Sequential Sampling Equilibrium

Duarte Gonçalves*

Abstract

This paper introduces an equilibrium framework based on sequential sampling in which players face strategic uncertainty over their opponents' behavior and acquire information to resolve it. Sequential sampling equilibrium delivers a disciplined model featuring an endogenous distribution of choices, beliefs, and decision times, that not only rationalizes well-known deviations from Nash equilibrium, but also makes novel predictions supported by existing data. It grounds a relationship between empirical learning and strategic sophistication, and generates stochastic choice through randomness inherent to sampling, without relying on indifference or choice mistakes. Further, it provides a rationale for Nash equilibrium when information costs vanish.

Keywords: Belief Formation; Game Theory; Information Acquisition; Sequential Sampling; Bayesian Learning; Statistical Decision Theory.
JEL Classifications: C70, D83, D84, C41.

* Department of Economics, University College London; duarte.goncalves@ucl.ac.uk.
I am very grateful to Yeon-Koo Che, Mark Dean, and Navin Kartik for the continued encouragement and advice. I also thank Elliot Lipnowski and Evan Sadler, as well as Martin Cripps, Teresa Esteban-Casanelles, Evan Friedman, Drew Fudenberg, Philippe Jehiel, Qingmin Liu, Jacopo Perego, Philip Reny, Ariel Rubinstein, Ran Spiegler, Yu Fu Wong and the participants at Columbia, NYU, Caltech, University of Pennsylvania, Tel Aviv, Essex, Surrey, Oxford, ESWM'20, SMYE'21, EWET'21, ESEM'21, BRIC'22 (CERGE-EI), City Theory Workshop '22, and Warwick Theory Icebreaking Workshop '22 for valuable feedback.
First posted draft: 25 November 2020. *This draft*: 15 November 2022.

1. Introduction

In many settings, individuals are uncertain about others' behavior and, prior to taking action, acquire data or rely on past experiences to address this strategic uncertainty. For instance, sellers survey consumers' willingness-to-pay to better price a product, infrequent bidders in online auctions look up data from other auctions, voters assess parties' past performance, and customers recall past experiences at a restaurant. In all such situations, choices entail different payoffs depending on others' behavior, which is the object of uncertainty agents try to address: the likelihood consumers are to purchase a good a given price, the distribution of bids, the tendency of representatives of a political party to pass specific legislation, the frequency with which the chef overcooks a particular dish.

Since acquiring information is time-consuming, costly, uncertainty is often not fully resolved when making a choice, and the time and effort committed to resolving it is itself endogenous to the environment. Such observation resonates with experimental evidence in strategic settings that is difficult to square with existing equilibrium concepts, namely that decision times are longer with stronger incentives and choices suggest greater strategic sophistication (Alós-Ferrer and Bruckenmaier, 2021; Esteban-Casanelles and Gonçalves, 2020). In contrast, otherwise puzzling behavior and its relation to decision time have been rationalized in individual decision-making settings by the lens of sequential sampling models, in which an agent sequentially acquires evidence prior to making a choice.

This paper introduces an equilibrium framework based on sequential sampling in which players face strategic uncertainty and acquire information to resolve it. Players have a prior belief about others' distribution of actions and before taking an action can sample from it at a cost, which can be interpreted as obtaining data or as an internal deliberation process based on past experiences. Optimal sequential sampling renders players' action distributions dependent on their opponents'. A *sequential sampling equilibrium* then corresponds to a fixed-point, a consistent distribution of actions of all players, being equivalently characterized as a steady-state distribution of actions when players sample from past evidence, sidestepping the apparent circularity of the solution concept. This delivers a disciplined model featuring an endogenous distribution of choices, beliefs, and decision times, that I show not only rationalizes well-known deviations from Nash equilibrium, but also makes novel predictions supported by existing data. Moreover, it provides a rationale for Nash equilibrium, which it approximates as costs to information vanish.

The solution concept builds on an individual decision-making foundation of sequential sampling in a rich environment of choice under uncertainty. Players effectively act as decision makers, taking as given others' uncertain behavior, characterized by an unknown distribution. Since players can sample at a cost from their opponents' choice distribution prior to choosing an action, they face an optimal stopping problem, trading-off informational gains and costs. Upon stopping, players choose an action to maximize their expected payoff, given their posterior beliefs. Optimal sequential sampling then generates stochastic choice through the randomness inherent to sampling, without relying on indifference or choice mistakes: actions chosen upon stopping depend on posterior beliefs, informed by the realized observations whose distribution depends on others' behavior. An equilibrium is then characterized by a fixed-point condition on the distribution over players' actions.

A sequential sampling equilibrium always exists and has a steady-state foundation. While players always believe they will stop sampling in finite time, it need not be the case when considering the true distribution of their samples (as given by others' behavior). The proof for existence follows the novel observation that each player's optimal stopping time is, more than finite, uniformly bounded with respect to opponents' distribution of actions. This allows us to obtain that players' distribution of choices are continuous with respect to others', and employ standard arguments. In order to ground the solution concept as a steady-state, I consider a sequence of populations of short-lived players who sequentially sample from data on past play. If the distribution of data on actions converges as the data accumulates, the limiting distribution is a sequential sampling equilibrium, and, conversely, any sequential sampling equilibrium can be taken to be one such limiting distribution of data. Moreover, convergence always occurs in 2×2 games with a unique Nash equilibrium.

Sequential sampling equilibrium provides a rationale for the relationship between higher incentives, longer decision times, and more sophisticated play. While sequential sampling equilibrium does allow for players to choose actions that are not rationalizable with positive probability, if one scales up players' payoffs by a sufficiently large factor, only k-rationalizable actions are chosen with positive probability at any equilibrium, where the scaling factor needs to be larger for higher orders of rationalizability. This relation between empirical learning and strategic sophistication follows exactly from the fact that higher payoffs entail longer decision times, thereby leading players to sample enough so as to learn to choose only k-rationalizable actions.

As an application, I explore the implications of sequential sampling in canonical binary action games and show how payoffs affect the joint distribution of choices and decision time. I establish comparative statics results for sequential sampling. First, that increasing the payoffs to a given action increases the probability that it is chosen more often and faster, and the other less often and slower, a result that generalizes beyond two-action settings. Second, that an increase in the underlying probability that an action is optimal entails an equivalent conclusion. Then, I use these to prove that sequential sampling equilibrium predicts that the frequency with which an action is chosen increases in its payoffs, and that the opponent chooses the best response to that action more often and *faster*. If the former is a pervasive pattern in experimental settings (e.g. Goeree and Holt, 2001), the latter provides a novel prediction on how time relates to choice, which I find is borne out by existing experimental evidence.

Sequential sampling equilibrium also has implications for players' equilibrium beliefs. Experimental evidence has suggested that beliefs about others' behavior are often biased (Costa-Gomes and Weizsäcker, 2008), appear stochastic (Friedman and Ward, 2022), and depend on own incentives even when others' behavior is held fixed (Esteban-Casanelles and Gonçalves, 2020). All these patterns are implied by sequential sampling, where beliefs upon stopping will typically be biased due to the dependence on the prior, stochastic, as they depend on the realized observations, and payoff-dependent, given these affect when players stop sampling. To go beyond these general properties, I look at the case in which players' priors to be Beta distributed and I uncover a systematic relation between beliefs and decision time in binary action games: time reveals preference intensity. Specifically, the longer the decision time, the closer is the player to being indifferent between taking either action. This result, which has been recently shown in different environments (Fudenberg et al., 2018; Alós-Ferrer et al., 2021), carries over to equilibrium analysis and is present in existing experimental data.

Sequential sampling also provides a new rationale for Nash equilibrium, based on costly information acquisition. While as sampling costs decrease, players acquire more information, conditional on stopping, their observations are neither independent nor identically distributed. I overcome this complication and establish that, as information costs vanish, players do nevertheless learn their opponents' action distribution, from which we obtain convergence of sequential sampling equilibria to Nash equilibria. But not all Nash equilibria are reached in this manner: Nash equilibria involving weakly dominated actions can never be reached, and any pure strategy Nash equilibrium not involving weakly dominated actions can.

Finally, I discuss a number of possible extensions to the model. It is possible to extend sequential sampling equilibrium to games of incomplete information, as well as to more general information structures. It is straightforward to adjust the 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 ensues: limit points of Bayesian sequential sampling equilibria as sampling costs vanish are Bayesian Nash equilibria. Furthermore, I provide an extension to more general information structures. Among these, I consider the case where players have limited observability, capturing situations in which players cannot learn about the behavior of all other players or those in which data does not allow them to differentiate between some action profiles. 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. 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 regarding choices that individuals make in strategic settings, but for timed-stochastic choice data, the joint distribution of choices, beliefs and decision times.

1.1. Related Literature

This paper is related to three broad literatures: information acquisition and learning in games, and sequential sampling.

Information Acquisition in Games. There is a growing literature on equilibrium solution concepts featuring information acquisition. Yang (2015) studies a coordination game in which 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 (2022) allows for players to obtain correlated information and for more general information cost functions (as in Caplin and Dean, 2015). Hébert and La'O (2020) study this solution concept in mean-field games.

This paper provides the first solution concept in which the cost of information acquisition is experimental (Denti et al., 2022), with information acquisition corresponding to costly sequentially sampling from an information structure. While the sequential information acquisition can be studied from a static, ex-ante perspective (Morris and Strack, 2019; Hébert and Woodford, 2022), there are two conceptual features distinguishing sequential sampling equilibrium beyond results specific to stopping time. First, in the aforementioned papers players hold beliefs and can learn about their opponents' *action realizations*. Second, players' equilibrium beliefs are correct, and so, whenever there is no uncertainty about exogenous parameters, equilibria in these papers correspond to Nash equilibria of the underlying normal-form game. Instead, at sequential sampling equilibrium players are uncertain — neither correct or incorrect — about the prevailing *distribution* of actions of their opponents.

Steady State Learning in Games. Osborne and Rubinstein (2003) suppose each player observes a fixed number of samples from their opponents' equilibrium distribution of actions, and the mapping from samples to actions is 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: players employ unbiased estimators and best-respond to the obtained estimate.[1] 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. More broadly, these correspond to a form of self-confirming equilibrium (Fudenberg and Levine, 1993; Battigalli et al., 1992) in which the feedback function is fixed. In contrast to these, sequential sampling equilibrium endogenizes the information acquisition by the players, allowing us to relate players' stopping time and choices. The endogeneity of the sampling process results, for instance, in players sampling more when payoffs are scaled up, thereby affecting equilibrium choices and beliefs, a phenomenon that cannot be captured with exogenous sampling. This has relevant implications on equilibrium behavior comparative statics, namely, that with higher payoffs we have longer stopping times, and a higher level of sophistication (in the sense of k-rationalizability) of the actions chosen in equilibrium.

It is natural to compare sequential sampling equilibria with dynamic learning processes in games. Following the original interpretation of equilibrium beliefs as given by a scenario where players "accumulate empirical evidence" (Nash, 1950), a large literature has studied convergence of learning processes to Nash equilibria.[2] Most relevant to our paper are models in which short-lived players best-responding to the observed frequency of past play, broadly known as fictitious play (Brown, 1951).[3] While not always delivering a convergent process (Shapley, 1964; Jordan, 1993),

[1]Related are solution concepts with noisy but unbiased beliefs, e.g. Friedman and Mezzetti (2005); Friedman (2022).

[2]See Fudenberg and Levine (1998) for a comprehensive overview, and Fudenberg and Levine (2009) and Fudenberg (2020) recent surveys.

[3]The other major strand in this literature considers long-lived, forward-looking players who learn from experimentation alone, e.g. Kalai and Lehrer (1993).

its steady states correspond to Nash equilibria (Fudenberg and Kreps, 1993). 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 main difference between fictitious play and the dynamic process I analyze is that, whereas data is freely observable in fictitious play, sequential sampling players face information acquisition costs. As these costs vanish, with limiting sequential sampling equilibria correspond to Nash equilibria.

Less related is the work on learning with misspecification. 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, which can be taken as arising as the limit case of Bayesian learning with potentially misspecified priors (see Fudenberg et al., 2021). Costly information acquisition in a sequential sampling equilibrium differs in two manners. First, in a normal-form game, absent misspecification about the opponents' action distribution, Berk-Nash corresponds to Nash equilibrium, whereas the focus on costly information acquisition allows us to obtain different predictions despite retaining full support priors. Second, again in contrast with sequential sampling equilibrium, equilibrium beliefs in Berk-Nash equilibrium in general preclude any strategic uncertainty as equilibrium beliefs are degenerate except in knife-edge cases.

Sequential Sampling. Optimal stopping has been studied at least since the seminal papers by Wald (1947) and Arrow et al. (1949). Sequential sampling has been used as a modeling device in cognitive psychology and neuroscience to ground a relation between choice and decision time,[4] and, in particular, to model choice based on memory retrieval (Gold and Shadlen, 2007; Shadlen and Shohamy, 2016; Bakkour et al., 2018). Alaoui and Penta (2022) provide an axiomatic foundation of sequential sampling as a representation of iterative of reasoning. Fudenberg et al. (2018) consider a binary-action problem in which a decision-maker sequentially acquire information about the difference in payoffs in an optimal manner. They show that at longer stopping times, the agent is closer to being indifferent between the two actions. Alós-Ferrer et al. (2021) examine the general relation between time-revealed indifference and stochastic choice primitives.

Sequential sampling equilibrium adopts this framework to model belief formation in strategic settings, providing a relation between stopping time, beliefs, and choices. I further contribute to

[4]The classic reference is Ratcliff (1978). See Ratcliff et al. (2016) and Forstmann et al. (2016) for a review of the literature and Krajbich et al. (2012); Clithero (2018); Chiong et al. (2020) for economic applications.

this literature with novel results in problems with multiple available actions: comparative statics results on how choices and stopping time relate to payoffs in general decision-problems with arbitrary payoff correlation across actions. In the binary-action case that has been the focus of much of the literature, I relate the true data generating process to the distribution of choices, stopping times, and posterior beliefs, and obtain a time-revealed indifference result in a tractable discrete-time environment.

2. Sequential Sampling Equilibrium

2.1. Setup

Preliminaries. Let $\Gamma = \langle I, A, u\rangle$ denote a normal-form game, where I denotes a finite set of players or roles, with generic elements i, j and where $-i$ denotes $I \setminus i$; $A := \times_{i\in I} A_i$, where A_i is i's finite set of feasible actions; and $u := (u_i)_{i\in I}$, with $u_i : A_i \times \Delta(A_{-i}) \to \mathbb{R}$ denoting player i's payoff function, where u_i is continuous and bounded, $\Delta(A_{-i})$ being endowed with the Euclidean norm. I extend u_i to the space of probability distributions over actions with $u_i(\sigma_i, \sigma_{-i}) = \mathbb{E}_{\sigma_i}[u_i(a_i, \sigma_{-i})]$, where $\mathbb{E}_{\sigma_i}[\cdot]$ corresponds to the expectation taken with respect to σ_i. While I focus throughout on normal-form games, I highlight that the framework extends naturally to Bayesian games, as discussed in Section 5.

Beliefs. In contrast to other solution concepts, each player i is uncertain about others' true distribution of actions, σ_{-i}, and holds beliefs given by a Borel probability measure $\mu_i \in \Delta(\Delta(A_{-i}))$, where $\Delta(\Delta(A_{-i}))$ is endowed with the topology of weak* convergence, metricized by Lévy-Prokhorov metric $\|\cdot\|_{LP}$. I require player i's beliefs to have as support, $\mathrm{supp}(\mu_i)$, the set of all distributions, *allowing for correlation* — $\mathrm{supp}(\mu_i) = \Delta(A_{-i})$ — or the set of all distributions *assuming independence* across opponents, in which case beliefs are given by a product measure $\mu_i = \times_{j\in -i}\mu_{ij}$, where each μ_{ij} is a probability measure on $\Delta(A_j)$ with full support. Results will hold in either unless explicitly mentioned.[5]

Information. Prior to making a choice, player i can acquire information about the unknown distribution σ_{-i} in a sequential and costly manner. Player i has access to an information structure $\pi_i : \Delta(A_{-i}) \to \Delta(Y_i)$, where Y_i is a finite signal space. Throughout the main text, I restrict attention to the case in which these signals are observations drawn from σ_{-i}, i.e. π_i corresponds

[5]It is also possible to extend this framework to accommodate other cases potentially of interest, e.g. ruling out opponents play strictly dominated actions; we omit these cases to simplify the presentation.

to the identity. More general signal structures are considered in Appendix B.

Sequential Sampling. As I mentioned, information acquisition is sequential. In other words, prior to taking an action, player i can sequentially observe signals $y_{i,t}$ and decide when to stop acquiring information.

I will write $y_i^t = (y_{i,\ell})_{\ell\in[1..t]}$ to stand for the sample path up to time t, where each realization $y_{i,\ell}$ is distributed according to σ_{-i}, with the understanding that $y_i^0 = \emptyset$. Formally, I denote $\mathbf{y}_i = \{y_{i,t}\}_{t\in\mathbb{N}}$ as a stochastic process defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with $\mathbb{F}_i$ denoting the natural filtration of $\mathbf{y}_i$. The set of sample paths of length t is denoted by Y_i^t and the set of all finite sample path realizations is denoted by $\mathscr{Y}_i := \bigcup_{t\in\mathbb{N}} Y_i^t$. Upon observing a given sample path up to time t, y_i^t, player i updates beliefs about σ_{-i} according to Bayes' rule, denoted by $\mu_i|y_i^t$.[6]

Cost of Information. Naturally, information acquisition is also costly as otherwise player i would never stop acquiring information. For the sake of convenience, I will throughout assume that the cost of each observation is given by $c_i > 0$. It is straightforward to adjust the model in order to accommodate costs that depend on the number of observations, insofar as they are eventually bounded away from zero from below,[7] which, for all purposes, subsumes cases in which there is an upper bound on the number of observations.

Extended Games. An *extended game* G is then a tuple comprising an underlying normal-form game Γ, each players' prior beliefs $\mu = (\pi_i)_{i\in I}$, and sampling costs $c = (c_i)_{i\in I}$.

2.2. Equilibrium

Having introduced all the primitives of the model, I will turn to the equilibrium definition.

Choice. Given a belief $\mu_i' \in \Delta(\Delta(A_{-i}))$, player i upon stopping acquiring information chooses an action in order to maximize their expected utility. I will denote the player's maximized utility by $v_i : \Delta(\Delta(A_{-i})) \to \mathbb{R}$

$$v_i(\mu_i') := \max_{\sigma_i \in \Delta(A_i)} \mathbb{E}_{\sigma_i}[\mathbb{E}_{\mu_i'}[u_i(a_i, \sigma_{-i})]],$$

and I write $\sigma_i^* : \Delta(A_{-i}) \to \Delta(A_i)$ to denote a selection of optimal choices given beliefs, $\sigma_i^*(\mu_i') \in \arg\max_{\sigma_i\in\Delta(A_i)} \mathbb{E}_{\mu_i}[u_i(\sigma_i, \sigma'_{-i})]$.

[6]Note that μ_i induces a measure on $\Delta(A_{-i}) \times \mathscr{Y}_i$.

[7] Formally: there is some N and $\underline{c}_i > 0$ such that player i's cost for any observation following the N-th is greater than $\underline{c}_i$.

Optimal Stopping. Player i acquires information optimally in order to maximize expected payoffs. That is, each player i faces an optimal stopping problem: based on accumulated evidence, decide whether to stop and make a choice or obtain another signal. Formally, player i chooses a stopping time t_i in the set $\mathbb{T}_i$ of all stopping times taking values in $\mathbb{N}_0 \cup \{\infty\}$ and adapted with respect to natural filtration associated to $\mathbf{y}_i$.

Given a prior $\mu_i \in \Delta(\Delta(A_{-i}))$, player i's value function $V_i : \Delta(\Delta(A_{-i})) \to \mathbb{R}$ can be written as

$$V_i(\mu_i) := \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i^{t_i}) - c_i \cdot t_i],$$

where $\mu_i \mid y_i^{t_i}$ denotes the player's posterior belief when, upon stopping according to stopping time t_i, the sample $y_i^{t_i}$ was observed.

It will be useful to consider the dynamic programming formulation of the optimal stopping problem, with V_i corresponding to a fixed point of an operator $B_i : C_b(\Delta(\Delta(A_{-i}))) \to C_b(\Delta(\Delta(A_{-i})))$,

$$B_i(\tilde{V}_i)(\mu_i') = \max\{v_i(\mu_i'), \mathbb{E}_{\mu_i'}[\tilde{V}_i(\mu_i' \mid y)] - c_i\},$$

which will be equivalent for our purposes. This lends our value function a clear interpretation:

$$\underbrace{V_i(\mu_i')}_{\text{value at belief } \mu_i'} = \max\{ \underbrace{v_i(\mu_i')}_{\text{value of stopping}}, \underbrace{\mathbb{E}_{\mu_i}[V_i(\mu_i' \mid y_i)] - c_i}_{\text{expected value of continuing sampling}} \}.$$

We focus on the earliest optimal stopping time

$$\tau_i(\omega) := \min\{t \in \mathbb{N}_0 \mid V_i(\mu_i' \mid y_i^t(\omega)) = v_i(\mu_i' \mid y_i^t(\omega))\},$$

where its optimality follows by standard arguments (Ferguson, 2008, Ch. 3, Theorem 3); while omitted, I note the dependence of τ_i on the prior μ_i'.

