_{-i},e_i\subseteq\tilde{P}_i(\theta_i)}\sum_{\theta\in e_i} \mathbb{P}_{\mu_i}\left(\{(a_{-i},\theta),\theta\in e_i\}\,|\,X_i^t,\theta_i\right)\frac{1}{|e_i|}u_i(\sigma_i,a_{-i},\theta) \\
&\overset{a.s.}{\to} \sum_{a_{-i}\in A_{-i},e_i\subseteq\tilde{P}_i(\theta_i)}\sum_{\theta\in e_i} \frac{1}{|e_i|}\frac{\sum_{\theta'\in e_i}\rho(\theta')\cdot\sigma_{-i,\theta'_{-i}}(a_{-i})}{\sum_{\theta'\in\tilde{P}_i(\theta_i)}\rho(\theta')}u_i(\sigma_i,a_{-i},\theta) \\
&= \sum_{a_{-i}\in A_{-i},e_i\subseteq\tilde{P}_i(\theta_i)} \frac{\sum_{\theta'\in e_i}\rho(\theta')}{\sum_{\theta'\in\tilde{P}_i(\theta_i)}\rho(\theta')}\frac{\sum_{\theta'\in e_i}\rho(\theta')\cdot\sigma_{-i,\theta'_{-i}}(a_{-i})}{\sum_{\theta'\in e_i}\rho(\theta')}u_i(\sigma_i,a_{-i},\theta) \\
&= \sum_{a_{-i}\in A_{-i}}\sum_{(\theta_i,\theta_{-i})\in\tilde{P}_i(\theta_i)} \rho(\theta_{-i}\,|\,\theta_i)\cdot\overline{\sigma}_{-i}(a_{-i}\,|\,\theta_i,\theta_{-i})\cdot u_i(\sigma_i,a_{-i},\theta_i,\theta_{-i}),
\end{aligned}
$$

establishing that, if the sample size grows large, the expected payoff converges to to the expected payoff where the opponents' gameplay is averaged within each element of the analogy partition, an analogy-based expected payoff. Again, adjusting the arguments in Lemmata 1 and 2 to obtain a similar no-regret condition and the arguments in Theorem 3, one obtains the result in claim (ii).

C. Other Proofs and Examples

Table of Contents of Appendix C

C.1. Uniform Bound on Optimal Stopping Time with Dirichlet Priors

Proposition 12. There is $T(c_i, u_i) = T_i \in \mathbb{N}$ such that for any $\alpha \in \mathbb{R}^{|A_{-i}|}_{++}$, $\forall \sigma_{-i} \in \Sigma_{-i}$, if μ_i is a Dirichlet distribution with parameters α, $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq T_i) = 1$.

Proof. Let μ_i denote the measure associated with a Dirichlet distribution with parameter $\alpha \in \mathbb{R}^{|A_{-i}|}_{++}$. For any $t \in \mathbb{N}$, for any $x_i^t \in \mathscr{X}_i$,

$$\begin{aligned}
\left\| \mathbb{E}_{\mu_i}\left[\sigma_{-i} \mid x_i^t\right] - \mathbb{E}_{\mu_i}\left[\sigma_{-i} \mid x_i^t, X_{i,t+1}\right] \right\|_\infty &= \left\| \mathbb{E}_{\mu_i}\left[\sigma_{-i} \mid x_i^t\right] - \frac{(\|\alpha\|_1 + t) \cdot \mathbb{E}_{\mu_i}\left[\sigma_{-i} \mid x_i^t\right] + X_{i,t+1}}{\|\alpha\|_1 + t + 1} \right\|_\infty \\
&= \left\| \frac{\mathbb{E}_{\mu_i}\left[\sigma_{-i} \mid x_i^t\right] - X_{i,t+1}}{\|\alpha\|_1 + t + 1} \right\|_\infty \\
&\leq \frac{1}{t+1}.
\end{aligned}$$

Then, as in Proposition 3, we have

$$
\begin{aligned}
0 &\le \mathbb{E}_{\mu_i|X_i^t}\left[v_i(\mu_i \mid X_i^{t+1}) - v_i(\mu_i \mid X_i^t)\right] \\
&\le \max_{a\in A}|u_i(a)| \cdot \mathbb{E}_{\mu_i|X_i^t}\left[\left\|\mathbb{E}_{\mu_i|X_i^{t+1}}[\sigma_{-i}] - \mathbb{E}_{\mu_i|X_i^t}[\sigma_{-i}]\right\|_\infty\right] \\
&\le \max_{a\in A}|u_i(a)| \cdot \frac{1}{t+1}.
\end{aligned}
$$

