

Rosenberg, Dinah; Solan, Eilon; Vieille, Nicolas

Working Paper

Social learning in one-arm bandit problems

Discussion Paper, No. 1396

Provided in Cooperation with:

Kellogg School of Management - Center for Mathematical Studies in Economics and Management Science, Northwestern University

Suggested Citation: Rosenberg, Dinah; Solan, Eilon; Vieille, Nicolas (2004) : Social learning in one-arm bandit problems, Discussion Paper, No. 1396, Northwestern University, Kellogg School of Management, Center for Mathematical Studies in Economics and Management Science, Evanston, IL

This Version is available at:

<https://hdl.handle.net/10419/31213>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Social Learning in One-Arm Bandit Problems*

Dinah Rosenberg[†], Eilon Solan[‡] and Nicolas Vieille[§]

December 20, 2004

Introduction

In the classical one-arm bandit problem, a decision maker receives consecutive payoffs, which are realizations of i.i.d. random variables whose distribution is not known. The decision maker's problem is to decide when to stop, where stopping implies switching to a known deterministic payoff. The optimal strategy in this problem has been characterized by Gittins (1979), and the problem is considered to be a paradigmatic example of the trade-off between experimentation and exploitation, that is, between gathering information and using it.

The bandit problem is typically studied in isolation. The only source of information available to the decision maker is her own experience. However,

*We are indebted to Itzhak Gilboa for his help. We thank Yisrael Aumann, Thomas Mariotti, Peter Sozou, and Shmuel Zamir for their suggestions, and seminar audiences at Caltech, Hebrew University of Jerusalem, LSE, Northwestern, Paris, Stanford, Tel Aviv University, Toulouse 1 University, and the second World Congress of the Game Theory Society.

[†]Laboratoire d'Analyse Géométrie et Applications, Institut Galilée, Université Paris Nord, avenue Jean-Baptiste Clément, 93430 Villetaneuse, France; and laboratoire d'Econométrie de l'Ecole Polytechnique, 1, rue Descartes 75005 Paris, France . e-mail: dinah@zeus.math.univ-paris13.fr

[‡]MEDS Department, Kellogg School of Management, Northwestern University, *and* the School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il

[§]Laboratoire d'économtrie, Ecole Polytechnique *and* Département Finance et Economie, HEC, 1, rue de la Libération, 78 351 Jouy-en-Josas, France. e-mail: vieille@hec.fr

in many real-life situation other decision makers may also be facing similar bandit problems, and their experience may also be useful. In particular, we are interested in the case in which each of several individuals faces a one-arm bandit, where the bandits' payoff are independent, but they are governed by the same unknown distribution.

For concreteness, imagine a casino in which several bandit machines, which share the same unknown distribution and produce independent pay-offs, are placed next to each other. A gambler who sits next to one of these machines may learn its attributes not only from her own experience, but also from the experience of other gamblers around her. If she can actually observe the payoffs that other gamblers receive, the problem is very similar to the original one-arm bandit. To be precise, if there are k gamblers at a given point, the decision to keep pulling the bandit's arm yields k i.i.d. realizations from the unknown distribution, rather than one such realization as in the classical case.

However, in many economic examples, the gambler cannot observe the payoff of other gamblers, but only their number. That is, the gambler only knows that there are $k - 1$ other gamblers who decide to keep pulling their machines' arms. This fact provides a signal about the payoffs that these gamblers received, but it falls short of knowledge of these payoffs. Following are some economic examples.

Example 1: Fishermen go out to sea at night. Each toss of a net can be modelled as an i.i.d. random variable. Since they all fish in the same area, we may assume that the random variables have the same distribution. Further, since the sea is large enough, we may assume that this distribution does not change throughout the night, and it does not depend on the fishing done by others. Each fisherman may observe the number of boats that are out at sea, but not the amount or type of fish found by the other boats.

Example 2: Firms have to decide whether to recruit potential employees in a given school. The school may be thought of as characterized by a distribution of the students' quality. Each student hired is a realization of a random variable having the same unknown distribution, but this realization is only observed by the firm who hired the students. Yet, each firm may observe how many other firms are hiring in the school, and attempt to learn something about the quality of the students in the school from the other firms' decisions whether or not to keep hiring in this school.

Example 3: A venture capital (VC) firm has to decide whether to invest in start-up companies in a given industry. Each new company would generate

profits that may, ex-ante, be considered a random variable, where all these random variables are i.i.d. Yet, the VC firm does not know distribution that characterizes future payoffs of new companies in the industry. For a relatively long period of time, the profitability of a new company is only known to the VC firm that invested in it. Thus, each VC firm does not know the profitability of companies it has not invested in, but it may know the number of other VC firms that are still investing in this industry.

All of these examples may be modelled as multiple identical one-arm bandit problems, played by distinct players whose only interaction is that they get to observe the stopping decision of each other. This is the game we analyze in this paper. Observe that the only interaction among the players is informational externalities: each player may learn something from the behavior of the others, but this behavior does not affect the player's payoff directly. Still, in order to perform correct inference from stopping decisions of others, we have to consider equilibria of this game. Indeed, when a player asks herself why have other players decided to remain in the game, she should take into account the fact that these players remain in the game because she did, and that they interpreted her decision as a signal that the common distribution they all face has a high expectation.

We therefore model the situation as a game. Each player can make an irreversible stopping decision at any point, after observing all her past payoff and the stopping decisions of the others, up to that point.

It is well known that in a single-gambler one-arm bandit the assumption that stopping is irreversible is not crucial. Since stopping to play the machine results in no new information, stopping remains optimal, and it will not be optimal to reverse it even if reversal were possible. By contrast, in the context of social learning this no longer holds. In fact, reversibility of the stopping decision would introduce a free-riding temptation into the problem. In this paper we focus on the problem in which stopping is irreversible. This assumption is quite plausible in applications where starting to play involves a non-trivial fixed cost, such as in Example 1 above. In other applications, the irreversible decision case is merely a benchmark model.

We now describe the game that we analyze. The game is based upon two-arm bandits. One arm is known, and yields a constant payoff (which we normalize to 0). The other arm is unknown, and can be either *High* or *Low*. Whatever be its type, the unknown arm outputs an i.i.d. sequence of payoffs; the expected payoff is positive if the type is *High*, and negative if it is *Low*.

The players operate replica two-arm bandits with all risky arms having the same type (either all *High*, or all *Low*). Each player has to decide when to switch from the risky arm to the safe arm. Switching from the safe to the risky arm is irreversible. We assume that payoffs are *privately* observed, while all exit decisions are *public* information. Equivalently, each player knows who dropped out in the past, and when.

After proving the existence of a symmetric equilibrium, we study qualitative properties of the equilibria. Our first result relates to the nature of equilibrium strategies. It is shown that *all* equilibria are pure, and moreover they are in cut-off strategies. That is, the inflow of (private and public) information is processed in the following simple way. At each stage, each player computes the conditional probability that the type of her machine is *High*, computed on the basis of her *private* information; this quantity is called the player's *private* belief, because it does not take into account public information given by the other players' actions. She then compares her private belief to a (time-dependent) cut-off. Whenever the news that someone dropped out is received, the player adapts by changing the cut-offs. Note that these equilibrium strategies are monotonic, in the sense that, everything else being equal, a player's decision is monotonic w.r.t. her private belief.

We then further investigate the equilibrium cut-offs. Such results are easier to state for two-player games. It turns out that the longer the other player is active the more optimistically a player behaves : it requires worse private beliefs to induce her switching to the safe arm. Indeed, casual intuition suggests that, the longer a player stays in, the better news it is for the other player. Therefore, the equilibrium cut-offs should be decreasing. While the conclusion indeed holds (at least for two players), the analysis is intricate. The reason is as follows. Consider the equilibrium cut-offs of player i , when player j is still active. These are computed by taking into account the belief over the common type, but also the option value of waiting. For a given value of her private belief, player i will indeed be more optimistic at stage $n + 1$ than at stage n . The option value involves the information that player i will later derive from her private payoffs and public decisions of player j . If say, player i finds it likely that player j will either drop out very soon or very late depending on her payoffs, she may wish to stay in few additional stages in order to get this piece of information. If moreover, this effect is comparatively stronger stage at than at stage $n + 1$, player i may be inclined to use a lower cut-off at stage n than at stage $n + 1$. We note that player i 's assessment of when player j might drop out depends both on (i) the shape

of player j 's future cut-offs and on (i) player i 's belief on player j 's private belief. More precisely, if the cut-offs of player j are not constant, the fact that player j drops out two stages after $n + 1$, or two stages after n , do not convey the same information about the quality of the arm.

In addition to the equilibrium cut-offs being decreasing, we prove that they converge to zero, as long as the other player is active. In other words, it takes really bad news to trigger exit when other players have remained active for a while. Since private beliefs converge to zero when the common type is *Low*,¹ this prompts the question of whether the convergence to zero of the equilibrium cut-offs is slow enough to ensure that all players will eventually drop out if the type is *Low*. We show that this is indeed the case, irrespective of the number of players. The basic idea is that if with positive probability the other players stay in forever, then from some stage on player i does not expect to learn any information from the other players' decisions, so that she actually faces a one-player problem. However, the one-player problem is equivalent to a standard one-arm bandit problem, and in this context when the type is *Low* the player eventually drops out. By contrast, if the type is *High*, there is a positive probability that no player will ever drop out. This result can be interpreted as a partial efficiency result.

Lastly, we analyze equilibrium behavior when the number of players is large. We provide an almost explicit characterization of the equilibrium outcome. In loose terms, we show that equilibria are asymptotically symmetric and deterministic. When the population is large enough, if the payoffs of all players were observed at stage 1, then *all* players would be able to use this information to deduce if the common type is *High* or *Low* and behave accordingly in stage 2. We prove that when the players do not observe each other's payoffs a fraction of the player drops out in stage 1, and in stage 2 all remaining players correctly infer the type of the machine and act accordingly. Therefore, after stage 2 the optimal behavior can be achieved and the fraction of the players that dropped out in stage 2 if the common type is *High* exactly measures the inefficiency related to imperfect observation. More precisely, the equilibrium strategies are as follows. A positive fraction of players drops out at stage 1, smaller when the common type is *High* than when it is *Low*. Hence, by observing the number of players that dropped out in the first stage all players will correctly infer the true common type in stage 2. Thus, in stage 1, all players compute the option value of waiting with the

¹Private beliefs are always the conditional probability that the type is *High*.

anticipation that they will be told the true state in stage 2.

Our model is closely related to the literature on strategic experimentation. In Bolton and Harris (1999) and Cripps *et al.* (2004), players dynamically allocate their time between a (bandit) machine that yields conditionally i.i.d. payoffs given the common type, and a machine with known output. All information (payoffs) is publicly disclosed. As a consequence, all agents share the same posterior belief over the common type, which serves as a state variable. In our model by contrast, different players have different posterior beliefs, since each player has two sources of information: her own past payoffs, and the behavior of the other players. Thus, if all players are active at a certain stage, player i cannot tell whether other players received good news on the common type, or whether they were encouraged to remain active by the fact that all other players were active.

In Bolton and Harris (1999), a player may be tempted to devote time to the known machine, in the hope of benefitting from the experimentation performed by others. This free-rider effect is mitigated by the so-called encouragement effect, according to which a player may choose to experiment, in order to encourage others to do so, and therefore to benefit from the induced positive informational externality. Bolton and Harris focus on the interplay between these two effects. In our game, there is no room for free-riding, since a player must either bear the cost of experimentation, or exit. Besides, the encouragement effect is rather defined by the fact that being active is usually interpreted as good news on the common type.

Models of strategic experimentation with non-pure informational externalities within various economic setups were studied in the literature, see, *e.g.*, Bergemann and Välimäki (1997, 2000), Décamps and Mariotti (forthcoming).

Our model also relates to the literature on strategic learning. In these latter models, players are initially endowed with private information on a state of nature, and decide on a course of action. In the seminal contributions to this literature (Banerjee, 1992, Bikhchandani *et al.*, 1992), players play in turn in some exogenously given order. In the contribution that is closest to ours, Chamley and Gale (1994) (see also Chamley (2004)) analyze a version with endogenous timing: players may act at most once, at the time of their choice. Their and our models have common features: both are games with pure informational externality, irreversible decisions, private informa-

tion and public decisions. However, private information is received only once in Chamley and Gale’s model, while in our model it keeps flowing in. Besides, Chamley and Gale’s paper is close to the real options framework, and payoffs incur only after the option is exercised. As we show, this difference turns out to be irrelevant, and our results are isomorphic to those that would obtain in a real options framework.

Part of Chamley and Gale’s focus is on efficiency, as measured by the investment delay at equilibrium. In our setup, efficiency is rather measured by the ability of players to discriminate between the two states of nature. We prove that some inefficiency remains at equilibrium, and prove a partial result related to efficiency.

A recent paper by Moscarini and Squintani (2004) studies a R&D race between firms with private information – received only once prior to the race – about the uncertain arrival date of an invention. Firms must pay costs in order to remain in the race. The first to obtain the invention wins the race. This creates a non-informational strategic interaction between the two firms.

Another relevant paper is Caplin and Leahy’s (1994) analysis of market crashes. In their (finite-horizon) model, players keep receiving private information on the final state of demand, as long as they do not drop out. Again, actions are publicly observed. In this respect, this paper is close to the model we study here. However, Caplin and Leahy assume a continuum of players, and focus on symmetric equilibria. Under this assumption, and conditional on the state of nature, the evolution of players’ beliefs is deterministic. By contrast, we focus on games with finitely many players, and prove results when the number of players grows to infinity. Our results on large games may be interpreted as providing a positive robustness result for Caplin and Leahy (1994)’s analysis.

Another trend of literature that is related to our model has to do with biology. Animals can learn some relevant information by observing the behavior of other animals of the same species. For instance a bird can try a new patch and then either probe it again or just leave it. By observing this behavior and not the quality of the patch (namely presence or absence of food) a fellow bird can try to infer if there is some food there or not. Such behavior has been studied by Valone and Templeton (2002), Giraldeau *et al.* (2002).

The paper is organized as follows. In Section 1 we present the model, and in Section 2 the main results of the paper. Section 3 is devoted to an example. In Section 4 we gain further intuition by studying the beliefs of the players. Sections 5 and 6 are devoted to the proofs of the results. Technical proofs are relegated to the appendix.

1 The model

1.1 Rules and assumptions

We here describe in detail the model we deal with. Time is discrete, and the set of players is a finite set I . We let \mathbf{N} denote the set of nonnegative integers, including zero.

The players operate replica one-arm bandit machines. The machines have two possible types: either all bandits are of type H , or all of them are of type L . Conditional on the type of the machine, payoffs are i.i.d. (over stages and players). The payoff to player i in stage n is X_n^i . At every stage $n \in \mathbf{N}$ each player has the option to drop out from the game, or to stay in. Dropping out is irreversible, so that the exit option can be exercised at most once: each player's strategic choice reduces to a timing decision. Once a player drops out, her payoff is zero forever. Finally, players discount future payoffs at the rate $\delta \in (0, 1)$.

A key feature is the information structure of the game. Payoffs are *private* information, while exit decisions are *publicly* observed. Specifically, the information available to player i at the beginning of stage n , if still active, consists of (a) the realized values of X_0^i, \dots, X_{n-1}^i and (b) the exit decisions of the other players, *i.e.*, *who* dropped out and *when*.

We let (Ω, \mathbf{P}) be the probability space on which all random variables are defined. We denote by $p_0 := \mathbf{P}(\Theta = H)$ the prior probability that the common type of the machines is H . We denote by $\mathbf{P}_H := \mathbf{P}(\cdot \mid \Theta = H)$ and $\mathbf{P}_L := \mathbf{P}(\cdot \mid \Theta = L)$ the conditional of \mathbf{P} , given the common type.

For a random variable X and $\theta = L, H$, we denote by $\mathbf{E}_\theta[X]$ the expectation of X under \mathbf{P}_θ .

Throughout the paper, we maintain the following assumptions:

- A1.** The variables (X_n^i) have a finite variance: $\mathbf{E}[(X_n^i)^2] < +\infty$.
- A2.** The inequalities $\mathbf{E}_H[X_n^i] > 0 > \mathbf{E}_L[X_n^i]$ hold.

- A3.** The conditional law of X_n^i given θ has a continuous density f_θ , whose support is independent of θ .
- A4.** For any $i \in I$ and $n \in \mathbf{N}$, the (random) likelihood ratio $f_H(X_n^i)/f_L(X_n^i)$ has a density.² In addition, letting $C \subset \mathbf{R}^+$ denote the support of the likelihood ratio, one has $\inf C = 0$ and $\sup C = +\infty$.

Assumption **A1** is only needed to make sure that the payoff function of the repeated game is well defined.

Assumption **A2** eliminates trivialities: if both expectations were non-negative (resp. non-positive), it would be a dominant strategy to remain active forever (resp. drop out at the outset of the game). By **A2**, one should think of H as the good type, and of L as the poor one.

Assumptions **A3** and **A4** are less innocuous. As for **A3**, the fact that the supports of f_L and f_H coincide implies that the common type is never perfectly inferred. The continuity assumption is made only for technical convenience. As for **A4**, the likelihood ratio $f_H(X_0^i)/f_L(X_0^i)$ directly relates to the private belief of player i in stage 1. The most significant implication is that this private belief has a density. One of our results – that all equilibria are in pure strategies – critically depends on it. Finally, the assumption $\inf C = 0$ ensures that this private belief takes arbitrarily low positive values with positive probability, while $\sup C = +\infty$ ensures that it takes arbitrarily high values.³

1.2 Comments

In this section we provide few comments and words of interpretation.

1.2.1 The interaction among players

Since payoffs depend on the common type they are correlated random variables. Therefore past payoffs contain valuable information on Θ , so that each player wishes to take into account past payoffs of the other players: the information held by other players matters. On the other hand, the payoff to a player does not depend on the other players' decisions or payoffs. In that sense, the interaction among players is purely of an informational nature.

²By **A3**, the denominator $f_L(X_n^i)$ is non-zero, \mathbf{P} -a.s., hence the likelihood ratio is well-defined.

³Up to some qualifications, our results are valid without this assumption.

Nevertheless, the information structure renders this interaction complex. Indeed, assume that all players are still active in some stage n . Player i cannot tell whether the other players decided to stay in because they received positive information on Θ , or because they themselves were encouraged by the fact that others were still active. Assume that players' strategies are given. Player i 's private information allows her to compute her belief over the private information received by player j , which in turn allows player i to compute her belief over the belief held by player j over her own private information, and so on. Plainly, this gives rise to an infinite hierarchy of beliefs over private information – a common feature of many games with incomplete information. This hierarchy is endogenously determined, as a function of the strategy choices of the players.

In other words, the only common knowledge events are the public ones: only past decisions are common knowledge. Our main result shows that, in all equilibrium, public and private information is incorporated in the equilibrium strategies in a particularly simple way: the dependency on public and private information can be separated easily. However, there is no public equilibrium – equilibrium that depends only on publicly available information.

1.2.2 The case of a single player

When there is only one player, our game reduces to an optimal stopping problem, in which the player has to choose when to switch from the “risky” action (remain in the game) to the “safe” one (drop out). This optimization problem is equivalent to the standard one-arm bandit problem, in which player i is allowed to alternate over time between the two arms. Indeed, it is well-known in the bandit literature that, if the decision maker finds it optimal in some stage to use the safe action, it is optimal to continue with it forever, see Ferguson (2004) and the references therein. Obviously, our multi-player game is *not* equivalent to the game in which players would be allowed to alternate between the two actions. Indeed, in the latter version players will have an incentive to “free-ride” over the experimentation performed by others. However, when all players but one dropped out from the game, the continuation game faced by the last player can be analyzed using the tools of the literature on bandit problems.

1.2.3 Timing games and real options

The basic model of real options (see, *e.g.*, McDonald and Siegel (1986), or Dixit and Pindyck (1994, p.136)) assumes that an investor has to choose an investing time, at which she pays a fixed amount C in return for a project, whose value V evolves stochastically.⁴

Consider now a situation with I investors and I projects. Each investor i can invest only in project i , and has to decide when to invest in that project. The payoff of project i at stage n is X_n^i . Thus, investing at stage n entitles the investor to receive X_n^i, X_{n+1}^i , etc. Suppose that with probability p_0 all projects are good, and with probability $1 - p_0$ they are all bad. Suppose further that (X_n^i) are i.i.d. given the quality of the project. W.l.o.g., assume that the fixed cost of investing C is zero.⁵

We assume that all investors observe past values of their own project, as well as the investing behavior of other investors, but not past values of other investors' projects.

Our analysis applies to this model as well. Indeed, let G' be the investment game we just described, and let G be the timing game described in Section 1 (with the same data). For every strategy profile $\phi = (\phi^i)_{i \in I}$, denote by $\gamma'^i(\phi)$ and $\gamma^i(\phi)$ the expected payoff under ϕ in the two games. We observe that

$$\gamma'^i(\phi) + \gamma^i(\phi) = (p_0 \mathbf{E}_H[X_0^i] + (1 - p_0) \mathbf{E}_L[X_0^i]) / (1 - \delta). \quad (1)$$

Indeed, if player i invests at stage t , her payoff in G is the sum of current payoffs received prior to stage t , while her payoff in G' is the sum of all stage payoffs, starting from stage t .