For ease of reference, I summarize properties of optimal sequential sampling in this proposition:

Proposition 1. *The following properties hold: (1) v_i and V_i are bounded, convex, and uniformly continuous. (2) For any prior μ_i, player i's optimal stopping time is finite μ_i-a.s. and satisfies $\mathbb{P}_{\mu_i}(\tau > T) \leq 2\|u_i\|_\infty / c_i T$.*

This and the remaining omitted proofs are in Appendix A.

Equilibrium Definition. Each player acquires information on their opponents' action distribution and their optimal stopping policy, τ_i, determines the sequences of signals following which they optimally stop at take an action:

$$\mathscr{Y}_i^{\tau_i} := \left\{ y_i^t \in \mathscr{Y}_i : V_i(\mu_i \mid y_i^t) = v_i(\mu_i \mid y_i^t) \text{ and } \forall 0 \leq \ell < t, V_i(\mu_i \mid y_i^\ell) > v_i(\mu_i \mid y_i^\ell) \right\}.$$

Different opponent action distributions σ_{-i} induce different distributions over the signals acquired $\mathbb{P}_{\sigma_{-i}}(y_i^{\tau_i} = y_i^t)$. Since different sequences of signals y_i^t induce different posteriors $\mu_i \mid y_i^t$ at which different actions a_i may be optimal, information acquisition implies a mapping from opponents' action distributions to the player's distribution of actions,

$$\mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})] = \underbrace{\sum_{y_i^t \in \mathscr{Y}_i^{\tau_i}}}_{\substack{\text{set of stopping} \\ \text{sequences}}} \underbrace{\prod_{\ell \in [1..t]} \sigma_{-i}(y_{i,\ell})}_{\substack{\text{probability} \\ \text{observing } y_i^t}} \underbrace{\sigma_i^*(\mu_i \mid y_i^t)}_{\substack{\text{best response} \\ \text{at posterior } \mu_i | y_i^t}}$$

That is, the probability of player i taking action a_i is given by the probability of taking such an action once player i after observing y_i^t, $\sigma_i^*(\mu_i \mid y_i^t)$, considering every sequence of signals y_i^t following which player i optimally stops, $y_i^t \in \mathscr{Y}_i^{\tau_i}$, and weighting it by the probability of its occurrence. The probability that a sequence of signals y_i^t is observed is then given by $\prod_{\ell \in [1..t]} \sigma_{-i}(y_{i,\ell})$, as each observation corresponds to an action profile $y_{i,\ell} \in A_{-i}$, sampled independently from i's opponents' action distribution, σ_{-i}.

I close the model by positing a consistency condition between every players' action distribution:

Definition 1. A sequential sampling equilibrium of an extended game $G = \langle \Gamma, \mu, c \rangle$ is a profile of action distributions σ such that, for every i, $\sigma_i = \mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})]$, where τ_i is player i's earliest optimal stopping time and $\sigma_i^*(\mu_i')$ is optimal given belief μ_i'.

Interpretation. Sequential sampling equilibrium can be interpreted as positing that, prior to taking an action, players can access existing information about others' past behavior to better ground their choices. As I show in Section 2.3, this interpretation of accessing past realizations is well-grounded in a steady-state foundation for the solution concept: given its fixed-point definition, sequential sampling equilibrium entails a self-enforcing distribution of action data.

Information acquisition can then refer to procuring hard information — such as data, experts' opinions, or reviews. Our setup can speak to many practical examples. For instance, a seller doing market research to better price its product, consumers parsing reviews on a product's quality, voters learning about candidates' platforms through their statements about different issues, or infrequent bidders in online auctions looking at data from other past auctions to reason how to bid. Alternatively, one take the sampling process as an underlying introspective process, whereby players reason about how others may act by reaching back in their memory and past experiences. As mentioned, this is in line with recent neuroscience literature, which has employed models of evidence accumulation to describe how choices relate to decision time in a variety of individual decision-making contexts.

Sequential sampling equilibrium can also be taken as relaxing the implicit epistemic assumption in Nash equilibrium that, in equilibrium, players come to know their opponents' distribution of actions. If players' priors did assign probability one to the same Nash equilibrium of the underlying game, that Nash equilibrium will coincide with a sequential sampling equilibrium of the game.[8] In our model, however, players are uncertain about the prevailing distribution of actions, and it is this uncertainty that drives their information acquisition. Further, it dispenses with the assumption of mutual knowledge of the game and of others' rationality, since all learning is driven by the procured information and players need not know others' payoff functions.

Existence. I briefly note that a sequential sampling equilibrium exists in all extended games.

Theorem 1. *Every extended game has a sequential sampling equilibrium.*

The proof proceeds by verifying that, for every player i, $\sigma_{-i} \mapsto b_i(\sigma_{-i}) := \mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})]$ (1) maps to a well-defined probability distribution of player i's actions, and (2) such mapping is continuous. The main difficulty is that, while we know that, by Proposition 1, τ_i is finite with probability 1 *with respect to the player's prior* ($\mathbb{P}_{\mu_i}(\tau_i < \infty) = 1$), we need player i's optimal stopping time to be finite with probability 1 *with respect to the actual distribution of opponents' actions*, ($\mathbb{P}_{\sigma_{-i}}(\tau_i < \infty) = 1$), as otherwise $b_i(\sigma_{-i}) \notin \Delta(A_i)$. If player i never stops sampling with positive probability (with respect to the true distribution of opponents' actions), then b_i does not define a probability distribution over player i's actions and no equilibrium exists.[9] The following lemma demonstrates that this condition on stopping time is also sufficient to guarantee the desired properties on b_i:

Lemma 1. *The following two statements are equivalent: (1) player i's optimal stopping time is finite with probability 1 with respect to σ_{-i}, $\mathbb{P}_{\sigma_{-i}}(\tau_i < \infty) = 1$ for any $\sigma_{-i} \in \Delta(A_{-i})$; (2) $b_i(\sigma_{-i}) \in \Delta(A_i)$ $\forall \sigma_{-i} \in \Delta(A_{-i})$. Moreover, if (1) holds, then b_i is also continuous.*

[8] As it is implicit in this statement, even though beliefs are degenerate and coincide on the same Nash equilibrium, not all best responses need to 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. Such non-uniqueness can occur even when the game has a unique Nash equilibrium, echoing 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.

[9] I provide an example in Online Appendix D to illustrate the potential non-existence of equilibria when priors do not have full support.

The proof can be found in Appendix A. With Lemma 1 in hand, it is then straightforward to show existence of a sequential sampling equilibrium.

Proof. Given Lemma 1, if, for any σ_{-i}, τ_i is finite with probability one with respect to σ_{-i}, then b_i is a continuous mapping from $\Delta(A_{-i})$ to $\Delta(A_i)$, and existence follows from Brouwer's fixed point theorem. By assumption, $\text{supp}(\mu_i) = \Delta(A_{-i})$ and, for any σ_{-i},

$$\mathbb{P}_{\sigma_{-i}}(\tau_i(\omega) \le T) = \mathbb{P}_{\sigma_{-i}}(\{\omega : \inf\{t \mid \mathbb{E}_{\mu_{i,t}(\omega)}[V_i(\mu_{i,t}(\omega)|y_{i,t+1}] - V_i(\mu_{i,t}(\omega)) \le c_i\} \le T\}).$$

As V_i is uniformly continuous, there is $\delta > 0$ such that, $\forall \mu_i, \mu_i' \in \Delta(P_i)$ satisfying $\|\mu_i - \mu_i'\|_{LP} < \delta$, $|V_i(\mu_i) - V_i(\mu_i')| < c$. Since, by Berk (1966), $\mu_{i,t}$ weak* converges to a Dirac on σ_{-i}, σ_{-i}-a.s., $\mathbb{P}_{\sigma_{-i}}(\lim_{t\to\infty} \mathbb{E}_{\mu_{i,t}}[V_i(\mu_{i,t}|y_{i,t+1}] - V_i(\mu_{i,t}) > c_i) = 0$. □

In fact, we can obtain an upper bound on the stopping time by combining uniform continuity of V_i and the fact that μ_i uniformly accumulates around the empirical frequency:

Remark 1. For every player i, $\exists \overline{T}_i < \infty$ such that $\tau_i \le \overline{T}_i$, where $\overline{T}_i$ depends on u_i, μ_i, and c_i.

This transforms optimal stopping into a finite horizon problem, a useful result that not only simplifies the analysis, but also makes our solution concept amenable to computational applications.

2.3. A Learning Foundation

For any equilibrium model, an important question is how players may come to behave according to the model's predictions. I previously stated that sequential sampling equilibria can be thought of as a steady state of a process where players sample from accumulated past data. This section formalizes that argument.

Dynamic Sequential Sampling. The dynamic process is as follows. Fix an extended game G. Every period, $n = 1, 2, \ldots$, a unit measure of agents plays the extended game G, evenly divided across the different roles I. Each agent believes they face a stationary distribution of opponents' actions, matching the empirical frequency of past actions, $\sigma^{n-1} \in \Delta(A)$, not knowing calendar time.

Within period n, each agent with role i leans about σ_{-i}^{n-1} by optimally sequentially samples according to τ_{θ_i}. Upon stopping, the agent best responds to their posterior beliefs.[10] This induces

[10]We keep fixed a selection of best responses σ_i^* used to break-ties.

a distribution of actions and types in period n given by $b(\sigma^{n-1})$, where $b : \Delta(A) \to \Delta(A)$ is such that $b(\sigma)(a) := \times_{i\in I} b_i(\sigma_{-i})(a_i)$, with $b_i(\sigma_{-i}) := \mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})]$, as before. After taking an action, agents then exit and are replaced by a new population as is standard in evolutionary models of learning in strategic settings. At the start of the following period, the empirical frequency is then $\sigma^n = \frac{1}{n+1} b(\sigma^{n-1}) + \frac{n}{n+1}\sigma^{n-1}$, with σ_0 given. I call any such $\{\sigma^n\}_n$ a **dynamic sequential sampling process** of G.

While akin to fictitious play (Brown, 1951), under dynamic sequential sampling, each agent observes but a sample of past play realizations and the sample itself is an endogenous object.

Equilibria and Steady States. I now show an equivalence between sequential sampling equilibria and steady states of dynamic sequential sampling processes.

Theorem 2. *Let G be an extended game. σ is a sequential sampling equilibrium of G if and only if there is some dynamic sequential process $\{\sigma^n\}_n$ of G such that $\sigma^n \to \sigma$.*

Proof. I focus on the if part, since the converse is immediate. Let $\bar{\sigma}$ denote the limit of σ^n. Then,

$$0 = \lim_{n\to\infty} \left\|\sigma^n - \bar{\sigma}\right\|_\infty = \left\|\lim_{n\to\infty} \sigma^n - \bar{\sigma}\right\|_\infty = \left\|\lim_{n\to\infty} \frac{1}{n+1}\sigma^0 + \frac{n}{n+1}\left(\frac{1}{n}\sum_{\ell\in[0..n-1]} b(\sigma^\ell)\right) - \bar{\sigma}\right\|_\infty .$$

As $\sigma^n \to \bar{\sigma}$ and b is continuous, then $b(\sigma^n) \to b(\bar{\sigma})$. Consequently, the Cesàro mean $\frac{1}{n}\sum_{\ell\in[0..n-1]} b(\sigma^\ell)$ also converges to $b(\bar{\sigma})$ and therefore $0 = \|b(\bar{\sigma}) - \bar{\sigma}\|_\infty \implies b(\bar{\sigma}) = \bar{\sigma}$. □

Theorem 2 establishes for sequential sampling equilibrium and the dynamic process I defined above an analogue to Fudenberg and Kreps's (1993) results relating Nash equilibria and fictitious play, in the sense that sequential sampling equilibria coincide with the limits of convergent dynamic processes. Below I discuss two ways in which the dynamic process can be generalized.

Remark 2. Often it may be the case that information about more recent events is more easily accessible. This can be modeled as a giving a different weight to each period, for instance, exponential discounting past data: $\sigma^n = \beta\sigma^{n-1} + (1-\beta)b(\sigma^{n-1})$, $\beta \in (0,1)$. Theorem 2 also holds under this alternative definition: as $\sigma^n \to \bar{\sigma} \implies b(\sigma^n) \to b(\bar{\sigma})$ and, for any fixed ℓ, $\beta^{n-1-\ell} b(\sigma^\ell) \to 0$, we have $\sigma^n = \beta^n\sigma^0 + (1-\beta)\cdot\sum_{\ell\in[0..n-1]} \beta^{n-1-\ell}\cdot b(\sigma^\ell) \to \bar{\sigma} = b(\bar{\sigma})$.

Remark 3. 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. Write a^n for the realized actions in period

n and σ^n for their empirical frequency (given a^0), with $a^n \sim b(\sigma^{n-1})$.[11] Note that $\sigma^n \to \bar{\sigma}$ still implies that $b(\sigma^n) \to b(\bar{\sigma})$, and the arguments above remain the same, with a^n converging in distribution to a sequential sampling equilibrium.

Convergence. While in general we cannot exclude dynamic sequential sampling from cycling and failing to converge — similarly to what occurs with fictitious play[12] — in specific classes of games, convergence and asymptotic stability are guaranteed.[13] This next proposition shows this is the case for binary action games, which we will discuss in more depth in the next section.

Proposition 2. *Let $G = \langle \Gamma, \mu, c \rangle$ be a two-player extended game. If Γ has a unique Nash equilibrium, the limit of dynamic sequential sampling is a globally asymptotically stable sequential sampling equilibrium.*

Since it makes use of results discussed below, I defer the proof to the appendix.

3. Behavioral Implications

In this section, I characterize different behavioral implications of sequential sampling equilibrium. First, I explore the relation between stopping time and action sophistication. Then, I relate incentives to the joint distribution of choices and stopping time. Finally, I focus on players' beliefs and their relation with stopping time.

3.1. Rationality and Sequential Sampling

There is evidence suggesting that when facing higher stakes, individuals take longer to make choices, and choices reflect higher sophistication (as given by their level of rationalizability).[14] This section shows how sequential sampling equilibrium can provide a rationale for such an association by relating higher incentives, longer decision times, and a lower bound on the level of rationalizability of action chosen in equilibrium. Further, this establishes a relation between

[11] If agents directly sample data with past actions, $\{a^\ell\}_{\ell<n}$, one may worry that about whether sampling without replacement affects the result; this is not the case — provided, of course, the starting dataset large enough (but still finite; cf. Remark 1) so that sequential sampling without replacement is well defined.

[12] The classical reference is Shapley (1964). Cycling can occur even with stochastic fictitious play: see Hommes and Ochea (2012).

[13] An equilibrium σ is asymptotically stable if for all $\epsilon > 0$, there is a $\delta > 0$ such that for any $\sigma^0 : \|\sigma^0 - \sigma\|_\infty < \delta$, $\|\sigma^n - \sigma\|_\infty < \epsilon$ for all n. That is, if the dynamic sequential sampling process starting close enough to the equilibrium remains closeby thereafter.

[14] See e.g. Rubinstein (2007); Esteban-Casanelles and Gonçalves (2020); Alós-Ferrer and Bruckenmaier (2021).

empirical learning (as given by sequential sampling) and introspective learning (as given by rationalizability).

I first observe that, in our context, higher incentives as given by scaling up a player's payoffs, is equivalent to scaling down the cost to sampling. Optimal sequential sampling naturally predicts an inverse relation between sampling cost and stopping time:

Remark 4. For lower sampling costs c_i, player i's optimal stopping time increases in first-order stochastic dominance with respect to the prior μ_i and to any true distribution of opponents' actions $\sigma_{-i} \in \Delta(A_{-i})$; that is, both $\mathbb{P}_{\mu_i}(\tau_i \leq t)$ and $\mathbb{P}_{\sigma_{-i}}(\tau_i \leq t)$ increase for any t.

Our main result of this section goes further in determining a relation between cost and the level of sophistication of actions chosen in equilibrium. Let us recall the definition of rationalizable actions.

Definition 2. An action $a_i \in A_i$ is *1-rationalizable* if there is some $\sigma_{-i} \in \Delta(A_{-i})$ such that, $\forall \sigma_i \in \Delta(A_i)$, $u_i(a_i, \sigma_{-i}) \geq u_i(\sigma_i, \sigma_{-i})$. An action $a_i \in A_i$ is *k-rationalizable*, for $k \geq 2$, if there is some $\sigma_{-i} \in \Delta(A_{-i}^{k-1})$ such that, $\forall \sigma_i \in \Delta(A_i)$, $u_i(a_i, \sigma_{-i}) \geq u_i(\sigma_i, \sigma_{-i})$, where $A_{-i}^{k-1} := \times_{j \neq i} A_j^{k-1}$ denotes the set of $(k-1)$-rationalizable action profiles of player i's opponents. An action a_i is *rationalizable* if $a_i \in \cap_{k \in \mathbb{N}} A_i^k$.

For presentation purposes — as implied by the above definition — I will focus on a definition of rationalizability allowing for correlation among opponents' actions, and require priors to have full support on $\Delta(A_{-i})$. The below result holds as well when considering a definition of rationalizable actions that requires independence across opponents' action distributions, provided beliefs also do not allow for correlation.

We now show that scaling up incentives enough — or, equivalently, for low enough sampling costs — only k-rationalizable actions are chosen at sequential sampling equilibria:

Theorem 3. *For any normal-form game Γ, priors μ, and $k \in \mathbb{N}$, there are cost thresholds $\overline{c}_i^k > 0$ such that, for any extended game $G = \langle \Gamma, \mu, c \rangle$ in which $c_i \leq \overline{c}_i^k$ for all i, in any sequential sampling equilibrium σ of G only k-rationalizable actions are chosen with positive probability.*

The result is obtained by combining three observations — the proofs for which can be found in Appendix A. First, if player i believes that, with high enough probability, their opponents only choose $(k-1)$-rationalizable actions, then player i will choose a k-rationalizable action:

Lemma 2. *For any $k \geq 2$, there are $\epsilon, \delta > 0$, such that, if $\mu_i(B_\delta(\Delta(A_{-i}^{k-1})) > 1-\epsilon$, then* $\arg\max_{a_i \in A_i} \mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] \subseteq A_i^k$.

Second, that if player i's opponents do indeed only choose $(k-1)$-rationalizable actions, then player i's beliefs uniformly accumulate on the event that opponents only choose $(k-1)$-rationalizable actions:

Lemma 3. *For any $\mu_i \in \Delta(\Delta(A_{-i}))$ with full support, and all $\epsilon, \delta > 0$, there is t such that, for any sequence of observations y_i^t for which $y_{i,\ell} \in A_{-i}^{k-1}$ for $\ell \in [1..t]$, $\mu_i \mid y_i^t(B_\delta(\Delta(A_{-i}^{k-1}))) > 1-\epsilon$.*

And third, that, when not all of player i's actions are rationalizable, it suffices that sampling costs are low enough to ensure that the player acquires a minimum number of signals:

Lemma 4. *Suppose that there is no action a_i that is a best response to all distribution of opponents' actions $\sigma_{-i} \in \Delta(A_{-i})$. Then, for any $T \in \mathbb{N}_0$ and any full support prior $\mu_i \in \Delta(A_{-i})$, there is $\overline{c}_i > 0$ such that for any sampling cost $c_i \leq \overline{c}_i$, the associated earliest optimal stopping time $\tau_i \geq T+1$.*

The proof of Theorem 3 then proceeds easily:

Proof. The proof follows an induction argument. First, observe that no player will choose actions that are not 1-rationalizable. Now, for $k \geq 1$, assume that players choose only $(k-1)$-rationalizable actions with positive probability. From Lemma 3, for any $\delta, \epsilon > 0$ there is a T such that, for all $t \geq T$, all $i \in I$, and any $y_i^t \in A_{-i}^{k-1}$, $\mu_{i,t}(B_\delta(\Delta(A_{-i}^{k-1}))) \geq \mu_{i,t}(B_\delta(\delta_{\overline{y}_i^t})) > 1-\epsilon$. By Lemma 2, this implies that if all players sample for at least T periods, they will only choose k-rationalizable actions with positive probability. Lemma 4 ensures that we can find $c^k > 0$ such that, if $c_i \leq c^k$ $\forall i$, all players sample at least T periods, i.e. that each player's earliest optimal stopping time is bounded below by T, $\tau_i \geq T$. This concludes the proof. □

Remark 5. It is possible to generalize the result to classes of priors that satisfy a condition akin to a lower bound on density:

Definition 3 (Diaconis and Freedman 1990)**.** Let $\phi : \mathbb{R}_{++} \to \mathbb{R}_{++}$. The set of ϕ-positive distributions on $\Delta(A_{-i})$ is given by $\mathscr{M}_i(\phi) := \{\mu_i \in \Delta(\Delta(A_{-i})) \mid \inf_{\sigma_{-i} \in \Delta(A_{-i})} \mu_i(B_\epsilon(\sigma_{-i})) \geq \phi(\epsilon), \forall \epsilon > 0\}$.

Since it is possible to obtain a uniform rate of accumulation around the empirical mean for any prior $\mu_i \in \mathscr{M}_i(\phi)$ that depends only on ϕ, we can then extend Theorem 3 so that the same cost thresholds holds for all ϕ-positive priors μ_i.

3.2. Comparative Statics in Binary Action Games

Binary Action Games. In this section I provide comparative statics results for binary actions games: normal-form games $\Gamma = \langle I, A, u \rangle$ with two players, $|I| = 2$, and such that each player $i \in I$ has two actions $A_i = \{0, 1\}$ and $u_i(1, \sigma_{-i}) - u_i(0, \sigma_{-i})$ is strictly monotone and continuous in σ_{-i}, the probability that i's opponent chooses action 1.[15] Consequently, I identify $\mu_i \in \Delta(\Delta(A_{-i}))$ with a distribution on the unit interval. An extended binary action game G is an extended game for which Γ is a binary action game.