This then implies that $\forall t \ge T_i = \lceil \frac{\max_{a\in A}|u_i(a)|-c}{c} \rceil$, the expected value of sampling information is lower than the sampling cost regardless of the realized sample paths and of α, by a similar argument as that in Proposition 3, we then have that the optimal stopping time τ_i is always lower than T_i, for any distribution of the samples. □

C.2. Misspecified Priors

I will now introduce a class of misspecified priors under which existence of a sequential sampling equilibrium is assured. I focus on the case of misspecified priors that allow for correlation, although the definitions and proofs are easily adjustable to the case where priors do not allow for correlation.

Bounded Optimal Stopping Time with Misspecified Priors

For $\sigma_{-i} \in \Sigma_{-i}$, let $B_\epsilon(\sigma_{-i})$ be defined as in the proof for Proposition 3 and let $B_\epsilon^{a_{-i}}(s_{-i}) := \left\{s'_{-i} \in B_\epsilon(s_{-i}) \,\middle|\, s'_{-i}(a'_{-i}) \ge s_{-i}(a'_{-i}),\ \forall a'_{-i} \in A_{-i} \setminus \{a_{-i}\}\right\}$.

Assumption 1. For any player $i \in I$,

(i) $\mu_i(\text{ int }(\Sigma_{-i})) > 0$;

(ii) supp (μ_i) is convex; and

(iii) $\forall \epsilon \in (0, 1/(2\cdot|A_{-i}|)$, $\inf_{\sigma_{-i}\in \text{supp }(\mu_i)} \max_{a_{-i}\in A_{-i}:\sigma_{-i}(a_{-i})\ge 1/|A_{-i}|} \mu_i\left(B_\epsilon^{a_{-i}}(\sigma_{-i})\right) =: \phi(\epsilon) > 0$.

Assumption 1(i) is simply to have Bayesian updating well-defined for any possible sample path; 1(ii) will imply that the posterior accumulates around a single point; and 1(iii) is a weakening of full support that accommodates misspecified priors. Specifically, $B_\epsilon^{a_{-i}}(s_{-i})$ denotes the set of distributions such that the action profile is chosen with lower probability — with a decrease up to ϵ — and all other action profiles are chosen with weakly greater probability. Condition 1(iii) posits that the agent always deems it possible that some action profile with probability greater than $1/|A_{-i}|$ is chosen less and all others are chosen with weakly greater probability. The counterpart of Assumption 1 for when μ_i does not allow for correlation is to impose that it holds for each μ_{ij}, $j \in -i$; the results follow from similar arguments in this case.

Let $q : \Sigma_{-i} \to \text{supp}\,(\mu_i)$ be such that $q(\sigma_{-i}) = \min_{\sigma'_{-i} \in \text{supp}\,(\mu_i)} H(\sigma_{-i}, \sigma'_{-i})$, where $H(\sigma_{-i}, \sigma'_{-i}) := \sum_{a_{-i} \in A_{-i}} \sigma_{-i}(a_{-i}) \ln(\sigma'_{-i}(a_{-i}))$, with the convention that $0 \cdot \infty = 0$. Furthermore, let

$$g(\epsilon) := \inf_{\sigma_{-i} \in \Sigma_{-i},\, \sigma'_{-i} \notin B_\epsilon(q(\sigma_{-i}))} D_{\text{KL}}(\sigma_{-i} \| \sigma'_{-i}),$$

where $D_{\text{KL}}(\sigma_{-i} \| \sigma'_{-i}) = H(\sigma_{-i}, \sigma'_{-i}) - H(\sigma_{-i}, \sigma_{-i})$. I will adjust the steps in Diaconis and Freedman (1990) to show the following result:

Lemma 24. Let $\mu_i \in \Delta(\Sigma_{-i})$ satisfy Assumption 1. Then, for any $\epsilon < 1/(2 \cdot |A_{-i}|)$, any $t \in \mathbb{N}$ and any $\overline{X}_i^t \in \Sigma_{-i}$,

$$\frac{\mu_i\left(B_\epsilon(q(\overline{X}_i^t)) \mid X_i^t\right)}{1 - \mu\left(B_\epsilon(q(\overline{X}_i^t)) \mid X_i^t\right)} \geq \psi_h(\epsilon) \cdot \exp\left(t \cdot 2\epsilon^2\right),$$

where $\psi_h(\epsilon) = \phi(\min\{\epsilon, 1/2 \cdot h \cdot g(\epsilon)\}) > 0$, $g(\epsilon) := \inf_{\sigma_{-i} \in \Sigma_{-i},\, \sigma'_{-i} \notin B_\epsilon(\sigma_{-i})} D_{\text{KL}}(\sigma_{-i} \| \sigma'_{-i}) > 0$, and $h \in (0,1)$.

Proof. First, note that

$$\begin{aligned}
\frac{\mu_i\left(B_\epsilon(q(\overline{X}_i^t)) \mid X_i^t\right)}{1 - \mu\left(B_\epsilon(q(\overline{X}_i^t)) \mid X_i^t\right)} &= \frac{\int_{B_\epsilon(q(\overline{X}_i^t))} \prod_{a_{-i} \in A_{-i}} \sigma_{-i}(a_{-i})^{\sum_{\ell=1}^t \mathbf{1}_{X_{i,\ell} = a_{-i}}} d\mu_i(\sigma_{-i})}{\int_{\Sigma_{-i} \setminus B_\epsilon(q(\overline{X}_i^t))} \prod_{a_{-i} \in A_{-i}} \sigma_{-i}(a_{-i})^{\sum_{\ell=1}^t \mathbf{1}_{X_{i,\ell} = a_{-i}}} d\mu_i(\sigma_{-i})} \\
&= \frac{\int_{B_\epsilon(q(\overline{X}_i^t))} \exp\left(t \cdot \sum_{a_{-i} \in A_{-i}} \overline{X}_i^t(a_{-i}) \cdot \ln(\sigma_{-i}(a_{-i}))\right) d\mu_i(\sigma_{-i})}{\int_{\Sigma_{-i} \setminus B_\epsilon(q(\overline{X}_i^t))} \exp\left(t \cdot \sum_{a_{-i} \in A_{-i}} \overline{X}_i^t(a_{-i}) \cdot \ln(\sigma_{-i}(a_{-i}))\right) d\mu_i(\sigma_{-i})} \\
&= \frac{\int_{B_\epsilon(q(\overline{X}_i^t))} \exp\left(-t \cdot H(\overline{X}_i^t, \sigma_{-i})\right) d\mu_i(\sigma_{-i})}{\int_{\Sigma_{-i} \setminus B_\epsilon(q(\overline{X}_i^t))} \exp\left(-t \cdot H(\overline{X}_i^t, \sigma_{-i})\right) d\mu_i(\sigma_{-i})} \\
&\geq \frac{\int_{B_\epsilon(q(\overline{X}_i^t))} \exp\left(-t \cdot H(\overline{X}_i^t, \sigma_{-i})\right) d\mu_i(\sigma_{-i})}{\exp\left(-t \cdot \left(H(\overline{X}_i^t, q(\overline{X}_i^t)) + g(\epsilon)\right)\right)},
\end{aligned}$$

where the inequality follows from the fact that $\forall \sigma_{-i} \notin B_\epsilon(q(\overline{X}_i^t))$, $H(\overline{X}_i^t, \sigma_{-i}) \geq H(\overline{X}_i^t, q(\overline{X}_i^t)) + g(\epsilon)$, $\forall \sigma_{-i} \in \text{supp}\,(\mu_i)$. Moreover, as H is convex and differentiable in the second argument, we have that, $\forall \sigma_{-i} \in B_h(q(\overline{X}_i^t))$,

$$\begin{aligned} H(\overline{X}_i^t, \sigma_{-i}) &\leq H(\overline{X}_i^t, q(\overline{X}_i^t)) + |\nabla_{\sigma_{-i}} H(\overline{X}_i^t, \sigma_{-i}) \cdot (q(\overline{X}_i^t) - \sigma_{-i})| \\ &\leq H(\overline{X}_i^t, q(\overline{X}_i^t)) + \|\nabla_{\sigma_{-i}} H(\overline{X}_i^t, \sigma_{-i})\|_\infty \cdot h \end{aligned}$$