This implies that ϕ is an equilibrium in the investment game G' if and only if it is an equilibrium in the timing game with payoffs (constant $- X_n^i$), where the constant is the right hand side of (1).

In particular in the real options framework, all equilibria will be in cut-off strategies; when the private belief is computed as the probability that the project is good, the cut-offs will be increasing with time as long as the other players did not invest.

⁴This is the discrete-time analog of the continuous-time model described in McDonald and Siegel (1986) and Dixit and Pindyck (1994, p.136).

⁵Alternatively, we may spread the cost over all stages, and subtract that amount from all pay-offs.

1.3 Information, Strategies and Payoffs

We collect here definitions relative to strategies. While concepts are standard, some technical care is needed, due to the generality of our framework. Since payoffs are continuous random variables, the language of partial histories/decision nodes is inappropriate, and we rely on stochastic processes notions.

For exposition sake, we consider first two-player games.

1.3.1 The information players have

Let a stage $n \in \mathbf{N}$ and a player $i = 1, 2$ be given. The private information of player i in stage n is the sequence of payoffs she has received so far. Accordingly, we denote by $\mathcal{F}_n^i := \sigma(X_0^i, \dots, X_{n-1}^i)$ the σ -algebra over Ω induced by these payoffs.

In addition to her private information, player i knows when the other player drops out. At stage n the other player, player j , is either still playing, or has already quit in a previous stage. This status is denoted by $t_n^j \in \mathbf{N} \cup \{\blacksquare\}$, where $t_n^j = \blacksquare$ means that player j is still active at the beginning of stage n , and for $k < n$, $t_n^j = k$ means that player j has dropped out at stage k . Generic values for t_n^j are denoted α .

1.3.2 Pure plans of actions

A *pure plan of action* of player i specifies when to drop out from the game, as a function of available information. In describing that rule, it is convenient to separate public from private information as follows.

Definition 1.1 *A pure plan of action of player i consists of a family $\phi^i = (\tau^i(\alpha), \alpha \in \mathbf{N} \cup \{\blacksquare\})$ of stopping times for the filtration $(\mathcal{F}_n^i)_{n \in \mathbf{N}}$, with the property that $\tau^i(k) > k$, \mathbf{P} -a.s. for each $k \in \mathbf{N}$.*

The argument α of τ^i stands for the currently available public information. The stopping time $\tau^i(\blacksquare)$ describes player i 's behavior, as long as player j is still active. It may only depend on private information. Specifically, given a stage $n \in \mathbf{N}$, the event $\{\tau^i(\blacksquare) = n\} \in \mathcal{F}_n^i$ consists of those realizations of the payoffs X_0^i, \dots, X_{n-1}^i after which player i would decide to drop out at stage n – but not earlier – if player j is still active by then.

For $k \in \mathbf{N}$, the stopping time $\tau^i(k)$ describes player i 's behavior after stage k , in the event player j drops out in stage k , hence the requirement $\tau^i(k) > k$.

Under such a plan ϕ^i , if active in some stage n , player i will act on the basis of the available information. If player j remained active in stage $n - 1$, player i will use $\tau^i(\blacksquare)$: she drops out if $\tau^i(\blacksquare) = n$, and she continues otherwise. If instead player j dropped out in stage $k < n$, player i will use $\tau^i(k)$ to decide on a course of action: namely, she will drop out at stage n if $\tau^i(k) = n$ and continue otherwise.

Our use of the term plan of action conforms with common terminology (see, *e.g.*, Osborne and Rubinstein (1999)). A plan ϕ^i is not a strategy according to game-theoretic terminology, since it does not describe how player i would play *after* she dropped out, in the counterfactual event where she would not have dropped out.⁶⁷ This being said, we will sometimes refer to plans of actions as strategies, with the hope that no confusion will arise.

1.3.3 Mixed plans of action

One of our results says that only pure strategies will be used in equilibrium. We thus need to introduce mixed strategies, only to be able to discard them later. In finite games, a mixed strategy is a probability distribution over pure ones. Such a definition is problematic for infinite games, and we follow the approach of Aumann (1964). Specifically, we introduce for each player $i = 1, 2$ a copy U_i of the unit interval, endowed with the Lebesgue measure λ_i . The pair (U_i, λ_i) serves as an external randomizing device for player i . The devices of the two players are independent. Thus, we enlarge the probability space (Ω, \mathbf{P}) to $(\Omega', \mathbf{P}') := (\Omega \times U_1 \times U_2, \mathbf{P} \otimes \lambda_1 \otimes \lambda_2)$. All random variables defined over Ω have a natural extension to Ω' . Let \mathcal{B}_i and \mathcal{C}_i denote respectively the Borel and the trivial σ -algebra over U_i . In the extended framework, the private information available to player i at stage n is the σ -algebra $\tilde{\mathcal{F}}_n^i = \mathcal{F}_n^i \otimes \mathcal{B}_i \otimes \mathcal{C}_j$ over $\Omega \times U_i \times U_j$. In words, player i now

⁶A pure strategy of player i would specify what to do at each decision node. It can be defined as a sequence $d_n^i(\blacksquare), d_n^i(k), k \in \mathbf{N}, n \in \mathbf{N}$ of \mathcal{F}_n^i -measurable functions, with values in the set {Drop, Continue}. Since we focus on Nash equilibrium analysis, plans of actions will suffice.

⁷This is not perfectly accurate, since $\tau^i(k)$ is defined on Ω , including those ω 's such that $\tau^i(\blacksquare) \leq k$. To be fully consistent, the stopping time $\tau^i(k)$ should only be defined on the event $\{\tau^i(\blacksquare) > k\}$. In particular, any two strategies ϕ^i and ϕ'^i of player i with $\tau^i(\blacksquare) = \tau'^i(\blacksquare) = 0$ are equivalent.

knows both her past payoffs (Ω -component) and the realization of her own randomizing device (U_i -component), but does not receive any information on the realization of the other player's randomizing device (U_j -component).

A *mixed plan of action* of player i is a family $\tau^i(k)$ ($k \in \mathbf{N} \cup \{\square\}$) of stopping times for the filtration $(\tilde{\mathcal{F}}_n^i)_{n \in \mathbf{N}}$, defined over (Ω', \mathbf{P}') , and such that $\tau^i(k) > k$, \mathbf{P}' -a.s. for each $k \in \mathbf{N}$.

Throughout the paper, we will assume w.l.o.g. that the probability space \mathbf{P} is rich enough to accommodate a randomizing device for each player, and take $\mathbf{P}' = \mathbf{P}$. Then mixed plans differ from pure plans only in that the latter consist of stopping times for the larger filtration $(\tilde{\mathcal{F}}_n^i)_{n \in \mathbf{N}}$.

Alternatively, one may wish to describe a mixed plan by the corresponding c.d.f.'s (cumulative distribution function). One would define $F_n^i(\square)$ to be the λ_i -probability that player i drops out at or before stage n , if player j has not dropped out earlier. $F_n^i(k)$ is defined analogously. Hence, both stochastic processes $(F_n^i(\square))_{n \in \mathbf{N}}$ and $(F_n^i(k))_{n \in \mathbf{N}}$ are non-decreasing, $[0, 1]$ -valued, adapted to the filtration $(\mathcal{F}_n^i)_{n \in \mathbf{N}}$. For the equivalence between the two descriptions the reader is referred to, *e.g.*, Touzi and Vieille (2001).

1.3.4 Payoffs

Let a strategy profile (ϕ^1, ϕ^2) be given, and denote by $t^i(\phi^1, \phi^2) \in \mathbf{N} \cup \{+\infty\}$ the time at which player i drops out. When there is no risk of confusion, we will abbreviate $t^i(\phi^1, \phi^2)$ to t^i .

Player i 's total payoff is

$$r^i(\phi^1, \phi^2) := \sum_{n=0}^{t^i(\phi^1, \phi^2)-1} \delta^n X_n^i,$$

and her expected payoff is $\gamma^i(\phi^1, \phi^2) := \mathbf{E}[r^i(\phi^1, \phi^2)]$, where the expectation is taken w.r.t. \mathbf{P} .

The existence of the latter expectation will have to be proven, for two reasons. First, one needs to check that the random series $\sum_{n=0}^{+\infty} \delta^n X_n^i$ converges \mathbf{P} -a.s., *i.e.*, that the total payoff to player i is well-defined, if she stays in forever. One then has to check that r^i is \mathbf{P} -integrable, for each strategy profile. This is done in Lemma B.2.

1.3.5 Cut-off strategies

We here describe a simple class of strategies, under which the joint flow of public and private information is processed in a particularly simple way. Given a stage $n \in \mathbf{N}$, we denote by $p_n^i = \mathbf{P}(\Theta = H \mid \mathcal{F}_n^i)$ the conditional probability that $\Theta = H$, given the private information available at stage n . We call p_n^i the *private belief* at stage n . Alternatively, this would be the belief of player i over Θ if she were playing the game alone. Note that $p_0^i = p_0$.

A *cut-off plan* of player i is a plan under which exit is triggered as soon as her private belief falls below a predetermined cut-off, which depends both on the current stage and on the activity status of the other players.

Definition 1.2 *A plan of action ϕ^i of player i is called a cut-off strategy if, for each $\alpha \in \mathbf{N} \cup \{\blacksquare\}$, there exists a map $\pi^i(\alpha) : \mathbf{N} \rightarrow [0, 1]$, such that $\tau^i(\blacksquare) = \inf\{n \geq 0 : p_n^i \leq \pi_n^i(\blacksquare)\}$ and $\tau^i(k) = \inf\{n > k : p_n^i \leq \pi_n^i(k)\}$, for each $k \in \mathbf{N}$.*

As is customary, we omit the qualifier “**P**-a.s.” when speaking of the equality of two r.v.’s. In particular, stopping times and the private belief p_n^i are only defined up to a **P**-null set, and the equalities in Definition 1.2 are meant to hold **P**-a.s.

Note that dropping out at stage 0 can be interpreted as a cut-off plan.

As long as player j is active, player i will compare her private belief at stage n to the cut-off $\pi_n^i(\blacksquare)$, and will drop out if it falls below the cut-off. If player j drops out of the game, say at stage k , player i adapts to the new situation by comparing her private belief to the different cut-offs $\pi_n^i(k)$. By definition, a cut-off plan is a pure plan. Note that the private information is incorporated in a cut-off strategy because the private belief is compared at each stage with the cut-off; the public information is incorporated in the strategy in two ways. Since $\pi_n^i(\blacksquare)$ and $\pi_{n+1}^i(\blacksquare)$ may differ, the cut-offs depend on the fact that the other player stays in one additional stage. Since $\pi_n^i(\blacksquare)$ and $\pi_n^i(k)$ may differ, the cut-offs depend on the fact that the other player dropped out.

We warn against a possible misinterpretation: it would not make sense to define a cut-off plan as depending only on player i ’s posterior belief over Θ , given *all* available information. Indeed, the posterior beliefs of player i cannot be computed without specifying first the *strategy* used by player j . Thus, player i ’s decision at stage n would depend on player j ’s strategy, and not only on the information set that is reached. Indeed, suppose player j

remains in the game at stage 1. If player j 's strategy is “drop out once the private belief is below 0.001”, the fact that she did not drop out is hardly informative, whereas if her strategy is “drop out once the private belief is below 0.99”, the fact that she did not drop out may be a very good news concerning the type of the bandit. Obviously, the posterior belief of player i will not be the same in the two cases.

In a sense, a cut-off strategy satisfies the Markov property. That is, a cut-off strategy depends on the triple (n, p_n^i, t_n^j) . It will later be shown that the sequence $(n, p_n^i, t_n^j)_{n \in \mathbf{N}}$ is a Markov chain, for every strategy profile, and may therefore serve as a state variable. When compared with Maskin and Tirole's (1999) definition, it may be argued that the private belief p_n^i summarizes the payoff relevant content of the private information, while the status t_n^j summarizes the payoff relevant content of the public information.

Note that at this point we just define a class of simple strategies. We will prove later that such strategies are rich enough to be used as equilibrium strategies.

1.3.6 Multi-player games

In more-than-two-player games, each player i has to keep track of the activity status of every other player. Hence, a (pure) plan of action is defined as a family $\tau^i(\vec{k})$ of stopping times, indexed by $I \setminus \{i\}$ -vectors $\vec{k} = (k_j)_{j \in I \setminus \{i\}}$, with components $k_j \in \{\blacksquare\} \cup \mathbf{N}$. The component k_j is the status variable relative to player j . In addition, whenever $\vec{k} \neq (\blacksquare, \dots, \blacksquare)$, the stopping time $\tau^i(\vec{k})$ must satisfy $\tau^i(\vec{k}) > \max\{k_j : k_j \neq \blacksquare\}$. The definition of mixed and cut-off plans extends in a straightforward way. The payoff can be computed as in two-player games. Details are omitted.

2 Main results

We here state and discuss our main results. To start with, and to serve as a benchmark, we quote a result relative to the one-player case.

Theorem 2.1 (one-player game) *There is a unique optimal stopping time τ^* . Moreover, $\tau^* := \inf\{n \geq 0 : p_n \leq \pi^*\}$, for some $\pi^* \in (0, 1)$.*

This theorem follows from classical results in the bandit literature, see, *e.g.*, Ferguson (2004) and the references therein.

2.1 General results

The results of this subsection are valid, irrespective of the number of players. We start with an equilibrium existence result.

Theorem 2.2 (existence) *The game has a symmetric equilibrium.*

As we argue below, there might be several symmetric equilibria. This stands in sharp contrast to one of the results of Bolton and Harris (1999) and Cripps et al. (2004), who prove that in their models there is a unique symmetric Markov equilibrium .

Throughout, we assume all players to be active in stage 0 at equilibrium. This assumption is automatically satisfied if $p_0 > \pi_*$. Note also that, if $p_0 \leq \pi_*$, it is an equilibrium for all players to drop out in stage 0. If $p_0 \leq \pi_*$, other equilibria may exist, in which all players enter the game.

We now state a first result of qualitative nature.

Theorem 2.3 (structure) *Under Assumptions A1-4, all Nash equilibria are in cut-off strategies.*

We will actually prove a stronger result: the best reply to any strategy profile is a cut-off strategy. Thus, it is always sufficient to consider cut-off strategies.

Theorem 2.3 calls for comments. Since cut-off strategies are pure, Theorem 2.3 asserts that *all* equilibria are pure. Moreover, in *all* equilibria the information is processed in a simple way. Thus, (for two-player games) as long as player j is active, player i compares her private beliefs to some cut-offs. If player j drops out, she adapts by updating the cut-offs to which her private beliefs are compared. The (informational) interaction is incorporated in the way cut-off values depend on the stage and on the public information.

Note also that the conclusion requires only weak assumptions on \mathbf{P}_H and \mathbf{P}_L . We do not assume that the distribution of payoffs under \mathbf{P}_H stochastically dominates the distribution of payoffs under \mathbf{P}_L , or that the likelihood ratio $f_H(\cdot)/f_L(\cdot)$ is monotonic increasing. This latter fact has a noteworthy implication. In the absence of a monotonic likelihood ratio, the private beliefs need not be monotonic in one's own payoffs: high payoffs need not be good news, see Milgrom (1982). As a consequence, equilibrium behavior need not be monotonic in payoffs: player i may choose to drop out in stage 1 with a very high payoff in stage 0, but not for some lower values.

2.2 Qualitative results

2.2.1 Two-player games

We now provide more precise statements on the equilibrium cut-offs. We start with two-player games, since statements are simpler and the results sharper. When i denotes a player, j is the other player. First, we focus on the monotonicity of equilibrium cut-offs.

Theorem 2.4 (non-increasing cut-offs) *Let an equilibrium with cut-offs (π_n^i) be given. Then the following properties hold for each of the two players.*

P1 *The cut-off sequence $(\pi_n^i(\blacksquare))_{n \in \mathbf{N}}$ is non-increasing;*

P2 *The cut-off sequence $(\pi_n^i(k))_{n \in \mathbf{N}}$ is constant, for each $k \in \mathbf{N}$.*

In light of the one-player case (see Theorem 2.1), statement **P2** is quite intuitive: once player j drops out from the game, player i faces a one-player problem with a new posterior that takes into account all her information so far, including the fact that player j dropped out. By Theorem 2.1, she will find it optimal to drop out once her posterior belief falls below π^* .⁸

We now discuss **P1**. If player i enters stage n with a private belief equal to $\pi_n^i(\blacksquare)$, she will be indifferent between dropping out and staying in. Assume she stays in and reaches stage $n + 1$ with the same private belief $\pi_n^i(\blacksquare)$, and sees that player j still did not drop out. Player i 's posterior belief that $\Theta = H$ is then higher in stage $n + 1$ than what it was in stage n : all else being equal, the more stages player j stays in the game, the better news it is on Θ (provided player j uses a cut-off strategy). However, the prospects of deriving in the future additional information over Θ through player j 's behavior are modified in a complex way, since it is not clear how the belief over player j 's stopping time depends on the private belief.

We take this opportunity to clarify the role of the second half of Assumption **A4** ($\inf C = 0$). Assume by contrast that the likelihood ratio $f_H(X_0^i)/f_L(X_0^i)$ only takes values close to one, so that information on Θ is disclosed very slowly. If the initial belief p_0 is high enough, a player will

⁸This intuition is incomplete in one respect, since that posterior belief differs from the private belief of player i . As will be shown, the private belief is in one-one relation with the posterior belief, and the constant cut-off π^* on the posterior belief translates to a constant cut-off on the private belief.

always find it optimal to remain active in stage 1, irrespective of the payoff received in stage 0, since that payoff cannot convey very bad news on Θ . The equilibrium cut-off $\pi_1^i(\blacksquare)$ must reflect this property and be such that $\mathbf{P}(p_1^i \leq \pi_1^i(\blacksquare)) = 0$. Obviously, the value of $\pi_1^i(\blacksquare)$ can then be lowered arbitrarily without affecting the equilibrium property. In particular, statement **P1** will not hold without proper qualifications. Moreover this assumption implies that for any private belief p_n^i at stage n the belief at stage $n + 1$ can be arbitrarily low so that for any cut-off strategy there is a positive probability of dropping out at each stage; this avoids situations in which a player dropping out at some stage k would be off-equilibrium-path.

We next provide some results on the cut-off values.

Theorem 2.5 (bounds on cut-offs) *Let an equilibrium with cut-offs (π_n^i) be given. Then the following statements hold:*

P3 $\lim_{n \rightarrow \infty} \pi_n^i(\blacksquare) = 0$ for at least one player i ;

P4 For each $n \in \mathbf{N}$ and $i = 1, 2$, one has $\pi_n^i(\blacksquare) < \pi^*$, where π^* is the optimal one-player cut-off level;

P5 For each player $i = 1, 2$ there is a stage $k \in \mathbf{N}$ such that $\pi^* < \pi_n^i(k)$.

Assume moreover that the equilibrium is symmetric (so that $\pi_n^i = \pi_n^j$). Then:

P6 there exist $0 < c_1 < c_2$, such that

$$c_1 \mathbf{P}_L(\tau^i(\blacksquare) \geq n) \leq \pi_n^i(\blacksquare) \leq c_2 \mathbf{P}_L(\tau^i(\blacksquare) \geq n),$$

for every $i \in I$, and every stage $n \in \mathbf{N}$.

Statement **P4** says that having an active opponent should be interpreted as good news on Θ . Moreover, the very fact that player j might drop out some time in the future creates a positive informational externality, that does not exist in the one-player case. Both effects add up, and explain why player i should be willing to accept lower private beliefs than if she were alone.

In the one-player case, when the discount rate goes to 1, the optimal cut-off π^* goes to 0. Hence, **P4** implies that as $\delta \rightarrow 1$, all equilibrium cut-offs $(\pi_n^i(\blacksquare))$ converge to zero. This is intuitive, since the cost of experimentation then drops to zero.

According to **P3**, at least one player must be asymptotically willing to accept arbitrarily low private beliefs. Assume, say, that all cut-offs of player j are higher than some $\alpha > 0$. If player j does not drop out as time goes by, player i deduces that all private beliefs of player j have constantly remained above α which is very good news on the state; she will therefore drop out only if *her* private beliefs are extremely low, *i.e.* if she got very precise and negative private information. Note that, in the case of a symmetric equilibrium, the common sequence $\pi_n^i(\blacksquare)$ of cut-offs then converges to zero.

We now discuss **P5**. After player j has dropped out from the game in stage k , player i will stay active until her posterior belief falls below π^* . Altogether, the fact that player j dropped out from the game some time in the past, is bad news on Θ . However, the news that player j dropped out *specifically* in stage k might be good news, since that implies that player j did not stop *before* stage k , which *is* good news. For such values of k , the posterior belief of player i will be higher than her private belief, and player i 's equilibrium cut-off will satisfy $\pi_n^i(k) \leq \pi^*$. According to **P5**, not all values of k can have this paradoxical feature: there is at least one stage k such that seeing the other player drop out at k can be interpreted as bad news on the state.