Actions and Stopping Time. Our object of interest will be the probability (according to σ_{-i}) that player i stops before time t and, upon stopping, action a_i is optimal, that is,

$$\mathbb{P}_{\sigma_j}(a_i \in A_i^*(\mu_i \mid y_i^{\tau_i}) \text{ and } \tau_i \leq t),$$

where $A_i^*(\mu_i') := \arg\max_{a_i \in A_i} \mathbb{E}_{\mu_i'}[u_i(a_i, \sigma_{-i}')]$ denotes the set of optimal choices at a given belief μ_i'. Our main result characterizes how the joint distribution of player i's choices and their stopping times changes along to three dimensions: (1) the player's payoffs, u_i, (2) their beliefs, μ_i, and (3) the true (unknown) distribution of their opponent's actions, σ_{-i}, taken as exogenous.

Ordering Payoffs and Beliefs. Let us introduce a partial order on player i's utility functions:

Definition 4. Let $u_i, u_i' : A_i \times \Delta(A_{-i}) \to \mathbb{R}$. u_i' is said to has higher incentives to action a_i than u_i, $u_i' \geq_{a_i} u_i$, if and only if there is $g : \Delta(A_{-i}) \to \mathbb{R}_+$ such that $u_i'(a_i', \sigma_{-i}') = u_i(a_i', \sigma_{-i}') + 1_{a_i' = a_i} g(\sigma_{-i})$.

Beliefs are ordered according to a generalized version of the monotone likelihood ratio property (cf. Lehrer and Wang, 2020):

Definition 5. Let $\mu_i, \mu_i' \in \Delta([0,1])$. μ_i' is said to strongly stochastic dominate μ_i, $\mu_i' \geq_{SSD} \mu_i$, if $\mu_i' \mid y_i^t$ first-order stochastically dominates $\mu_i \mid y_i^t$ for any $y_i^t \in \mathscr{Y}_i$.

Note that, when μ_i and μ_i' are mutually absolutely continuous, $\geq_{SSD}$ corresponds to the monotone likelihood ratio property, i.e. $d\mu_i'/d\mu_i$ is increasing μ_i'-a.e.

Monotone Comparative Statics. The next result characterizes the behavior induced by optimal sequential information acquisition taking σ_{-i} are exogenous:

[15] Monotonicity in σ_{-i} is automatically satisfied when $u_i(a_i, \sigma_{-i}) = \mathbb{E}_{\sigma_{-i}}[u_i(a_i, a_{-i})]$. I require strict monotonicity to prevent the case in which players are always indifferent between both actions ($u_i(1, \sigma_{-i}) = u_i(0, \sigma_{-i})$, $\forall \sigma_{-i}$), which is a trivial case. Since I will allude to extensions that may require non-linearity in σ_{-i}, I impose only minimal conditions on payoffs.

Theorem 4. *Let G be an extended binary action game and let $a_i \in \arg\max_{a_i' \in A_i} u_i(a_i', 1)$. Then, $\mathbb{P}_{\sigma_{-i}}(a_i \in A_i^*(\mu_i \mid y_i^{\tau_i})$ and $\tau_i \le t)$ increases (i) in u_i with respect to $\ge_{a_i}$, (ii) in μ_i with respect $\ge_{SSD}$, and (iii) in σ_{-i}. Moreover, it is $\mathscr{C}^\infty$ in σ_{-i}.*

Let us discuss the intuition behind the theorem (the proof is deferred to the Appendix A).

Claim (i) shows that increasing the payoff associated to action a_i, $u_i' \ge_{a_i} u_i$, makes the player not only more likely to take that action under the true distribution of actions of the opponent, but to take it *faster* and to choose the other action less often and *slower*. While an increase in payoffs does increase the value of acquiring information at some posterior beliefs — which could lead the player to learn more about the true σ_{-i} and find out that perhaps action a_i is not optimal after all — this additional information acquisition occurs only when before the player was stopping and taking an action other than a_i. In other words, player i requires now less information to be convinced to stop and take action a_i and more information to stop and choose another action. This result is not particular to binary action games: claim (i) is shown for general settings with arbitrary finitely many actions and general payoff functions.[16]

Claim (ii) can be interpreted as stating that player i is more likely to stop earlier and take action a_i the greater the probability their prior assigns to action a_i being optimal. The main difficulty is again to show that this seemingly tautological statement holds with respect to the actual, unknown, distribution of the opponent's actions; importantly, note the claim does not depend on whether or how correct player i's beliefs are. Such monotonicity in beliefs allows one to make predictions on how behavior changes with, for instance, the provision of information that shifts beliefs in the stochastic dominance order (e.g. $\mu_i|1 \ge_{SSD} \mu_i|0$).

Finally, the argument for why claim (iii) should hold is straightforward: higher σ_{-i} means that player i is more likely to observe higher signals and therefore becoming convinced that action a_i is the better alternative. The proof follows from claim (ii) and an induction argument. The fact that the probability of action a_i being optimal when stopping before time t is a polynomial with respect to σ_{-i} implies the claim on differentiability.

Theorem 4 provides comparative statics on the optimality of a given action, but leaves open the possibility that more than one action is optimal. The next lemma closes this gap by showing that, in binary action games, a player is never indifferent between the two actions at any belief held upon stopping, provided the player samples at least once or is not indifferent under the prior μ_i.

[16]See Proposition 7 in the Appendix A.

		Clasher	
		a	b
Matcher	a	δ_M , 0	0 , 1
	b	0 , γ_C	1 , 0

Figure 1: Generalized Matching Pennies

Note: $\delta_M, \gamma_C > 0$.

Lemma 5. *Let G be an extended binary action game. Then, for any player i, $A_i^*(\mu_i \mid y_i^{\tau_i})$ is a singleton or $\tau_i = 0$. Moreover, if $\tau_i > 0$, $A_i^*(\mu_i \mid y_i^{\tau_i}) = \arg\max_{a_i \in A_i} u_i(a_i, y_{i,\tau_i})$.*

The reasoning underlying the proof is simple. Without loss of generality, assume that player i's best-response to a_{-i} is to choose $a_i = a_{-i}$. Suppose that player i stops sampling after observing a 0-valued signal leaving player i indifferent between the two actions (the argument is symmetric if the last signal is 1-valued). Then, before sampling the last observation, action 1 was already 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 signal is still optimal regardless of the realization of the signal. Therefore, given that the player will not sample any further, the last signal bears no informational value to the player. As the signal is costly, then it is suboptimal to take it.

Applications. One immediate implication of Theorem 4 and Lemma 5 is in establishing a strong connection between uniqueness of sequential sampling equilibrium in an extended game and uniqueness of a Nash equilibrium of the underlying binary action game:

Proposition 3. *A binary action game Γ has a unique Nash equilibrium if and only if any extended game $G = \langle \Gamma, \mu, c \rangle$ in which players with no weakly dominant actions sample at least once there is a unique sequential sampling equilibrium.*

An analogous result holds when, in symmetric extended binary action games (same payoff functions, same prior, same sampling cost), one restricts to symmetric Nash equilibria and symmetric sequential sampling equilibria. While uniqueness of a Nash equilibrium implies uniqueness of a sequential sampling equilibrium, it is not the case that the two coincide.

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 1. When the payoffs to action a_i of player i increase, Nash equilibrium predicts that the probability with which action a_i is chosen remains the same and it is, instead, the opponent's mixed strategy that adjusts to make player i indifferent between choosing any of the two actions — what one could call the *opponent-payoff choice effect*. However, experimental evidence shows that increasing player i's payoffs to an action leads that player to choose that action more often, an *own-payoff choice effect*.[17] This motivated the emergence of different models, one of the most successful of which quantal response equilibrium (McKelvey and Palfrey, 1995), which directly embeds monotonicity of choices with respect to payoffs in the assumptions for players' behavior (Goeree et al., 2005).

Sequential sampling equilibrium not only rationalizes this empirical regularity via comparative statics pertaining to behavior induced by optimal information acquisition, it delivers novel behavior implications regarding stopping times. Increasing player i's payoffs to action a_i, (1) increases the *equilibrium* probability that player i chooses action a_i, and (2) leads their opponent, player j, in equilibrium, choosing the best response to action a_i more often and faster, and their other action less often and slower, in the sense of Theorem 4. If the first observation states sequential sampling equilibrium predicts the own-payoff choice effect,[18] the second uncovers an entirely novel prediction relating equilibrium choices and stopping time. Both follow directly from combining Proposition 3, Theorem 4, and Lemma 5.

Supporting Evidence. To investigate whether these predictions find support in existing data, I will rely on experimental data generously made available by Friedman and Ward (2022) who collected data on choices and decision times for six different generalized matching pennies games. The goal of this exercise is not to fit data or claim that sequential sampling equilibrium perfectly describes subjects' behavior or that it does so better than other existing models, but rather to present suggestive evidence supporting its novel behavioral implications. No feedback or information was provided throughout the experiment; details on the experiment, the data, and further analysis can be found in Appendix C.

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

[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.

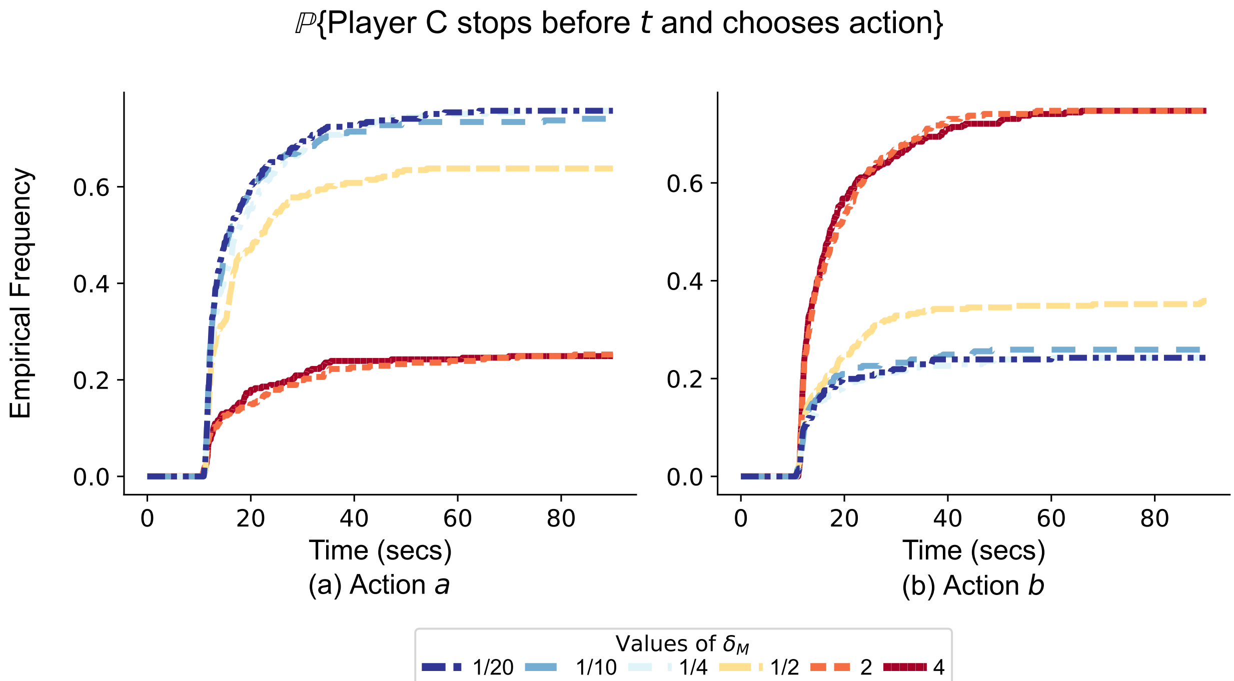

Figure 2: Opponent-Payoff Time Effect

Notes: The figure compares choices and decision times in generalized matching pennies games as given in Figure 1, for $\gamma_C = 1$ (and scaled by 20). The data is from Friedman and Ward (2022). The panels exhibit the frequency with which subjects in the player C's role take a given action (a in panel (a); b in panel (b)) before time t (in seconds). Different lines correspond to games in which the player M has different payoffs to action a. This figure uses only choice data for instances where beliefs were not elicited. The same patterns are present when beliefs are elicited. See Appendix C for further details on the data.

As shown in Figure 2, if one is to interpret stopping time as a proxy for decision time, the data supports our predictions: when increasing δ_M subjects in the Clasher's role do tend to choose action b not only more often but also faster. Moreover, they choose action a less often and slower.

3.3. Time-Revealed Preference Intensity

In this section I characterize how stopping time relates to players' posterior beliefs by considering a general family of priors in binary action games. For this section, I restrict attention to games in which payoffs are linear in the opponent's distribution of actions, i.e. $u_i(a_i, \sigma_{-i}) = \mathbb{E}_{\sigma_{-i}}[u_i(a_i, a_{-i})]$.

Beta Beliefs. For tractability, I focus on priors that are linear in new information in a manner that mimics Bayesian updating for Gaussian priors:

Definition 6. A prior μ_i is said to be linear in the accumulated information if it is non-degenerate and there are constants $\alpha_t, b_t \in \mathbb{R}$ such that for any $y_i^t \in \mathscr{Y}_i$ the posterior mean satisfies $\mathbb{E}_{\mu_i}[\sigma_{-i} \mid y_i^t] = \alpha_t \sum_{\ell=1}^t y_{i,\ell}^t + b_t$.

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, $\mathbb{E}_{\mu_i}[\sigma_{-i} \mid y_i^t] = \alpha_t/t \cdot \sum_{\ell=1}^t y_{i,\ell}^t + (1-\alpha_t) \cdot \mathbb{E}_{\mu_i}[\sigma_{-i}]$, 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.

Collapsing Boundaries. When beliefs are linear in the accumulated information, we have the following characterization of the set of beliefs at which player i optimally stops:

Proposition 4. *Let Γ be a binary action game. Suppose that there is $\tilde{\sigma}_{-i} \in \Delta(A_{-i})$ such that $u_i(1, \tilde{\sigma}_{-i}) = u_i(0, \tilde{\sigma}_{-i})$. For any $c_i > 0$, there are continuous functions $\overline{\sigma}_{-i}, \underline{\sigma}_{-i} : \mathbb{R}_{++} \to [0,1]$ such for any Beta distributed prior μ_i with parameters (α, β) player i does not optimally stop at μ_i if and only if $\mathbb{E}_{\mu_i}[\sigma_{-i}] \in (\underline{\sigma}_{-i}(\alpha+\beta), \overline{\sigma}_{-i}(\alpha+\beta))$. Furthermore, $\overline{\sigma}_{-i}$ is decreasing and $\underline{\sigma}_{-i}$ is increasing, and $\exists T_i$ such that $\forall t \geq T_i$ $\overline{\sigma}_{-i}(t) = \underline{\sigma}_{-i}(t) = \tilde{\sigma}_{-i}$.*

The proof of the result is in Appendix A.

Proposition 4 shows that when beliefs are linear in accumulated information, 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. Note that if μ_i is a Beta distribution with parameters summing to t, then $\mu_i \mid y_i$ has parameters summing to $t+1$. The continuation region is then characterized by an upper and lower threshold that delimit a decreasing interval that "collapses" to a single point: the distribution at which player i is indifferent between either action. This translates to our setting what is commonly known in the neuroscience literature as "collapsing boundaries".[19]

[19] See Hawkins et al. (2015) for a discussion on the evidence of collapsing boundaries and Bhui (2019) for supporting experimental evidence in an environment in which, as in our model, there is uncertainty about the difference in the binary actions' expected payoffs.

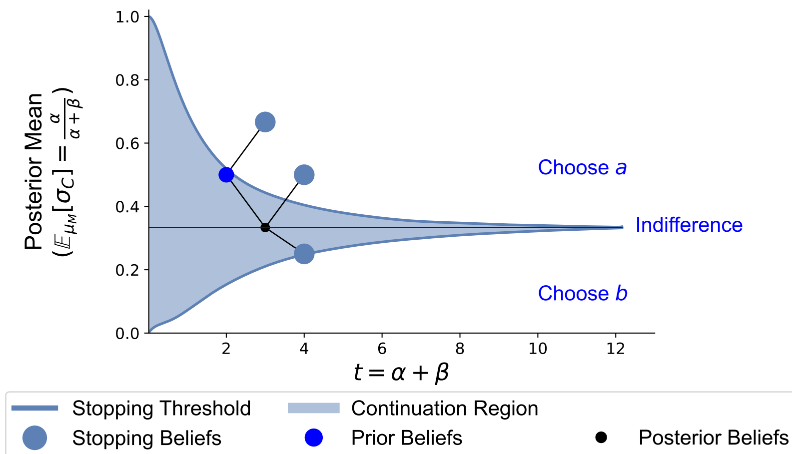

Figure 3: Stopping Regions for Beta Priors

Notes: The figure exhibits the continuation region (shaded area) and the stopping thresholds (darker blue lines) for posterior means at which player i with a Beta stops. The figure also illustrates the possible realizations of the sampling process for a player with a uniform prior (Beta(1,1)), with the posterior means indicated by circles.

One can then interpret the stopping time as an indicator of the intensity of player i's preference for one action over another: player i samples for longer if and only if the player is sufficiently close to being indifferent between the two alternatives, a phenomenon that resembles existing experimental evidence in individual decision-making (e.g. Konovalov and Krajbich, 2019). In other words, Proposition 4 entails a behavior marker in the form of time-revealed preference intensity, akin to results in Alós-Ferrer et al. (2021).

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. A clear parallel emerges between our setup and that in Fudenberg et al. (2018), 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 our setup they hold even without these qualifications.

Comparative Statics in Stopping Beliefs. From Proposition 4 and Theorem 4, we obtain that the distribution of beliefs shifts monotonically with respect to the true distribution. Specifically, approximating the stopping posterior mean by the threshold, $\mathbb{E}_{\mu_i}[\sigma_{-i} \mid y_i^{\tau_i}]$,[20] and labeling actions so that $u_i(1,\sigma_{-i}) - u_i(0,\sigma_{-i})$ is increasing in σ_{-i}, then player i's stopping (threshold) beliefs increase in a first-order stochastic dominance sense as σ_{-i} increases. This is because a higher σ_{-i} leads to a higher probability that player i chooses action 1 more often and faster (resp. action 0 less often and slower), implying that the posterior mean has to exceed a higher threshold when the player stops earlier (resp. later), as the upper (resp. lower) bound characterizing the continuation region is decreasing (resp. increasing) in the stopping time.

Supporting Evidence. Relying again on Friedman and Ward's (2022) data, I find support for both these predictions: (1) decision time is significantly negatively related to the distance between the reported mean belief and the indifference point, and (2) increasing a player's payoff to an action shifts the opponent's beliefs in the predicted first-order stochastic dominance sense — see Online Appendix C.

4. Relation to Nash Equilibrium

One initial interpretation of Nash equilibrium posits that equilibrium beliefs are reached as players "accumulate empirical information" (Nash, 1950). In a sequential sampling equilibrium, players accumulate empirical information but at a cost. A natural question is whether, as these costs vanish, sequential sampling equilibria converge to a Nash equilibrium. In this section I show this is the case. Formally,

Theorem 5. *Let Γ be a normal-form game, μ a collection of priors, and $\{c^n\}_n$ be a sequence of sampling costs such that $c^n \to 0$. For any sequence $\{\sigma^n\}_n$ such that each σ_n is a sequential sampling equilibrium of extended game $G^n = \langle \Gamma, \mu, c^n \rangle$, the limit points of $\{\sigma^n\}_n$ are Nash equilibria of Γ.*

The claim is conventional in form: players best-respond to their beliefs and their beliefs converge to the true distribution of actions of their opponents. The main complication comes from the fact that, conditional on stopping, the observations $y_i^{\tau_i}$ are not independent nor independently distributed according to player i's opponents' action distribution. To overcome this issue, the proof (see Appendix A) relies on three arguments. First, from Lemma 4 one has that as sampling costs

[20] This is so as to avoid discreteness issues inherent to the sampling procedure.

vanish, players acquire a minimum number of observations T, and, for that minimum number, each observation y_i^T is iid according to the opponents' action distribution. Second, I note beliefs accumulate at a uniform rate around the empirical mean of the observed signals. Finally, I use the optional stopping theorem to show that beliefs upon stopping converge to the true underlying distribution in an appropriate manner.

Some comments on which Nash equilibria can be selected in this manner are in order. First, let us define the concept of reachability of a Nash equilibrium:

Definition 7. A Nash equilibrium σ of a normal-form game Γ is *reachable* if there is a collection of priors μ, a sequence of costs $\{c^n\}_n \subset \mathbb{R}_{++}$ such that $c^n \to 0$, and a sequence $\{\sigma^n\}^n$, where for each n, σ^n is a sequential sampling equilibrium of the extended game $G^n = \langle \Gamma, \mu, c^n \rangle$, such that $\sigma^n \to \sigma$. A Nash equilibrium if *robustly reachable* if it is reachable for any collection of priors μ.

In the remainder of the section, I will restrict player's payoffs to be linear in distributions as usual. In other words, I require that, for every player i, $u_i(a_i, \sigma_{-i}) = \mathbb{E}_{\sigma_{-i}}[u_i(a_i, a_{-i})]$, as conventional. This will be a maintained assumption throughout the rest of this section.

Our first result provides, separately, necessary and sufficient conditions for reachability of a Nash equilibrium.