Fix $h \in (0,1)$, let $\epsilon^* = \min\left\{\epsilon, \frac{1}{2} h \cdot g(\epsilon)\right\}$ and $a''_{-i} \in A_{-i}$ be such that $\mu_i\left(B_{\epsilon^*}^{a''_{-i}}(q(\overline{X}_i^t))\right) \geq \phi(\epsilon^*)$ and $\overline{X}_i^t(a''_{-i}) \geq 1/|A_{-i}|$. Then, $\forall \sigma_{-i} \in B_\epsilon^{a''_{-i}}(q(\overline{X}_i^t))$

$$\left\|\nabla_{\sigma_{-i}} H(\overline{X}_i^t, \sigma_{-i})\right\|_\infty = \left\|\left(-\frac{\overline{X}_i^t(a_{-i})}{\sigma_{-i}(a_{-i})} + \frac{1 - \sum_{a'_{-i} \in A_{-i} \setminus \{a''_{-i}\}} \overline{X}_i^t(a'_{-i})}{1 - \sum_{a'_{-i} \in A_{-i} \setminus \{a''_{-i}\}} \sigma_{-i}(a'_{-i})}\right)_{a_{-i} \in A_{-i} \setminus \{a''_{-i}\}}\right\|_\infty$$

and as $-1 \leq \frac{\overline{X}_i^t(a_{-i})}{\sigma_{-i}(a_{-i})} \leq -\frac{\overline{X}_i^t(a_{-i})}{\overline{X}_i^t(a_{-i}) + h} \leq 0 \ \forall a_{-i} \neq a''_{-i}$ and as $\frac{\partial}{\partial a} \frac{1-a}{1-a-b} > 0$ for any $1 - a \geq 1/|A_{-i}| > 1/(2 \cdot |A_{-i}|) \geq b > 0$ and as $1 - \sum_{a'_{-i} \in A_{-i} \setminus \{a''_{-i}\}} \overline{X}_i^t(a'_{-i}) = \overline{X}_i^t(a''_{-i}) \geq 1/|A_{-i}|$ and $1 - \sum_{a'_{-i} \in A_{-i} \setminus \{a''_{-i}\}} \sigma_{-i}(a'_{-i}) = \overline{X}_i^t(a''_{-i}) - b$, $1/(2 \cdot |A_{-i}|) \geq b > 0$, we have that

$$\left|\frac{1 - \sum_{a'_{-i} \in A_{-i} \setminus \{a''_{-i}\}} \overline{X}_i^t(a'_{-i})}{1 - \sum_{a'_{-i} \in A_{-i} \setminus \{a''_{-i}\}} \sigma_{-i}(a'_{-i})}\right| \leq \left|\frac{1/|A_{-i}|}{1/|A_{-i}| - b}\right| \leq 2$$

and thus $\|\nabla_{\sigma_{-i}} H(\overline{X}_i^t, \sigma_{-i})\|_\infty \leq 2$ and $H(\overline{X}_i^t, \sigma_{-i}) \leq H(\overline{X}_i^t, q(\overline{X}_i^t)) + 2\epsilon^* \leq H(\overline{X}_i^t, q(\overline{X}_i^t)) + h \cdot g(\epsilon)$ and therefore

$$\begin{aligned} \frac{\mu_i\left(B_\epsilon(q(\overline{X}_i^t)) \mid X_i^t\right)}{1 - \mu\left(B_\epsilon(q(\overline{X}_i^t)) \mid X_i^t\right)} &\geq \frac{\int_{B_\epsilon(q(\overline{X}_i^t))} \exp\left(-t \cdot H(\overline{X}_i^t, \sigma_{-i})\right) d\mu_i(\sigma_{-i})}{\exp\left(-t \cdot \left(H(\overline{X}_i^t, q(\overline{X}_i^t)) + g(\epsilon)\right)\right)} \\ &\geq \frac{\int_{B_{\epsilon^*}^{a''_{-i}}(q(\overline{X}_i^t))} \exp\left(-t \cdot H(\overline{X}_i^t, \sigma_{-i})\right) d\mu_i(\sigma_{-i})}{\exp\left(-t \cdot \left(H(\overline{X}_i^t, q(\overline{X}_i^t)) + g(\epsilon)\right)\right)} \\ &\geq \frac{\int_{B_{\epsilon^*}^{a''_{-i}}(q(\overline{X}_i^t))} \exp\left(-t \cdot \left(H(\overline{X}_i^t, q(\overline{X}_i^t)) + hg(\epsilon)\right)\right) d\mu_i(\sigma_{-i})}{\exp\left(-t \cdot \left(H(\overline{X}_i^t, q(\overline{X}_i^t)) + g(\epsilon)\right)\right)} \\ &= \mu_i\left(B_{\epsilon^*}^{a''_{-i}}(q(\overline{X}_i^t))\right) \cdot \exp(t \cdot (1-h) g(\epsilon)) \\ &\geq \phi(\epsilon^*) \cdot \exp(t \cdot (1-h) g(\epsilon)). \end{aligned}$$