Statement **P6** provides implicit information on the rate of convergence of the cut-offs to zero. Note that the probability $\mathbf{P}_L(\tau^i(\blacksquare) \geq n)$ of survival depends on past cut-offs $\pi_1^i(\blacksquare), \dots, \pi_{n-1}^i(\blacksquare)$, and on the payoff distributions. The higher the cut-offs, the smaller this probability. If the sequence $(\pi_n^i(\blacksquare))$ decreases very fast to zero, the probability of survival would decrease slowly, and the left hand-side inequality would fail to hold. If the cut-offs decrease very slowly, the survival probability would decrease very fast, and the right hand-side equality would fail to hold.

We next provide a first step in the analysis of the efficiency of the equilibrium. Plainly, first best considerations require that players drop as early as possible if $\Theta = L$, and stay active forever if $\Theta = H$. We prove that all players stay active forever with a positive probability if $\Theta = H$ and that all players drop out in finite time if $\Theta = L$.

Recall that $t^i \in \mathbf{N} \cup \{+\infty\}$ is the stage at which player i drops out.

Theorem 2.6 (efficiency) *Let an equilibrium with cut-offs (π_n^i) be given. Then the following statements hold:*

P7 Let $t := \min\{t^1, t^2\}$. The law of t under \mathbf{P}_H stochastically dominates the law of t under \mathbf{P}_L .

P8 $\mathbf{P}_H(t_i = +\infty, \forall i \in I) > 0$.

P9 $\mathbf{P}_L(t^i < +\infty, \forall i \in I) = 1$.

These three statements provide qualitatively similar information. According to **P7**, the first player to stop will stop earlier if the state is L than if it is H .

According to **P8-P9**, all players drop out in finite time if $\Theta = L$ whereas, if $\Theta = H$, there is a positive probability that all players will stay active forever. Thus, it cannot happen that the common type is L , but both players stay in because each observes that the other stays in.

2.2.2 N -player games

A profile is symmetric if, for each stage $n \in \mathbf{N}$, and every two players i and j , one has $\pi_n^i(\vec{\alpha}) = \pi_n^j(\vec{\alpha}')$ whenever the vector $\vec{\alpha}'$ is obtained by from the vector $\vec{\alpha}$ by swapping coordinates i and j .

Once the obvious modifications on the cut-offs are made, all the results stated above are valid irrespective of the number N of players, with the exception of **P1** and **P2**. Indeed, as will be clear in Section 6.1, our proofs in this case require that the number of players be equal to two. Statement **P6** has to be amended to $c_1 \mathbf{P}_L(\tau^i(\blacksquare) \geq n)^{N-1} \leq \pi_n^i(\blacksquare) \leq c_2 \mathbf{P}_L(\tau^i(\blacksquare) \geq n)^{N-1}$, for some $c_1, c_2 \in (0, +\infty)$.

2.3 Large games

When the number of players is large, equilibria can be described almost explicitly.

Assuming all players enter the game in stage 0, the only information available in stage 1 is the payoff received in stage 0. Assume player i remains active in stage 1 and then discovers, for some reason, the true common type. Depending on the state, she will choose to drop out at once, or to stay in forever. Hence, player i 's expected payoff, as viewed from stage 1, is $p_1^i \mathbf{E}_H[X_n^i] + (1 - \delta)(1 - p_1^i) \mathbf{E}_L[X_n^i] = 0$. This is the highest payoff player i may hope for, if she stays in.

We define p_* to be the value of p_1^i at which a player is indifferent between dropping out or staying in one additional stage and then being told Θ , *i.e.*, the following equality is satisfied:

$$p_* \mathbf{E}_H[X_n^i] + (1 - \delta)(1 - p_*) \mathbf{E}_L[X_n^i] = 0. \quad (2)$$

Since $\mathbf{E}_H[X_n^i] > 0 > \mathbf{E}_L[X_n^i]$, we have $0 < p_* < 1$.

Theorem 2.7 describes the equilibrium behavior in the first two stages, as the number of players gets large. In the first stage, all players use a cut-off that gets closer to p_* as N increases. In stage 2, a player drops out or not, depending on the fraction of players who dropped out in the previous stage, but (almost) irrespective of her own private belief, and herding takes place. Similar results can be obtained for subsequent stages.

Theorem 2.7 (*large games*) *For each $N \in \mathbf{N}$, let ϕ^N be an equilibrium of the N -player game, with cut-offs $(\pi_n^{N,i})$. The following statements hold.*

- *One has*

$$\lim_{N \rightarrow +\infty} \sup_{i=1, \dots, N} |\pi_1^{N,i}(\blacksquare) - p_*| = 0 \text{ and } \lim_{N \rightarrow +\infty} \sup_{i=1, \dots, N} \pi_n^{N,i}(\blacksquare) = 0, \text{ for every } n > 1.$$

- *Let $\vec{\alpha}^N$ be a possible status vector of all players $j \neq i$, and let $\rho^N := \frac{1}{N-1} |\{j: \alpha_j^N = 1\}|$ be the fraction of players who stayed in at stage 1. Assume that $\lim_{N \rightarrow +\infty} \rho^N = \rho \in [0, 1]$. There is a threshold $\rho_* \in (0, 1)$ such that the equilibrium cut-offs in stage 2 satisfy the following.*

- *If $\rho < \rho_*$, then $\lim_{N \rightarrow +\infty} \left\{ \sup_{i=1, \dots, N} \pi_2^{N,i}(\vec{\alpha}^N) \right\} = 0$.*
- *If $\rho > \rho_*$, then $\lim_{N \rightarrow +\infty} \left\{ \inf_{i=1, \dots, N} \pi_2^{N,i}(\vec{\alpha}^N) \right\} = 1$.*

The driving force behind the first statement is a large number intuition. In large games, given $\Theta = \theta$, it is very likely that the proportion of players who drop out at stage 1 be close to a constant ρ_θ . It is higher when the common type is L than when it is H : $\rho_L > \rho_H$. Thus, in stage 2, all active players will be able to infer Θ from the proportion of players who dropped out at stage 1. Knowing this, all players at stage 1 will compute the option value of waiting, as if the common type is going to be revealed at the next stage. In addition, all players that are still active in stage 2 drop out if they deduce that $\theta = L$ and stay in for ever if they conclude that $\theta = H$.

Roughly, if $\Theta = L$, all players will either drop out in stage 1 or in stage 2. On the other hand, if $\Theta = H$, a fixed proportion will drop out in stage 1 and other players will remain active forever. In particular, when the population is large (i) full learning of the state occurs after one stage (for all the still active players), and (ii) the fraction of players dropping out in stage 1 if $\Theta = H$ precisely describes the efficiency loss inherent to the learning problem.

Thus, with high probability, the fraction of players ρ dropping out after a single stage is either close to ρ_L or to ρ_H . The statement 2 describes what happens in the rare event where this does not occur. As stated, there is a threshold ρ_* on ρ below which (resp. above which) player i will drop out (resp. stay in), irrespective of her own private belief.

Caplin and Leahy (1994) consider a similar setup with a continuum of players and directly analyze the deterministic equilibrium learning process, focusing on *symmetric* equilibrium profiles. By contrast, our result does not impose any symmetry restriction on the equilibrium, and studies the convergence properties, as the set of players converges to the continuum. In a sense, Theorem 2.7 may be seen as supporting Caplin and Leahy's analysis in two respects: (i) as $N \rightarrow +\infty$, all equilibria become asymptotically symmetric and (ii) the continuum-player case appears as the limit of large games.

2.4 Discussion

The statements above provide only few qualitative insights. Additional comparative statics would be desirable. For instance, it would be interesting to understand the effect on equilibrium cut-offs of the variance of the payoff distributions, of the number of players, etc. Such results would necessitate a fine analysis of the optimal continuation payoff. We briefly describe below the determinants of the optimal continuation payoff, in order to understand the inherent difficulties.

We let a cut-off strategy ϕ^2 of player 2 be given. In stage n , player 1 will choose to drop out if her optimal continuation payoff does not exceed zero, and will otherwise choose to continue. Assume player 1 reaches stage n with a private belief equal to p^1 . The continuation payoff induced by a given (cut-off) strategy ϕ^1 against ϕ^2 can be computed as follows (assuming player 2 is still alive).

Step 1: Assess the joint (conditional) law \mathbf{Q} of the common type Θ and of the private belief p_n^2 held by player 2, given all available information.

Observe that \mathbf{Q} depends on p^1 , on n , and on the cut-offs $\pi_k^2(\blacksquare)$ ($k < n$) associated with ϕ^2 in previous stages.

Step 2: For each $\theta = L, H$ and $p^2 \in (0, 1)$, find the continuation payoff λ^1 to player 1, in the event where the common type is θ and player 2 reached stage n with a private belief p^2 . To be specific: the first player to drop out will do so at stage $t := \inf\{k \geq n : p_k^i \leq \pi_k^i(\blacksquare), \text{ for some } i = 1, 2\}$. If $p_t^1 \leq \pi_t^1(\blacksquare)$, player 1 drops out and her total continuation payoff (discounted back to stage n) is $\delta^{-n} \sum_{k=n}^{t-1} \delta^k X_k^1$. Otherwise, if $p_t^1 > \pi_t^1(\blacksquare)$, player 1 continues until stage $t^1 := \inf\{k > t : p_k^1 \leq \pi_k^1(t)\}$. Thus, the computation of λ^1 involves the (conditional) joint law of t and of the payoffs $(X_k^1)_{k \geq n}$, given that $(\Theta, p_n^2) = (\theta, p^2)$. Observe that λ^1 depends on θ , on private beliefs' levels p^1, p^2 , and on the thresholds $\pi_k^i(\blacksquare)$ ($k \geq n$) associated with ϕ^i in the present and following stages.

The continuation payoff induced by the strategy pair (ϕ^1, ϕ^2) is the expectation of λ^1 under \mathbf{Q} .

For comparative statics purposes, it is instructive to analyze the dependency of \mathbf{Q} and λ^1 on each parameters.

As for \mathbf{Q} , the case is clear in most respects. Indeed,⁹

- \mathbf{Q} is increasing in p^1 (in the sense of first-order stochastic dominance). The intuition is that the posterior belief that $\Theta = H$ is increasing in p^1 , while the private belief p_n^2 will tend to be higher if $\Theta = H$ than if $\Theta = L$.
- \mathbf{Q} is increasing with the cut-offs $\pi_k^2(\blacksquare)$, $k < n$. The intuition is the same as above.
- Finally, for fixed p^1 and ϕ^2 , the posterior belief p^2 that $\Theta = H$ increases with n (this relies on ϕ^2 being a cut-off strategy). On the other hand, private beliefs of player 2 tend to *decrease* with time if $\Theta = L$. If we assume moreover that player 2 is active, it is unclear how the distribution of p_n^2 changes with n if the state is L .¹⁰ The overall change in \mathbf{Q} is unclear.

⁹The first and third statements are formally proven below.

¹⁰If the state is H , the conditional distribution of p_n^2 , given that player 2 is still active, unambiguously increases with n .

As for λ^1 , the situation is less clear. When p^2 is lowered to $\tilde{p}^2 < p^2$, it is more likely that player 2 will be the first one to drop out, and public information arrives earlier. In many cases – but not all – player 1 will then drop out earlier in the new situation. Observe that this conclusion holds irrespective of θ . Consequently, the continuation payoff will increase if $\Theta = L$, but decrease if $\Theta = H$.

When cut-offs associated with ϕ^2 are increased, this analysis remains valid. Finally, the change of λ^1 with n is ambiguous.

3 An Example

To illustrate some aspects of the model, we now analyze a simple three-stage example. Stages are numbered 0, 1 and 2. The payoff X_n^i takes only the values¹¹ $+1$ and -1 , and, for simplicity, the discount factor $\delta = 1$. The probability $\mathbf{P}_\theta(X_n^i = +1)$ is equal to q if $\theta = H$, and to $1 - q$ if $\theta = L$, where $q \in (1/2, 1)$.

We analyze several versions. First, we compute the optimal strategy in the one-player case. Next, we deal with a two-player game, in which all information is public – that is, payoffs are publicly observed. Finally, we describe equilibria, assuming payoffs are privately observed. Computations are omitted.

3.1 The one-player problem

We analyse the one-player version, and omit all superscripts i . Let p_n be the belief at stage n that $\Theta = H$, given the payoffs received in the previous stages.

The expected payoff in stage n , given past payoffs, is therefore

$$p_n(2q - 1) + (1 - p_n)(1 - 2q) = (2p_n - 1)(2q - 1).$$

The optimal strategy τ_* is found using backward induction. If still active at the beginning of stage 2, the decision-maker will drop out if $(2p_2 - 1)(2q - 1) < 0$, *i.e.* $p_2 < 1/2$, and will otherwise remain active.

¹¹Because of this assumption, this example does not fit into the general model.

Straightforward computations show that: (i) if active until stage 1, the decision-maker drops out if and only if $p_1 < \frac{2-q}{3}$, and (ii) the decision-maker does not enter the game if and only if¹² $p_0 < p_* := \frac{3-3q+q^2}{4-2q+2q^2}$.

3.2 Two players, observable payoffs

In this variant, the information available to both players is identical. Therefore, at a symmetric profile, both players drop out at the same time.

There is a unique symmetric equilibrium. At equilibrium, both players drop out simultaneously, depending on their (common) posterior belief over Θ . This equilibrium is also computed using backward induction.

We denote by q_n the posterior belief that $\Theta = H$, given all information available at stage n , that is, given the payoffs received by both players, in previous stages.

It turns out that: (i) if active until stage 2, the players drop out if and only if $q_2 < 1/2$, (ii) if active until stage 1, the players drop out if and only if $q_1 < \frac{1 + (1-q)^2}{2 + q^2 + (1-q)^2}$, and (iii) the players enter the game if and only if $q_0 = p_0 \geq p_{**} := \frac{1}{2} \times \frac{3 - 4q + q^2 + 2q^3 - q^4}{2 - 2q + q^2 + 2q^3 - q^4}$.

3.3 Two players, non-observable payoffs, observable actions

We now assume that payoffs are privately observed, while decisions are public. If $p_0 \leq p_*$, it is an equilibrium for both players to drop out at stage 0. We describe below symmetric equilibria in which both players enter the game. Since players may receive different payoffs, they need not drop out simultaneously.

As a first step, we focus on equilibria in which a player's decision in stage 1 is independent of the payoff received in stage 0. Then, player 1 cannot infer anything from player 2's behavior, neither at stage 1 nor at stage 2. As a consequence, the equilibrium strategy of both players must coincide with τ_* , the equilibrium strategy in the one-player problem.

It remains to be checked when (τ_*, τ_*) is an equilibrium. Since τ_* is independent of the player's payoffs one must have $p_1^i \geq \frac{2-q}{3}$: the player

¹²Note that the sequence of optimal cut-offs is increasing. This is a last-stage effect.

remains active in stage 1 even if her payoff in stage 0 is -1 . This reduces to $p_0 > \frac{q(2-q)}{1+2q-2q^2}$. In words, such an equilibrium exists only if the prior probability that $\Theta = H$ is high enough.

We now turn to equilibria in which a player's decision in stage 1 depends on her payoff in stage 0. If player 1 remains active in stage 1 – that is, if $X_0^1 = 1$ – she will infer player 2's payoff X_0^2 in stage 0 from her decision in stage 1. As a consequence, the information available to player 1 in stage 1, is X_0^1 , while her information in stage 2 consists of $X_0^1 = 1$, X_1^1 and X_0^2 .¹³

One can check that there is only one possible such equilibrium. In this (symmetric) profile τ , player 1 remains active in stage 2 only if at least one out of the two payoffs X_1^1 and X_0^2 is equal to one. Thus, she will remain active if either player 2 did not drop out at stage 1, or if her own payoff in stage 1 was one. For τ to be an equilibrium, the following conditions should be met. (i) It is optimal for player 1 to stay in in stage 1 only if $X_0^1 = 1$, and (ii) it is optimal for player 1 to stay in in stage 2 unless $X_1^1 = X_0^2 = -1$. These conditions amount to

$$1 - q \leq p_0 \leq \frac{q(2 - q^2)}{q(2 - q^2) + (1 - q)(2 - (1 - q)^2)}.$$

3.4 Comments

Some comments are in place.

1. **Multiplicity of equilibria.** In the third variant, there need not be a unique equilibrium. For p_0 in the (non-empty) interval $[\frac{q(2-q)}{1+2q-2q^2}, \frac{q(2-q^2)}{q(2-q^2)+(1-q)(2-(1-q)^2)}]$, there are two symmetric equilibria in which both players enter the game. Besides, for $p_0 \in [1 - q, p_*]$ there are two equilibria: the profile in which no player enters the game, and the strategy described in the end of section 3.3.
2. **Dropping out in stage 0.** In each of the three variants, dropping out in stage 0 is the unique equilibrium, as soon as p_0 is below a certain level. This level is p_* in the one-player case, p_{**} in the two-player game with public payoffs, and $1 - q$ in the two-player game with private

¹³Since the action of player 1 in stage 1 reveals her payoff in stage 0, if player 1 stayed in in stage 1, we must have $X_0^1 = 1$.

payoffs. The positive externality is strongest when payoffs are publicly observed. This is reflected by the inequality $p_{**} < 1 - q < p_*$.

3. **Payoff comparison.** The equilibrium payoff in the two-player games is (weakly) higher than the optimal payoff in the one-player problem. Indeed, since the interaction is purely informational, playing the one-player optimal strategy in the two-player game yields a payoff equal to the one-player optimal payoff, so the player's best reply yields at least the optimal payoff in the one-player problem.

Surprisingly, the (common) equilibrium payoff when payoffs are public may be strictly *lower* than the equilibrium payoff when payoffs are private information, *e.g.* for $p_0 = 0.9$ and $q = 0.787$.

4 Beliefs

We here define the two kinds of beliefs. The private belief of a player incorporates only her (private) payoffs. The posterior belief includes all available information. We focus on the evolution of the private belief through time, and on the relation between the two beliefs. In particular, it is instructive to know when the available information on player j 's does induce player i to optimism – so that player i 's posterior belief will be higher than her private belief. Technical proofs are relegated to Appendix A.

4.1 Private beliefs

4.1.1 The private belief

The private belief $p_n^i := \mathbf{P}(\Theta = H \mid \mathcal{F}_n^i)$ of player i in stage n , is the conditional probability that $\Theta = H$, based on her private information. Thus, $p_0^i = p_0$ for each $i \in I$.

A version of the private belief p_n^i , obtained by Bayes rule, is given by the equality

$$\frac{p_n^i}{1 - p_n^i} = \frac{p_0}{1 - p_0} \times \prod_{k=0}^{n-1} \frac{f_H(X_k^i)}{f_L(X_k^i)}. \quad (3)$$

It expresses how the likelihood ratio of H and L (left-hand side) changes with the initial likelihood ratio and the likelihood ratio of the payoffs.

By assumption **A4** and (3) we have:

Lemma 4.1 *The r.v. p_n^i has a density, for every player $i \in I$ and every stage $n \in \mathbf{N}$.*

4.1.2 The evolution of the private belief

We recall without proof the following well-known properties of the sequence $(p_n^i)_{n \in \mathbf{N}}$: beliefs tend to increase (resp. to decrease) if the state is H (resp. L), and are asymptotically correct.

Proposition 4.2 *Under \mathbf{P} , the sequence (p_n^i) is a martingale (for the filtration (\mathcal{F}_n^i)).*

Under \mathbf{P}_H , (p_n^i) is a submartingale that converges a.s. to one.

Under \mathbf{P}_L , (p_n^i) is a supermartingale that converges a.s. to zero.

In addition, for a fixed stage n , the private belief p_n^i is likely to be higher if the state is actually H than if it is L . The basic result in this vein is Lemma 4.3 below.

Lemma 4.3 *The law of p_1^i under \mathbf{P}_H stochastically dominates (in the first-order sense) the law of p_1^i under \mathbf{P}_L .¹⁴*

We now extend this result in two ways. First, we show that a sharper result holds in our setup. It relies on the fact that p_1^i has a density, and that both distributions \mathbf{P}_H and \mathbf{P}_L differ.

Proposition 4.4 *For $\theta \in \{L, H\}$, let F_θ denote the c.d.f. of p_1^i under \mathbf{P}_θ . For each $p \in (0, 1]$ such that $F_H(p) < 1$ and $F_L(p) > 0$, one has*

$$F_H(p) < F_L(p).$$

Next, we prove that the result of Lemma 4.3 generalizes to the sequence of private beliefs.

Proposition 4.5 *For each $n \in \mathbf{N}$, the law of the vector (p_1^i, \dots, p_n^i) under \mathbf{P}_H stochastically dominates the law of (p_1^i, \dots, p_n^i) under \mathbf{P}_L .¹⁵*

Moreover, for each $x_1, \dots, x_n \in [0, 1]$, one has $\mathbf{P}_L(p_1^i > x_1, \dots, p_n^i > x_n) \leq \mathbf{P}_H(p_1^i > x_1, \dots, p_n^i > x_n)$.

¹⁴Let μ and λ be two distributions over the real line. We say that μ *stochastically dominates* λ if $\mu((-\infty, x)) \leq \lambda((-\infty, x))$ for every $x \in \mathbf{R}$.

¹⁵Let μ and λ be two distributions over \mathbf{R}^n . We say that μ *stochastically dominates* λ if $\mu(C) \leq \lambda(C)$ for every set C of the form $(-\infty, x_1) \times (-\infty, x_2) \times \dots \times (-\infty, x_n)$.