Proposition 5. *Let Γ be a normal-form game. (1) If σ is a Nash equilibrium of Γ involving weakly dominated actions, then σ is not reachable. (2) If a is a pure-strategy Nash equilibrium of Γ not involving weakly dominated actions, then a is reachable.*

(1) holds since for any prior, no player will ever choose weakly dominated actions — recall that priors have full support. For (2), note that if a does not involve weakly dominated strategies, then, by Pearce's (1984) Lemma 4, for each player i there is $\sigma_{-i}^0 \in \operatorname{int} \Delta(A_{-i})$ such that a_i is a best response to σ_{-i}^0. If we endow each player i with prior $\mu_i \in \Delta(\Delta(A_{-i}))$ corresponding to a Dirichlet distribution with mean σ_{-i}^0, then a_i is a best response to any posterior belief $\mu_i \mid y_i^t$ when $y_{i,t} = a_{-i}$. Hence, for any costs c^n, σ is sequential sampling equilibrium of $\langle \Gamma, \mu, c^n \rangle$. Note that I require the Nash equilibrium to be in pure-strategies in order to control posterior beliefs exactly, as otherwise, with some probability, σ_i may not be a best response to the posterior belief held upon stopping.

For a Nash equilibrium to be reachable with any priors, we obtain a sufficient condition:

Proposition 6. *If a is a pure-strategy Nash equilibrium in undominated strategies of the normal-form game Γ such that, for any player i, a_i is a best response to any $\sigma'_{-i} \in B_{\epsilon_i}(\delta_{a_{-i}})$ for some $\epsilon_i > 0$, then a is robustly reachable.*

The intuition for the proof (in Appendix A) can be summarized as follows: for any prior, if player i samples enough a_{-i} observations, their posterior mean will lie within ϵ_i of a_{-i} and choosing a_i is optimal. Lemma 4 then guarantees that players sample enough. Our requirement that a_i is a best response to any distribution of opponents' actions that assigns high enough probability to a_{-i} is at the same time more relaxed than strictness of Nash equilibria, and more restrictive than trembling hand perfection.

5. Extensions and Discussion

I conclude with a discussion of possible extensions of sequential sampling equilibrium.

Types and Bayesian Games. Sequential sampling equilibrium can be easily extended to accommodate Bayesian games. This aligns with the idea that sequential sampling equilibria corresponds to the case in which players don't know their opponents' payoffs, as given by their type.

In particular, consider games described by $\Gamma := \langle I, A, \Theta, \rho, u\rangle$, such that I denote the finite set of players, $A := \times_{i\in I} A_i$ the (finite) set of action profiles, $\Theta = \times_{i\in I}\Theta_i$ the (finite) set of type profiles, where Θ_i are player i's possible types, $u_i : A\times\Theta \to \mathbb{R}$ payoff functions, and $\rho \in \Delta(\Theta)$ a distribution over types. Endowing each player i with a prior $\mu_i \in \Delta(\Delta(A_{-i}\times\Theta))$ and a sampling cost c_i as before we have extended games $G = \langle\Gamma, \mu, c\rangle$. Each player i with type θ_i now learns about the joint distribution of opponents' action profiles and type profiles, $q_i \in \Delta(A_{-i}\times\Theta)$, sequentially sampling from q_i at cost c_i and stopping according to the earliest optimal stopping time τ_{i,θ_i}.

A sequential sampling equilibrium σ would then correspond to a fixed point such that $\sigma_{i,\theta_i} = \mathbb{E}_{q_i}[\sigma^*_{i,\theta_i}(\mu_i \mid y_i^{\tau_{i,\theta_i}})]$, where $\sigma^*_{i,\theta_i}(\mu_i)$ is a selection of best responses given belief μ_i, and $q_{i,\theta_i}(a_{-i},\theta) = \rho(\theta)\times_{j\neq i}\sigma_{j,\theta_j}(a_j)$ for every $a_{-i} = (a_j)_{j\neq i} \in A_{-i}$ and $\theta = (\theta_j)_{j\in I} \in \Theta$.

Different assumptions on players beliefs will give rise to different equilibria. To apply similar arguments to obtain existence of an equilibrium, we need but to require that players know the distribution of their own types and that μ_i has full support on the set of distributions $q_i \in \Delta(A_{-i}\times\Theta)$ satisfying $q_i(\theta_i) = \rho(\theta_i)$ for any $\theta_i \in \Theta_i$.[21] Differently, one could assume players know the true

[21]This renders their expected payoff given their type, $\mathbb{E}_{q_i}[u_i(a_i, a_{-i}, \theta) \mid \theta_i]$, to be continuous in $q_i \in \operatorname{supp}\mu_i$.

distribution of types, or, when types are independent, that they know so.

With the proper adjustments, behavioral implications can also be obtained, now comparing across types. For instance, if the payoff to action a_i is higher for type θ_i than for θ_i', everything else equal, in every sequential sampling equilibrium type θ_i chooses action a_i more often and faster (in the sense of Theorem 4) than type θ_i'. Finally, convergence to Bayesian Nash equilibria when sampling costs vanish can be similarly obtained.

General Information Structures and Analogy Partitions. Throughout, it was assumed that players observe action profiles drawn from a steady state distribution. Often, of course, information — and even memory — is fuzzier, and it is not possible to perfectly distinguish between certain actions taken by others, or even to observe what some other players do at all. In Online Appendix B, I provide sufficient conditions under which is it possible to generalize sequential sampling equilibrium to cases under which players observe not action profiles of their opponents, but a garbling, thereby accommodating situations such as noisy recollections, or missing or misrecorded data. When considering information structures under which players are unable to distinguish between specific action profiles (or types) of their opponents, as sampling costs vanish, sequential sampling equilibria reach not Nash equilibria but analogy-based expectation equilibria (Jehiel, 2005; Jehiel and Koessler, 2008).

Cost of Information and Discounting. In this paper, I considered a constant additive cost per observation. One could have defined this cost of information in a more general manner, allowing it to depend on the number of observations already acquired, or allowing a finite number of observations at no cost. Alternatively, one could rely on discounting payoffs instead. It is indeed possible to extend the setup to accommodate either, posing no problem for existence of an equilibrium.

Misspecified Priors. Another maintained assumption was that priors are not misspecified, i.e. players are able to learn the true data generating process. This assumption was crucial to obtain existence of an equilibrium: it is the full support of players' prior that guarantees that, as they acquire more and more observations, their beliefs accumulate around a degenerate distribution. When, instead, priors are misspecified, it is possible that players never stop sampling (according to the true data generating process), even though they believe they will (according to their posterior beliefs). I provide one such example of nonexistence in Online Appendix D.

Myopic Sequential Sampling. Finally, a comment on a simpler, alternative, version of sequential sampling equilibrium, in which information acquisition is myopic. This is closely related to a recent paper by Alaoui and Penta (2022), which provides an axiomatization of the deliberation process as one of myopic sequential information acquisition (cf. their Theorem 4). In other words, the myopic stopping time would be given by $\tau_i^M(\omega) := \min\{t \mid \mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i^{t+1}) \mid y_i^t(\omega)] - v_i(\mu_i \mid y_i^t(\omega)) \leq c_i\}$, with players stopping whenever the expected value of sampling information is smaller than the cost of one more observation.

One could then define myopic sequential sampling equilibria simply by replacing the optimal stopping time τ_i with the myopic one τ_i^M. While appealing for its simplicity, one immediate implication is that myopic sequential sampling equilibria do not generically reach Nash equilibria as the sampling costs vanish. This is because the expected value of sampling information is zero whenever the optimal choices under belief μ_i are still optimal under the posterior $\mu_i \mid y_i$, no matter the realization of the observation. It is therefore immediate that any prior that is sufficiently concentrated around a distribution of opponents' actions for which 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 it is optimal to play such an action with probability 1, myopic sequential sampling will fail to converge to a Nash equilibrium.

References

Alaoui, L., and A. Penta. 2022. "Cost-Benefit Analysis in Reasoning." *Journal of Political Economy* 130 (4): 881–925. 10.1086/718378.

Aliprantis, C. D., and K. C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhiker's Guide.* Springer.

Alós-Ferrer, C., and J. Bruckenmaier. 2021. "Cognitive sophistication and deliberation times." *Experimental Economics* 24 558–592. 10.1007/s10683-020-09672-w.

Alós-Ferrer, C., E. Fehr, and N. Netzer. 2021. "Time Will Tell: Recovering Preferences When Choices Are Noisy." *Journal of Political Economy* 129 (6): 1828–1877. 10.1086/713732.

Arrow, K., D. Blackwell, and M. Girshick. 1949. "Bayes and Minimax Solutions of Sequential Decision Problems." *Econometrica* 17 (3/4): 213–244. 10.2307/1905525.

Aumann, R., and A. Brandenburger. 1995. "Epistemic Conditions for Nash Equilibrium." *Econometrica* 63 (5): 1161–1180. 10.2307/2171725.

Bakkour, A., A. Zylberberg, M. Shadlen, and D. Shohamy. 2018. "Value-based decisions involve sequential sampling from memory." *Working Paper.* 10.1101/269290.

Battigalli, P., M. Gilli, and M. C. Molinari. 1992. "Learning and Covergence to Equilibrium in Repeated Strategic Interactions: An Introductory Survey." *Ricerche Economiche* 46 335–378.

Berk, R. 1966. "Limiting Behavior of Posterior Distributions when the Model is Incorrect." *Annals of Mathematical Statistics* 37 (1): 51–58. 10.1214/aoms/1177699597.

Bhui, R. 2019. "Testing Optimal Timing in Value-Linked Decision Making." *Computational Brain and Behavior* 2 (2): 85–94. 10.1007/s42113-019-0025-9.

Brown, George W. 1951. *Activity Analysis of Production and Allocation.* Chap. XXIV. Iterative Solution of Games by Fictitious Play 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. 10.1257/aer.20140117.

Che, Y.-K., and K. Mierendorff. 2019. "Optimal Dynamic Allocation of Attention." *American Economic Review* 109 (8): 2993–3029. 10.1257/aer.20171000.

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. 10.1007/s101140000098.

Chernoff, H. 1961. "Sequential Tests for the Mean of a Normal Distribution." *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* 1 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): 1–11. 10.1038/s41598-018-21962-1.

Chiong, K., M. Shum, R. Webb, and R. Chen. 2020. "Combining Choice and Response Time Data: A Drift-Diffusion Model of Mobile Advertisements." *Working Paper* 1–30. 10.2139/ssrn.3289386.

Clithero, J. 2018. "Response Times in Economics: Looking Through the Lens of Sequential Sampling Models." *Journal of Economic Psychology* 69 61–86. 10.1016/j.joep.2018.09.008.

Costa-Gomes, M., and G. Weizsäcker. 2008. "Random belief equilibrium in normal form games." *Review of Economic Studies* 75 (3): 729–762. 10.1111/j.1467-937X.2008.00498.x.

Denti, T. 2022. "Unrestricted information acquisition." *Theoretical Economics (Forthcoming)* 1–29.

Denti, T., M. Marinacci, and A. Rustichini. 2022. "Experimental Cost of Information." *American Economic Review* 112 (9): 3106–3123. 10.1257/aer.20210879.

Diaconis, P., and D. Freedman. 1990. "On the Uniform Consistency of Bayes Estimates for Multinomial Probabilities." *Annals of Statistics* 18 (3): 1317–1327. 10.1214/aos/1176347751.

Diaconis, P., and D. Ylvisaker. 1979. "Conjugate Priors for Exponential Families." *The Annals* *of Statistics* 7 (2): 269–281. 10.1214/aos/1176344611.

Esponda, I., and D. Pouzo. 2016. "Berk-Nash Equilibrium: A Framework for Modeling Agents With Misspecified Models." *Econometrica* 84 (3): 1093–1130. 10.3982/ecta12609.

Esteban-Casanelles, T., and D. Gonçalves. 2020. "The Effect of Incentives on Choices and Beliefs in Games: An Experiment." *Working Paper.*

Evans, L. C., and R. F. Gariepy. 2015. *Measure Theory and Fine Properties of Functions.* Chapman and Hall.

Ferguson, T. 2008. *Optimal Stopping and Applications.* UCLA.

Forstmann, B., 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. 10.1146/annurev-psych-122414-033645.

Freedman, D. 1963. "On the Asymptotic Behavior of Bayes' Estimates in the Discrete Case." *Annals of Mathematical Statistics* 34 (4): 1386–1403. 10.1214/aoms/1177703871.

Friedman, E. 2022. "Stochastic Equilibria: Noise in Actions or Beliefs?" *American Economic Journal: Microeconomics* 14 (1): 1–65. 10.1257/mic.20190013.

Friedman, E., and J. Ward. 2022. "Stochastic Choice and Noisy Beliefs in Games: an Experiment." *Working Paper* 1–78. 10.2139/ssrn.4190338.

Friedman, J., and C. Mezzetti. 2005. "Random belief equilibrium in normal form games." *Games and Economic Behavior* 51 (2): 296–323. 10.1016/j.geb.2003.08.004.

Fudenberg, D. 2020. "Learning and Equilibrium Refinements." *Working Paper* 1–14.

Fudenberg, D., and D. Kreps. 1993. "Learning Mixed Equilibria." *Games and Economic Behavior* 5 (3): 320–367. 10.1006/game.1993.1021.

Fudenberg, D., G. Lanzani, and P. Strack. 2021. "Limit Points of Endogenous Misspecified Learning." *Econometrica* 89 (3): 1065–1098. 10.3982/ECTA18508.

Fudenberg, D., and D. Levine. 1993. "Self-Confirming Equilibrium." *Econometrica* 61 (3): 523–545. 10.2307/2951716.

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

Fudenberg, D., and D. Levine. 2009. "Learning and Equilibrium." *Annual Review of Economics* 1 385–419. 10.1146/annurev.economics.050708.142930.

Fudenberg, D., P. Strack, and T. Strzalecki. 2018. "Speed, Accuracy, and the Optimal Timing of Choices." *American Economic Review* 108 (12): 3651–3684. 10.1257/aer.20150742.

Goeree, J., and C. Holt. 2001. "Ten Little Treasures of Game Theory and Ten Intuitive Contradictions." *American Economic Review* 91 (5): 1402–1422. 10.1257/aer.91.5.1402.

Goeree, J., C. Holt, and T. Palfrey. 2005. “Regular Quantal Response Equilibrium.” *Experimental Economics* 8 (4): 347–367. 10.1007/s10683-005-5374-7.

Gold, J., and M. Shadlen. 2007. “The Neural Basis of Decision Making.” *Annual Review of Neuroscience* 30 535–574. 10.1146/annurev.neuro.29.051605.113038.

Hawkins, G., B. Forstmann, E.-J. Wagenmakers, R. Ratcliff, and S. Brown. 2015. “Revisiting the Evidence for Collapsing Boundaries and Urgency Signals in Perceptual Decision-Making.” *Journal of Neuroscience* 35 (6): 2476–2484. 10.1523/jneurosci.2410-14.2015.

Hébert, B., and J. La'O. 2020. “Information Acquisition, Efficiency, and Non-Fundamental Volatility.” *NBER Working Paper* 1–92. 10.3386/w26771.

Hébert, B., and M. Woodford. 2022. “Rational Inattention when Decisions Take Time.” *Working Paper* 1–111.

Hommes, C., and M. Ochea. 2012. “Multiple equilibria and limit cycles in evolutionary games with Logit Dynamics.” *Games and Economic Behavior* 74 (1): 434–441. 10.1016/j.geb.2011.05.014.

Jehiel, P. 2005. “Analogy-based expectation equilibrium.” *Journal of Economic Theory* 123 (2): 81–104. 10.1016/j.jet.2003.12.003.

Jehiel, P. 2021. “Analogy-based expectation equilibrium and related concepts: Theory, applications, and beyond.” *Working Paper* 1–107.

Jehiel, P., and F. Koessler. 2008. “Revisiting games of incomplete information with analogy-based expectations.” *Games and Economic Behavior* 62 (2): 533–557. 10.1016/j.geb.2007.06.006.

Jordan, J. S. 1993. “Three Problems in Learning Mixed-Strategy Nash Equilibria.” *Games and Economic Behavior* 5 (3): 368–386. 10.1006/game.1993.1022.

Kalai, E., and E. Lehrer. 1993. “Rational Learning Leads to Nash Equilibrium.” *Econometrica* 61 (5): 1019–1045. 10.2307/2951492.

Konovalov, A., and I. Krajbich. 2019. “Revealed strength of preference: Inference from response times.” *Judgment and Decision Making* 14 (4): 381–394.

Krajbich, I., D. Lu, C. Camerer, and A. Rangel. 2012. “The attentional drift-diffusion model extends to simple purchasing decisions.” *Frontiers in Psychology* 3 1–18. 10.3389/fpsyg.2012.00193.

Lehrer, E., and T. Wang. 2020. “Strong Stochastic Dominance.” *Working Paper* 1–31.

Massart, P. 1990. “The Tight Constant in the Dvoretzky-Kiefer-Wolfowitz Inequality.” *Annals of Probability* 18 (3): 1269–1283. 10.1214/aop/1176990746.

Matějka, Filip, and Alisdair McKay. 2015. “Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model.” *American Economic Review* 105 (1): 272–298. 10.1257/aer.20130047.

McKelvey, R., and T. Palfrey. 1995. "Quantal Response Equilibrium." *Games and Economic Behavior* 10 (1): 6–38. 10.1006/game.1995.1023.

McKelvey, R., T. Palfrey, and R. 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. 10.1016/S0167-2681(00)00102-5.

Morris, S., and P. Strack. 2019. "The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs." *Working Paper.* 10.2139/ssrn.2991567.

Moscarini, G., and L. Smith. 1963. "The Optimal Level of Experimentation." *Econometrica* 69 (6): 1629–1644. 10.1111/1468-0262.00259.

Nash, J. 1950. *Non-Cooperative Games.* Ph.D. dissertation, Princeton University.

Ochs, J. 1995. "Games with Unique, Mixed Strategy Equilibria: An Experimental Study." *Games and Economic Behavior* 10 (1): 202–217. 10.1006/game.1995.1030.

Osborne, M., and A. Rubinstein. 1998. "Games with Procedurally Rational Players." *American Economic Review* 88 (4): 834–847.

Osborne, M., and A. Rubinstein. 2003. "Sampling equilibrium, with an application to strategic voting." *Games and Economic Behavior* 45 (2): 434–441. 10.1016/s0899-8256(03)00147-7.

Pearce, D. 1984. "Rationalizable Strategic Behavior and the Problem of Perfection." *Econometrica* 52 (4): 1029–1050. 10.2307/1911197.

Ratcliff, R. 1978. "A theory of memory retrieval." *Psychological Review* 85 (2): 59–108. 10.1037/0033-295X.85.2.59.

Ratcliff, R., P. Smith, S. Brown, and G. McKoon. 2016. "Diffusion Decision Model: Current Issues and History." *Trends in Cognitive Sciences* 20 (4): 260–281. 10.1016/j.tics.2016.01.007.

Rubinstein, A. 2007. "Instinctive and Cognitive Reasoning: A Study of Response Times." *The Economic Journal* 117 (523): 1243–1259. 10.1111/j.1468-0297.2007.02081.x.

Rudin, W. 1973. *Functional Analysis.* McGraw-Hill.

Salant, Y., and J. Cherry. 2020. "Statistical Inference in Games." *Econometrica* 88 (4): 1725–1752. 10.3982/ECTA17105.

Shadlen, M., and D. Shohamy. 2016. "Decision Making and Sequential Sampling from Memory." *Neuron* 90 (5): 927–939. 10.1016/j.neuron.2016.04.036.

Shapley, L. S. 1964. *Some Topics in Two-Person Games.* 1–28, Princeton University Press, . 10.1515/9781400882014-002.

Sims, C. 2003. "Implications of rational inattention." *Journal of Monetary Economics* 50 (3): 665–690. 10.1016/S0304-3932(03)00029-1.

Wald, A. 1947. "Foundations of a General Theory of Sequential Decision Functions." *Econometrica* 15 (4): 279–313. 10.2307/1905331.
Yang, M. 2015. "Coordination with flexible information acquisition." *Journal of Economic Theory* 158 (Part B): 721–738. 10.1016/j.jet.2014.11.017.

Appendix A. Omitted Proofs

Proof of Proposition 1. For ease of notation, I'll write $Y_i := A_{-i}$, $P_i := \Delta(Y_i)$, and $\mu_i \in \Delta(P_i)$. I consider general continuous utility functions $u_i : A_i \times P_i \to \mathbb{R}$. Boundedness of V_i and v_i follows immediately from boundedness of u_i. Below I prove the remaining properties.