The proof concludes by replacing $h = g(\epsilon) - 2\epsilon^2 > 0$, by Proposition 3.4 and Corollary 3.5 in Diaconis and Freedman (1990), noting $g(\epsilon) < 1$ for $\epsilon < 1/(2 \cdot |A_{-i}|)$. □

Existence of a Sequential Sampling Equilibrium under Misspecified Priors

From Lemma 24, an analogous result to Proposition 3 follows:

Proposition 13. Suppose that μ_i satisfies Assumption 1. Then, $\exists T(u_i, \mu_i, c_i) = T_i \in \mathbb{N}_0$ such that $\forall \sigma_{-i} \in \Sigma_{-i}$, $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq T_i) = 1$.

The proof follows the exact same steps as that of Proposition 3, replacing $h(\epsilon, t)$ with $\phi(\epsilon^*) \cdot \exp(t \cdot (1-h) g(\epsilon))$ from Lemma 24, using $q(\overline{X}_i^t)$ instead of $\overline{X}_i^t$, and $d(2/(t+1))$ instead of $2/(t+2)$, where $d(e) := \max_{\sigma_{-i} \in \Sigma_{-i}, \sigma'_{-i} \in B_e(\sigma_{-i})} |q(\sigma_{-i}) - q(\sigma'_{-i})|$ is continuous and $\lim_{e \to \infty} d(e) = 0$.

Finally, the result on sufficient conditions for existence under non-degenerate misspecified priors ensues:

Proposition 14. Let G be an extended game such that Assumption 1 holds. Then, a sequential sampling equilibrium exists.

The proof is analogous to that of Theorem 1 and therefore omitted.

C.3. Relation Between Strong Robustness and Singleton Stable Sets

Trembling-hand perfect Nash equilibria in normal-form games are defined as the limit of Nash equilibria of some sequence of perturbed games. In this sequence of perturbed games, players are constrained to choosing from the set of probability distributions which place strictly positive but vanishing probability on every action.[26]

As is well-known, this definition is equivalent to another: Trembling-hand perfect Nash equilibria are those that, for each player, the Nash equilibrium strategy is a best-response to some sequence of totally mixed strategies of the opponents' that converges to the opponents' Nash equilibrium strategy. It is then immediate that strong robustness relaxes this second definition of trembling-hand perfection in requiring that, for each player, the Nash

[26] In this, I follow the textbook treatment given by Mas-Colell et al. (1995, Section 8.F) rather than the original paper, Selten (1975), which focuses on extensive-form games.

equilibrium strategy is a best-response to any sequence of totally mixed strategies of the opponents' that converges to the opponents' Nash equilibrium strategy, provided that all elements of the sequence remain close enough to their limit.

Kohlberg and Mertens's (1986) definition of stable set relates to the first definition of trembling-hard perfection above. A stable set S corresponds to a closed set of Nash equilibria satisfying the following condition: for any $\epsilon > 0$ there is $\delta_0 > 0$ such that for any $\sigma \in \text{int}(\Sigma)$ and any $\delta_1, ..., \delta_n$, $0 < \delta_k < \delta_0$, the perturbed game where players are constrained to choosing strategies in the set $\{(1 - \delta_k) \cdot \sigma'_i + \delta_k \sigma_i, \sigma'_i \in \Sigma_i\}$ has a Nash equilibrium that is ϵ-close to S. It should be clear that if a strategy profile is a singleton stable set, then it is also trembling-hand perfect.

While strong robustness and singleton stable sets refine trembling-hand perfect equilibria in the same spirit, they are not equivalent. In fact, the former is strictly stronger than the latter.

That a strongly robust Nash equilibrium corresponds to a singleton stable set is immediate as every players' equilibrium strategy is a best response to *any* strategy profile of the opponents that is close enough to the opponents' equilibrium strategy profile.