The second assertion of proposition 4.5 does not follow from the first.

An immediate consequence of Proposition 4.5 is that, assuming the other player is still active, a player will tend to drop out earlier when $\Theta = L$ than when $\Theta = H$.

Corollary 4.6 *For every cut-off strategy $\phi^i = (\tau^i(\alpha))_{\alpha \in \mathbf{N} \cup \{\blacksquare\}}$ we have*

$$\mathbf{P}_H(\tau^i(\blacksquare) \geq n) \geq \mathbf{P}_L(\tau^i(\blacksquare) \geq n), \quad \forall n \in \mathbf{N}.$$

Another simple consequence is that when the common type is H there is a positive probability that a player who uses a fixed-cut-off strategy will stay in forever.

Corollary 4.7 *Let $p < p_0$ and $i = 1, 2$ be given, and set $t = \inf\{n \in \mathbf{N} : p_n^i \leq p\}$. Then one has $\mathbf{P}_H(t < +\infty) < 1$.*

Proof. By Proposition 4.2, (p_n^i) is a submartingale under \mathbf{P}_H , bounded by one. Fix $N \in \mathbf{N}$. By the optional sampling theorem applied to the stopping time $\min\{t, N\}$ we obtain

$$p_0 \leq \mathbf{E}_H [p_{\min\{t, N\}}^i] \leq \mathbf{P}_H(t > N) + p \mathbf{P}_H(t \leq N).$$

When N increases to infinity, $\mathbf{P}_H(t \leq N)$ and $\mathbf{P}_H(t > N)$ converge to $\mathbf{P}_H(t < \infty)$ and $1 - \mathbf{P}_H(t < \infty)$ respectively. This yields $\mathbf{P}_H(t < +\infty) \leq \frac{1 - p_0}{1 - p} < 1$. ■

4.2 Posterior beliefs

4.2.1 The posterior belief

We now introduce the posterior belief that takes into account both private and public information. As emphasized earlier, computing the posterior belief requires the knowledge of the other player's strategy. We therefore let an arbitrary strategy $\phi^j = (\tau^j(k))_{k \in \mathbf{N} \cup \{\blacksquare\}}$ of player j be given. The public information at stage n is the status of the players, which is described by the status variables (t_n^i) .

As long as player i is active, her information at stage n is described by the σ -algebra $\mathcal{G}_n^i := \sigma(\mathcal{F}_n^i, t_n^j)$. The *posterior belief* of player i at stage n is defined to be $q_n^i := \mathbf{P}(\Theta = H \mid \mathcal{G}_n^i)$.

Lemma 4.8 *Let a stage $n \in \mathbf{N}$ be given. On the event $\{\tau^j(\blacksquare) \geq n\}$, one has*

$$\frac{q_n^i}{1 - q_n^i} = \frac{p_0}{1 - p_0} \times \prod_{k=0}^{n-1} \frac{f_H(X_k^i)}{f_L(X_k^i)} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)} = \frac{p_n^i}{1 - p_n^i} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)}. \quad (4)$$

On the event $\{\tau^j(\blacksquare) = k\}$ ($k < n$), one has

$$\frac{q_n^i}{1 - q_n^i} = \frac{p_n^i}{1 - p_n^i} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) = k)}{\mathbf{P}_L(\tau^j(\blacksquare) = k)}. \quad (5)$$

The lemma follows by Bayes' rule, using the fact that the σ -algebra \mathcal{F}_n^i and the random variable t_n^j are conditionally independent given Θ . Details are omitted.

By Lemma 4.8, the posterior belief q_n^i depends only on p_n^i and on the status t_n^j of player j . We denote by $Q_n^i(p_n^i, t_n^j)$ this version of q_n^i . On the event $t_n^j = \blacksquare$, the value of $Q_n^i(p_n^i, t_n^j) = Q_n^i(p_n^i, \blacksquare)$ is given by equation (4). On the event $t_n^j = k$, the value of $Q_n^i(p_n^i, t_n^j)$ is given by (5). In particular in this explicit version of q_n^i , the posterior belief is continuous and monotonic increasing in the private belief. Moreover, the dependency of the posterior belief with private belief and public information is fairly simple. We study the dependency between q_n^i and p_n^i more carefully in the next subsection.

4.2.2 Posterior vs. private belief

We here focus on the interplay between the private and posterior beliefs of player i , assuming player j follows a cut-off strategy. We use Lemma 4.8 to disentangle the effect of private and public information on the posterior belief.

Equations (4) and (5) directly imply that if we fix the status of the other player, the posterior belief is increasing in the private belief. This property is stated in the following proposition.

Proposition 4.9 *For each $n \in \mathbf{N}$, the function $Q_n^i(p, \alpha)$ is continuous and increasing in p .*

We now argue that having an active opponent is always good news on Θ , and the posterior belief is always higher than the private belief.

Proposition 4.10 *For each $n \in \mathbf{N}$ and $p \in [0, 1]$, one has $Q_n^i(p, \blacksquare) \geq p$.*

Proof. Let $n \in \mathbf{N}$, and $p \in [0, 1]$ be given. By Corollary 4.6, $\mathbf{P}_H(\tau^j(\blacksquare) \geq n) \geq \mathbf{P}_L(\tau^j(\blacksquare) \geq n)$. The result follows from (4). ■

All we used in the proof is the inequality $\mathbf{P}_H(\tau^j(\blacksquare) \geq n) \geq \mathbf{P}_L(\tau^j(\blacksquare) \geq n)$. It is equivalent to $\mathbf{P}_H(\tau^j(\blacksquare) < n) \leq \mathbf{P}_L(\tau^j(\blacksquare) < n)$, so that $\mathbf{P}_H(\tau^j(\blacksquare) = k) \leq \mathbf{P}_L(\tau^j(\blacksquare) = k)$, for at least one stage $k < n$. For such a k , by (5), one has $Q_n^i(p, k) \leq p$. Thus, we have proven Proposition 4.11.

Proposition 4.11 *For each $n \geq 2$, there is $k < n$ such that $Q_n^i(p, k) \leq p$, for every $p \in [0, 1]$.*

The following proposition asserts that the posterior belief when the opponent is active is always higher than when she has already dropped out.

Proposition 4.12 *For every two stages $k < n \in \mathbf{N}$, and every $p \in [0, 1]$, one has $Q_n^i(p, \blacksquare) > Q_n^i(p, k)$.*

Proof. Let $m, n \in \mathbf{N}$ be given. To serve as a useful technical device, we introduce the belief $p_{n,m}^{i,j} := \mathbf{P}(\Theta = H \mid \mathcal{F}_n^i, \mathcal{F}_m^j)$ over Θ , obtained by collecting the private information of player i at stage n and of player j at stage m . Since the term $\frac{p_0}{1-p_0}$ appears in the calculation of both p_n^i and p_m^j , one has

$$\frac{p_{n,m}^{i,j}}{1 - p_{n,m}^{i,j}} = \frac{p_n^i}{1 - p_n^i} \times \frac{p_m^j}{1 - p_m^j} \times \frac{1 - p_0}{p_0}.$$

Hence, $p_{n,m}^{i,j}$ coincides with a continuous and increasing function of p_n^i and p_m^j , that we denote $p_{n,m}^{i,j}(\cdot, \cdot)$.

We let $Q_{n,m}^i(p_n^i, t_m^j) := \mathbf{P}_H(\Theta = H \mid p_n^i, t_m^j)$, and we proceed to compare $Q_{n,m}^i(p_n^i, t_m^j)$ with $p_{n,m}^{i,j}$.

By the law of iterated conditional expectations, and since t_m^j is a function of p_1^j, \dots, p_m^j ,

$$Q_{n,m}^i(p_n^i, t_m^j) = \mathbf{E} [\mathbf{P}(\Theta = H \mid p_n^i, p_1^j, \dots, p_m^j) \mid p_n^i, t_m^j] = \mathbf{E} [p_{n,m}^{i,j} \mid p_n^i, t_m^j].$$

If $t_m^j = \blacksquare$, one has $p_m^j > \pi_m^j(\blacksquare)$. Since $p_{n,m}^{i,j}$ is monotonic in p_m^j , $Q_{n,m}^i(p_n^i, t_m^j) > p_{n,m}^{i,j}(p_n^i, \pi_m^j(\blacksquare))$. By contrast, if $t_m^j = m - 1$, one has $p_m^j \leq \pi_m^j(\blacksquare)$, hence $Q_{n,m}^i(p_n^i, t_m^j) \leq p_{n,m}^{i,j}(p_n^i, \pi_m^j(\blacksquare))$. Combining the two inequalities yield

$$Q_{n,m}^i(p, m - 1) < Q_{n,m}^i(p, \blacksquare). \quad (6)$$

Next, we compare $Q_{n,m}^i(p, \blacksquare)$ with $Q_{n,m+1}^i(p, \blacksquare)$. Using once more the law of iterated conditional expectations, one has

$$Q_{n,m}^i(p_n^i, t_m^j) = \mathbf{E} [\mathbf{P}(\Theta = H \mid p_n^i, t_{m+1}^j) \mid p_n^i, t_m^j] = \mathbf{E}[Q_{n,m+1}^i \mid p_n^i, t_m^j].$$

It follows that $Q_{n,m}^i(p, \blacksquare)$ is in the convex hull of (i) $Q_{n,m+1}^i(p, \blacksquare)$, of (ii) $Q_{n,m+1}^i(p, m)$, and of (iii) $Q_{n,m}^i(p, k) = Q_{n,m+1}^i(p, k)$ for each $k < m$.

Using now (6), one gets

$$Q_{n,m+1}^i(p, m-1) = Q_{n,m}^i(p, m-1) < Q_{n,m}^i(p, \blacksquare) < Q_{n,m+1}^i(p, \blacksquare).$$

The result follows inductively, since $Q_{n,n}^i$ is equal to Q_n^i , the posterior belief of player i at stage n . ■

4.3 Markov properties of the beliefs

Markov chains are stochastic sequences whose current value contain all relevant information when it comes to predicting future values of the sequence.

Some of the sequences we introduced turn out to be Markov chains. These properties will be decisive when proving that all equilibria are in cut-off strategies.

We first recall the definition of a Markov chain, see Shiryaev (1984).¹⁶

Definition 4.13 *Let $(\mathcal{G}_n)_{n \in \mathbf{N}}$ be a filtration over a probability space (Ω, \mathbf{P}) , and let $(A_n)_{n \in \mathbf{N}}$ be a sequence of random variables, adapted to $(\mathcal{G}_n)_{n \in \mathbf{N}}$, and with values in \mathbf{R}^d . The sequence (A_n) is a Markov chain for the filtration $(\mathcal{G}_n)_{n \in \mathbf{N}}$ if, for each Borel set $B \subset \mathbf{R}^d$ and $n \in \mathbf{N}$, one has $\mathbf{P}(A_{n+1} \in B \mid \mathcal{G}_n) = \mathbf{P}(A_{n+1} \in B \mid A_n)$, \mathbf{P} -a.s.*

The next proposition lists the sequences which form Markov chains in our setup. The proof is provided in Appendix B.2.

Proposition 4.14 *Under \mathbf{P} , the sequence $(p_n^i, t_n^j)_{n \in \mathbf{N}}$ is a Markov chain for $(\mathcal{G}_n^i)_{n \in \mathbf{N}}$.*

Both under \mathbf{P} and under \mathbf{P}_θ , the sequence $(p_n^i)_{n \in \mathbf{N}}$ is a Markov chain for $(\mathcal{F}_n^i)_{n \in \mathbf{N}}$.

Under \mathbf{P}_θ , both sequences (p_n^i) and (t_n^j) are Markov chains for the filtration (\mathcal{G}_n^i) .

¹⁶By Theorems 9.5 and 9.6 in Doob (1953), this definition is equivalent to saying that the conditional law of A_{n+1} given \mathcal{G}_n is one version of the conditional law of A_{n+1} given A_n .

The first statement means that all the information available to player i can be summarized by her private belief and by the status of player j .

The second statement simply says that the private belief adequately summarizes the privately available information. By contrast, since the status of the other player provides important information on the common type, the private belief does not fully summarize the information available to a player. Therefore, the sequence $(p_n^i)_{n \in \mathbf{N}}$ is not a Markov chain for $(\mathcal{G}_n^i)_{n \in \mathbf{N}}$ under \mathbf{P} . For a similar reason, the sequence $(t_n^j)_{n \in \mathbf{N}}$ is not a Markov chain for $(\mathcal{G}_n^j)_{n \in \mathbf{N}}$ under \mathbf{P} either.

Given the common type, the signals of the players are independent. Therefore, in this case, the private information or the status of one player do not help in predicting future private belief or the future status of the other player. This is the content of the third statement.

5 Proofs of Theorems 2.2 and 2.3

This section is devoted to Theorems 2.2 and 2.3, which assert the existence of a symmetric equilibrium, and that all equilibria are in cut-off strategies. Here we provide the main steps and the intuition. Most proofs are relegated to appendix B.

We fix an arbitrary mixed plan $\phi^j = \{\tau^j(\alpha), \alpha \in \mathbf{N} \cup \{\blacksquare\}\}$ of player j be given, where $\tau^j(k)$ is a stopping time for the extended filtration $(\tilde{\mathcal{F}}_n^i)_{n \in \mathbf{N}}$ that contains the randomization devices of the two players.

For every plan ϕ^i of player i , player i 's payoff $\gamma^i(\phi^i, \phi^j)$ depends on ϕ^j only through player j 's behavior until i drops out. Indeed, player j 's behavior once player i drops out does not affect i 's payoff. Hence, as far as only player i 's best reply problem is under study, we may assume that $\tau^j(k) = \tau^j(\blacksquare)$, for each $k \in \mathbf{N}$. In other words, we will assume that player j 's behavior depends only on her private information, and abbreviate $\tau^j(\blacksquare)$ to τ^j .

The best-reply problem faced by player i can be rephrased as an optimal stopping problem: given ϕ^j , there is a one-to-one relationship between mixed plans of actions and stopping times for $(\mathcal{G}_n^i)_{n \in \mathbf{N}}$, which is made explicit in appendix B.2. We emphasize that this identification is contingent on the choice of ϕ^j .

If player i stops at stage n , her realized payoff is

$$Y_n^i := \sum_{k=0}^{n-1} \delta^k X_k^i.$$

Hence, when using the stopping time σ , her expected payoff is $\mathbf{E}[Y_\sigma^i]$, and the best-reply problem of player i reduces to that of solving

$$\text{Problem } \mathcal{P}: \text{ find a solution to } \sup_{\sigma} \mathbf{E}[Y_\sigma^i],$$

where the supremum is taken over all stopping times σ for (\mathcal{G}_n^i) .

We will assert that there is a unique solution to Problem \mathcal{P} , that moreover is a cut-off plan. Let

$$Y_{n \rightarrow \sigma}^i = \sum_{k=n}^{\sigma-1} \delta^{k-n} X_k^i$$

be the discounted sum of payoffs that incur between stages n and σ , discounted to stage n .

Let Λ_n denote the set of stopping times σ such that $\sigma \geq n$. Denote by W_n^i the optimal continuation payoff of player i at stage n , if she *must* stay in at stage n . Formally it is given by¹⁷

$$W_n^i = \text{ess sup}_{\sigma \in \Lambda_{n+1}} \mathbf{E} [Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i]. \quad (7)$$

If $W_n^i > 0$, staying in is the optimal action at stage n , if $W_n^i < 0$, dropping out is the optimal action at stage n , while if $W_n^i = 0$, player i is indifferent between staying in and dropping out. This assertion, stated formally in Lemma 5.1 below, is a very general result in optimal stopping. It follows from Chow and Robbins (1963), or Ferguson (2004, ch. 3).

Lemma 5.1 *Define*

$$\sigma^* = \inf\{n \geq 0 : W_n^i \leq 0\} \text{ (with } \inf \emptyset = +\infty \text{)}.$$

Then σ^ is a solution to problem \mathcal{P} . Moreover, if σ is an optimal stopping time, then $\sigma \geq \sigma^*$, \mathbf{P} -a.s.*

More generally, let a stage $n \in \mathbf{N}$ be given, and set $\sigma_{n+1}^ := \inf\{k \geq n+1 : W_k^i \leq 0\}$. Then $W_n^i = \mathbf{E} [Y_{n \rightarrow \sigma_{n+1}^*}^i \mid \mathcal{G}_n^i]$.*

¹⁷The essential supremum $\text{ess sup}_{x \in X} Y_x$ of a family $(Y_x)_{x \in X}$ of r.v.'s is the r.v. Y such that (i) for each x , $Y \geq Y_x$ \mathbf{P} -a.s. and (ii) for any variable Z that satisfies $Z \geq Y_x$ \mathbf{P} -a.s. for every x one has $Z \geq Y$, \mathbf{P} -a.s.

We will next prove that σ^* depends only on the sequence (p_n^i, t_n^j) .

Lemma 5.2 *There is a sequence v_k^i of measurable maps, such that*

$$\sigma_n^* = \inf\{k \geq n : v_k^i(p_k^i, t_k^j) = 0\}.$$

Accordingly, the information of player i can be summarized by her private belief and the status variable of player j . The proof of Lemma 5.2 is a variant on standard tools in the optimal stopping of Markov chains.

Using Lemma 5.2 and the relation between posterior and private beliefs (see Lemma 4.8), we proceed with finding an explicit version of W_n^i as a function $w_n^i(p_n^i, t_n^j)$. To this end, we compute

$$W_n^i = \sum_{\theta=L,H} \mathbf{P}(\Theta = \theta | \mathcal{G}_n^i) \mathbf{E}_\theta \left[Y_{n \rightarrow \sigma_{n+1}^*}^i | \mathcal{G}_n^i \right]$$

and obtain

$$\begin{aligned} w_n^i(p_n^i, t_n^j) = & \sum_{k \in \mathbf{N}} \left(\mathbf{E}_H[X_0^i] \delta^k Q_n^i(p_n^i, t_n^j) \mathbf{P}_H(\sigma_{n+1}^* > n+k \mid (p_n^i, t_n^j)) \right. \\ & \left. + \mathbf{E}_L[X_0^i] \delta^k (1 - Q_n^i(p_n^i, t_n^j)) \mathbf{P}_L(\sigma_{n+1}^* > n+k \mid (p_n^i, t_n^j)) \right). \end{aligned} \quad (8)$$

In addition, we will choose versions of the conditional probability $\mathbf{P}_H(\sigma_{n+1}^* > n+k \mid (p_n^i, t_n^j))$ such that w_n^i has the desirable properties listed in Lemma 5.3 below.

Lemma 5.3 *For every fixed status variable $\alpha \in \mathbf{N} \cup \{\blacksquare\}$, $w_n^i(p, \alpha)$ is an increasing and continuous function of p on $(0, 1)$.*

When the type of the bandit is known, the player ignores the information received from the other player, and either drops out as soon as possible or stays in forever, depending on this type. We therefore have the following result.

Lemma 5.4 $\lim_{p \rightarrow 0} w_n^i(p, \alpha) = \mathbf{E}_L[X_0^i] < 0 < \frac{\mathbf{E}_H[X_0^i]}{1-\delta} = \lim_{p \rightarrow 1} w_n^i(p, \alpha).$

Given α and by Lemmas 5.3 and 5.4, there is a unique p such that $w_n^i(p, \alpha) = 0$. Therefore, the optimal stopping time σ^* is a cut-off plan.

By our non-atomicity conditions, the probability that the private belief is ever equal to the cut-off is 0.

Corollary 5.5 *Player i has a unique best reply to ϕ^j . This best-reply is a cut-off plan.*

We now turn to the existence of an equilibrium. We apply a fixed point theorem. The set Φ^i of cut-off strategies of player i is the set of sequences $(\pi_n^i(k))$ indexed by $n \in \mathbf{N}$ and $k \in \{\blacksquare, 1, \dots, n-1\}$, and with values in $[0, 1]$. We endow Φ^i with the product topology, and the set $\Phi^1 \times \Phi^2$ of cut-off profiles with the product topology. Then Φ^i is compact.

Lemma 5.6 *The payoff function γ^i is continuous over $\Phi^i \times \Phi^j$.*

Proof. Given a cut-off profile (ϕ^i, ϕ^j) , the payoff to player i is Y_n^i if player i drops out at stage n .

The event that player i drops out at stage n is the union of finitely many events: (i) player i drops out at stage n , and player j is active at stage $n-1$, and (ii) player i drops out at stage n , and player j dropped out at stage k , with $k < n$.

As a consequence, the realized payoff to player i is

$$\begin{aligned} r^i(\phi^i, \phi^j) = & \sum_{n \in \mathbf{N}} Y_n^i \times \left\{ \prod_{l=1}^{n-1} \left(1_{p_l^i > \pi_l^i(\blacksquare)} \times 1_{p_l^j > \pi_l^j(\blacksquare)} \right) 1_{p_n^i \leq \pi_n^i(\blacksquare)} \right\} \\ & + \sum_{n \in \mathbf{N}} Y_n^i \times \sum_{k=1}^{n-1} \left[\prod_{l=1}^{k-1} \left(1_{p_l^j > \pi_l^j(\blacksquare)} \right) \times 1_{p_k^j \leq \pi_k^j(\blacksquare)} \right] \times \prod_{l=1}^k \left(1_{p_l^i > \pi_l^i(\blacksquare)} \right) \\ & \times \prod_{l=k+1}^{n-1} \left(1_{p_l^i > \pi_l^i(k)} \right) 1_{p_n^i \leq \pi_n^i(k)}. \end{aligned}$$

and the expected payoff to player i is $\gamma^i(\phi^i, \phi^j) := \mathbf{E}[r^i(\phi^i, \phi^j)]$.