Claim 1. *For any $\mu_i \in \Delta(P_i)$, and any optimal stopping time τ_i, $\mathbb{P}_{\mu_i}(\tau_i > T) \le 2\|u_i\|_\infty / c_i T$.*

Proof. Let $\|u_i\|_\infty = \max_{(a_i,p_i)\in A_i\times P_i} |u(a_i,p_i)| < \infty$. We then have that, for any $\mu_i \in \Delta(P_i)$ and $T \in \mathbb{N}$,

$$-\|u_i\|_\infty \le V_i(\mu_i) \le \mathbb{P}_{\mu_i}(\tau_i \le T)\|u_i\|_\infty + \mathbb{P}_{\mu_i}(\tau_i > T)(\|u_i\|_\infty - c_i T) \Longrightarrow \mathbb{P}_{\mu_i}(\tau_i > T) \le 2\|u_i\|_\infty / c_i T.$$

□

Claim 2. *v_i is uniformly continuous.*

Proof. $\mathbb{E}_{\mu_i}[u_i(a_i,p_i)]$ is jointly continuous in (a_i,μ_i) with respect to product topology. Let $\mu_{i,n} \to \mu_i$, $a_{i,n} \to a_i$. Note that $\forall \epsilon > 0$, $\exists N$ such that $\forall n \ge N$, $\forall p_i$, $|u_i(a_{i,n},p_i) - u_i(a_i,p_i)| < \epsilon/2$ and $|\mathbb{E}_{\mu_{i,n}}[u_i(a_i,p_i)] - \mathbb{E}_{\mu_i}[u_i(a_i,p_i)]| < \epsilon/2$. Hence,

$$|\mathbb{E}_{\mu_{i,n}}[u_i(a_{i,n},p_i)] - \mathbb{E}_{\mu_i}[u_i(a_i,p_i)]| \le |\mathbb{E}_{\mu_{i,n}}[u_i(a_{i,n},p) - u_i(a_i,p_i)]| + |\mathbb{E}_{\mu_{i,n}}[u_i(a_i,p_i)] - \mathbb{E}_{\mu_i}[u_i(a_i,p_i)]| < \epsilon.$$

Continuity of v_i follows from Berge's maximum theorem and uniform continuity from Heine–Cantor theorem. □

Claim 3. *V_i is uniformly continuous.*

Proof. Let $\mathbb{T}_{i,T}$ denote the set of stopping times $\tau' \in \mathbb{T}_i$ that are bounded above by T and, for every $T \in \mathbb{N}$, $V_{i,T} : \Delta(P_i) \to \mathbb{R}$ be given by

$$V_{i,T}(\mu) := \sup_{\tau' \in \mathbb{T}_{i,T}} \mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i^{\tau'}) - c_i \cdot \tau'].$$

Note that, as $\mathbb{T}_{i,T}$ is finite, it is compact with respect to the discrete topology, and an application of Berge's maximum theorem implies $V_{i,T}$ is continuous.

Note that, for any $\mu_i \in \Delta(P_i)$, $T \in \mathbb{N}$, $0 \le V_i(\mu_i) - V_{i,T}(\mu_i) \le \mathbb{P}_{\mu_i}(\tau_i > T)\|u_i\|_\infty \le 2\|u_i\|_\infty^2/c_i T$. Hence, $\|V_i - V_{i,T}\| \le 2\|u_i\|_\infty^2/c_i T$, and $V_{i,T}$ converges uniformly to V_i. Since for any T, $V_{i,T}$ is in the space of bounded continuous functions $C_b^0(\Delta(P_i))$, which, endowed with the sup-norm is a Banach space, V_i is continuous; by the Heine–Cantor theorem, it is uniformly continuous. □

Claim 4. *v_i, V_i, and $V_{i,T}$ are convex, for any $T \in \mathbb{N}$.*

Proof. This follows since each of these can be seen as the pointwise supremum over a family of convex functions over $\Delta(P_i)$, which is compact with respect to $\|\cdot\|_{LP}$. □

Proof of Lemma 1. By contrapositive, that (2) implies (1) is straightforward. I prove a more general claim that implies the converse. Let $\Sigma_{-i}^{\tau_i} := \{\sigma_{-i} \in \Delta(A_{-i}) \mid \mathbb{P}_{\sigma_{-i}}(\tau_i < \infty) = 1\}$ denote the opponents' distribution actions with respect to which player i's optimal stopping time is finite with probability 1.

Lemma 6. *On $\Sigma_{-i}^{\tau_i}$, b_i is a continuous mapping to $\Delta(A_i)$.*

Proof. Fix a selection of optimal choices $\sigma_i^*(\mu_i') \in \arg\max_{\sigma_i \in \Delta(A_i)} \mathbb{E}_{\mu_i'}[u_i(\sigma_i, \sigma_{-i}')]$. For $t \in \mathbb{N}$, let $b_{i,t} : \Sigma_{-i}^{\tau_i} \to [0,1]^{|A_i|}$ be given by $b_{i,t}(\sigma_{-i}) := \mathbb{E}_{\sigma_{-i}}[1_{\tau_i < t}\sigma_i^*(\mu_i \mid y_i^{\tau_i})]$, and $p_{i,t} : \Sigma_{-i}^{\tau_i} \to [0,1]$ denote $p_{i,t}(\sigma_{-i}) := \mathbb{P}_{\sigma_{-i}}(\tau_i < t)$. Note that, for every $t' \ge t$, $b_{i,t} \le b_{i,t'}$ and $p_{i,t} \le p_{i,t'}$, ensuring pointwise convergence to b_i and 1, respectively. Further, both $b_{i,t}$ and $p_{i,t}$ are continuous.

Since $p_{i,t} \to 1$, by Dino's theorem $p_{i,t}$ converges uniformly on $\Sigma_{-i}^{\tau_i}$. We then have $\|b_i(\sigma_{-i}) - b_{i,t}(\sigma_{-i})\|_1 \le 1 - {}_{i,t}(\sigma_{-i})$, and $b_{i,t}$ also converges uniformly to $b_i\,|_{\Sigma_{-i}^{\tau_i}} : \Sigma_{-i}^{\tau_i} \to [0,1]^{|A_i|}$, ensuring that, on $\Sigma_{-i}^{\tau_i}$, b_i is continuous. To see that $b_i(\sigma_{-i}) \in \Delta(A_i)$ for any $\sigma_{-i} \in \Sigma_{-i}^{\tau_i}$, note that $1 \ge \|b_i(\sigma_{-i})\|_1 \ge \|b_{i,t}(\sigma_{-i})\|_1 \ge p_{i,t}(\sigma_{-i}) \to 1$. □

Proof of Remark 1. Let $\sum_{\ell \in [1..t]} \delta_{y_{i,\ell}^t}/t =: \overline{y_i}^t \in \Delta(A_{-i})$ denote the empirical frequency. When μ_i has full support on $\Delta(A_{-i})$, then there is $\phi : \mathbb{R}_{++} \to \mathbb{R}_{++}$ such that $\inf_{\sigma_{-i} \in \Delta(A_{-i})} \mu(B_\epsilon(\sigma_{-i})) \ge \phi(\epsilon)$. Then, by Diaconis and Freedman (1990), for every $\epsilon > 0$, there is T such that for all $t \ge T$, $\mu_{i,t}(B_{\epsilon/2}(\overline{y_i}^t))/(1 - \mu_{i,t}(B_{\epsilon/2}(\overline{y_i}^t)) \ge \frac{2-\epsilon}{\epsilon}$, which implies $\|\mu_{i,t} - \delta_{\overline{y_i}^t}\|_{LP} \le \epsilon$. Immediately, for every $\epsilon > 0$, there is T such that for all $t \ge T$,

$$\|\mu_{i,t} - \mu_{i,t+1}\|_{LP} \le \|\mu_{i,t} - \delta_{\overline{y_i}^t}\|_{LP} + \|\mu_{i,t+1} - \delta_{\overline{y_i}^{t+1}}\|_{LP} + \|\delta_{\overline{y_i}^t} - \delta_{\overline{y_i}^{t+1}}\|_{LP} \le 2/3\epsilon + 2/t \le \epsilon.$$

As V_i is uniformly continuous (cf. Proposition 1), this implies $\exists \overline{T}_i$ such that $\forall t \geq T$, $\forall y_i^{t+1}$, $V_i(\mu_i \mid y_i^{t+1}) - V_i(\mu_i \mid y_i^t) \leq c_i \Longrightarrow \mathbb{E}_{\mu_i}[V_i(\mu_i \mid y_i^{t+1}) - V_i(\mu_i \mid y_i^t) \mid y_i^t] \leq c_i \Longrightarrow \tau_i \leq \overline{T}_i$.

When the prior does not allow for correlation, $\mu_i = \times_{j \neq i} \mu_{ij}$ and each marginal μ_{ij} uniformly accumulates around the empirical frequency projected on $\Delta(A_j)$. As I is finite, one can similarly obtain a uniform rate of convergence that depends only on t. It is then straightforward to adjust the proof to obtain the result.

Proof of Proposition 2. Let $A_i = \{0,1\}$, $i \in I$; denote the probability that player i chooses action 1 by σ_i. By manner of a continuous-time approximation as in Fudenberg and Levine (1998, Ch. 2), the dynamic system can be written as $\dot{\sigma}_i = b_i(\sigma_j) - \sigma_i$, $i, j \in I$, $i \neq j$. The Jacobian of the dynamic system is given by

$$\begin{pmatrix} -1 & b_i'(\sigma_j) \\ b_j'(\sigma_i) & -1 \end{pmatrix}$$

and its eigenvalues are given by $\lambda = -1 \pm \sqrt{b_i'(\sigma_j) b_j'(\sigma_i)}$, where differentiability of b_i, b_j is ensured by Theorem 4. If $b_i'(\sigma_j) b_j'(\sigma_i) \leq 0$, there is a unique σ such that $b_i(\sigma_j) = \sigma_i$ and, since the real parts of the eigenvalues of the Jacobian matrix are strictly negative, by the Jacobian conjecture on global asymptotic stability — proved to hold on the plane (Chen et al., 2001) — σ is globally asymptotically stable.

In particular, if there is a unique Nash equilibrium, either one player has a dominant strategy (and then $b_i' = 0$ for some player i) or neither does. If some player does not sample, i.e. $\tau_i = 0$ for some player i, then again $b_i' = 0$. If both players sample, we must then have $b_i'(\sigma_j) b_j'(\sigma_i) < 0$. In any of these cases, $b_i'(\sigma_j) b_j'(\sigma_i) \leq 0$.

Proof of Remark 4. Let $c_i' \geq c_i$ and denote V_i' and V_i the value functions associated with c_i' and c_i, respectively. Since $v_i(\mu_i) \leq V_i'(\mu_i) \leq V_i(\mu_i)$, it is immediate that $V_i(\mu_i) = v_i(\mu_i) \Longrightarrow V_i'(\mu_i) = v_i(\mu_i)$. Let τ_i' and τ_i the earliest optimal stopping times associated with c_i' and c_i, respectively. Then

$$\begin{aligned}
\left\{\omega \in \Omega \,\middle|\, \tau_i(\omega) \leq t\right\} &= \left\{\omega \in \Omega \,\middle|\, \exists t' \leq t : \begin{array}{l} V_i(\mu_i \mid y_i^{t'}(\omega)) = v_i(\mu_i \mid y_i^{t'}(\omega)) \\ V_i(\mu_i \mid y_i^{\ell}(\omega)) > v_i(\mu_i \mid y_i^{\ell}(\omega)), \forall \ell < t' \end{array}\right\} \\
&\subseteq \left\{\omega \in \Omega \,\middle|\, \exists t' \leq t : \begin{array}{l} V_i'(\mu_i \mid y_i^{t'}(\omega)) = v_i(\mu_i \mid y_i^{t'}(\omega)) \\ V_i'(\mu_i \mid y_i^{\ell}(\omega)) > v_i(\mu_i \mid y_i^{\ell}(\omega)), \forall \ell < t' \end{array}\right\} = \left\{\omega \in \Omega : \tau_i'(\omega) \leq t\right\}.
\end{aligned}$$

Proof of Lemma 2. Take any $a_i \in A_i \setminus A_i^k$. For $\delta \geq 0$ define

$$\overline{B}_\delta(\Delta(A_{-i}^{k-1})) := \{\sigma_{-i} \in \Delta(A_{-i}) \mid \exists \sigma'_{-i} \in \Delta(A_{-i}^{k-1}) : \|\sigma_{-i} - \sigma'_{-i}\| \leq \delta\}$$

and let

$$h_i^k(\delta) := \max_{\sigma_{-i} \in \overline{B}_\delta(\Delta(A_{-i}^{k-1}))} \left(u_i(a_i, \sigma_{-i}) - \max_{\sigma'_i \in \Delta(A_i)} u_i(\sigma'_i, \sigma_{-i}) \right).$$

Since u_i is continuous and $\delta \mapsto \overline{B}_\delta(\Delta(A_{-i}^{k-1}))$ is a continuous, convex-, compact-, and nonempty-valued correspondence, h_i^k is continuous by Berge's maximum theorem. As $\overline{B}_\delta(\Delta(A_{-i}^{k-1}))$ increases in subset order with δ, h_i^k is nondecreasing.

By definition of k-rationalizability, $h_i^k(0) < 0$. This implies that there is $\delta' > 0$ such that, $\forall \delta \leq \delta'$, $h_i^k(\delta) \leq h_i^k(0)/2 < 0$. Then, for any $\delta < \delta'$, and any $\sigma_{-i} \in B_\delta(\Delta(A_{-i}^{k-1}))$, $u_i(a_i, \sigma_{-i}) - \max_{\sigma'_i \in \Delta(A_i)} u_i(\sigma'_i, \sigma_{-i}) \leq h_i^k(0)/2 < 0$. Finally, observe that $\max_{\sigma_{-i} \in \Delta(A_{-i}) \setminus \Delta(A_{-i}^{k-1})} \max_{\sigma_i \in \Delta(A_i)} u_i(a_i, \sigma_{-i}) - u_i(\sigma_i, \sigma_{-i}) < 2\|u_i\|_\infty$. Let $\epsilon < -h_i^k(0)/(4\|u_i\|_\infty - h_i^k(0))$ and $\mu_i(B_\delta(\Delta(A_{-i}^{k-1}))) > 1 - \epsilon$. It then follows that

$$\begin{aligned}\mathbb{E}_{\mu_i}[u_i(a_i, \sigma_{-i})] - \max_{\sigma_i \in \Delta(A_i)} \mathbb{E}_{\mu_i}[u_i(\sigma_i, \sigma_{-i})] &\leq \mu_i(B_\delta(\Delta(A_{-i}^{k-1})))h_i^k(0)/2 + (1 - \mu_i(B_\delta(\Delta(A_{-i}^{k-1}))))2\|u_i\|_\infty \\ &< (1-\epsilon)h_i^k(0)/2 + \epsilon 2\|u_i\|_\infty < 0.\end{aligned}$$

Proof of Lemma 3. By assumption, $\overline{y}_i^t \in \Delta(A_{-i}^{k-1})$. Since every player's prior has full support, by Diaconis and Freedman (1990), each player's prior concentrates on an δ-ball around the empirical frequency of at a uniform rate. This implies that, for any $\delta, \epsilon > 0$ there is a T such that, for all $t \geq T$ and any $y_i^t \in A_{-i}^{k-1}$, $\mu_{i,t}(B_\delta(\Delta(A_{-i}^{k-1}))) \geq \mu_{i,t}(B_\delta(\delta_{\overline{y}_i^t})) > 1 - \epsilon$.

Auxiliary Results.

Lemma 7. *For any $\epsilon > 0$, $\exists \delta > 0$ such that for any $T \in \mathbb{N}$, $c_i > 0$, $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})] - V_i(\mu_i) \leq (1 - 2\exp(-2T\delta^2))\epsilon/4 + 8\exp(-2T\delta^2)\|u_i\|_\infty + c_i T$.*

Proof. Let $\hat{T}_i$ be the stopping time such that player i stops after T periods and let $\overline{y}_i^T$ denote the empirical frequency of y_i^T, i.e. $\overline{y}_i^T := \sum_{t \in [1..T]} \delta_{y_{i,t}}/T$. Let $\hat{a}_i : \Delta(A_{-i}) \to A_i$ be such that $\hat{a}_i(\mu'_i) \in \arg\max \mathbb{E}_{\mu'_i}[u_i(a_i, \sigma_{-i})]$. Since $u_i(a_i(\sigma_{-i}), \sigma_{-i}) = \max_{a_i \in A_i} u_i(a_i, \sigma_{-i})$ is continuous by Berge's maximum theorem and uniformly so by Heine–Cantor theorem, let $\delta > 0$ be such that for any $\|\sigma_{-i} - \sigma'_{-i}\| < \delta \Longrightarrow |u_i(a_i(\sigma_{-i}), \sigma_{-i}) - u_i(a_i(\sigma'_{-i}), \sigma'_{-i})| + |u_i(a_i(\sigma'_{-i}), \sigma_{-i}) - u_i(a_i(\sigma'_{-i}), \sigma'_{-i})| < \epsilon/4$. Since $\hat{T}_i$ and $\hat{a}_i(\overline{y}_i^{\hat{T}_i})$ are potentially suboptimal stopping time and choices for player i, $V_i(\mu_i) \geq$

$\mathbb{E}_{\mu_i}[u_i(\hat{a}_i(\overline{y}_i^{\hat{T}_i}),\sigma_{-i})-c_i\hat{T}_i]$. For $c_i=k/(4T)>0$,

$$\begin{aligned}\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-V_i(\mu_i)&\leq\mathbb{E}_{\mu_i}[u_i(\hat{a}_i(\sigma_{-i}),\sigma_{-i})-u_i(\hat{a}_i(\overline{y}^{\hat{T}_i}),\sigma_{-i})+c_i\hat{T}_i]\\&=\mathbb{E}_{\mu_i}[u_i(\hat{a}_i(\sigma_{-i}),\sigma_{-i})-u_i(\hat{a}_i(\overline{y}^{\hat{T}_i}),\overline{y}_i^{\hat{T}_i})+u_i(\hat{a}_i(\overline{y}_i^{\hat{T}_i}),\overline{y}_i^{\hat{T}_i})-u_i(\hat{a}_i(\overline{y}_i^{\hat{T}_i}),\sigma_{-i})]+c_iT\\&\leq(1-2\exp(-2T\delta^2))\epsilon/4+8\exp(-2T\delta^2)\|u_i\|_\infty+c_iT,\end{aligned}$$

where the last inequality follows by the Dvoretzky–Kiefer–Wolfowitz–Massart inequality (Massart, 1990), which delivers that, for any $\delta>0$ and σ_{-i}, $\mathbb{P}_{\sigma_{-i}}(\|\sigma_{-i}-\overline{y}_i^T\|>\delta)<2\exp(-2T\delta^2)$. □

Proof of Lemma 4. I first prove the weaker statement:

Claim 5. *Suppose that there is no action a_i that is a best response to all distribution of opponents' actions $\sigma_{-i}\in\Delta(A_{-i})$. Then, for any full support prior $\mu_i\in\Delta(A_{-i})$, there is a sampling cost $c_i>0$ such that the associated earliest optimal stopping time $\tau_i\geq 1$.*

Proof. I start by showing that $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-v_i(\mu_i)>0$. Since there is no action a_i that is a best response to all distribution of opponents' actions $\sigma_{-i}\in\Delta(A_{-i})$, for any μ_i with full support and any $a_i(\mu_i)\in\arg\max_{a_i\in A_i}\mathbb{E}_{\mu_i}[u_i(a_i,\sigma_{-i})]$, there is $\sigma'_{-i}\in\Delta(A_{-i})$ such that $v_i(\delta_{\sigma'_{-i}})=\max_{a_i\in A_i}u_i(a_i,\sigma'_{-i})>u_i(a_i(\mu_i),\sigma'_{-i})$. By continuity, there is an $\epsilon>0$ such that for all $\sigma''_{-i}\in B_\epsilon(\sigma'_{-i})$, $v_i(\delta_{\sigma''_{-i}})-u_i(a_i(\mu_i),\sigma''_{-i})\geq(v_i(\delta'_{\sigma_{-i}})-u_i(a_i(\mu_i),\sigma'_{-i}))/2>0$. Hence, $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-v_i(\mu_i)\geq\mu_i(B_\epsilon(\delta_{\sigma'_{-i}}))(v_i(\delta'_{\sigma_{-i}})-u_i(a_i(\mu_i),\sigma'_{-i}))/2>0$.

Next, we show that, if c_i is low enough, $V_i(\mu_i)>v_i(\mu_i)$ (implying $\tau_i\geq 1$) by proving that $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-V_i(\mu_i)<\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-v_i(\mu_i)=:k$. By Lemma 7, for any $\epsilon>0$, $\exists\delta>0$ such that for any $T\in\mathbb{N}$, $c_i>0$, $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-V_i(\mu_i)\leq(1-2\exp(-2T\delta^2))\epsilon/4+8\exp(-2T\delta^2)\|u_i\|_\infty+c_iT$. Letting $\epsilon=k$, $c_i=k/(4T)$, we have $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-V_i(\mu_i)\leq k/2+8\exp(-2T\delta^2)\|u_i\|_\infty$. It is then straightforward to see that, for T large enough (and $c_i>0$ small enough), $\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-V_i(\mu_i)<k=\mathbb{E}_{\mu_i}[v_i(\delta_{\sigma_{-i}})]-v_i(\mu_i)$, proving the result.

□

Now consider all possible posteriors following any possible t realizations, $\{\mu_i\mid y_i^t,y_i^t\in\cup_{t\in[1..T]}Y^t\}$, observing that this is a finite set. Given the nature of the information process, since the prior μ_i has full support, so do the posterior beliefs. By the above claim, for each $\mu_i\mid y_i^t$, there is a cost $c_i>0$ such that player i would find it optimal to acquire at least one more signal. Taking the lowest of all such costs implies that under such cost, player i would deem it optimal to acquire at least T signals, concluding the proof.

Proof of Theorem 4. Label player i's actions so that $u_i(1,\sigma_{-i}) - u_i(0,\sigma_{-i})$ is increasing in σ_{-i}.