To see that the converse does not hold, consider the game in Figure 10. Player 2 is indifferent between any strategy, whereas player 1 strictly prefers to match actions whenever player 2 chooses a or b and is indifferent between the two actions when player 2 chooses c. It is immediate that (A, c) is a Nash equilibrium that is not strongly robust as for any ϵ, A is not a best-response to any σ_2 such that $\epsilon/2 > \sigma_2(b) > \sigma_2(a)$. However, (A, c) is a singleton stable set. This follows because, for any $\epsilon > 0$, there is a small enough $\delta_0 < \epsilon$ such that, for any perturbed choice set for player 2 as specified above, there is always some strategy σ_2 to which A is a strict best-response that is within ϵ of $\sigma_2^* : \sigma_2^*(c) = 1$. Given that player 2 is indifferent between any action, player 2 choosing σ_2 and player 1 choosing the strategy that places the largest probability possible on A is an equilibrium of the constrained game and such an equilibrium is ϵ-close to (A, c).

		Player 2		
		a	b	c
Player 1	A	1 , 1	0 , 1	1 , 1
	B	0 , 1	1 , 1	1 , 1

Figure 10.

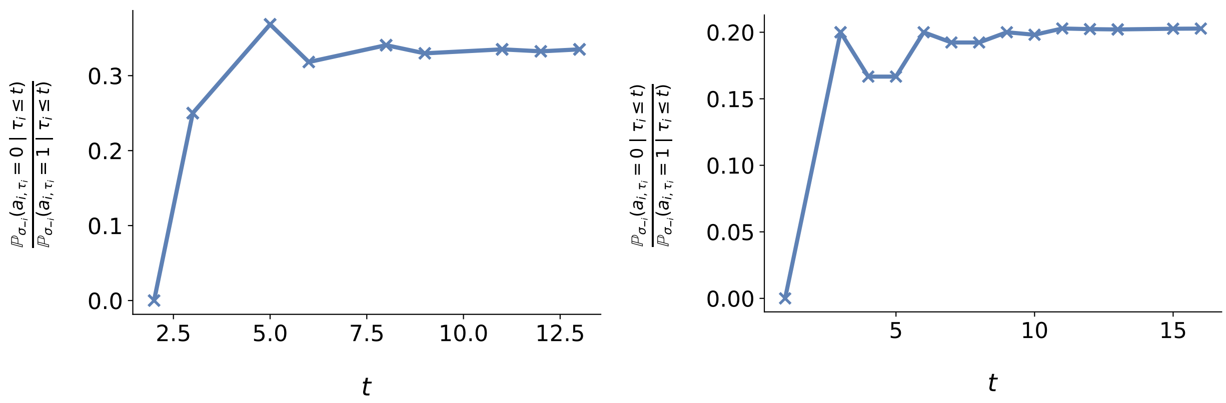
C.4. Examples of Non-Monotone Speed-Accuracy Relation

I will now present two examples illustrating how speed-accuracy relation may be non-monotone. In the first example, the prior mean coincides with the indifference point, that is, given the indifference point $\tilde{\sigma}_{-i} : u_i(1, \tilde{\sigma}_{-i}) = u_i(0, \tilde{\sigma}_{-i})$, we have that $\mathbb{E}_{\mu_i}[s_{-i}] = \tilde{\sigma}_{-i}$. In the second example, the prior is correct on average, that is, the prior mean coincides with the true probability, $\mathbb{E}_{\mu_i}[s_{-i}] = \sigma_{-i}$. In both, the relation between speed and accuracy is non-monotone, that is,

$$\frac{\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) \notin \arg\max_{\sigma_i \in [0,1]} u_i(\sigma_i, \sigma_{-i}) \mid \tau_i \leq t\right)}{\mathbb{P}_{\sigma_{-i}}\left(b_i(\mu_i \mid X_i^{\tau_i}) \in \arg\max_{\sigma_i \in [0,1]} u_i(\sigma_i, \sigma_{-i}) \mid \tau_i \leq t\right)}$$

is not monotone in t.