Let now a sequence (ϕ_m^i, ϕ_m^j) of cut-off profiles be given, that converges to (ϕ^i, ϕ^j) in the product topology.

Note that $\lim_{m \rightarrow +\infty} r^i(\phi_m^i, \phi_m^j) = r^i(\phi^i, \phi^j)$ except possibly on the event

$$\{p_l^i = \pi_l^i(k), \text{ for some } l \in \mathbf{N}, k = \blacksquare, 1, \dots, l-1\} \cup \{p_l^j = \pi_l^j(\blacksquare), \text{ for some } l \in \mathbf{N}\}.$$

Since the laws of p_n^i and of p_n^j have a density for each $n \in \mathbf{N}$, this event has \mathbf{P} -measure zero.

Since $|r^i(\phi_m^i, \phi_m^j)| \leq \sup_{n \in \mathbf{N}} |Y_n^i|$ and $|r^i(\phi^i, \phi^j)| \leq \sup_{n \in \mathbf{N}} |Y_n^i|$, the dominated convergence theorem applies, and $\lim_{m \rightarrow \infty} \gamma^i(\phi_m^i, \phi_m^j) = \gamma^i(\phi^i, \phi^j)$. ■

Lemma 5.7 *There is a symmetric equilibrium.*

Proof. Note that $\Phi^i = \Phi^j$. For $\phi' \in \Phi^i$, we define

$$B(\phi') := \left\{ \phi \in \Phi^i : \gamma^i(\phi, \phi') = \max_{\Phi^i} \gamma^i(\cdot, \phi') \right\}.$$

By Corollary 5.5, $B(\phi')$ is singleton-valued, hence may be viewed as a function from Φ^i into itself. By Lemma 5.6, B is continuous over Φ^i . Since Φ^i and Φ^j are compact, by Ky Fan's theorem B has a fixed point ϕ^* . Plainly, the strategy profile (ϕ^*, ϕ^*) is a symmetric equilibrium of the game. ■

Remark 5.8

We end this section by observing that there might be multiple symmetric equilibria. Indeed, suppose that the initial belief is $p_0 = \pi^*$, the cut-off in the one-player problem. As said earlier, it is then an equilibrium for both players to drop out immediately. Besides, each player has a strategy which stays in at stage 0 and ensures a payoff 0.

We will now informally argue that there is another symmetric equilibrium. Let Ψ^i be the space of cut-off strategies of player i in which she stays in at stage 0. Define Ψ^j analogously.

As Ψ^i and Ψ^j are compact, Lemma 5.6 implies the existence of a symmetric fixed-point to the best reply function. Corollary 5.5 implies that this fixed point is in cut-off strategies, and by **P1** the sequence of cut-offs is non-increasing.

Since each player can ensure 0 by ignoring the actions of the other player, the expected payoff in this fixed point is non-negative. In particular, this fixed point is a symmetric equilibrium. Since in this equilibrium both players continue in stage 0, this equilibrium differs from the previous one.

We now argue that the expected payoff in this equilibrium is strictly positive, so the two equilibria yield different payoffs.

The probability that player j will drop out in stage 1 is positive (by **A4**), and less than one (by Proposition 4.14 and since cut-offs are non-increasing). Therefore, if player i stays in until stage 2, player j 's decision in stage 1 reveals meaningful information about Θ . Therefore, the *ex ante* optimal payoff to player i is strictly higher than what she would get on the basis of private information only, and hence higher than zero, as desired.

This sketch suggests that there are two equilibria, one in which both players drop out in stage zero, and the other in which both players are active in stage zero. Whether or not there is a unique equilibrium of the latter type is an open problem.

6 Qualitative features

We now proceed to the proof of Theorems 2.4 through 2.6. The proof of **P2** will be omitted. All proofs are logically independent, with the exception of the proof of **P8**, which relies on **P2**. The section is organized as follows. We start with **P1**, then insert a basic observation. We write the proofs for the case of two players. For statements **P3** through **P9**, the extension to more-than-two-player games adaptation presents no difficulty. In the case of **P1**, we will mention why this is not so.

6.1 Proof of P1

We here prove Proposition 6.1 below, which implies **P1** in Theorem 2.4.

Proposition 6.1 *Let ϕ^2 be a cut-off plan of player 2, and let ϕ^1 be the unique best reply to ϕ^2 , with cut-offs (π_n^1) . Then the sequence $(\pi_n^1(\blacksquare))$ is non-increasing.*

Proof. The formal proof involves a long list of inequalities. We provide a detailed sketch, which can be easily transformed into the formal proof.

We let a cut-off strategy ϕ^2 of player 2 be given, and look at the optimization problem of player 1. We proceed by comparing the following situations player 1 can face in the game:

- (A) $p_n^1 = p$ and $t_n^2 = \blacksquare$: this is the game in stage n .
- (B) $p_n^1 = p$, $t_n^2 = \blacksquare$, and there is an additional interim stage $n - \frac{1}{2}$ (between stages $n - 1$ and n) in which *only* player 1 observes a signal (but the players make no choice) – this situation is purely fictitious.
- (C) $p_{n+1}^1 = p$, $t_n^2 = \blacksquare$, and from stage n on player 1 observes the status of player 2 with a delay of one stage.
- (D) $p_{n+1}^1 = p$ and $t_n^2 = \blacksquare$.

(E) $p_{n+1}^1 = p$ and $t_{n+1}^2 = \blacksquare$: this is the game in stage $n + 1$.

Our goal is to compare situations (A) and (E). That is, to prove that player 1's optimal continuation payoff (OCP in short) at stage n in situation (A) does not exceed player 1's OCP at stage $n + 1$ in situation (E). This will show that $w_n^1(p, \blacksquare) \leq w_{n+1}^1(p, \blacksquare)$ for every p . In particular, $0 = w_n^1(\pi_n^1(\blacksquare), \blacksquare) \leq w_{n+1}^1(\pi_n^1(\blacksquare), \blacksquare)$. Since w_{n+1}^1 is non-decreasing, this implies that $\pi_{n+1}^1(\blacksquare)$, which is defined by the equality $w_{n+1}^1(\pi_{n+1}^1(\blacksquare), \blacksquare) = 0$, is below $\pi_n^1(\blacksquare)$, as desired.

We will do that by comparing the various situations.

Step 1: Variations (A) and (B).

All relevant information contained in past payoffs is summarized in the private belief: it is irrelevant that in (A) and (B) different payoffs and a different number of payoffs lead to the same private belief. Since from stage n the two situations coincide, player 1's OCP at stage n in both situations is the same.

Step 2: Variations (B) and (C).

The continuation game in situation (B) from stage n is equivalent to the continuation game in situation (C) from stage $n + 1$ and on. Therefore player 1's OCP at stage n in situation (B) is equal to player 1's OCP at stage $n + 1$ in situation (C).

Step 3: Variations (C) and (D).

The only difference between the continuation games from stage $n + 1$ in the two situations is that in (C) information is delayed for player 1. Hence in (C) player 1 has fewer available plans, so that player 1's OCP at stage $n + 1$ in situation (C) does not exceed her expected OCP at stage $n + 1$ in situation (D).

Step 4: Variations (D) and (E).

Here we define yet another situation:

(F) $p_{n+1}^1 = p$ and $t_{n+1}^2 = n$.

this is the case in which player 2 drops out at stage n .

Player 1's expected OCP at stage $n + 1$ in situation (D), in which player 1 knows that player 2 was still active at the beginning of stage n , is a weighted average of her OCPs in situations (E) and (F). Indeed, in the former she has just learned that player 2 is still active in stage $n + 1$, while in the latter she has just learned that player 2 has dropped out in stage n .

Therefore, to prove that the OCP in situation (D) is at most her OCP in situation (E), it is sufficient to prove that the OCP in situation (F) does not exceed that in continuation (E).

Let σ be an optimal strategy of player 1 in situation (F), starting from stage $n + 1$. This is an optimal strategy of player 1 in a one-player problem, hence, by Theorem 2.1 (and (5)), σ is a cut-off plan, with a fixed cut-off. Suppose that player 1 follows σ in situation (E), starting from stage $n + 1$. Conditional on Θ , the continuation payoffs in situations (E) and (F) coincide. Besides, this continuation payoff is positive if $\Theta = H$, and negative if $\Theta = L$. In addition, the posterior belief that $\Theta = H$ is higher in (E) than in (F). Altogether, this implies that the expected continuation payoff induced by σ is higher in (E) than in (F). Since σ is an available strategy for player 1 from stage $n + 1$ in situation (E), the OCP in situation (E) is at least equal to the continuation payoff induced by σ in (F). Hence the OCP in (E) is larger than the OCP in (F) so that the OCP in (D) is smaller than the OCP in (E). ■

6.2 Simple bounds

We here point out simple bounds on the optimal continuation payoff to player i . These bounds will help to derive basic, yet useful, observations on equilibrium cut-offs.

Lemma 6.2 *Let a strategy ϕ^j , and a stage $n \in \mathbf{N}$ be given. The optimal continuation payoff W_n^i satisfies the following inequalities*

$$q_n^i \mathbf{E}_H[X_0^i] + (1 - q_n^i) \mathbf{E}_L[X_0^i] \leq (1 - \delta) W_n^i \leq q_n^i \mathbf{E}_H[X_n^i] + (1 - \delta)(1 - q_n^i) \mathbf{E}_L[X_n^i]. \quad (9)$$

Proof. We start with the lower bound. If player i – when active in stage n – decides to stay in *forever*, her expected payoff in every stage coincides with the left-hand side in (9). Since staying in forever is one of many options, this proves the first inequality in (9).

We now prove the upper bound. The argument is valid since by **A1**, X_n^i has finite expectation and variance. If the type of the bandit is H , the maximal payoff a player can get is $\mathbf{E}_H[X_n^i]/(1 - \delta)$. Since W_n^i is the payoff when player i must stay in for at least one stage, and since $\mathbf{E}_L[X_0^i]$ is negative, when the type of the bandit is L the optional sampling theorem implies that

the payoff is at most $\mathbf{E}_L[X_0^i]$. Therefore, the payoff is bounded by

$$q_n^i \frac{\mathbf{E}_H[X_n^i]}{(1-\delta)} + (1-q_n^i)\mathbf{E}_L[X_0^i],$$

and the right-hand side inequality follows. ■

We derive two straightforward implications of Lemma 6.2. Let $\bar{p} = \frac{\mathbf{E}_L[X_0^i]}{\mathbf{E}_H[X_0^i] - \mathbf{E}_L[X_0^i]} \in (0, 1)$ be the value of q_n^i for which the left-hand side in (9) is zero, and recall that $p_* \in (0, 1)$ has been defined as the value of q_n^i for which the right-hand side is zero.

Corollary 6.3 *Let ϕ be an equilibrium.*

If player i finds it optimal to drop out at stage n , then $q_n^i \leq \bar{p}$.

If player i finds it optimal to remain active in stage n , then $q_n^i \geq p_$.*

Indeed, one has $W_n^i \leq 0$ in the former case, and $W_n^i \geq 0$ in the latter.

In stage 1 – and only in that stage – one has $q_n^i = p_n^i$, since all players are anyway active in stage 0. Besides, one has $q_n^i \geq p_n^i$ whenever player j is active. Thus, Corollary 6.3 implies the following.

Corollary 6.4 *Let ϕ be an equilibrium with cut-offs (π_n^i) . Then (i) for each player i one has $p_* \leq \pi_1^i(\blacksquare) \leq \bar{p}$, and (ii) for $n \geq 1$, one has $\pi_n^i(\blacksquare) \leq \bar{p}$.*

6.3 Proof of P3

We here prove **P3**. Plainly, the result holds if $\lim_{n \rightarrow +\infty} \pi_n^j(\blacksquare) = 0$. Hence, we may assume that $\limsup_{n \rightarrow +\infty} \pi_n^j(\blacksquare) > 0$. As we will show, this implies that $\lim_{n \rightarrow +\infty} \pi_n^i(\blacksquare) = 0$.

If the players are active in stage 0, then $\pi_0^j(\blacksquare) < p_0$; since cutoffs are decreasing, $\pi_n^j(\blacksquare) < p_0$, for any stage n . Therefore, one has $\mathbf{P}_H(\tau^j(\blacksquare) = +\infty) > 0$ by Corollary 4.7. On the other hand, (p_n^j) converges \mathbf{P}_L -a.s. to zero. Since cut-offs $(\pi_n^j(\blacksquare))$ do not converge to zero, one has $\tau^j(\blacksquare) < +\infty$, \mathbf{P}_L -a.s., so that

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)} = +\infty. \quad (10)$$

We now deduce that $\limsup_{n \rightarrow \infty} \pi_n^i(\blacksquare) \leq p$ for every positive p , which implies that the limit exists and is 0, as desired.

Fix $p \in (0, 1]$, and let $n \in \mathbf{N}$ be an arbitrary stage. Suppose that $t_n^j = \blacksquare$ and $p_n^i \geq p$. By Lemma 4.8, one then has

$$\frac{q_n^i}{1 - q_n^i} = \frac{p_n^i}{1 - p_n^i} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)} \geq \frac{p}{1 - p} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)}. \quad (11)$$

In the light of (10), this implies that $q_n^i > \bar{p}$, provided n is large enough. Using Corollary 6.3, we have therefore proven that $w_n^i(p_n^i, \blacksquare) > 0$ \mathbf{P} -a.s. whenever $p_n^i \geq p$ and n is large enough. In particular, $\pi_n^i(\blacksquare) \leq p$, and the result follows.

6.4 Proof of P4 and P5

We limit ourselves to a heuristic proof. The corresponding formal proof is a simple but tedious elaboration, and is therefore omitted.

We start with **P4**. The analysis of the optimal stopping problem in Section 5 applies to the one-player problem. Dropping superscripts, one has, in the case of 1 player, $W_n = \mathbf{E}[Y_{n-\sigma_{n+1}} \mid \mathcal{F}_n] = w_n(p_n)$, where $\sigma_{n+1} := \inf\{k \geq n+1 : p_k \leq \pi^*\}$.

Let a stage $n \in \mathbf{N}$ be given, and consider any player i . If player j is still active at stage n , then $q_n^i \geq p_n^i$, and player i has more strategies available in the continuation game, than if she were in the one-player case. As a consequence, her OCP is higher than in the one-player case.

We continue with **P5**. If player j drops out at stage k , the belief of player i at stage $n > k$ satisfies $\frac{q_n^i}{1 - q_n^i} = \frac{p_n^i}{1 - p_n^i} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) = k)}{\mathbf{P}_L(\tau^j(\blacksquare) = k)}$. Once player j drops out, player i faces a one-player problem with a fixed cut-off π^* . Hence, in the continuation game, player i drops out once $q_n^i \leq \pi^*$. This implies that $\frac{\pi_n^i(k)}{1 - \pi_n^i(k)} = \frac{\pi^*}{1 - \pi^*} \times \frac{\mathbf{P}_L(\tau^j(\blacksquare) = k)}{\mathbf{P}_H(\tau^j(\blacksquare) = k)}$. In particular, $\pi_n^i(k) > \pi^*$ if and only if $\mathbf{P}_H(\tau^j(\blacksquare) = k) < \mathbf{P}_L(\tau^j(\blacksquare) = k)$.

To prove the existence of such a k , it suffices to choose $k = 1$ and to observe that $\mathbf{P}_L(\tau^j(\blacksquare) = 1) > 0$ (since p_1^i takes arbitrary low values with positive values) while $\mathbf{P}_H(\tau^j(\blacksquare) = 1) < 1$. Therefore, by Proposition 4.4, one has $\mathbf{P}_H(\tau^j(\blacksquare) = 1) < \mathbf{P}_L(\tau^j(\blacksquare) = 1)$.

6.5 Proof of P7, P8 and P9

We start with **P7**:

Lemma 6.5 *For every cut-off profiles $(\tau^i(\alpha))$ and $(\tau^j(\alpha))$, the law of $\min\{\tau^i(\blacksquare), \tau^j(\blacksquare)\}$ under \mathbf{P}_H stochastically dominates the law of the same variable under \mathbf{P}_L .*

Proof. The variables $\tau^i(\blacksquare)$ and $\tau^j(\blacksquare)$ are independent under \mathbf{P}_θ , so that $\mathbf{P}_\theta(\min\{\tau^i(\blacksquare), \tau^j(\blacksquare)\} \geq n) = \mathbf{P}_\theta(\tau^i(\blacksquare) \geq n) \times \mathbf{P}_\theta(\tau^j(\blacksquare) \geq n)$. The result follows from Corollary 4.6. ■

We proceed with **P9** claiming that there is a positive probability under \mathbf{P}_H that no player ever stops.

We first argue that

$$\mathbf{P}_H(\tau^i(\blacksquare) < +\infty) < 1, \quad \forall i \in I. \quad (12)$$

Indeed, by assumption **A4** there is a positive probability that the private belief of both players at stage 1 is strictly higher than p_0 . Since the cut-offs are non-increasing, Eq. (12) follows from Corollary 4.7 applied to p_1^i and $\pi_1^i(\blacksquare)$.

Since $\tau^i(\blacksquare)$ and $\tau^j(\blacksquare)$ are independent given Θ , Eq. (12) implies that $\mathbf{P}_H(\min(\tau^i(\blacksquare), \tau^j(\blacksquare)) < +\infty) < 1$, and the claim follows, since $\min(t^i, t^j) = \min(\tau^i(\blacksquare), \tau^j(\blacksquare))$.

We conclude with **P8** claiming that all players stop in finite time under \mathbf{P}_L . The basic intuition is as follows. First, observe that once one player drops out, the other faces a one-player problem. Since in the one-player problem the cut-off sequence is positive and constant, while if the type is L the private belief converges to 0 with probability 1, this player will drop out in finite time.

Now, if both players continue forever with positive probability when the type is L , then by **P7** they also do so when the type is H , so that there is a stage such that if both players stayed in until that stage, they would continue forever with high probability. Therefore, from that stage on each of the players essentially faces a one-player problem, so that her cut-offs are positive and roughly constant. Since in the one-player problem the cut-off sequence is positive and constant, while if the type is L the private belief converges to 0 with probability 1, at some point the private belief will fall below the cut-off, and one of the players will drop out.

We elaborate on these two steps in the next two lemmas.

Lemma 6.6 *One has $\mathbf{P}_L(\min\{\tau^i(\blacksquare), \tau^j(\blacksquare)\} = +\infty) = 0$.*

Proof. If $\mathbf{P}_L(\tau^i(\blacksquare) = +\infty) = 0$, the result is proven. Assume now that $\mathbf{P}_L(\tau^i(\blacksquare) = +\infty) > 0$, so that the ratio $\frac{\mathbf{P}_H(\tau^i(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^i(\blacksquare) \geq n)}$ has a positive finite limit, when n goes to $+\infty$. Since $\lim_{n \rightarrow +\infty} p_n^j = 0$, \mathbf{P}_L -a.s., it follows by Lemma 4.8, that $\lim_{n \rightarrow +\infty} q_n^j = 0$, \mathbf{P}_L -a.s. Hence, by Corollary 6.3, one has $\tau^j(\blacksquare) < +\infty$ \mathbf{P}_L -a.s., and the result follows. ■

Lemma 6.7 *One has $\mathbf{P}_L(\max\{t^i, t^j\} = +\infty) = 0$.*

Proof. It is enough to prove that $t^i < +\infty$, \mathbf{P}_L -a.s. The same proof will apply to player j as well. Observe first that by Lemma 6.6 one has $t^i < +\infty$ whenever $t^i \leq t^j$. Let $k \in \mathbf{N}$ be an arbitrary stage. On the event $\{\tau^j(\blacksquare) = k < \tau^i(\blacksquare)\}$, the exit time t^i of player i coincides with $\inf\{n > k : p_n^i < \pi_n^i(k)\}$, where $\pi_n^i(k)$ are the cut-offs associated with the plan ϕ^i . By **P2** the sequence $(\pi_n^i(k))_n$ is constant, and since $\lim_{n \rightarrow +\infty} p_n^i = 0$, \mathbf{P}_L -a.s. and positive, it follows that $t^i < +\infty$ on the event $\{\tau^j(\blacksquare) = k < \tau^i(\blacksquare)\}$. The result follows. ■

6.6 Proof of P6

We proceed with the proof of **P6**. Let ϕ be a symmetric equilibrium, with cut-offs (π_n^i) , and let a stage $n \in \mathbf{N}$ be given. Whenever $p_n^i = \pi_n^i(\blacksquare)$ and $t_n^j = \blacksquare$, player i is indifferent between dropping out or not. Hence, by Corollary 6.3, her posterior belief satisfies $p_* \leq q_n^i \leq \bar{p}$. Since $t_n^j = \blacksquare$, private and posterior beliefs are related through the equality

$$\frac{q_n^i}{1 - q_n^i} = \frac{p_n^i}{1 - p_n^i} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)} = \frac{\pi_n^i(\blacksquare)}{1 - \pi_n^i(\blacksquare)} \times \frac{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)}.$$

Since the function $\frac{x}{1-x}$ is increasing, and since $p_* \leq q_n^i \leq \bar{p}$, simple algebraic manipulations yield

$$\frac{p_*}{1 - p_*} \times \frac{1 - \pi_n^i(\blacksquare)}{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)} \leq \frac{\pi_n^i(\blacksquare)}{\mathbf{P}_L(\tau^j(\blacksquare) \geq n)} \leq \frac{\bar{p}}{1 - \bar{p}} \times \frac{1 - \pi_n^i(\blacksquare)}{\mathbf{P}_H(\tau^j(\blacksquare) \geq n)}. \quad (13)$$

Set

$$c_1 = \frac{p_*(1 - \pi_1^i(\blacksquare))}{1 - p_*}.$$

By **P2** the sequence $(\pi_n^i(\blacksquare))_{n \in \mathbf{N}}$ is decreasing, so that from (13) we obtain

$$\pi_n^i(\blacksquare) \geq c_1 \mathbf{P}_L(\tau^j(\blacksquare) \geq n).$$

Since $\pi_1^i(\blacksquare) \leq p_* < 1$, c_1 is positive.