Proof of Theorem 4(i). I first prove a more general comparative statics result:

Proposition 7. *$\mathbb{P}_{\sigma_{-i}}(a_i \in A_i^*(\mu_i \mid y_i^{\tau_i})$ and $\tau_i \le t)$ is increasing with respect to $\ge_{a_i}$, and $\mathbb{P}_{\sigma_{-i}}(a_i \notin A_i^*(\mu_i \mid y_i^{\tau_i})$ and $\tau_i \le t)$ is decreasing with respect to $\ge_{a_i}$, for any $t \in \mathbb{N}$ and $\sigma_{-i} \in \Delta(A_{-i})$.*

Proof. Let $\tilde{u}_i \ge_{a_i} u_i$ and denote the respective (i) value functions, (ii) earliest optimal stopping times, (iii) optimal choices at given beliefs, and (iv) selections of optimal choices by (i) $\tilde{V}_i$ and V_i, (ii) $\tilde{\tau}_i$ and τ_i, (iii) $\tilde{A}_i^*$ and A_i^*, and (iv) $\tilde{\sigma}_i^*$ and σ_i^*, respectively. With $g := \tilde{u}_i(a_i,\cdot) - u_i(a_i,\cdot)$, by definition we obtain

$$V_i(\mu_i) + \mathbb{E}_{\mu_i}[\tilde{\sigma}_i^*(\mu_i \mid y_i^{\tilde{\tau}_i})(a_i) g(\sigma'_{-i})] \ge \tilde{V}_i(\mu_i) \ge V_i(\mu_i) + \mathbb{E}_{\mu_i}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})(a_i) g(\sigma'_{-i})] \ge V_i(\mu_i)$$

Lemma 8. *$V_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a_i,\sigma'_{-i})] \Longrightarrow \tilde{V}_i(\mu_i) = \mathbb{E}_{\mu_i}[\tilde{u}_i(a_i,\sigma'_{-i})]$ and, for $a'_i \ne a_i$ $\tilde{V}_i(\mu_i) = \mathbb{E}_{\mu_i}[\tilde{u}_i(a'_i,\sigma'_{-i})] \Longrightarrow V_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i})]$.*

Proof. Since $\mathbb{E}_{\mu_i}[\tilde{\sigma}_i^*(\mu_i \mid y_i^{\tilde{\tau}_i})(a_i) g(\sigma'_{-i})] \le \mathbb{E}_{\mu_i}[g(\sigma'_{-i})]$, if $V_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(a_i,\sigma'_{-i})]$, then $\mathbb{E}_{\mu_i}[\tilde{u}_i(a_i,\sigma'_{-i})] = \mathbb{E}_{\mu_i}[u_i(a_i,\sigma'_{-i})] + \mathbb{E}_{\mu_i}[g(\sigma'_{-i})] \ge V_i(\mu_i) + \mathbb{E}_{\mu_i}[\tilde{\sigma}_i^*(\mu_i \mid y_i^{\tilde{\tau}_i})(a_i) g(\sigma'_{-i})] \ge \tilde{V}_i(\mu_i) \ge \mathbb{E}_{\mu_i}[\tilde{u}_i(a_i,\sigma'_{-i})]$. Moreover, if $\tilde{V}_i(\mu_i) = \mathbb{E}_{\mu_i}[\tilde{u}_i(a'_i,\sigma'_{-i})] = \mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i})]$, then $\mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i})] = \tilde{V}_i(\mu_i) \ge V_i(\mu_i) \ge \mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i})]$. □

Note that, by the contrapositive of Lemma 8, $\tilde{V}_i(\mu_i) > \mathbb{E}_{\mu_i}[\tilde{u}_i(a_i,\sigma'_{-i})] \Longrightarrow V_i(\mu_i) > \mathbb{E}_{\mu_i}[u_i(a_i,\sigma'_{-i})]$ and $V_i(\mu_i) > \mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i})] \Longrightarrow \tilde{V}_i(\mu_i) > \mathbb{E}_{\mu_i}[\tilde{u}_i(a'_i,\sigma'_{-i})]$ for $a'_i \ne a_i$. This implies

$$\begin{aligned}
&\left\{\omega \in \Omega \,\middle|\, \tau_i(\omega) \le t \text{ and } a_i \in A_i^*(\mu_i \mid y_i^{\tau_i}(\omega))\right\} \\
&= \left\{\omega \in \Omega \,\middle|\, \exists t' \le t: \begin{array}{l} V_i(\mu_i \mid y_i^{t'}(\omega)) = \mathbb{E}_{\mu_i}[u_i(a_i,\sigma'_{-i}) \mid y_i^{t'}(\omega)] \\ V_i(\mu_i \mid y_i^{\ell}(\omega)) > \mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i}) \mid y_i^{t'}(\omega)],\ \forall \ell < t', \forall a'_i \end{array}\right\} \\
&\subseteq \left\{\omega \in \Omega \,\middle|\, \exists t' \le t: \begin{array}{l} \tilde{V}_i(\mu_i \mid y_i^{t'}(\omega)) = \mathbb{E}_{\mu_i}[\tilde{u}_i(a_i,\sigma'_{-i}) \mid y_i^{t'}(\omega)] \\ \tilde{V}_i(\mu_i \mid y_i^{\ell}(\omega)) > \mathbb{E}_{\mu_i}[u_i(a'_i,\sigma'_{-i}) \mid y_i^{t'}(\omega)],\ \forall \ell < t', \forall a'_i \ne a_i \end{array}\right\} \\
&= \left\{\omega \in \Omega \,\middle|\, \tilde{\tau}_i(\omega) \le t \text{ and } a_i \in \tilde{A}_i^*(\mu_i \mid y_i^{\tilde{\tau}_i}(\omega))\right\}
\end{aligned}$$

and

$$
\begin{aligned}
&\left\{\omega \in \Omega \,\middle|\, \tilde{\tau}_i(\omega) \le t \text{ and } a_i \notin \tilde{A}_i^*(\mu_i \mid y_i^{\tilde{\tau}_i}(\omega))\right\} \\
&= \left\{\omega \in \Omega \,\middle|\, \exists t' \le t: \begin{array}{l} \tilde{V}_i(\mu_i \mid y_i^{t'}(\omega)) = \tilde{v}_i(\mu_i \mid y_i^{t'}(\omega)) > \mathbb{E}_{\mu_i}[\tilde{u}_i(a_i, \sigma'_{-i}) \mid y_i^{t'}(\omega)] \\ \tilde{V}_i(\mu_i \mid y_i^{\ell}(\omega)) > \mathbb{E}_{\mu_i}[u_i(a'_i, \sigma'_{-i}) \mid y_i^{t'}(\omega)], \forall \ell < t', \forall a'_i \end{array}\right\} \\
&\subseteq \left\{\omega \in \Omega \,\middle|\, \exists t' \le t: \begin{array}{l} V_i(\mu_i \mid y_i^{t'}(\omega)) = v_i(\mu_i \mid y_i^{t'}(\omega)) > \mathbb{E}_{\mu_i}[u_i(a_i, \sigma'_{-i}) \mid y_i^{t'}(\omega)] \\ V_i(\mu_i \mid y_i^{\ell}(\omega)) > \mathbb{E}_{\mu_i}[u_i(a_i, \sigma'_{-i}) \mid y_i^{t'}(\omega)], \forall \ell < t' \end{array}\right\} \\
&= \left\{\omega \in \Omega \,\middle|\, \tau_i(\omega) \le t \text{ and } a_i \notin A_i^*(\mu_i \mid y_i^{\tau_i}(\omega))\right\}.
\end{aligned}
$$

□

The above, together with Lemma 5 — which is proved independently from Theorem 4(i) — delivers the result. I note Proposition 7 further implies

Corollary 1. *$\mathbb{P}_{\sigma_{-i}}(a_i \in$ (resp. $\notin) A_i^*(\mu_i \mid y_i^{\tau_i}) \mid \tau_i \le t)$ is increasing (resp. decreasing) with respect to $\ge_{a_i}$, $\forall \sigma_{-i} \in \Delta(A_{-i})$ and $t \in \mathbb{N}$.*

Proof of Theorem 4(ii). For $a_i \in \{0,1\}$, define (i) $u_i^{a_i}(a'_i, \sigma_{-i}) := u_i(a'_i, \sigma_{-i}) - u_i(1-a_i, \sigma_{-i})$, (ii) $v_i^{a_i}(\mu_i) := \max_{a'_i \in A_i} \mathbb{E}_{\mu_i}[u_i^{a_i}(a'_i, \sigma_{-i})]$, and (iii) $V_i^{a_i}(\mu_i) := \sup_{t_i \in \mathbb{T}_i} \mathbb{E}_{\mu_i}[v_i^{a_i}(\mu_i \mid y_i^{t_i}) - c_i \cdot t_i]$. Note that, by definition, $u_i^{a_i}(1-a_i, \sigma_{-i}) = 0$. Moreover, $v_i^{a_i}(\mu_i) = v_i(\mu_i) - \mathbb{E}_{\mu_i}[u_i(1-a_i, \sigma_{-i})]$, which also implies that $V_i^{a_i}(\mu_i) = V_i(\mu_i) - \mathbb{E}_{\mu_i}[u_i(1-a_i, \sigma_{-i})]$. A useful property of $V_i^{a_i}$ is as follows:

Lemma 9. *For any $\mu'_i \ge_{SSD} \mu_i$ $V_i^1(\mu'_i) \ge V_i^1(\mu_i)$ and $V_i^0(\mu'_i) \le V_i^0(\mu_i)$.*

Proof. Let $B_i^{a_i} : \mathscr{C}^0(\Delta(\Delta(A_{-i}))) \to \mathscr{C}^0(\Delta(\Delta(A_{-i})))$ be such that $B_i^{a_i}(w)(\mu_i) := \max\{v_i^{a_i}(\mu_i), \mathbb{E}_{\mu_i}[w(\mu_i \mid y_i)] - c_i\}$. As argued in Section 2, $V_i^{a_i}$ is a fixed-point of $B_i^{a_i}$. Moreover, by Remark 1, there is a finite $n \in \mathbb{N}$, such that $V_i^{a_i} = B_i^{a_i(n)}(v_i^{a_i})$, where $B_i^{a_i(1)} = B_i^{a_i}$ and, for $n \ge 1$, $B_i^{a_i(n+1)} = B_i^{a_i} \circ B_i^{a_i(n)}$.

Note that v_i^1 (resp. v_i^0 is increasing (resp. decreasing) in $\ge_{SSD}$. If $w \in \mathscr{C}^0(\Delta(\Delta(A_{-i})))$ is increasing in $\ge_{SSD}$, then so is $B_i^1(w)$ — a symmetric argument applies to B_i^0. To see this, note that

$$
\begin{aligned}
B_i^1(w)(\mu'_i) &= \max\{v_i^1(\mu'_i), \mathbb{E}_{\mu'_i}[\sigma_{-i}]w(\mu'_i \mid 1) + \mathbb{E}_{\mu'_i}[1-\sigma_{-i}]w(\mu'_i \mid 0) - c_i\} \\
&\ge \max\{v_i^1(\mu_i), \mathbb{E}_{\mu'_i}[\sigma_{-i}]w(\mu_i \mid 1) + \mathbb{E}_{\mu'_i}[1-\sigma_{-i}]w(\mu_i \mid 0) - c_i\} \\
&\ge \max\{v_i^1(\mu_i), \mathbb{E}_{\mu_i}[\sigma_{-i}]w(\mu_i \mid 1) + \mathbb{E}_{\mu_i}[1-\sigma_{-i}]w(\mu_i \mid 0) - c_i\} = B_i^1(w)(\mu_i),
\end{aligned}
$$

where the first inequality follows from monotonicity of v_i^1 with respect to $\ge_{SSD}$, by monotonicity of w and the fact that $\mu'_i \ge_{SSD} \mu_i \implies \mu'_i|y_i \ge_{SSD} \mu_i|y_i$ for $y_i \in \{0,1\}$, and the second because

$\mathbb{E}_{\mu_i'}[\sigma_{-i}] \geq \mathbb{E}_{\mu_i}[\sigma_{-i}]$ (by FOSD) and, as can be shown, $\mu_i|1 \geq_{SSD} \mu_i|0$, implying $w(\mu_i|1) \geq w(\mu_i|0)$. □

Lemma 9 implies:

Corollary 2. *Let* $\mu_i' \geq_{SSD} \mu_i$. $V_i(\mu_i' \mid y_i^t) = \mathbb{E}_{\mu_i'}[u_i(0,\sigma_{-i}) \mid y_i^t] \implies 0 = V_i^1(\mu_i' \mid y_i^t) \geq V_i^1(\mu_i \mid y_i^t) \geq 0 \implies V_i(\mu_i \mid y_i^t) = \mathbb{E}_{\mu_i}[u_i(0,\sigma_{-i}) \mid y_i^t]$ *and* $V_i(\mu_i \mid y_i^t) = \mathbb{E}_{\mu_i}[u_i(1,\sigma_{-i}) \mid y_i^t] \implies 0 = V_i^0(\mu_i \mid y_i^t) \geq V_i^0(\mu_i' \mid y_i^t) \geq 0 \implies V_i(\mu_i' \mid y_i^t) = \mathbb{E}_{\mu_i'}[u_i(1,\sigma_{-i}) \mid y_i^t]$.

In order to conclude the proof of Theorem 4(ii), let τ_i and τ_i' denote the earliest optimal stopping times associated with μ_i and μ_i'. Then, by Corollary 2, $\left\{\omega \in \Omega \,\middle|\, \tau_i(\omega) \leq t \text{ and } 1 \in (\text{resp.} \notin) A_i^*(\mu_i \mid y_i^{\tau_i}(\omega))\right\} \subseteq$ (resp. $\supseteq$) $\left\{\omega \in \Omega \,\middle|\, \tau_i'(\omega) \leq t \text{ and } 1 \in (\text{resp.} \notin) A_i^*(\mu_i' \mid y_i^{\tau_i'}(\omega))\right\}$.

Proof of Theorem 4(iii). Let

$$\mathcal{N} := \left\{ n \in \mathbb{N}_0^2 \mid \exists y_i^t : \text{(i) } t = n_0 + n_1, \;\; \text{(ii)} \sum_{\ell \in [1..t]} y_{i,\ell} = n_1, \;\; \text{and} \;\; \text{(iii) } \forall \ell \leq t, V_i(\mu_i \mid y_i^\ell) > v_i(\mu_i \mid y_i^\ell) \right\},$$

and, for $j \in \{0,1\}$, let $\mathcal{N}_j := \{n \in \mathbb{N}_0^2 \mid n - (j, 1-j) \in \mathcal{N}\}$. Note that, if $n \in \mathcal{N}_j$, then there is some sequence y_i^t satisfying $t = n_0 + n_1$, $\sum_{\ell \in [1..t]} y_{i,\ell} = n_1$, and along which player i decides to keep sampling every period (according to τ_i), i.e. $V_i(\mu_i \mid y_i^\ell) > v_i(\mu_i \mid y_i^\ell)$ for all $\ell < t$, and decides to stop at y_i^t and take action j — a consequence of Lemma 5. Let $T_i := \sup \operatorname{supp}\{\tau_i\}$ (where supp is defined with respect to μ_i). By Remark 1, $T_i < \infty$ and thus, $\forall n \in \mathcal{N}$, $n_0 + n_1 < T_i$ and $\mathcal{N}$ is finite. Below I implicitly rely on the fact that, if $(n_0, n_1), (n_0', n_1') \in \mathcal{N}_j$ and $n_{1-j}' > n_{1-j}$, then $n_j' \geq n_j$, which is implied by Corollary 2.

I recursively define the probability of stopping and choosing action 1. Define the asymmetric part of a linear order on $\mathcal{N}$ given by $n \rhd n'$ if and only if $n_1' > n_1$ or $n_1' = n_1$ and $n_0' > n_0$. Let $p : \mathcal{N} \times [0,1] \to [0,1]$ be given by $p(n;\sigma_{-i}) := \sigma_{-i}$ if $n + (0,1) \in \mathcal{N}_1$ and $p(n;\sigma_{-i}) := \sigma_{-i} p(n + (0,1);\sigma_{-i}) + (1 - \sigma_{-i}) p(n + (1,0);\sigma_{-i})$ if otherwise; p can be recursively defined on $n \in \mathcal{N}$ increasing with respect to $\rhd$. Extend p to $n \in \mathcal{N}_j$ by letting $p(n;\sigma_{-i}) = j$ if $n \in \mathcal{N}_j$. Note that $p((0,0);\sigma_{-i}) = \mathbb{P}_{\sigma_{-i}}(1 = A_i(\mu_i \mid y_i^{\tau_i})) = \mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})]$.

I now show by induction that, if $(0,0) \in \mathcal{N}$, $p((0,0);\bullet)$ is $\mathscr{C}^\infty$ and strictly increasing. Note that for $n : n' \rhd n$ for all $n' \neq n$ in $\mathcal{N}$, $n_0 + n_1 = T_i - 1$ and $p(n;\sigma_{-i}) = \sigma_{-i}$ is $\mathscr{C}^\infty$ and strictly increasing in σ_{-i} and $0 = p(n + (0,1);\sigma_{-i}) < p(n + (0,1);\sigma_{-i}) = 1$, $\forall \sigma_{-i} \in [0,1)$.

Suppose that, for all $n' \in \mathcal{N} : n \rhd n'$, $p(n';\bullet)$ is $\mathscr{C}^\infty$ and strictly increasing, and that $p(n' +$

$(1,0);\bullet) < p(n'+(0,1);\bullet)$. As $p(n;\sigma_{-i}) = \sigma_{-i}p(n+(0,1);\sigma_{-i}) + (1-\sigma_{-i})p(n+(1,0);\sigma_{-i})$ and $n \rhd n+(1,0), n+(0,1)$, then $p(n+(0,1);\bullet)$ and $p(n+(1,0);\bullet)$ are $\mathscr{C}^\infty$, strictly increasing, and satisfy $p(n+(1,0);\bullet) < p(n+(0,1);\bullet)$. Then $p(n;\bullet) \in \mathscr{C}^\infty$ and $\frac{\partial}{\partial\sigma_{-i}}p(n;\sigma_{-i}) = p(n+(0,1);\sigma_{-i}) - p(n+(1,0);\sigma_{-i}) + \sigma_{-i}\frac{\partial}{\partial\sigma_{-i}}p(n+(0,1);\sigma_{-i}) + (1-\sigma_{-i})\frac{\partial}{\partial\sigma_{-i}}p(n+(1,0);\sigma_{-i}) > 0$ for all $\sigma_{-i} \in [0,1)$.

To obtain that $\mathbb{P}_{\sigma_{-i}}(1 = A_i^*(\mu_i \mid y_i^{\tau_i})$ and $\tau_i \le t) = \mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})1_{\tau_i \le t}]$ is also $\mathscr{C}^\infty$ and strictly increasing in $\sigma_{-i} \in [0,1)$ for $t \ge n_1(0)$, it is necessary to restrict $\mathscr{N}$. Define $\underline{n}_1(m) := n_1$ if $(m,n_1) \in \mathscr{N}_1$ and let $\mathscr{N}^t := \{n \in \mathscr{N} \mid n_0 + \underline{n}_1(n_0) \le t\}$, $\mathscr{N}_j^t := \{n \mid n-(1-j,j) \in \mathscr{N}\}$. Let $p^t : \mathscr{N} \times [0,1] \to [0,1]$ be given by $p^t(n;\sigma_{-i}) := \sigma_{-i}$ if $n+(0,1) \in \mathscr{N}_1^t$ and $p^t(n;\sigma_{-i}) := \sigma_{-i}p(n+(0,1);\sigma_{-i}) + (1-\sigma_{-i})p(n+(1,0);\sigma_{-i})$ if otherwise. Extend p^t to $n \in \mathscr{N}_j^t$ by letting $p^t(n;\sigma_{-i}) = j$ if $n \in \mathscr{N}_j^t$. An analogous inductive argument applied to p^t delivers the result. For $t < n_1(0)$, $\mathbb{P}_{\sigma_{-i}}(1 = A_i^*(\mu_i \mid y_i^{\tau_i})$ and $\tau_i \le t) = 0$, for all $\sigma_{-i} \in [0,1]$.

Proof of Lemma 5. Label player i's actions so that $u_i(1,\sigma_{-i}) - u_i(0,\sigma_{-i})$ is increasing in σ_{-i}. If, (i) $u_i(1,\sigma_{-i}) - u_i(0,\sigma_{-i}) \le 0 \ \forall \sigma_{-i}$ or (ii) $u_i(1,\sigma_{-i}) - u_i(0,\sigma_{-i}) \ge 0 \ \forall \sigma_{-i}$, then $\tau_i = 0$. Suppose then that $\tau_i > 0$ and note $a_{-i} = \arg\max_{a_i \in A_i} u_i(a_i, a_{-i})$. Let $V_i(\mu_i) > v_i(\mu_i)$ and $V_i(\mu_i \mid a_{-i}) = \mathbb{E}_{\mu_i}[u_i(1-a_{-i},\sigma_{-i}) \mid a_{-i}]$ and observe that, from Lemma 9, $0 = V_i^{a_{-i}}(\mu_i \mid a_{-i}) \ge V_i^{a_{-i}}(\mu_i) \ge 0$, which implies $V_i(\mu_i) = \mathbb{E}_{\mu_i}[u_i(1-a_{-i},\sigma_{-i})] \le v_i(\mu_i)$, a contradiction.

Proof of Proposition 3. Note that in a binary action game there are multiple unique Nash equilibria if and only if for any $i \in I$, $u_i(1,1) - u_i(0,1), u_i(0,0) - u_i(1,0) > 0$. This implies that, for both players, both actions are undominated and so, by Lemma 4, players sample at least once whenever the sampling costs are sufficiently low. By Lemma 5, if $\sigma_{-i} = 1$ (=0) and $\tau_i > 0$, then $\mathbb{E}_{\sigma_{-i}}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})] = 1$ (=0). Hence, (0,0) and (1,1) are both Nash equilibria and sequential sampling equilibria.