Figure 11 shows the mentioned ratio for the two examples. In particular, I set the true probability distribution at $\sigma_{-i} = 1/2$, payoffs $u_i(1, s_{-i}) = s_{-i} - 7/20$ and $u_i(0, s_{-i}) = 0$, sampling cost $c_i = 1/200$, with the prior given by a Beta distribution with parameters $(\hat{\sigma}_{-i}\,,\, 1 - \hat{\sigma}_{-i})$. In the first example (panel 11a), $\hat{\sigma}_{-i} = 7/20 = \tilde{\sigma}_{-i}$ and in the second (panel 11b) $\sigma_{-i} = \hat{\sigma}_{-i} = 1/2$. As it is possible to observe, in both cases, the ratio is non-monotone, with alternating increasing and decreasing spans.



(a) (b)

Figure 11. **Examples of Non-Monotone Speed-Accuracy Relation**

C.5. Myopic Sequential Sampling and Non-Convergence to Nash Equilibrium

In this appendix, I consider the case where players acquire information in a myopic manner and briefly sketch an argument as to why convergence to Nash equilibrium fails to hold as sampling costs vanish. The earliest myopic stopping rule for player i, $\underline{\tau}_i^M$, would then correspond to:

$$\underline{\tau}_i^M := \inf\left\{t \in \mathbb{N}_0 \,\middle|\, v_i(\mu_i \mid X_i^t) \geq \mathbb{E}_{\mu_i}[v_i(\mu_i \mid X_i^{t+1}) \mid X_i^t] - c_i\right\}$$

while the latest stopping rule, $\overline{\tau}_i^M$, is obtained by just replacing the weak inequality with a strict one.

A myopic sequential sampling equilibrium of an extended game G is then a distribution of actions $\sigma \in \Sigma$ whereby, for all $i \in I$, there is a some selection of best responses $b_i(\mu_i \mid x_i) \in \arg\max_{\sigma_i \in \Sigma_i} \mathbb{E}_{\mu_i}[u_i(\sigma_i, s_{-i}) \mid x_i]$, for all $x_i \in \mathscr{X}_i$, such that $\sigma_i = \mathbb{E}_{\sigma_{-i}}[b_i(\mu_i \mid X_i^{\tau_i^M})]$, with τ_i^M being some myopic stopping rule.

For any $c_i > 0$, a necessary condition for player i to keep sampling up to time T is player i faces a sample path $x_i^T \in \mathscr{X}_i$ such that, $\forall t = 1, \dots, T-1$, $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i^t] \in M_i(a_i) \implies \exists a_{-i} \in A_{-i}$ such that $\mathbb{E}_{\mu_i}[s_{-i} \mid x_i^t, a_{-i}] \notin M_i(a_i)$. That is, it is only worthwhile to keep sampling under a myopic information acquisition if there is a possibility of obtaining evidence that would change the player's beliefs in a way that optimal actions under the prior would no longer be optimal under the posterior. It is therefore immediate that any full-support prior that is sufficiently concentrated around a point in the interior of some region of Σ_{-i} where action a_i is a strict best response, *for any* $c_i > 0$, player i will not see it worthwhile to acquire information. Thus, if in no Nash equilibrium is it optimal to play such an action with probability 1, myopic sequential sampling will fail to converge to a Nash equilibrium.

A concrete example is as follows. Suppose that the underlying game is a 2×2 game with $A_1 = A_2 = \{0,1\}$ and with a unique Nash equilibrium in mixed strategies given by $\{\sigma_1^*, \sigma_2^*\}$, $\sigma_i^* \in \operatorname{int}(\Sigma_i)$, $i = 1,2$. Then, σ_2^* is such that $u_1(0,\sigma_2^*) = u_1(1,\sigma_2^*)$. If player 1's prior is given by a Beta(α,β), where $(\alpha,\beta) = t \cdot \sigma_2^0$ satisfying $\sigma_2^0(1) > \sigma_2^*(1)\frac{t+1}{t}$, then for any $c_i > 0$, $\mathbb{E}_{\mu_1}[v_1(\mu_1 \mid X_1)] = v_1(\mu_1) = u_1(1,\sigma_2^0)$. Thus, the value of a single sample is 0 and strictly below the sampling cost for any $c_i > 0$, which then implies that player 1 will not sample and just choose to take action 1. Hence, at any myopic sequential sampling equilibrium, $\sigma_1(1) = 1$, which therefore implies that, in the limit, as sampling costs vanish, $\sigma_1(1) = 1 \neq \sigma_1^*(1)$, which is interior, by assumption.