Set

$$c_2 = \frac{\bar{p}}{(1 - \bar{p})\mathbf{P}_H(\tau^j(\blacksquare) = +\infty)}.$$

By **P8** we have

$$\mathbf{P}_H(\tau^j(\blacksquare) \geq n) \geq \mathbf{P}_H(\tau^j(\blacksquare) = +\infty) > 0,$$

so that from (13) we obtain

$$\pi_n^i(\blacksquare) \leq c_2 \mathbf{P}_L(\tau^j(\blacksquare) \geq n).$$

The proof is complete since in a symmetric equilibrium $\tau^i = \tau^j$.

7 Large games

This section is devoted to the analysis of large games. We here provide a detailed intuition for the proof of the first statement in Theorem 2.7. The proof itself is given in appendix C. Given any $N \in \mathbf{N}$, we consider the N -player game with player set $\{1, \dots, N\}$. For any $N \in \mathbf{N}$ we fix an equilibrium $(\phi^{N,i})_{i=1, \dots, N}$ be given, with cut-offs $(\pi_n^{N,i}(\vec{\alpha}^N))$, where $\vec{\alpha}^N$ are the $(N - 1)$ -dimensional status vectors.

We first focus on stage 1. Since no public information is available at that stage, the relevant cut-offs are abbreviated to $\pi_1^{N,i}$. Let ρ_θ^N be the expected proportion of players who drop out in stage 1 when the common type is θ . By taking a subsequence, we may assume that (ρ_L^N) and (ρ_H^N) converge to some constants ρ_L and ρ_H respectively. Using Proposition 4.4 we will prove that $\rho_L > \rho_H$: when the common type is L more players drop out at stage 1 than when it is H . By (some version of) the law of large numbers, the actual number of players who drop out at stage 1 converges to ρ_θ , as N goes to infinity. As a consequence, all players know that at stage 2 they will (almost perfectly) discriminate between the two types H and L , on the basis of the proportion of players who drop out in stage 1. Let us adopt the viewpoint of some player i in stage 1, and consider the continuation strategy, according to which (i) player i will continue forever if this proportion is close to ρ_H , (ii) will exit in stage 2 if it is close to ρ_L , and (iii) play in some arbitrary way otherwise. This strategy yields a payoff close to $\mathbf{E}_H[X_0^i]/(1 - \delta)$ if $\Theta = H$, and close to $\mathbf{E}_L[X_0^i]$ if $\Theta = L$.

Observe now that in stage 1, player i 's posterior belief on Θ coincides with her private belief p_1^i . As a consequence, player i 's OCP in stage 1 is close to $p_1^i \mathbf{E}_H[X_0^i]/(1 - \delta) + (1 - p_1^i) \mathbf{E}_L[X_0^i]$. Hence, the cut-off $\pi_1^{N,i}$ used by player i in the equilibrium, must converge to p_* (which is defined in Eq. (2)), and the first assertion follows.

8 Concluding remarks

8.1 The finite-stage game

The game we have studied lasts for infinitely many stages. Some of our results hold for finite-stage games as well.

The three-stage example we have studied provides two observations.

- Whereas in the one-player problem, the fixed cut-off is independent of the initial belief, in the example the cut-offs depend on the initial belief.
- There are multiple equilibria in the finite-stage game.

8.2 Conditions for the existence of an equilibrium

Some of our results hold in a much more general framework. Assume that the payoffs $(X_n^i)_{n \in \mathbf{N}}$ are arbitrary random variables defined over a probability space (Ω, \mathbf{P}) . Assume that the inflow of private information is described by arbitrary filtrations $(\mathcal{F}_n^i)_{n \in \mathbf{N}}$ over Ω , for $i \in I$. Let a timing game proceed as in Section 1.1 (each player i receives her private information, and observes in addition all past decisions). Then, provided $\sup_{n \in \mathbf{N}} |X_n^i|$ is \mathbf{P} -integrable, the timing game has an equilibrium, not necessarily in pure strategies.¹⁸

This more general framework allows for (i) asymmetric games, (ii) non-exchangeable payoffs, and (iii) games in which the acquisition of private information is disentangled from one's own payoffs.

Plainly, the existence of a *symmetric* equilibrium requires a symmetric framework. Since the information structure is general, the posterior belief is not necessarily a sufficient statistic, and the payoff need not be monotonic in the private belief, so that a cut-off equilibrium need not exist.

¹⁸The existence proof for this framework differs from the one given here.

8.3 Variations of the model

Our model is specific in many aspects, and there are many interesting variations that are worth studying.

In some cases (e.g., financial markets), players can strategically decide whether to reveal the private information they have learned.

In other cases, at every given stage players can either acquire new information, or learn the status of the other players, but they cannot learn both types of information simultaneously. The decision which information to learn, though, is strategic. Such cases occur, e.g., among animals looking for a patch of berries, who can either taste the berries or look at their fellow birds, but cannot do both activities at the same time (see, e.g., Valone and Templeton (2002) and Giraldeau et al. (2002), for additional variations that arise from animal behavior).

References

- [1] Aumann, R.J., Mixed and behavior strategies in infinite extensive games, in *Advances in Game Theory*, Dresher, M., Shapley, L.S. and A.W. Tucker (eds), Annals of Mathematics Study 52, Princeton University Press, 1964.
- [2] Bergemann, D. and J. Välimäki (1997), Market Diffusion with two-sided Learning, *RAND J. Econ.*, **28**, 773–795.
- [3] Bergemann, D. and J. Välimäki (2000), Experimentation in markets, *Rev. Econ. Stud.*, **67**, 213–234.
- [4] Berry, D.A. and B. Fristedt (1985) *Bandit Problems*. New York: Chapman and Hall.
- [5] Bikhchandani, S., Hirshleifer, D. and I. Welch (1992), A theory of fads, fashions, custom, and cultural change as informational cascades, *J. Polit. Econ.*, **100**, 992–1026.
- [6] Billingsley, P. (1995), *Probability and Measure*, Wiley.
- [7] Bolton, P. and C. Harris (1999), Strategic Experimentation, *Econometrica*, **67**, 349–374.

- [8] Border, K. *Fixed point Theorems with applications to economics and game theory*, Cambridge U.P, 1985.
- [9] Bulow J. and Klemperer P. (1994), Rational Frenzies and Crashes. *J. Polit. Econ.*, **102**, 1–22.
- [10] Caplin, A. and J. Leahy (1994), Business as Usual, Market Crashes and Wisdom after the Fact, *Amer. Econ. Rev.*, **84**, 547–564.
- [11] Chamley, C. (2004), Delays and equilibria with large and small information in social learning, *Europ. Econ. Rev.*, **48**, 477–501.
- [12] Chamley, C. and D. Gale (1994), Information revelation and strategic delay in a model of investment, *Econometrica*, **62**, 175–189.
- [13] Cripps, M., G. Keller and S. Rady (2002), Strategic Experimentation: The Case of Poisson Bandits, CESifo group working papers, #737.
- [14] Cripps, M., G. Keller and S. Rady (2004), Strategic Experimentation with exponential bandits, *Econometrica*, forthcoming.
- [15] Décamps, J.P., T. Mariotti and S. Villeneuve (2001), Investment timing under incomplete information, *Maths. of Op. Research.*, forthcoming.
- [16] Décamps, J.P. and T. Mariotti, Investment timing and learning externalities, *J. Econ. Theory*, forthcoming.
- [17] Dixit, A.K. and R.S. Pindyck, *Investment under uncertainty*, Princeton University Press, Princeton, 1994.
- [18] Doob, J.L. *Stochastic Processes*, 1953.
- [19] Feller, W. *An introduction to probability theory and its applications*, 2nd Edition, Wiley & Sons, Inc., 1971.
- [20] Ferguson, T., *Optimal Stopping and Applications*, <http://www.math.ucla.edu/~tom/stopping/contents.html>.
- [21] Giraldeau, L.-A., T.J. Valone and J.J. Templeton (2002), Potential disadvantages of using socially acquired information, *Phil. Trans. R. Soc. Lond. B*, **357**, 1559–1566.

- [22] Gittins, J.C. (1979), Bandit problems and Dynamic Allocation Indices, *J. R. Statist. Soc. B*, **41**, 148-177.
- [23] Keller, G. and S. Rady (1999), Optimal experimentation in a changing environment, *Rev. Econ. Stud.*, **66**, 475–507.
- [24] Milgrom, P. (1981), Representation theorems and applications, *Bell J. Econ.*, **12**, 380–391.
- [25] Moscarini, G. and F. Squintani (2004), Competitive experimentation with private information, *Cowles Foundation DP 1489*.
- [26] Neveu, J., *Martingales à temps discret*, Masson, Paris, 1972.
- [27] Osborne M.J. and Rubinstein, A., *A course in Game Theory*, MIT Press, 1999.
- [28] Rotschild, M. (1974), A two-armed bandit theory of market pricing, *J. Econ. Theory*, **9**, 185–202.
- [29] Shiryaev, A. *Probability*. Springer, 1984.
- [30] Valone, T.J. and Templeton, J.J., *Phil. Trans. R. Soc. Lond. B*, **357**, 1549-1557.

Appendix

For the reader’s convenience, the organization of the appendix mirrors the organization of the main text.

A Beliefs

A.1 The private belief

We start with the proof of Lemma 4.3.

Proof of Lemma 4.3. We need to prove that $F_H(p) \leq F_L(p)$, for each $p \in [0, 1]$. By equality (3), one has $p_1^i \leq p$ if and only if $\frac{f_H(X_0^i)}{f_L(X_0^i)} \leq r$, with

$r := \frac{p(1-p_0)}{p_0(1-p)}$. For $r \leq 1$ one has

$$\begin{aligned} F_H(p) &= \int_{\left\{\frac{f_H(\cdot)}{f_L(\cdot)} \leq r\right\}} f_H(x) dx = \int_{\left\{\frac{f_H(\cdot)}{f_L(\cdot)} \leq r\right\}} \frac{f_H(x)}{f_L(x)} f_L(x) dx \\ &\leq r \int_{\left\{\frac{f_H(\cdot)}{f_L(\cdot)} \leq r\right\}} f_L(x) dx \leq r F_L(p) \leq F_L(p). \end{aligned} \quad (14)$$

For $r > 1$ we adapt the previous computation. For $\tilde{p} \in (p, 1]$, we set $\tilde{r} := \frac{\tilde{p}(1-p_0)}{(1-\tilde{p})p_0} > r$, so that $p_1^i \geq \tilde{p}$ if and only if $\frac{f_L(X_0^i)}{f_H(X_0^i)} \leq \frac{1}{\tilde{r}}$. Exchanging the roles of the two states in the computation, we obtain

$$\mathbf{P}_L(p_1^i \geq \tilde{p}) \leq \frac{1}{\tilde{r}} \mathbf{P}_H(p_1^i \geq p).$$

Letting $\tilde{p} \rightarrow p$ and using the right-continuity of F_H and F_L , we get

$$\mathbf{P}_L(p_1^i > p) \leq \frac{1}{r} \mathbf{P}_H(p_1^i > p),$$

which implies that

$$F_H(p) = \mathbf{P}_H(p_1^i \leq p) \leq \mathbf{P}_L(p_1^i \leq p) = F_L(p).$$

This concludes the proof of Proposition 4.3. ■

We now proceed with the extensions of Lemma 4.3.

Proof of Proposition 4.4. Under the assumption that $F_L(p) < 1$ and $F_H(p) > 0$, the computation in the above proof of Lemma 4.3 delivers $F_H(p) < F_L(p)$ unless $r = 1$. Therefore, we may assume w.l.o.g. that $r = 1$, which corresponds to $p = p_0$.

Since $\mathbf{E}[p_1^i] = p_0$ and since p_1^i has a density, one has $\mathbf{P}(p_1^i < p_0) > 0$. Hence, there is $\tilde{p} < p_0$ with $F_H(\tilde{p}) > 0$. Set now $\tilde{r} = \frac{\tilde{p}(1-p)}{p(1-\tilde{p})}$. Following the steps of the proof of Lemma 4.3,

$$\mathbf{P}_H(p_1^i \in (\tilde{p}, p]) = \int_{\left\{\frac{f_H(\cdot)}{f_L(\cdot)} \in (\tilde{r}, r]\right\}} f_L(x) dx \leq \mathbf{P}_L(p_1^i \in (\tilde{p}, p]).$$

On the other hand, since $\tilde{r} < r$ one has $F_H(\tilde{p}) < F_L(\tilde{p})$. Therefore,

$$F_H(p) = F_H(\tilde{p}) + \mathbf{P}_H(p_1^i \in (\tilde{p}, p]) < F_L(\tilde{p}) + \mathbf{P}_L(p_1^i \in (\tilde{p}, p]) = F_L(p).$$

this concludes the proof of Proposition 4.4. ■

We conclude with the proof of Proposition 4.5. It makes use of the following lemma.

Lemma A.1 *Let $K \in \mathbf{N}$ be given, and let X_1, \dots, X_K and Y_1, \dots, Y_K be two collections of independent real-valued variables, with continuous c.d.f.'s. Assume that X_k stochastically dominates Y_k , for each $k = 1, \dots, K$. For each k , set $\tilde{X}_k = X_1 + \dots + X_k$ and $\tilde{Y}_k = Y_1 + \dots + Y_k$. Then the vector $(\tilde{X}_k)_{k=1, \dots, K}$ stochastically dominates the vector $(\tilde{Y}_k)_{k=1, \dots, K}$. Moreover, one has $\mathbf{P}(\tilde{Y}_1 > y_1, \dots, \tilde{Y}_K > y_K) \leq \mathbf{P}(\tilde{X}_1 > y_1, \dots, \tilde{X}_K > y_K)$, for each $y_1, \dots, y_K \in \mathbf{R}$.*

Proof. For each k , we define a measurable and non-decreasing function $g_k : \mathbf{R} \rightarrow \mathbf{R}$ by

$$g_k(x) = \inf\{z : F_{Y_k}(z) \geq F_{X_k}(x)\}.$$

Observe that Y_k and $g_k(X_k)$ have the same law. Since X_k dominates Y_k , one has $g_k(x) \leq x$, for each $x \in \mathbf{R}$. The conclusions follow easily. ■

Proof of Proposition 4.5. For $n \in \mathbf{N}$, we define $Z_n^i := \ln \frac{f_H(X_n^i)}{f_L(X_n^i)}$ to be the log-likelihood ratio of the payoff in stage n . For each θ , the variables $(Z_n^i)_n$ are i.i.d. under \mathbf{P}_θ . By (3), one has

$$\ln \frac{p_n^i}{1 - p_n^i} = \ln \frac{p_0}{1 - p_0} + \sum_{k=0}^{n-1} Z_k^i. \quad (15)$$

By Lemma 4.3, the law of Z_n^i under \mathbf{P}_H stochastically dominates the law of Z_n^i under \mathbf{P}_L . By (15) and Lemma A.1, the law of the vector $(\ln \frac{p_1^i}{1 - p_1^i}, \dots, \ln \frac{p_n^i}{1 - p_n^i})$ under \mathbf{P}_H dominates the law of the same vector under \mathbf{P}_L . The result follows since the map $x \mapsto \ln \frac{x}{1-x}$ is an increasing bijection. ■

A.2 Markov properties of the beliefs

We here prove Proposition 4.14 relative to Markov chains. We will use the following technical observation.

Lemma A.2 *Let \mathcal{H}^1 and \mathcal{H}^2 be two independent σ -algebras on a probability space (Ω, \mathbf{P}) and, for $i = 1, 2$, let \mathcal{A}^i be a sub- σ -field of \mathcal{H}^i . For each $C^1 \in \mathcal{H}^1$ and each $C^2 \in \mathcal{H}^2$, one has*

$$\mathbf{P}(C^1 \cap C^2 \mid \sigma(\mathcal{A}^1, \mathcal{A}^2)) = \mathbf{P}(C^1 \mid \mathcal{A}^1) \times \mathbf{P}(C^2 \mid \mathcal{A}^2).$$

Proof. One has to prove that, for each $C \in \sigma(\mathcal{A}^1, \mathcal{A}^2)$, the equality $\mathbf{E}[1_{C^1 \cap C^2} 1_C] = \mathbf{E}[\mathbf{P}(C^1 \mid \mathcal{A}^1) \mathbf{P}(C^2 \mid \mathcal{A}^2) 1_C]$ holds. It is sufficient to prove the equality for

$C \in \mathcal{A}^1 \cup \mathcal{A}^2$. For concreteness, let $C \in \mathcal{A}^1$ be given. Note that

$$\begin{aligned} \mathbf{E} [\mathbf{P}(C^1 | \mathcal{A}^1) \mathbf{P}(C^2 | \mathcal{A}^2) 1_C] &= \mathbf{E} [\mathbf{P}(C^1 | \mathcal{A}^1) 1_C] \times \mathbf{E} [\mathbf{P}(C^2 | \mathcal{A}^2)] \\ &= \mathbf{E} [1_{C^1 \cap C}] \mathbf{E} [\mathbf{P}(C^2 | \mathcal{A}^2)] \\ &= \mathbf{P}(C^1 \cap C^2) \mathbf{P}(C^2) \end{aligned}$$

where the first equality follows by independence, and the other two from basic properties of conditional expectations. ■

We prove all three statements in Proposition 4.14 in turn. We start with the second one, since it is used in the proof of the other two.

Lemma A.3 *The sequence $(p_n^i)_{n \in \mathbf{N}}$ is a Markov chain for $(\mathcal{F}_n^i)_{n \in \mathbf{N}}$, both under \mathbf{P} and under \mathbf{P}_θ .*

Proof. We prove the claim for \mathbf{P} . Let a stage $n \in \mathbf{N}$ be given. The belief p_{n+1}^i is a function of p_n^i and $Z_n^i = \ln \frac{f_H(X_n^i)}{f_L(X_n^i)}$. Therefore, by Dynkin's lemma, it is sufficient to prove that, for any two Borel sets B_1 and B_2 in \mathbf{R} , one has $\mathbf{P}(p_n^i \in B_1, Z_n^i \in B_2 | \mathcal{F}_n^i) = \mathbf{P}(p_n^i \in B_1, Z_n^i \in B_2 | p_n^i)$.

Write

$$\begin{aligned} \mathbf{P}(p_n^i \in B_1, Z_n^i \in B_2 | \mathcal{F}_n^i) &= \sum_{\theta} \mathbf{P}(\Theta = \theta | \mathcal{F}_n^i) \mathbf{P}_\theta(p_n^i \in B_1, Z_n^i \in B_2 | \mathcal{F}_n^i) \\ &= \sum_{\theta} \mathbf{P}(\Theta = \theta | p_n^i) \mathbf{P}_\theta(p_n^i \in B_1, Z_n^i \in B_2 | \mathcal{F}_n^i). \end{aligned}$$

Since Z_n^i and \mathcal{F}_n^i are conditionally independent given Θ , and since p_n^i is \mathcal{F}_n^i -measurable, this equality reduces to

$$\mathbf{P}(p_n^i \in B_1, Z_n^i \in B_2 | \mathcal{F}_n^i) = \sum_{\theta} \mathbf{P}(\Theta = \theta | p_n^i) 1_{p_n^i \in B_1} \mathbf{P}_\theta(Z_n^i \in B_2),$$

and the result follows. ■

We proceed with the third statement.

Lemma A.4 *Under \mathbf{P}_θ , both sequences (p_n^i) and (t_n^j) are Markov chains for the filtration (\mathcal{G}_n^i) .*

Proof. Let a stage $n \in \mathbf{N}$ be given. Since p_{n+1}^i is \mathcal{F}_{n+1}^i -measurable, one has

$$\mathbf{P}_\theta(p_{n+1}^i \in B | \mathcal{G}_n^i) = \mathbf{P}_\theta(p_{n+1}^i \in B | \mathcal{F}_n^i) = \mathbf{P}_\theta(p_{n+1}^i \in B | p_n^i),$$

where the first equality follows by Lemma A.2 and the second one by Lemma A.3. Therefore, the sequence (p_n^i) is a Markov chain for (\mathcal{G}_n^i) .