Recall that, by assumption, $u_i(1,\sigma_{-i}) - u_i(0,\sigma_{-i})$ is strictly monotone. If there is a unique Nash equilibrium, either (i) both players have a weakly dominant action; (ii) one player has both actions undominated, and the other has a weakly dominant action; or (iii) both players have undominated actions. In (i), uniqueness of a sequential sampling equilibrium follows as both players always choose their weakly dominant action. In (ii), the player with the weakly dominant action chooses it with probability 1, and the opponent, whenever they sample at least once, by Lemma 5, will choose the best response to the weakly dominant action with probability 1, thus entailing a unique sequential sampling equilibrium. In (iii), uniqueness of a Nash equilibrium implies a payoff structure akin to a matching pennies game: one player, i, wants to match the other, i.e. $u_i(1,\sigma_j)-u_i(0,\sigma_j)$ is strictly increasing, and their opponent j seeks to mismatch, i.e. $u_j(1,\sigma_i)-u_j(0,\sigma_i)$ is strictly decreasing. Consequently, by Theorem 4, $\sigma_j \mapsto b_i(\sigma_j)=\mathbb{E}_{\sigma_j}[\sigma_i^*(\mu_i \mid y_i^{\tau_i})] \in [0,1]$ is increasing, and $\sigma_i \mapsto b_j(\sigma_i)=\mathbb{E}_{\sigma_i}[\sigma_j^*(\mu_j \mid y_j^{\tau_j})] \in [0,1]$ is decreasing, both are continuous, and therefore their graph has a unique intersection. Lemma 5 clarifies that the sequential sampling equilibrium is unaffected by the selection of tie-breaking σ_i^*, σ_j^*.

Proof of Proposition 4. Let $B(\sigma_{-i},t)$ to denote a Beta distribution with parameters $\alpha,\beta \geq 0$, such that $t=\alpha+\beta>0$ and $\sigma_{-i}=\alpha/t$, with the convention that $B(1,t)$ and $B(0,t)$ correspond to Dirac measures on 1 and 0, respectively. Label player i's actions so that $u_i(1,\sigma_{-i})-u_i(0,\sigma_{-i})$ is increasing in σ_{-i}.

Let $V_i^{a_i}$ be as defined in the proof of Lemma 9. I show the following properties of $V_i^{a_i}$:

Lemma 10. *$V_i^1(B(\sigma_{-i},t))$ (resp. $V_i^0(B(\sigma_{-i},t))$) is (1) increasing (resp. decreasing) in σ_{-i}; (2) convex in σ_{-i}; (3) decreasing in t; (4) continuous in (σ_{-i},t).*

Proof. (1) follows immediately from Lemma 9, since for $\sigma_{-i}>\sigma'_{-i}$, $B(\sigma_{-i},t) \geq_{SSD} B(\sigma'_{-i},t)$.

For (2), note that $V_i^1(B(\sigma_{-i},t))=V_i^0(B(\sigma_{-i},t))+u_i(1,\sigma_{-i})-u_i(0,\sigma_{-i})$, and thus it suffices to show convexity of $V_i^1(B(\sigma_{-i},t))$ in σ_{-i}. Let $z(\sigma_{-i},t)$ be a random variable that delivers $\frac{1}{t+1}(t\sigma_{-i}+1)$ with probability σ_{-i}, and $\frac{1}{t+1}(t\sigma_{-i}+0)$ with probability $1-\sigma_{-i}$. Note that, for any $w:[0,1]\times\mathbb{R}_{++}$ that is increasing and convex in the first argument, for any $\sigma_{-i}\geq\sigma'_{-i}$, $\lambda\in(0,1)$, and $\sigma''_{-i}:=\lambda\sigma_{-i}+(1-\lambda)\sigma'_{-i}$, straightforward algebra shows that $\mathbb{E}[w(z(\sigma''_{-i},t)] \leq \lambda\mathbb{E}[w(z(\sigma_{-i},t)]+(1-\lambda)\mathbb{E}[w(z(\sigma'_{-i},t)]$.

I observe that, by definition, $v_i^1(B(\sigma_{-i},t))=\max_{a_i\in A_i} u_i(a_i,\sigma_{-i})-u_i(0,\sigma_{-i})$, which is increasing and convex in σ_{-i} and invariant with respect to t.

Now I show that if $w(B(\sigma_{-i},t))$ is increasing and convex in σ_{-i}, so is $B_i^1(w)(B(\sigma_{-i},t))$:

$$B_i^1(w)(B(\sigma''_{-i},t))=\max\left\{v_i^1(B(\sigma''_{-i},t)),\mathbb{E}[w(z(\sigma''_{-i},t+1))]-c_i\right\}$$
$$\leq \max\{\lambda v_i^1(B(\sigma_{-i},t))+(1-\lambda)v_i^1(B(\sigma'_{-i},t)),\lambda\mathbb{E}[w(z(\sigma_{-i},t+1))]+(1-\lambda)\mathbb{E}[w(z(\sigma'_{-i},t+1))]-c_i\}$$
$$\leq \lambda B_i^1(w)(B(\sigma_{-i},t))+(1-\lambda)B_i^1(w)(B(\sigma'_{-i},t)).$$

By similar arguments as in Lemma 9 — V_i^1 is a fixed point of B_i^1 which can be obtained by applying the n-th composition of B_i^1 with itself to v_i^1 — we have that $V_i^1(B(\sigma_{-i},t))$ is convex in σ_{-i}, and thus so is $V_i(B(\sigma_{-i},t))$.

For (3), note that $\mathbb{E}[z(\sigma_{-i},t)]=\sigma_{-i}$ and that, for $t<t'$, $z(\sigma_{-i},t)$ is a mean-preserving spread of $z(\sigma_{-i},t')$; hence for any convex function f, $\mathbb{E}[f(z(\sigma_{-i},t))]\geq\mathbb{E}[f(z(\sigma_{-i},t'))]$. Take any w such that

$w(B(\sigma_{-i}, t))$ is convex in σ_{-i} and decreasing in t. Then, for any $t' > t$,

$$B_i(w)(B(\sigma_{-i}, t') = \max\left\{v_i(B(\sigma_{-i}, t')), \mathbb{E}[w(B(z(\sigma_{-i}, t'), t'))] - c_i\right\}$$
$$= \max\left\{v_i(B(\sigma_{-i}, t)), \mathbb{E}[w(B(z(\sigma_{-i}, t'), t'))] - c_i\right\} \leq \max\left\{v_i(B(\sigma_{-i}, t)), \mathbb{E}[w(B(z(\sigma_{-i}, t'), t))] - c_i\right\}$$
$$\leq \max\{v_i(B(\sigma_{-i}, t)), \mathbb{E}[w(B(z(\sigma_{-i}, t), t))] - c_i\} = B_i(w)(B(\sigma_{-i}, t)).$$

By the same argument as before, $V_i(B(\sigma_{-i}, t))$ is then decreasing in t.

Finally, (4) follows immediately from the fact that for any $(\sigma^n_{-i}, t^n) \to (\sigma_{-i}, t)$, $B(\sigma^n_{-i}, t^n) \to B(\sigma_{-i}, t)$ (with respect to $\|\cdot\|_{LP}$), and therefore by Proposition 1, $V_i(B(\sigma_{-i}, t))$ is continuous in (σ_{-i}, t). □

By Corollary 2, if $V_i(B(\sigma_{-i}, t)) = u_i(a_i, \sigma_{-i})$ for $a_i = 1$ (resp. $a_i = 0$), then $V_i(B(\sigma'_{-i}, t)) = u_i(a_i, \sigma'_{-i})$ for any $\sigma'_{-i} \geq \sigma_{-i}$ (resp. $\leq$), since $B(\sigma'_{-i}, t) \geq_{SSD} B(\sigma_{-i}, t)$. Define $\overline{\sigma}_{-i}(t) := \min\{\sigma_{-i} \in [0,1] \mid V_i(B(\sigma_{-i}, t)) = u_i(1, \sigma_{-i})\}$ and $\underline{\sigma}_{-i}(t) := \max\{\sigma_{-i} \in [0,1] \mid V_i(B(\sigma_{-i}, t)) = u_i(0, \sigma_{-i})\}$.

That $\overline{\sigma}_{-i}(t)$ is continuous and decreasing in t follows from continuity of $V_i(B(\sigma_{-i}, t))$ in (σ_{-i}, t), continuity of $u_i(1, \sigma_{-i})$ in σ_{-i}, and the fact that $V_i(B(\sigma_{-i}, t))$ is decreasing in t and $u_i(1, \sigma_{-i}) - u_i(0, \sigma_{-i})$ is increasing in σ_{-i}. An analogous argument applies to show that $\underline{\sigma}_{-i}(t)$ is continuous and increasing in t.

Finally, I show that both these functions converge to $\tilde{\sigma}_{-i}$. To see this, note that for μ_i given by $B(\sigma_{-i}, t)$, simple algebra shows that $\mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i)] - v_i(\mu_i) = \mathbb{E}[v_i(z(\sigma_{-i}, t))] - v_i(\sigma_{-i})$ is maximized for any t at $\sigma_{-i} = \tilde{\sigma}_{-i} : \tilde{\sigma}_{-i} u_i(1,1) + (1 - \tilde{\sigma}_{-i}) u_i(1,0) = \tilde{\sigma}_{-i} u_i(0,1) + (1 - \tilde{\sigma}_{-i}) u_i(0,0)$, with maximum value $\frac{1}{t+1}(u_i(1,1) - u_i(0,1) + u_i(0,0) - u_i(1,0))(1 - \tilde{\sigma}_{-i})\tilde{\sigma}_{-i}$. Therefore, it is always optimal to keep sampling at belief $B(\tilde{\sigma}_{-i}, t)$ if $t < T := (u_i(1,1) - u_i(0,1) + u_i(0,0) - u_i(1,0))(1 - \tilde{\sigma}_{-i})\tilde{\sigma}_{-i}/c_i - 1$. From here, one can deduce that $\tau_i \leq T$ for any Beta prior and that $\overline{\sigma}_{-i}(t) > \tilde{\sigma}_{-i} > \underline{\sigma}_{-i}(t)$ for all $t < T$ and $\overline{\sigma}_{-i}(t) = \tilde{\sigma}_{-i} = \underline{\sigma}_{-i}(t)$, for all $t \geq T$.

Proof of Theorem 5. I prove the result when priors allow for correlation; adjusting the proof to accommodate the case in which they do not is tedious but straightforward.

Let $\Sigma^*_i(\mu_i) := \arg\max_{\sigma_i \in \Delta(A_i)} \mathbb{E}_{\mu_i}[u_i(\sigma_i, \sigma_{-i})]$ denote the set of maximizers at belief μ_i.

I first prove that if a sequence of probability measures $\mu^m_i \in \Delta(\Delta(A_{-i}))$ weak* converges to $\delta_{\sigma_{-i}}$ and $a_i \in A_i$ is not a best response to σ_{-i}, then for any sequence of distributions $\sigma^m_i \in \Sigma^*_i(\mu^m_i)$, $\sigma^m_i(a_i) \to 0$. Note that $\mathbb{E}_{\mu_i}[u_i(\sigma_i, \sigma_{-i})]$ is jointly continuous in (σ_i, μ_i) with respect to the product metric, where $\Delta(A_i)$ is endowed with the standard Euclidean metric and $\Delta(\Delta(A_{-i}))$ with the Lévy-Prokhorov metric. Then, by Berge's maximum theorem, Σ^*_i is upper-hemicontinuous and compact-valued. Supposing that $\sigma_i^m(a_i)$ does not converge to zero, implies that for any convergent subsequence of σ_i^m, its limit assigns strictly positive probability to a_i being chosen, while also belonging to $\Sigma_i^*(\delta_{\sigma_{-i}})$, a contradiction.

Now take any sequence of profiles of action distributions $\{\sigma^n\}_n$ as in the statement of the theorem and let $\{\sigma^m\}_m$ be a convergent subsequence of $\{\sigma^n\}_n$ with limit σ^k. Suppose that σ_i^k is not a best response to σ_{-i}^k. This implies that $\exists a_i \in A_i$ such that $\sigma_i^k(a_i) \geq \delta > 0$ for some $\delta > 0$ and a_i is not a best response to σ_{-i}^k. By continuity of u_i, note that if a_i is not a best response to σ_{-i}^k, then it is not a best response to any $\sigma_{-i} \in B_\epsilon(\sigma_{-i}^k)$ for small enough $\epsilon > 0$. By convergence of σ^m, we then obtain that for all large enough m, a_i is not a best response to σ_{-i}^m and $\sigma_i^m(a_i) \geq \delta/2$. That is, $\mathbb{E}_{\sigma_{-i}^m}[\sigma_i^*(\mu_i \mid y_i^{\tau_i^m})(a_i)] \geq \delta/2$ $\forall$ large enough m, from which we deduce $\mathbb{P}_{\sigma_{-i}^m}(\sigma_i^*(\mu_i \mid y_i^{\tau_i^m})(a_i) \geq \delta/4) \geq \delta/4$. In turn, from the above, this implies that there must be $\epsilon > 0$ such that $\mathbb{P}_{\sigma_{-i}^m}(\|\mu_i \mid y_i^{\tau_i^m} - \delta_{\sigma_{-i}^m}\|_{LP} \geq \epsilon) \geq \delta/4$. I now prove that this cannot be the case, that is, I show that $\lim_{m\to\infty} \mathbb{P}_{\sigma_{-i}^k}(\|\mu_i \mid {y_i}^{\tau_i^m} - \delta_{\sigma_{-i}^k}\|_{LP} > \epsilon) = 0$.

It would be natural to expect that, with sampling costs going to zero, optimal stopping time grows unboundedly and, by the law of large numbers, players learn the true distribution of actions of their opponents, best respond to it, and sequential sampling equilibrium converges to a Nash equilibrium. But, *conditional on stopping*, the set of signals are neither independent or identically distributed, so we cannot apply the law of large number directly. We then need to take a detour.

Denote player i's the associated earliest optimal stopping time by τ_i^m and their value function (which depends on c_i^m) as V_i^m. From Lemma 4, there is $\{T^m\}_m$ such that $\tau_i^m \geq T^m$ and $T^m \uparrow \infty$.

By Diaconis and Freedman (1990), there is $\epsilon(t)$ nonincreasing and such that $\epsilon(t) \to 0$ as $t \to \infty$ such that $\|\mu_i \mid y_i^t - \delta_{\overline{y_i}^t}\|_{LP} \leq \epsilon(t)$ uniformly over sequences of t observations, $y_i^t \in \mathscr{Y}_i$. Since, taking $y_{i,\ell} \sim \sigma_{-i}$, $\|\delta_{\overline{y_i}^t} - \delta_{\sigma_{-i}}\|_{LP} = \|\overline{y_i}^t - \sigma_{-i}\|$ is a (bounded) supermartingale with respect to σ_{-i}, by the optional stopping theorem, for $\tau_i \geq t$, $\mathbb{E}_{\sigma_{-i}}[\|\delta_{\overline{y_i}^{\tau_i}} - \delta_{\sigma_{-i}}\|_{LP}] \leq \mathbb{E}_{\sigma_{-i}}[\|\delta_{\overline{y_i}^t} - \delta_{\sigma_{-i}}\|_{LP}]$. Hence, for any σ_{-i}, $\mathbb{E}_{\sigma_{-i}}[\|\mu_i \mid {y_i}^{\tau_i^m} - \delta_{\sigma_{-i}}\|_{LP}] \leq \mathbb{E}_{\sigma_{-i}}[\|\delta_{\overline{y_i}^{\tau_i^m}} - \delta_{\sigma_{-i}}\|_{LP} \mid y_i^{T^m}] + \epsilon(T^m) \leq \mathbb{E}_{\sigma_{-i}}[\|\delta_{\overline{y_i}^{T^m}} - \delta_{\sigma_{-i}}\|_{LP}] + \epsilon(T^m) \leq \mathbb{E}_{\sigma_{-i}}[\|\overline{y_i}^{T^m} - \sigma_{-i}\|] + \epsilon(T^m)$.

Let $x_m := \mathbb{E}_{\sigma_{-i}^m}[\|\overline{y_i}^{T^m} - \sigma_{-i}^m\|] \in [0,2]$. Suppose that $\{x_m\}_m$ does not converge to 0. Take any convergent subsequence $x_\ell \to \gamma > 0$. For all large enough ℓ, $x_\ell \geq \gamma/2$. That is, $\mathbb{E}_{\sigma_{-i}^\ell}[\|\overline{y_i}^{T^\ell} - \sigma_{-i}^\ell\|] \geq \gamma/2$, implying that $\mathbb{P}_{\sigma_{-i}^\ell}(\|\overline{y_i}^{T^\ell} - \sigma_{-i}^\ell\| \geq \gamma/4) \geq \gamma/4$, as otherwise $\mathbb{E}_{\sigma_{-i}^\ell}[\|\overline{y_i}^{T^\ell} - \sigma_{-i}^\ell\|] \leq (1-\gamma/4)\gamma/4 + \gamma/4 \leq \gamma/4 < \gamma/2$. However, by the Dvoretzky–Kiefer–Wolfowitz–Massart inequality (Massart, 1990) $\mathbb{P}_{\sigma_{-i}^\ell}(\|\overline{y_i}^{T^\ell} - \sigma_{-i}^\ell\| \geq \gamma/4) \leq 2\exp(-T^\ell\gamma^2/8) \to 0$, a contradiction.

We conclude that $\lim_{m\to\infty}\mathbb{E}_{\sigma^m_{-i}}[\|\mu_i \mid {y_i}^{\tau_i^m} - \delta_{\sigma^m_{-i}}\|_{LP}] \leq \mathbb{E}_{\sigma^m_{-i}}[\|\overline{y_i}^{T^m} - \sigma^m_{-i}\|] + \epsilon(T^m) \to 0$.

Proof of Proposition 6. By Diaconis and Freedman (1990), for any μ_i, there is $T_i < \infty$ such that $\forall t \geq T_i$, $\mathbb{E}_{\mu_i}[\sigma'_{-i} \mid a^t_{-i}] \in B_{\epsilon_i}(\delta_{a_{-i}})$. Take $T := \max_{i\in I} T_i$. By Lemma 4, there is $\overline{c}$ such that, $\tau_i \geq T$ for every player i for which a_i is not always a best response (where $\overline{c}$ may depend on μ). Hence, for any μ there is an N such that $\forall n \geq N$, a is a sequential sampling equilibrium of $\langle \Gamma, \mu, c^n \rangle$.

Online Appendix

Appendix B. General Information Structures

In this section, I extend sequential sampling equilibrium to accommodate analogy partitions and more general information structures. For brevity, I'll focus on the case in which players' beliefs allows for correlation.

B.1. Existence of a Sequential Sampling Equilibrium

Let us simplify the notation in our baseline setup: $Y_i := A_{-i}$, $P_i := \Delta(Y_i)$, $\mu_i \in \Delta(P_i)$ with full support, and $u_i : A_i \times P_i \to \mathbb{R}$, a continuous function.

An analogy partition for player i can be represented by a surjective function $f_i : Y_i \to Z_i$, where $|Z_i| < |Y_i| < \infty$. Naturally, it defines a garbling: instead of observing y_i, the player has access to coarser information $f_i(y_i)$. More generally, one can consider $|Z_i| \times |Y_i|$ stochastic matrices B_i, where $B_i(z, y) \in [0, 1]$ and $\sum_{z \in Z} B_i(z, y) = 1$, and such that B_i has rank $|Z_i| < |Y_i|$.[22]

Let $Q_i := \Delta(Z_i)$ and ν_i be the pushforward measure on Q_i given μ_i and B_i, where for every measurable set $S \subseteq Q_i$, $\nu_i(S) := \mu_i(\{p_i \in P_i \mid B_i p_i =\in S\})$. We assume that the player now has access to iid draws from a fixed $B_i p_{i,0} \in Q_i$. It is straightforward to adjust the definition of the optimal stopping problem, expand the definition of an extended game with the additional primitives $\{\pi_i\}_{i \in I}$, where π_i denotes the information structure defined by (B_i, Z_i), and have sequential sampling equilibrium accommodate such more general information structures.

I provide the following sufficient condition for existence of a sequential sampling equilibrium:

Theorem 6. *Let $G := \langle \Gamma, \mu, c, \pi \rangle$ be an extended game such that for every player i, μ_i admits a continuous density, and rank$(B_i) = |Z_i| < |A_{-i}|$. Then G admits a sequential sampling equilibrium.*

Proof. Let $J_{B_i} := \det(B_i B_i^T)$, which is strictly positive, as rank$(B_i) = |Z_i|$. For convenience, define $\pi_i : P_i \to Q_i$ as $\pi_i(p_i) := B_i p_i$. Denote by g_{μ_i} the density of μ_i. Denoting λ^n the n-dimensional

[22] Allowing for $|Z_i| = |Y_i|$ is possible, but makes the proofs more cumbersome.