If at some stage n , the other player is no longer active, both \mathcal{G}_n^i and t_n^j reveal that. Therefore for $k < n$, both $\mathbf{P}_\theta(t_{n+1}^j = k \mid \mathcal{G}_n^i)$ and $\mathbf{P}_\theta(t_{n+1}^j = k \mid t_n^j)$ are one if $t_n^j = k$, and zero otherwise: hence $\mathbf{P}_\theta(t_{n+1}^j = k \mid \mathcal{G}_n^i) = \mathbf{P}_\theta(t_{n+1}^j = k \mid t_n^j)$. For each of the two remaining values $\alpha = \blacksquare, n$ of t_{n+1}^j , one has, by Lemma A.2,

$$\mathbf{P}_\theta(t_{n+1}^j = \alpha \mid \mathcal{G}_n^i) = \mathbf{P}_\theta(t_{n+1}^j = \alpha \mid t_n^j),$$

and the result follows. ■

We conclude with the first assertion. We will use the equality

$$\mathbf{P}(G \mid \mathcal{G}_n^i) = \sum_{\theta \in \{L, H\}} \mathbf{P}(\Theta = \theta \mid \mathcal{G}_n^i) \times \mathbf{P}_\theta(G \mid \mathcal{G}_n^i), \quad (16)$$

that holds for each $G \in \mathcal{G}_{n+1}^i$. For $G \in \mathcal{F}_{n+1}^i$, by Lemma A.2, it reduces to

$$\mathbf{P}(G \mid \mathcal{G}_n^i) = q_n^i \mathbf{P}_H(G \mid \mathcal{F}_n^i) + (1 - q_n^i) \mathbf{P}_L(G \mid \mathcal{F}_n^i).$$

Lemma A.5 *Under \mathbf{P} , the sequence $(p_n^i, t_n^j)_{n \in \mathbf{N}}$ is a Markov chain for $(\mathcal{G}_n^i)_{n \in \mathbf{N}}$.*

Proof. Let a stage $n \in \mathbf{N}$ be given. By Theorem 9.6 in Doob (1953), it is sufficient to prove that, for each $\alpha \in \{\blacksquare, 1, \dots, n\}$, and for each Borel set $B \in \mathbf{R}$, one has $\mathbf{P}(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid \mathcal{G}_n^i) = \mathbf{P}(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid p_n^i, t_n^j)$. Let such an α and such a B be given. Using (16), write

$$\mathbf{P}(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid \mathcal{G}_n^i) = \sum_{\theta \in \{L, H\}} \mathbf{P}(\Theta = \theta \mid \mathcal{G}_n^i) \times \mathbf{P}_\theta(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid \mathcal{G}_n^i). \quad (17)$$

By Lemmas A.2 and A.3, for $\theta \in \{L, H\}$ one has

$$\begin{aligned} \mathbf{P}_\theta(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid \mathcal{G}_n^i) &= \mathbf{P}_\theta(p_{n+1}^i \in B \mid \mathcal{F}_n^i) \times \mathbf{P}_\theta(t_{n+1}^j = \alpha \mid t_n^j) \\ &= \mathbf{P}_\theta(p_{n+1}^i \in B \mid p_n^i) \times \mathbf{P}_\theta(t_{n+1}^j = \alpha \mid t_n^j). \end{aligned}$$

On the other hand, $\mathbf{P}(\Theta = H \mid \mathcal{G}_n^i) = q_n^i$ is only a function of p_n^i and of t_n^j , hence $\mathbf{P}(\Theta = H \mid \mathcal{G}_n^i) = \mathbf{P}(\Theta = H \mid p_n^i, t_n^j)$. Thus, (17) yields

$$\begin{aligned} &\mathbf{P}_\theta(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid \mathcal{G}_n^i) \\ &= \sum_{\theta \in \{L, H\}} \mathbf{P}(\Theta = \theta \mid p_n^i, t_n^j) \times \mathbf{P}_\theta(p_{n+1}^i \in B \mid p_n^i) \times \mathbf{P}_\theta(t_{n+1}^j = \alpha \mid t_n^j) \\ &= \mathbf{P}(p_{n+1}^i \in B, t_{n+1}^j = \alpha \mid p_n^i, t_n^j). \end{aligned}$$

this concludes the proof of the lemma. ■

B The optimal stopping problem

B.1 About the definition

We start by making explicit the equivalence between plans of player i and stopping times for $(\mathcal{G}_n^i)_{n \in \mathbf{N}}$. Let a plan ϕ^j be given. Given a plan $(\tau^i(\alpha), \alpha \in \mathbf{N} \cup \{\blacksquare\})$, define a stopping time σ^i by: $\sigma^i = \tau^i(t_n^j)$. Conversely, let σ^i be a stopping time for $(\mathcal{G}_n^i)_n$. We construct an equivalent plan ϕ^i as follows. For each $n \in \mathbf{N}$, the event $\{\sigma^i = n\}$ belongs to \mathcal{G}_n^i , hence there is a set $A \subset \mathbf{R}^n$, such that $\{\sigma^i = n\} = \{(X_1^i, \dots, X_{n-1}^i, t_n^j) \in A\}$, up to a \mathbf{P} -null set. For $k \in \{\blacksquare\} \cup \mathbf{N}$, define A_k as the set $\{(x_1^i, \dots, x_{n-1}^i) : (x_1^i, \dots, x_{n-1}^i, k) \in A\}$, and (for $n > k$) define $\{\tau^i(k) = n\}$ as the event $\{(X_1^i, \dots, X_{n-1}^i) \in A_k\}$.

The next two lemmas imply that the expected payoff $\mathbf{E}[Y_\sigma^i]$ is always well-defined. Recall that $Y_n^i := \sum_{k=0}^{n-1} \delta^k X_k^i$.

Lemma B.1 *The sequence $(Y_n^i)_{n \in \mathbf{N}}$ converges \mathbf{P} -a.s.*

Proof. Since $\mathbf{P} = \sum_{\theta} \mathbf{P}(\Theta = \theta) \mathbf{P}_{\theta}$, it is sufficient to prove that (Y_n^i) converges \mathbf{P}_{θ} -a.s., for each θ . Under \mathbf{P}_{θ} , the variables $(X_n^i)_{n \in \mathbf{N}}$ are independent, with the same finite variance. Thus, the two series $\sum \mathbf{E}_{\theta}[\delta^n X_n^i]$ and $\sum \text{Var}_{\theta}[\delta^n X_n^i]$ converge. By Kolmogorov's two series criterion (see, *e.g.*, Shiryaev, 1984, p. 386), the sequence $\sum_n \delta^n X_n^i$ converges \mathbf{P} -a.s., as desired. ■

Lemma B.2 *The variable $\sup_{n \in \mathbf{N}} |Y_n^i|$ is \mathbf{P} -integrable.*

Proof. It is sufficient to prove that $\mathbf{E}_{\theta}[\sup_n |Y_n^i|] < +\infty$, for each θ . For $n \in \mathbf{N}$, define $\tilde{X}_n^i := X_n^i - \mathbf{E}_{\theta}[X_n^i]$. Under \mathbf{P}_{θ} , the variables (\tilde{X}_n^i) are centered and independent. Therefore, the sequence $\tilde{Y}_n^i := \sum_{k=0}^{n-1} \delta^k \tilde{X}_k^i$ of partial sums is a martingale under \mathbf{P}_{θ} , with $\mathbf{E}_{\theta}[\tilde{Y}_n^i] = 0$. Observe that

$$Y_n^i = \tilde{Y}_n^i + \sum_{k=0}^{n-1} \delta^k \mathbf{E}_{\theta}[X_k^i],$$

so that $\sup_{n \in \mathbf{N}} |Y_n^i| \leq \sup_{n \in \mathbf{N}} |\tilde{Y}_n^i| + \mathbf{E}_{\theta}[X_0^i]/(1 - \delta)$.

Since $\mathbf{E}_{\theta}[\tilde{X}_n^2] \leq \text{Var}_{\theta}(X_n^i)$, one has

$$\mathbf{E}_{\theta} \left[\left(\tilde{X}_n^i \right)^2 \right] < K := \frac{1}{1 - \delta^2} \text{Var}_{\theta}(X_0^i), \text{ for each } n \in \mathbf{N}.$$

By Kolmogorov's inequality (see, *e.g.*, Shiryaev, 1984, p. 384), this yields

$$\mathbf{P}_\theta \left(\sup_{n \leq k} |\tilde{Y}_n^i| > c \right) < \frac{K}{c^2}, \text{ for each } c > 0 \text{ and } k \in \mathbf{N}.$$

The result follows, since

$$\mathbf{E}_\theta \left[\sup |\tilde{Y}_n^i| \right] \leq 1 + \sum_{c \in \mathbf{N}} \mathbf{P}_\theta \left(\sup_{n \in \mathbf{N}} |\tilde{Y}_n^i| \geq c \right) \leq 1 + \sum_{c \in \mathbf{N}} \frac{K}{c^2} < +\infty.$$

■

B.2 The Snell envelope

The quantity $V_n^i := Y_n^i + \delta^n \max(0, W_n^i) = \text{ess sup}_{\sigma \in \Lambda_n} \mathbf{E}[Y_\sigma^i \mid \mathcal{G}_n^i]$ is the optimal payoff of player i when she must stay in for at least n stages. It is called the *Snell envelope* of the sequence (Y_n^i) .

It satisfies the following dynamic programming principle (see Chow and Robbins (1963), or Ferguson (2004, ch. 3) for a proof).

Lemma B.3 *One has*

$$V_n^i = \max \{ Y_n^i, \mathbf{E} [V_{n+1}^i \mid \mathcal{G}_n^i] \} \text{ and } \mathbf{E} [V_n^i] = \sup_{\sigma \in \Lambda_n} \mathbf{E}[Y_\sigma^i].$$

Besides, $(V_n^i)_{n \in \mathbf{N}}$ is the least supermartingale that majorizes $(Y_n^i)_{n \in \mathbf{N}}$.

Indeed, the optimal continuation payoff is the maximum between dropping out at stage $n+1$ and the optimal continuation payoff if the player stays in at stage $n+1$.

Lemma 5.2 follows from Lemma B.4.

Lemma B.4 *For each stage $n \in \mathbf{N}$, there is a measurable map v_n^i such that $V_n^i = Y_n^i + v_n^i(p_n^i, t_n^j)$.*

Proof. We obtain v_n^i as the limit of the sequence $(\nu_n^k)_{k \in \mathbf{N}}$, which is defined inductively as follows. We set $\nu_n^0 = 0$, for each $n \in \mathbf{N}$. Given an index $k \in \mathbf{N}$ and a stage $n \in \mathbf{N}$, we define

$$\nu_n^k(p_n^i, t_n^j) := \max \left\{ 0, \delta^n X_n^i + \mathbf{E} \left[\nu_{n+1}^{k-1}(p_{n+1}^i, t_{n+1}^j) \mid \mathcal{G}_n^i \right] \right\}. \quad (18)$$

Since $(p_n^i, t_n^j)_n$ is a Markov chain for $(\mathcal{G}_n^i)_n$ under \mathbf{P} , the maximum on the right-hand side is indeed a function of (p_n^i, t_n^j) . For every fixed n we have $\nu_n^0 = 0 \leq \nu_n^1$,

and therefore by induction the sequence $(\nu_n^k)_k$ is non-decreasing. In particular, $v_n^i := \lim_{k \rightarrow +\infty} \nu_n^k$ exists.

It remains to prove that $(Y_n^i + v_n^i(p_n^i, t_n^j))_n$ is the Snell envelope of $(Y_n^i)_n$. By the dominated convergence for conditional expectations, letting $k \rightarrow +\infty$ in (18) yields

$$v_n^i(p_n^i, t_n^j) := \max \left\{ 0, \delta^n X_n^i + \mathbf{E} \left[v_{n+1}^i(p_{n+1}^i, t_{n+1}^j) \mid \mathcal{G}_n^i \right] \right\}.$$

Thus, $(Y_n^i + v_n^i(p_n^i, t_n^j))_{n \in \mathbf{N}}$ is a supermartingale which majorizes $(Y_n^i)_{n \in \mathbf{N}}$, so that $Y_n^i + v_n^i(p_n^i, t_n^j) \geq V_n^i$, \mathbf{P} -a.s. for each stage $n \in \mathbf{N}$.

Conversely, we first argue by induction that

$$V_n^i \geq Y_n^i + \nu_n^k(p_n^i, t_n^j), \quad (19)$$

\mathbf{P} -a.s. for each stage n and k . Indeed, since $\nu_n^0 = 0$, and since $V_n^i \geq Y_n^i$ (19) holds for $k = 0$ and every $n \in \mathbf{N}$. Since $V_n^i = \max\{Y_n^i, \mathbf{E}[V_{n+1}^i \mid \mathcal{G}_n^i]\}$, and using (19) with $n + 1$, this yields $V_n^i \geq Y_n^i \max\left\{0, \mathbf{E}[\nu_{n+1}^k(p_{n+1}^i, t_{n+1}^j) \mid \mathcal{G}_n^i]\right\} = \nu_n^{k+1}(p_n^i, t_n^j)$, for each $n \in \mathbf{N}$, so that (19) holds for $k + 1$.

Finally, letting $k \rightarrow +\infty$ yields $V_n^i \geq Y_n^i + v_n^i(p_n^i, t_n^j)$. this concludes the proof. \blacksquare

B.3 An explicit form for w_n^i

In this section we construct a version of $W_n^i = E[Y_{n \rightarrow \sigma_{n+1}^*}^i \mid \mathcal{G}_n^i]$, as a function of p_n^i and t_n^j . It relates to the explicit version of q_n^i as a function of p_n^i and t_n^j , and to the expression of the optimal stopping time σ^* in terms of $w_n^i(p_n^i, t_n^j)$. In particular, we prove that Eq. (8) holds.

Observe first that

$$W_n^i = \sum_{\theta \in \{L, H\}} \mathbf{P}(\Theta = \theta \mid \mathcal{G}_n^i) \mathbf{E}_\theta \left[Y_{n \rightarrow \sigma_{n+1}^*}^i \mid \mathcal{G}_n^i \right], \quad (20)$$

and recall that $Q_n^i(p_n^i, t_n^j) = \mathbf{P}(\Theta = H \mid \mathcal{G}_n^i)$. We claim that for fixed θ

$$\mathbf{E}_\theta \left[Y_{n \rightarrow \sigma_{n+1}^*}^i \mid \mathcal{G}_n^i \right] = \sum_{k \in \mathbf{N}} \mathbf{E}_\theta[X_0^i] \sum_{k \in \mathbf{N}} \delta^k \mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i). \quad (21)$$

In words, if active in stage $n + k$, which occurs when $\sigma_{n+1}^* > n + k$, player i receives the payoff $\mathbf{E}_\theta[X_0^i]$ in that stage, which is discounted back to stage n .

Formally, let $k \in \mathbf{N}$ be given. Observe that the event $\sigma_{n+1}^* > n + k$ is \mathcal{G}_{n+k}^i -measurable, while X_{n+k}^i is independent of \mathcal{G}_{n+k}^i given Θ . Therefore,

$$\begin{aligned} \mathbf{E}_\theta \left[X_{n+k}^i 1_{\sigma_{n+1}^* > n+k} \mid \mathcal{G}_n^i \right] &= \mathbf{E}_\theta \left[\mathbf{E} \left[X_{n+k}^i 1_{\sigma_{n+1}^* > n+k} \mid \mathcal{G}_{n+k}^i \right] \mid \mathcal{G}_n^i \right] \\ &= \mathbf{E}_\theta[X_0^i] \times \mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i). \end{aligned}$$

The equality (21) follows by summation over $k \in \mathbf{N}$.

In the light of (20) and (21), we need to find a version for the conditional probability $\mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i)$.

By definition of σ_{n+1}^* , one has $\{\sigma_{n+1}^* > n + k\} = \bigcap_{l=1}^k \{w_{n+l}^i(p_{n+l}^i, t_{n+l}^j) > 0\}$.

Let $l = 1, \dots, k$ be given. By (3), the private belief p_{n+l}^i in stage $n + l$ is related to the private belief in stage n through the equality

$$\frac{p_{n+l}^i}{1 - p_{n+l}^i} = \frac{p_n^i}{1 - p_n^i} \times \prod_{m=n}^{n+l-1} \frac{f_H(X_m^i)}{f_L(X_m^i)}, \quad (22)$$

so that p_{n+l}^i coincides with a function of the vector $(p_n^i, X_n^i, \dots, X_{n+l-1}^i)$.

Therefore, the indicator function of the set $\{\sigma_{n+1}^* > n + k\}$ also coincides with a function of $(p_n^i, X_n^i, \dots, X_{n+l-1}^i)$, which we denote ε_k .

Under \mathbf{P}_θ , the vectors $\vec{t}^j = (t_{n+1}^j, \dots, t_{n+k}^j)$ and $\vec{X}^i = (X_n^i, \dots, X_{n+k-1}^i)$ are conditionally independent given \mathcal{G}_n^i .

Consequently, the conditional probability $\mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i)$ can be expanded into

$$\mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i) = \sum_{\vec{\alpha}} \mathbf{P}_\theta(\vec{t}^j = \vec{\alpha} \mid \mathcal{G}_n^i) \mathbf{E}_\theta[\varepsilon_k(p_n^i, \vec{X}^i, \vec{\alpha}^j) \mid \mathcal{G}_n^i].$$

Since (t_n^j) is a Markov chain for (\mathcal{G}_n^i) under \mathbf{P}_θ , since p_n^i is \mathcal{G}_n^i -measurable, and since \vec{X}^i is independent of \mathcal{G}_n^i under \mathbf{P}_θ , this yields

$$\mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i) = \sum_{\vec{\alpha}} \mathbf{P}_\theta(\vec{t}^j = \vec{\alpha} \mid t_n^j) \times \int_{\mathbf{R}^k} \varepsilon_k(p_n^i, x_1, \dots, x_k, \vec{\alpha}) \prod_{l=1}^k f_\theta(x_l) dx_l. \quad (23)$$

Thus, $\mathbf{P}_\theta(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i)$ is a function of (t_n^j, p_n^i) .

Define now

$$\begin{aligned} w_n^i(p, \alpha) &= \sum_{k \in \mathbf{N}} \left\{ Q_n^i(p, \alpha) \mathbf{E}_H[X_0^i] \delta^k \mathbf{P}_H(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i) \right. \\ &\quad \left. + (1 - Q_n^i(p, \alpha)) \mathbf{E}_L[X_0^i] \delta^k \mathbf{P}_L(\sigma_{n+1}^* > n + k \mid \mathcal{G}_n^i) \right\}. \end{aligned} \quad (24)$$

By Eqs. (20), (23) and (21) $w_n^i(p_n^i, t_n^j) = W_n^i$ \mathbf{P} -a.s., as desired.

B.4 Regularity properties of the Snell envelope

We prove here Lemmas 5.3. The proof makes use of the following technical lemma.

Lemma B.5 *Let ν be a probability measure over \mathbf{R} , absolutely continuous w.r.t. the Lebesgue measure, and let $B \subseteq \mathbf{R}$ be a Borel set. Then the map $x \in \mathbf{R} \mapsto \nu(x + B)$ is continuous.*

Proof. Let a sequence $(x_n)_{n \in \mathbf{N}}$ be given, converging to some $x \in \mathbf{R}$. Set $B_x = B + x$, and, for $n \in \mathbf{N}$, define a measure ν_n by $\nu_n(C) = \nu(C + x_n - x)$. The sequence $(\nu_n)_n$ converges weakly to ν . Since ν is absolutely continuous w.r.t. the Lebesgue measure, one has $\nu(\partial B_x) = 0$, where ∂B_x is the topological boundary of B_x . Therefore, by Theorem 25.8 in Billingsley (1995), one has $\lim_{n \rightarrow +\infty} \nu_n(B_x) = \nu(B_x)$, and the result follows. ■

Lemma B.6 *For every fixed status variable $\alpha \in \mathbf{N} \cup \{\blacksquare\}$, the function $p \mapsto w_n^i(p, \alpha)$ is continuous in $[0, 1]$.*

Proof. Let a stage $n \in \mathbf{N}$, and $k \in \mathbf{N}$ be given. Fix a vector $\vec{\alpha} = (\alpha_{n+1}, \dots, \alpha_{n+k})$, with components in $\mathbf{N} \cup \{\blacksquare\}$. By Eqs. (23) and (24), it is sufficient to prove that the map

$$p \mapsto \int_{\mathbf{R}^k} \varepsilon_k(p, \vec{x}, \vec{\alpha}) \prod_{l=1}^k f_\theta(x_l) dx_l$$

is continuous in p , for each choice of $k \in \mathbf{N}$.