Lebesgue measure, by the coarea formula (see Evans and Gariepy, 2015, Theorem 3.10) we have

$$\mathbb{E}_{\mu_i}[u_i(a_i,p_i)\mid z_i^t] = \int_{P_i}\prod_{\ell\in[1..t]}(B_ip_i)(z_{i,\ell}^t)u_i(a_i,p_i)g_{\mu_i}(p_i)d\lambda^{|Y_i|-1}(p_i)$$
$$= \int_{Q_i}\prod_{\ell\in[1..t]}q_i(z_{i,\ell}^t)\int_{\pi_i^{-1}(q_i)}u_i(a_i,p_i)g_{\mu_i}(p_i)J_{B_i}{}^{-1/2}d\lambda^{|Y_i|-|Z_i|}(p_i)d\lambda^{|Z_i|-1}(q_i).$$

Define

$$u_i(a_i,q_i) := \int_{\pi_i^{-1}(q_i)}u_i(a_i,p_i)g_{\mu_i}(p_i)d\lambda^{|Y_i|-|Z_i|}(p_i)J_{B_i}{}^{-1/2}.$$

I will now show the following:

Lemma 11. *$u_i(a_i,q_i)$ is continuous in q_i.*

Proof. I first show that π_i^{-1} is a continuous correspondence.

Let $K(P_i)$ denote the set of nonempty, compact, and convex subsets of P_i. Take any $q_{i,n}\to q_i$ and $p_{i,n}\in\pi_i^{-1}(q_{i,n})$ converging to p_i and note that $B_ip_i = \lim_n B_ip_{i,n} = \lim_n q_{i,n} = q_i$, and thus π_i^{-1} is upper-hemicontinuous (uhc).

To show that it is lower-hemicontinuous (lhc) take any open set $U\subseteq P_i$ such that $U\cap\pi_i^{-1}(q_i)\neq\emptyset$. This implies that there is $p_i\in\text{int}(U\cap\pi_i^{-1}(q_i))$ and $\epsilon>0$ such that $B_\epsilon(p_i)\subset U\cap\pi_i^{-1}(q_i)$. As π_i is a linear mapping, then by the open mapping theorem (Rudin, 1973, Theorem 2.11), $\pi_i(B_\epsilon(p_i))$ is open and, therefore, $\exists\delta>0$ such that $B_\delta(q_i)\subseteq\pi_i(B_\epsilon(p_i))$. Consequently, $\forall q_i'\in B_\delta(q_i)$, $\emptyset\neq\pi_i^{-1}(q_i')\cap B_\epsilon(p_i)\subseteq\pi_i^{-1}(q_i')\cap U$.

I remark that π_i^{-1} is not only continuous, but also compact- and convex-valued correspondence, from Q_i to P_i. When restricted to $\pi_i(P_i)$, it is also nonempty, and thus then $\pi_i^{-1}:\pi_i(P_i)\to K(P_i)$ is continuous with respect to the Hausdorff metric (see Aliprantis and Border, 2006, Theorem 7.15).

Let $h: P_i\to\pi_i^{-1}(q_i)$ be such that $h(p_i) := \arg\min_{p_i'\in\pi_i^{-1}(q_i)}\|p_i-p_i'\|_\infty$. By continuity of π_i^{-1}, for any ϵ, there is an N such that $\forall n\geq N$, $\pi_i^{-1}(q_{i,n})\subset B_\epsilon(\pi_i^{-1}(q_i))$. Hence, for large n, for any point in $\pi_i^{-1}(q_{i,n})$, there is a point in $\pi_i^{-1}(q_i)$ that is at most ϵ away. As $u(a_i,p_i)g_{\mu_i}(p_i)$ is continuous in p_i and, by Heine-Cantor theorem, uniformly so (P_i is compact). Hence, for any q_i' close enough to q_i, the difference in the payoff function will be well-approximated by the the difference in measure (up to a constant scaling factor), $|u_i(a_i,q_i')-u_i(a_i,q_i)|\approx|\lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_i'))-\lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_i))|$.

In the sequel, I show that the measure is continuous in q_i. Take any sequence $(q_{i,n})_n\subseteq\pi_i(P_i)$ (a compact set) satisfying $q_{i,n} \to q_i$. As π_i^{-1} is continuous, and, in particular, upper hemicontinuous,

$$\limsup_n \lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_{i,n})) \leq \lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_i)),$$

since for any open set containing $\pi_i^{-1}(q_i)$, it will contain $\pi_i^{-1}(q_{i,n})$ for $q_{i,n}$ sufficiently close to q_i.

We now argue that the above inequality holds with equality. To see this, fix an arbitrary $\epsilon > 0$ and take a collection of points on the boundary of $\pi_i^{-1}(q)$ such that any two points are not closer than $\epsilon/2$ and not farther away than $\epsilon > 0$. This implies that we have a finite collection of such points. As π_i^{-1} is lhc, there is a $\delta > 0$ such that for every $q_i' \in B_\delta(q_i)$, $\pi_i^{-1}(q_i')$ contains a point in an $\epsilon/4$ neighborhood of every point in our collection, and by convexity, their convex hull. This implies that we can approximate arbitrarily well the interior of $\pi^{-1}(q_i)$ taking any n sufficiently large; i.e. there is a $\gamma(\epsilon) > 0$ such that $\lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_i)) - \gamma(\epsilon) \leq \lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_i'))$, with $\gamma(\epsilon) \to 0$ as $\epsilon \to 0$.

Hence, as $q_{i,n} \to q_i$, for any $q_{i,n} \neq q_i$, $\lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_{i,n})) \to \lambda^{|Y_i|-|Z_i|}(\pi_i^{-1}(q_i))$.

□

We can then redefine the problem by considering ν_i to be uniform on Q_i and take $u_i(a_i, q_i)$ as the utility function.

Redefining

- $v_i : \Delta(Q_i) \to \mathbb{R}$ as $v_i(\nu_i') := \max_{a_i \in A_i} \mathbb{E}_{\nu_i}[u_i(a_i, q_i)]$;
- $V_i : \Delta(Q_i) \to \mathbb{R}$ as $V_i(\nu_i) := \sup_{\tau' \in \mathbb{T}_i} \mathbb{E}_{\nu_i}[v_i(\nu_i \mid z_i^{\tau'} - c_i \tau_i]$;
- $\tau_i(\omega) := \inf\{t \mid V_i(\nu_i \mid z_i^t(\omega)) = v_i(\nu_i \mid z_i^t(\omega))\}$;
- $b_i(p_i) := \mathbb{E}_{\pi_i(p_i)}[\sigma_i^*(z_i^{\tau_i})]$, for some fixed $z_i^t \mapsto \sigma_i^*(z_i^t) \in \arg\max_{\sigma_i \in \Delta(A_i)} \mathbb{E}_{\nu_i}[u_i(a_i, q_i) \mid z_i^t]$.

we obtain — by analogous arguments to Proposition 1 — that V_i is continuous and, by Berk (1966), that $\nu_{i,t}$ weak* converges to $B_i p_i$, p_i-a.s., when $z_{i,t} \sim \pi_i(p_i)$, for all t. Therefore, τ_i is finite p_i-a.s., for any $p_i \in P_i$. Finally, an analogous version of Lemma 1 applies and b_i is continuous in p_i and maps to $\Delta(A_i)$. Hence, by essentially the same arguments as in Theorem 1, a sequential sampling equilibrium exists. □

In the above, we restricted to the case in which $|Z_i| < |Y_i|$. If instead $\text{rank}(B_i) = |Z_i| = |Y_i|$, we have the following:

Proposition 8. *Let $G := \langle \Gamma, \mu, c, B, Z \rangle$ be an extended game such that for every player i, rank$(B_i) = |Z_i| = |A_{-i}|$. Then G admits a sequential sampling equilibrium.*

Proof. Note that now B_i is invertible and $\pi_i(p_i) := B_i p_i$ is bijective when restricting its domain to $\pi_i(P_i)$. Hence, π_i admits a continuous inverse (note it is a linear mapping). Since for any $p_i \in \Delta(A_{-i})$, and μ_i has full support, $\pi_i(p_i)$ is in the support of the pushforward measure $\nu_i := \pi_i \# \mu_i \in \Delta(Q_i)$. Thus, for any p_i, by Berk (1966), $\nu_{i,t}$ weak* converges to a Dirac on $\pi_i(p_i)$. Moreover, if $\nu_i \in \Delta(Q_i)$ is the pushforward measure given $\mu_i \in \Delta(P_i)$ and $\pi_i : P_i \to Q_i$, we now also have that μ_i is ν_i's pushforward measure given $\pi_i^{-1} : Q_i \to P_i$. Then, weak*convergence of $\nu_{i,t}$ to δ_{q_i} implies weak* convergence of the $\mu_{i,t}$ to $\delta_{\pi_i^{-1}(q_i)}$. Uniform continuity of $V_i : \Delta(P_i) \to \mathbb{R}$ (as originally defined on the main text) delivers existence of an equilibrium as in Theorem 1. □

Naturally, all the above can also be extended to Bayesian games.

B.2. Relation to Analogy-Based Expectation Equilibrium

Finally, I discuss convergence to analogy-based expectation equilibria — see Jehiel (2021) for a survey. In line with the literature, I consider payoffs that are linear in the distribution of actions.

This solution concept when applied to normal-form games (including Bayesian games) can be readily translate to our setup: σ is an analogy-based expectation equilibrium if, for each player $i \in I$,

(1) $q_i(z_i) = \sum_{a_{-i} \in f_i^{-1}(z_i)} \sigma_{-i}(a_{-i})$ for every $z_i \in Z_i$;

(2) $u_i(a_i, z_i) := \sum_{a_{-i} \in f_i^{-1}(z_i)} \frac{1}{|f_i^{-1}(z_i)|} u_i(a_i, a_{-i})$; and

(3) $\sigma_i \in \arg\max_{\sigma_i' \in \Delta(A_i)} \mathbb{E}_{q_i} u_i(a_i, z_i)$.

Recalling that f_i is a surjective mapping from A_{-i} to Z_i, condition (1) states that each player i bundles difference action profiles (or players, or types, contingencies) a_{-i} into the same 'analogy class' z_i. Condition (2) can be read as a simplification device by player i: the player cannot distinguish across the different a_{-i} within the same analogy class z_i, they consider the average behavior, as if the probability of each a_{-i} within the same analogy class were the same. Then (condition (3)), they best respond to the expected payoff given the actual distribution over analogy classes, but assuming that, within the analogy class, distribution over contingencies is uniform.

In the above setup, this is achieved whenever μ_i is uniform. The result follows from arguments

	{Player M chooses a}		{Player C chooses a}	
	OLS	Logit	OLS	Logit
	(1)	(2)	(3)	(4)
δ_M	0.230***	0.949***	-0.772***	-3.430***
	(0.041)	(0.169)	(0.036)	(0.197)
Intercept	0.329***	-0.702***	0.842***	1.522***
	(0.018)	(0.079)	(0.017)	(0.090)
(Pseudo) R^2	0.02	0.01	0.20	0.15
Observations	1782	1782	1806	1806

Heteroskedasticity-robust standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 1: Payoffs and Choices: Own- and Opponent-Payoff Choice Effect

Notes: The table exhibits the association between player M's payoff to action a and the frequency with which subjects in each role choose action a. δ_M parametrizes player M's payoff to action a. The games in question are generalized matching pennies games as given in Figure 1, for $\gamma_C = 1$ (and scaled by 20). The sample includes data from rounds in which there is no belief elicitation. The data is from Friedman and Ward (2022).

analogous to those in the proof of Theorem 5, but with a crucial simplification: the linearity of payoffs in distributions, the posterior means pin-down the set of best responses. And, upon stopping, $\mathbb{E}_{\mu_i}[q_i' \mid z_i^{\tau_i}] = \frac{\tau_i}{\tau_i + |A_{-i}|}\overline{z_i}^{\tau_i} + \frac{1}{\tau_i + |A_{-i}|}$. Since, for any player i for which no action is always a best response, there is a lower bound to the stopping time $\tau_i \geq T^n$, that grows unboundedly as sampling costs vanish, $c_i^n \to 0$ (Lemma 4), by a similar application of the optional stopping theorem as in the proof of Theorem 5, $\mathbb{E}_{q_i}\left[\left\|\mathbb{E}_{\mu_i}[q_i' \mid z_i^{\tau_i^m}] - q_i\right\|\right] \leq \mathbb{E}_{q_i}\left[\left\|\mathbb{E}_{\mu_i}[q_i' \mid z_i^{T^m}] - q_i\right\|\right] \leq \frac{T^m}{T^m + |A_{-i}|}\mathbb{E}_{q_i}[\|\overline{z_i}^{T^m} - q_i\|] + 2\frac{1}{T^m + |A_{-i}|}$. Then, by the law of large numbers, $\mathbb{E}_{q_i}[\|\overline{z_i}^{T^m} - q_i\|] \to 0$. This provides a shorter route to show that players' posterior means converge to the underlying true distribution.

Appendix C. Experimental Data and Analysis Details

In this appendix, I provide details on the experimental data used along with additional analysis. A total of 164 subjects were recruited for sessions run in the Columbia Experimental Laboratory in the Social Sciences (CELSS) to play matching pennies games as the one depicted in Figure 1. Subjects are randomly and anonymously matched, but their roles are fixed throughout. Player M's payoff to action a, δ_M, took one of six values (here rescaled by a factor of 20 for convenience): 4, 2, 1/2, 1/4, 1/10, and 1/20.

	Distance to Indifference		
	Player M	Player C	Both
	(1)	(2)	(3)
Log Decision Time	-3.682***	-2.021**	-2.961***
	(1.225)	(0.881)	(0.790)
Intercept	42.314***	45.365***	48.185***
	(4.181)	(2.670)	(2.435)
Fixed Effects	Game	Game	Role × Game
R-Squared	0.08	0.27	0.18
Observations	1620	1680	3300

Heteroskedasticity-robust standard errors in parentheses.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2: Decision Time and Reported Beliefs: Time-Revealed Preference Intensity

Notes: The table presents regression results on the relation between log decision times (in seconds) and the distance between reported beliefs to indifference points with data from Friedman and Ward (2022). Reported beliefs refer to the elicited beliefs about the probability the opponents would play action a, and the indifference point refers to the posterior mean that would make the player indifferent between taking either action. The games in question are generalized matching pennies games as given in Figure 1, for $\gamma_C = 1$ (and scaled by 20). δ_M parametrizes player M's payoff to action a. Game (as indexed by δ_M) and role and game fixed effects are used where indicated. Columns (1) and (2) only use data for subjects in the roles of player M and C, respectively; the right-most column uses both. The data is from Friedman and Ward (2022).

The experiment consisted of two stages. In the first stage, actions elicited and each game is played twice. In the second stage, each game is played 5 times and either both actions and beliefs about the probability that the opponent chooses action a are elicited or only actions are elicited. Beliefs here refer to point estimates reported by the subjects, neglecting any strategic uncertainty. In other words, belief reports would correspond to posterior means in our framework. Elicitation of actions and beliefs is incentive-compatible and robust to risk attitudes and game payoffs correspond to probability points towards prizes of $10. Throughout, no feedback was provided, game order was randomized and, importantly for our purposes, decision times are recorded. Other details on the experimental design can be found in Friedman and Ward (2022).

There are some important caveats to note. First, beliefs elicited in the second stage refer to opponent's actions from the first stage. This, together with the fact that elicitation of actions and beliefs is sequential instead of simultaneous, with beliefs being elicited first, may raise concerns of whether reported beliefs are a good proxy for the beliefs that subjects hold when taking an

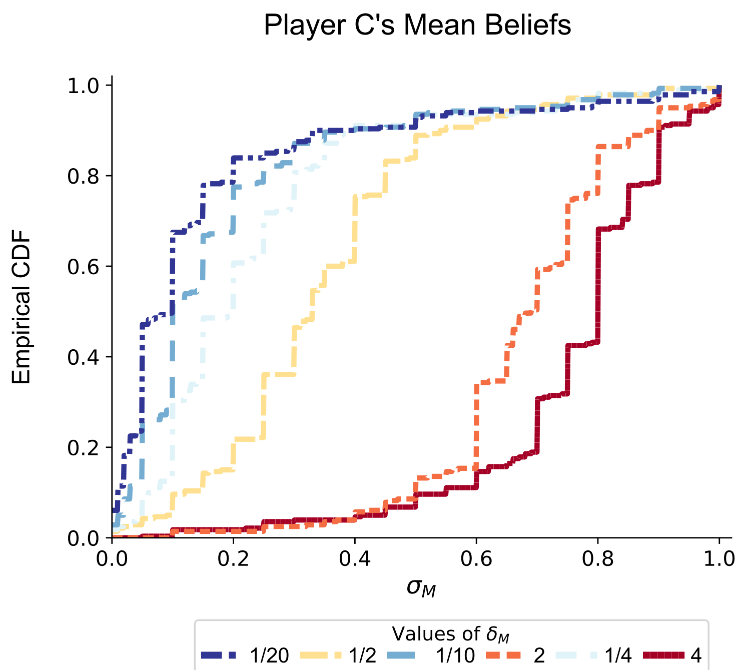

Figure 4: Opponent Payoff and Beliefs

Notes: The figure compares the distribution of beliefs in generalized matching pennies games as given in Figure 1, for $\gamma_C = 1$ (and scaled by 20). The figure exhibits the empirical CDF of reported (mean) beliefs about the probability with which subjects in the role of player C believe their opponent (in the role of player M) will take action a. Different lines correspond to games in which the player M has different payoffs to action a; δ_M parametrizes player M's payoff to action a. The data is from Friedman and Ward (2022).

action. Second, while decision time was recorded, subjects were forced to wait a minimum of 10 seconds before reporting their beliefs. As the subjects' decision times will be used as a proxy to test sequential sampling equilibrium's predictions for stopping times, the forced minimum decision time may undermine the exercise. Finally, the authors highlight there being evidence of "no-feedback learning" as the same subject plays the same game multiple times. This is especially worrying when comparing instances where only actions are elicited with those where both actions and beliefs are. In order to avoid issues due to experience or learning, and focus on initial response as much as possible, I will focus on choice data when beliefs are not elicited.

Table 1 documents the own- and opponent-payoff choice effects mentioned on Section 3.2: as player M's payoffs to action a increase, subjects in that tend to choose the action more often and action b less often, while the opposite is true for subjects in the opponent's role, player C. Table 2 provides support for the collapsing boundaries result presented in Section ??: a negative

Player C Beliefs		
$\sigma_M^{\tau_C} \mid \delta_M^{High} \geq_{FOSD} \sigma_M^{\tau_C} \mid \delta_M^{Low}$		
High	Low	KS-Statistic
(1)	(2)	(3)
4	2	0.33***
2	1/2	0.76***
1/2	1/4	0.40***
1/4	1/10	0.23***
1/10	1/20	0.23***

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Opponent Payoff and Beliefs: FOSD Tests

Notes: The table exhibits the results of two-sample Kolmogorov-Smirnov tests for first-order stochastic dominance of the distribution of reported beliefs by subjects in the role of player C in games with different values of δ_M, High (col (1)) and Low (col (2)). Column (3) presents the test statistic, with number of observations $(n, m) = (280, 280)$. δ_M parametrizes player M's payoff to action a. The games in question are generalized matching pennies games as given in Figure 1, for $\gamma_C = 1$ (and scaled by 20). The data is from Friedman and Ward (2022).

association between the distance of subjects' reported beliefs to their indifference point and the decision time. Related to this result, the main text discussed a first-order stochastic dominance shift of beliefs for player C as δ_M (player M's payoffs to action a) increases; Figure 4 exhibits such lawful relation. Appropriate statistical tests are given in Table 3.

Appendix D. Misspecified Priors

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 1 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.

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 observations 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, and optimal stopping time is not bounded.[23] Further, it stands in sharp contrast to the case in which beliefs are misspecified.

I now provide an example in which misspecification leads to a player never stopping with probability 1 (with respect to the true distribution of opponents' actions), and sequential sampling equilibrium fails to exist.

Let Γ be a two-player game in which player i's opponent has three possible actions, a, b, and c, and always chooses c (e.g. because c is dominant, or because their sampling cost is too high and c is uniquely optimal under their prior). Denote $\sigma_{-i} = (\sigma_{-i}(a), \sigma_{-i}(b), \sigma_{-i}(c)) \in \Delta(\{a, b, c\})$. Suppose player i's prior beliefs about σ_{-i}, μ_i, are such that player i assigns probability 1/2 to $(1/2, 1/6, 1/3)$ and probability 1/2 to $(1/6, 1/2, 1/3)$. Player i can choose either a or b and player i's payoffs are given by $u_i(a_i, a_{-i}) = 1$ if $a_i = a_{-i}$, and 0 if otherwise. Then, if y_i^t is such that $y_{i,\ell} = c$ for all $\ell \in [1..t]$, $\mu_i \mid y_i^t = \mu_i$. Under their prior, $v_i(\mu_i) = 1/4$, $v_i(\mu_i \mid a) = v_i(\mu_i \mid b) = 3/4$, and $v_i(\mu_i \mid c) = v_i(\mu_i) = 1/4$, hence $\mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i)] = \frac{2}{3}\frac{3}{4} + \frac{1}{3}\frac{1}{4} = \frac{7}{12}$. Note that, a necessary condition for player i to stop is that $\mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i)] - v_i(\mu_i) \leq c_i$. But, since at any sequential sampling equilibrium player i's opponent chooses c with probability 1, we have that $\mu_i \mid y_i^t = \mu_i$ and, for any $c_i < 1/3$, we always obtain $\mathbb{E}_{\mu_i}[v_i(\mu_i \mid y_i)] - v_i(\mu_i) = \frac{1}{3} > c_i$. Therefore, since $\sigma_{-i}(c) = 1$, $\mathbb{P}_{\sigma_{-i}}(\tau_i = \infty) = 1$.

[23] 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 Mierendorff, 2019).