We proceed as follows. By Fubini's Theorem, we first integrate w.r.t. x_1 , and write

$$\int_{\mathbf{R}^k} \varepsilon_k(p, \vec{x}, \vec{\alpha}) \prod_{l=1}^k f_\theta(x_l) dx_l = \int_{\mathbf{R}^{k-1}} \left\{ \int_{\mathbf{R}} \varepsilon_k(p, x, \vec{x}_{-1}, \vec{\alpha}) f_\theta(x) dx \right\} \prod_{l=2}^k f_\theta(x_l) dx_l.$$

We will prove that the inner integral is continuous in p , for each choice of (x_2, \dots, x_k) . Once this claim is established, the continuity property follows by the dominated convergence theorem, applied to the integral over \mathbf{R}^{k-1} in (B.4).

Thus, let (x_2, \dots, x_k) be given. We will describe the set of values (p, x) such that $\varepsilon_k(p, x, \vec{x}_{-1}, \vec{\alpha}) = 1$. In effect, fixing the payoffs x_2, \dots, x_k to player 1 in stages n through $n + k - 2$, and the status $\vec{\alpha}$ of player j in stages $n + 1$ through $n + k$, this amounts to characterizing the values of p_n^i and X_n^i for which player i will still be active in stage $n + k$ (included).

By definition, $\varepsilon_k(p, x, \vec{x}_{-1}, \vec{\alpha})$ is equal to one if and only if, for each $l = 1, \dots, k$, one has $g_{n+l}^i(p_{n+l}^i, \alpha_{n+l}) > 0$, where p_{n+l}^i is the private belief in stage $n+l$, as computed from $p_n^i = p$ and the payoffs x, x_2, \dots, x_{l-1} received in stages $n, n+1, \dots, n+l-1$.

As a function of (p, x) and of (x_2, \dots, x_k) , this private belief is given by (see the equality (22)).

$$\ln \frac{p_{n+l}^i}{1 - p_{n+l}^i} = \ln \frac{p}{1 - p} + \ln \frac{f_H(x)}{f_L(x)} + \sum_{m=2}^{l-1} \ln \frac{f_H(x_m)}{f_L(x_m)}, \quad (25)$$

In the light of (25), the condition $g_{n+l}^i(p_{n+l}^i, \alpha_{n+l}) > 0$ rewrites

$$\ln \frac{p}{1 - p} + \ln \frac{f_H(x)}{f_L(x)} \in F_l,$$

for some Borel set F_l in \mathbf{R} .

Define now $F := \bigcap_{l=1}^k F_l$. According to the above discussion,

$$\int_{\mathbf{R}} \varepsilon_k(p, x, \vec{x}_{-1}, \vec{\alpha}) f_{\theta}(x) dx = \int_{\mathbf{R}} 1_{\{\ln \frac{p}{1-p} + \ln \frac{f_H(x)}{f_L(x)} \in F\}} f_{\theta}(x) dx. \quad (26)$$

Introduce the function $\omega(t) := 1_{\ln \frac{p}{1-p} + t}$. By (26),

$$\begin{aligned} \int_{\mathbf{R}} \varepsilon_k(p, x, \vec{x}_{-1}, \vec{\alpha}) f_{\theta}(x) dx &= \mathbf{E}_{\theta} \left[\omega \left(\ln \frac{f_H(X_0^i)}{f_L(X_0^i)} \right) \right] \\ &= \int_{\mathbf{R}} \omega(y) d\nu(y) = \nu(F - \ln \frac{p}{1-p}), \end{aligned}$$

where ν is the law under \mathbf{P}_{θ} of the random variable $\ln \frac{f_H(X_0^i)}{f_L(X_0^i)}$. By assumption **A4**, ν is absolutely continuous w.r.t. the Lebesgue measure. Therefore, the continuity claim follows from Lemma B.5. ■

Lemma B.7 *For every fixed status variable $\alpha \in \mathbf{N} \cup \{\blacksquare\}$, the function $p \mapsto w_n^i(p, \alpha)$ is increasing.*

Proof. The monotonicity assertion will follow if, for any given $p \in [0, 1]$, one has $w_n^i(p_n^i, t_n^j) > w_n^i(p, \alpha)$ on the event $p_n^i > p, t_n^j = \alpha$.

Let $p \in [0, 1]$ be given. We define a stopping time σ by

$$\sigma = \inf \left\{ k \geq n+1 : w_{n+k}^i(p_{n+k}^i(p, X_n^i, \dots, X_{n+k-1}^i), t_{n+k}^j) = 0 \right\}.$$

Note that σ stops when player i would have stopped for the same payoffs from stage n to stage $n + k - 1$ and if her belief at stage n was p . When $p_n^i = p$, σ is equal to σ_{n+1}^* .

Decompose

$$\mathbf{E}[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i] = q_n^i \mathbf{E}_H[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i] + (1 - q_n^i) \mathbf{E}_L[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i].$$

Since σ does not depend on p_n^i , we obtain

$$\mathbf{E}[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i] = q_n^i \mathbf{E}_H[Y_{n \rightarrow \sigma}^i \mid t_n^j] + (1 - q_n^i) \mathbf{E}_L[Y_{n \rightarrow \sigma}^i \mid t_n^j].$$

Thus, p_n^i influences $\mathbf{E}[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i]$ only through q_n^i .

By Lemma 4.9, $q_n^i > Q_n^i(p, \alpha)$, if $p_n^i > p$ and $t_n^j = \alpha$. Since $\mathbf{E}_H[Y_{n \rightarrow \sigma}^i \mid t_n^j] > 0 > \mathbf{E}_L[Y_{n \rightarrow \sigma}^i \mid t_n^j]$, the quantity $\mathbf{E}[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i]$ (which depends on p) increases with p . The result follows by (7). ■

As a consequence, given α , there is a unique $p \in (0, 1)$ such that $w_n^i(p, \alpha) = 0$. We denote it by $\pi_n^i(\alpha)$.

B.5 Uniqueness of the best reply

We now show that σ^* is the unique solution to \mathcal{P} , thereby establishing Corollary 5.5. The intuition is simple. At stage n , player i is indifferent between dropping or not if $W_n^i = 0$, that is, if $p_n^i = \pi_n^i(t_n^j)$. Since the law of p_n^i has a density, and since t_n^j takes at most finitely many values, this occurs with probability zero. Thus, there is a zero probability that player i will ever be indifferent between the two alternatives.

Proposition B.8 *Let σ be an optimal stopping time for \mathcal{P} . One has $\sigma = \sigma^*$, \mathbf{P} -a.s.*

Proof. Let σ be an optimal stopping time. In the light of Lemma 5.1, we already know that $\sigma \geq \sigma^*$, \mathbf{P} -a.s.

We argue by contradiction, and assume that, for some stage $n \in \mathbf{N}$, and some α , one has $\mathbf{P}(\sigma^* = n < \sigma, t_n^j = \alpha) > 0$. Recall that $\mathbf{E}[Y_{n \rightarrow \sigma}^i \mid \mathcal{G}_n^i] \leq W_n^i = w_n^i(p_n^i, t_n^j)$. If $\{\sigma^* = n < \sigma, t_n^j = \alpha\}$, one has $p_n^i < \pi_n^i(\alpha)$, hence $w_n^i(p_n^i, \alpha) < 0$. Therefore, $\mathbf{E}[Y_{n \rightarrow \sigma}^i 1_{n=\sigma^* < \sigma, t_n^j=\alpha}] < 0$. Summing over n and α yields $\mathbf{E}[Y_{\sigma^* \rightarrow \sigma}^i 1_{\sigma^* < \sigma}] < 0$. One obtains

$$\mathbf{E}[Y_\sigma^i] = \mathbf{E}[Y_{\sigma^*}^i + Y_{\sigma^* \rightarrow \sigma}^i] < \mathbf{E}[Y_{\sigma^*}^i],$$

a contradiction. This concludes the proof. ■

C Large games

Given $N \in \mathbf{N}$, we consider a version of the game with player set $I_N = \{1, \dots, N\}$. For each $N \in \mathbf{N}$, we let a cut-off equilibrium be given, with cut-offs $\pi_n^{N,i}$. The argument of $(\pi_n^{N,i})$ is an $N - 1$ -dimensional vector $\vec{\alpha}^{N-1}$. Since all players are active at stage 0, only the cut-off $\pi_1^{N,i}(\vec{\alpha}^{N-1})$ is meaningful at stage 1, which we abbreviate to $\pi_1^{N,i}$.

Adopting the notations of Section 5, we denote by $w_1^{N,i}(p)$ the OCP of player i in the N -player game, when she reaches stage 1 with a private belief equal to p .¹⁹ The function $w_1^{N,i}$ is increasing and continuous, and $\pi_1^{N,i} = \inf\{p: w_1^{N,i}(p) \geq 0\}$.

We prove Proposition C.1, which readily implies the first statement in Theorem 2.7.

Proposition C.1 *One has*

$$\lim_{N \rightarrow +\infty} \sup_{p \in [0,1], i \in I_N} \left| w_1^{N,i}(p) - \left(p \frac{\mathbf{E}_H[X_0^i]}{1-\delta} + (1-p)\mathbf{E}_L[X_0^i] \right) \right| = 0.$$

Proof. We start with some preparation, and set $o(p) := p \frac{\mathbf{E}_H[X_0^i]}{1-\delta} + (1-p)\mathbf{E}_L[X_0^i]$. By Lemma B.7 all functions $w_1^{N,i}$ ($N \in \mathbf{N}, i \in I_N$) are increasing. Therefore it is enough to prove that for each fixed $p \in [0, 1]$ one has

$$\lim_{N \rightarrow +\infty} \sup_{i \in I_N} \left| w_1^{N,i}(p) - o(p) \right| = 0.$$

We will prove the stronger claim that for any sequence $(i_N)_{N \in \mathbf{N}}$ of players such that $i_N \in \{1, \dots, N\}$ one has $\lim_{N \rightarrow +\infty} \left| w_1^{N,i_N}(p) - o(p) \right| = 0$. We therefore let such a sequence be given. For notational ease, it is convenient to relabel players so that $i_N = 1$ for each $N \in \mathbf{N}$. We will prove below that

$$\lim_{N \rightarrow +\infty} \left| w_1^{N,1}(p) - o(p) \right| = 0. \quad (27)$$

Each N -player game involves a probability space (Ω^N, \mathbf{P}^N) that is rich enough to accommodate payoff sequences and randomizing devices for all players. It is convenient to proceed under the assumption that all these probability spaces coincide. This is w.l.o.g., since one may always embed the sequence (Ω^N, \mathbf{P}^N) into the product space $(\times_{N \in \mathbf{N}} \Omega_N, \otimes_{N \in \mathbf{N}} \mathbf{P}_N)$. More constructively, we may take (Ω, \mathbf{P}) to be rich enough to support countably many sequences of payoffs and randomizing devices with the required properties.

¹⁹under the constraint that she remains active in stage 1.

The proof is organized along the lines of the sketch provided in Section 7. We first prove that the (random) proportion of players who drop out in stage 1 converge \mathbf{P} -a.s. to some constant ρ_θ as the number of players increases to infinity, with $\rho_L > \rho_H$.

For each $N \geq 2$, we denote by μ^N the empirical distribution of the cut-offs $\pi_1^{N,i}$, $i > 1$, used by player 1's opponents in stage 1. That is, μ^N is the atomic distribution on $[0, 1]$ defined by $\mu^N(\{x\}) = \frac{1}{N-1} |\{i > 1 : \pi_1^{N,i} = x\}|$. By Corollary 6.4, μ^N assigns probability one to the interval $[p_*, p_0]$. Since the set of probability distributions on $[0, 1]$, endowed with the topology of weak convergence, is compact and metric, there is a subsequence of $(\mu^N)_{N \in \mathbf{N}}$ that converges weakly, say to μ . We will prove that (27) holds along this subsequence. Since the converging subsequence is arbitrary, the result also holds along the original sequence. We relabel the subsequence as $(\mu^N)_{N \in \mathbf{N}}$.

For $\theta = L, H$, we set $\rho_\theta = \int_{[0,1]} F_\theta(x) d\mu(x)$, where F_θ is the c.d.f. of p_1^i under \mathbf{P}_θ . Observe that F_θ does not depend on i . Thus, ρ_θ is the limit probability (as the number N of players increases) that a randomly selected player will drop out at stage 1. Lemma C.2 below substantiates this interpretation.

Given a number $N \in \mathbf{N}$ of players, we denote by ω^N the (random) proportion of players (not including player 1) who drop out at stage 1. That is, $\omega^N = \frac{1}{N-1} \left| \left\{ i > 1 : p_1^i \leq \pi_1^{N,i} \right\} \right|$.

Lemma C.2 *For $\theta = L, H$, one has $\lim_{N \rightarrow +\infty} \omega^N = \rho_\theta$, \mathbf{P}_θ -a.s.*

Proof. Fix θ , set $T_{N,i} := 1_{\{p_1^i \leq \pi_1^{N,i}\}} - F_\theta(\pi_1^{N,i})$, and define

$$\tilde{\omega}^N := \frac{1}{N-1} \sum_{i=2}^N T_{N,i}.$$

For each N , the variables $T_{N,i}$ ($i > 1$) are centered, independent, and bounded by 1. From Chebyshev's inequality it follows that $(\tilde{\omega}^N)_{N \geq 2}$ converges \mathbf{P}_θ -a.s. to zero (see Cantelli's proof of the strong law of large numbers, Shiryaev, 1984, p.388), and the result follows. ■

Lemma C.3 *One has $\rho_L > \rho_H$.*

Proof. We first prove that $F_\theta(x) \in (0, 1)$ whenever $x \in [p_*, p_0]$. By assumption A4, $\inf C = 0$, where C is the support of p_1^1 . This implies $F_\theta(p_*) > 0$. Since $\ln \frac{p_1^1}{1-p_1^1} = \ln \frac{p_0}{1-p_0} + \ln \frac{f_H(X_0^1)}{f_L(X_0^1)}$, we have $\mathbf{E}_H[\ln \frac{p_1^1}{1-p_1^1}] > \ln \frac{p_0}{1-p_0}$. Since p_1^1 is non-constant, this implies $F_H(p_0) < 1$, and $F_L(p_0) < 1$ as well. Therefore $F_\theta(x) \in (0, 1)$ for each $x \in [p_*, p_0]$.

By Proposition 4.4 this yields $F_H(x) < F_L(x)$ for each $x \in [p_*, p_0]$. Since μ is concentrated on $[p_*, p_0]$, it follows that $\int_{[0,1]} F_H(x) d\mu(x) < \int_{[0,1]} F_L(x) d\mu(x)$. ■

We now conclude by providing a lower bound for the OCP. For each $N \in \mathbf{N}$, define a stopping time σ^N by $\sigma^N = 2$ if $\omega^N \geq (\rho_L + \rho_H)/2$, and $\sigma^N = +\infty$ otherwise. According to σ^N , if the proportion of the number of players who stay in at stage 1 is closer to ρ_H , player 1 stays in forever, whereas if it is closer to ρ_L , she drops out at stage 2.

By Lemma C.2 the proportion of the number of players who stay in at stage 1 reveals the common type, so that $\lim_{N \rightarrow +\infty} \sigma^N = +\infty$ \mathbf{P}_H -a.s., and $\lim_{N \rightarrow +\infty} \sigma^N = 2$ \mathbf{P}_L -a.s. By dominated convergence, it follows that $\lim_{N \rightarrow +\infty} \mathbf{E}_H[Y_{1 \rightarrow \sigma^N}^1] = \frac{\mathbf{E}_H[X_0^i]}{1-\delta}$, while $\lim_{N \rightarrow +\infty} \mathbf{E}_L[Y_{1 \rightarrow \sigma^N}^1] = \mathbf{E}_L[X_0^i]$. Consider now the conditional continuation payoff given \mathcal{G}_1^1 . By the choice of σ^N , the continuation payoff $Y_{1 \rightarrow \sigma^N}^1$ is conditionally independent of \mathcal{G}_1^1 , given θ , hence $\mathbf{E}_\theta[Y_{1 \rightarrow \sigma^N}^1 \mid \mathcal{G}_1^1] = \mathbf{E}_\theta[Y_{1 \rightarrow \sigma^N}^1]$. Therefore,

$$\mathbf{E}[Y_{1 \rightarrow \sigma^N}^1 \mid \mathcal{G}_1^1] = p_1^1 \mathbf{E}_H[Y_{1 \rightarrow \sigma^N}^1] + (1 - p_1^1) \mathbf{E}_L[Y_{1 \rightarrow \sigma^N}^1],$$

so that $w_1^{N,1}(p) \geq p \mathbf{E}_H[Y_{1 \rightarrow \sigma^N}^1] + (1 - p) \mathbf{E}_L[Y_{1 \rightarrow \sigma^N}^1]$. Letting $N \rightarrow +\infty$ yields

$$\liminf_{N \rightarrow +\infty} w_1^{N,1}(p) \geq o(p).$$

■

We proceed to the limit behavior in stage 2. We first set up some notation. Recall that ρ_θ is the (limit) probability that an arbitrary player will stop in stage 1. Define $\rho_* \in (0, 1)$ by the equality

$$\rho_* \ln \frac{\rho_H}{\rho_L} + (1 - \rho_*) \ln \frac{1 - \rho_H}{1 - \rho_L} = 0.$$

We will prove Proposition C.4 below, which concludes the proof of Theorem 2.7.

Proposition C.4 *For $N \in \mathbf{N}$, let $\vec{k}^N = (k_j^N)_{1 \leq j \leq N} \in \{\blacksquare, 1\}^{N-1}$ be a $N - 1$ -dimensional vector. Assume that the sequence $\rho^N := \frac{1}{N-1} \left| \left\{ 1 \leq j \leq N : k_j^N = 1 \right\} \right|$ has a limit $\rho \in [0, 1]$.*

If $\rho < \rho_$, then $\lim_{N \rightarrow \infty} \pi_2^{N,1}(\vec{k}^N) = 0$.*

If $\rho > \rho_$, then $\lim_{N \rightarrow \infty} \pi_2^{N,1}(\vec{k}^N) = 1$.*

Lemma 4.8 generalizes to N -player games in an obvious way and the relation between the posterior and the private beliefs of player 1 at stage 2 is expressed by the equality

$$\frac{q_2^{N,1}}{1 - q_2^{N,1}} = \frac{p_2^1}{1 - p_2^1} \times \prod_{j \neq 1, k_j^N = 1} \frac{F_H(\pi_1^{N,j})}{F_L(\pi_1^{N,j})} \times \prod_{j \neq 1, k_j^N = \blacksquare} \frac{1 - F_H(\pi_1^{N,j})}{1 - F_L(\pi_1^{N,j})} \quad (28)$$

in the event D where the set of players who actually dropped out in stage 1 coincides with $\{1 < j \leq N : k_j^N = 1\}$.

We denote by \prod_N^1 and \prod_N^\blacksquare the two products that appear on the right-hand side of (28). The following lemma implies that $\prod_N^1 \times \prod_N^\blacksquare$ converges either to $+\infty$ or to 0, depending on whether $\rho < \rho^*$ or $\rho > \rho^*$.

Lemma C.5 *One has*

$$\lim_{N \rightarrow \infty} \frac{1}{N-1} (\ln \prod_N^1 + \ln \prod_N^\blacksquare) = \rho \ln \frac{\rho_H}{\rho_L} + (1 - \rho) \ln \frac{1 - \rho_H}{1 - \rho_L}.$$

Proof. Since F_θ is continuous, and by Proposition C.1, one has

$$\lim_{N \rightarrow +\infty} \sup_{j \neq 1} |F_\theta(\pi_1^{N,1}) - \rho_\theta| = 0.$$

Thus, $\lim_{N \rightarrow +\infty} \left| \frac{1}{N-1} (\ln \prod_N^1 + \ln \prod_N^\blacksquare) - \left(\rho^N \ln \frac{\rho_H}{\rho_L} + (1 - \rho^N) \ln \frac{1 - \rho_H}{1 - \rho_L} \right) \right| = 0$, and the result follows, since $\lim_{N \rightarrow +\infty} \rho^N = \rho$. ■

Proof of the second assertion. Let us assume, say, that $\rho < \rho_*$. By Lemma C.5, $\lim_{N \rightarrow \infty} \prod_N^1 \times \prod_N^\blacksquare = +\infty$. Let $\varepsilon > 0$ be given, choose λ large enough so that $\frac{\varepsilon}{1-\varepsilon} \lambda \geq \frac{p_0}{1-p_0}$, and N large enough so that $\prod_N^1 \times \prod_N^\blacksquare \geq \lambda$. By the choice of N and λ , and from (28), it follows that $q_2^{N,1} \geq p_0$ on the event $\{p_2^1 \geq \varepsilon\} \cap D$. By Corollary 6.3, this implies $\pi_2^{N,1}(\vec{k}^N) \leq \varepsilon$. Since ε is arbitrary, one gets $\lim_N \pi_2^{N,1}(\vec{k}^N) = 0$, as desired.

Assume now that $\rho > \rho^*$. so that $\lim_{N \rightarrow \infty} \prod_N^2 \times \prod_N^* = 0$ by Lemma C.5. Let $a \in (0, 1)$ be given, and choose λ small enough so that $\frac{a}{1-a} \lambda \leq \frac{p_*}{1-p_*}$. Choose N large enough so that $\prod_N^1 \times \prod_N^\blacksquare \leq \lambda$. As above, one has $q_2^{N,1} \leq p_*$ on the event $\{p_2^1 \leq a\} \cap D$. By Corollary 6.3, this implies $\pi_2^{N,1}(\vec{k}^N) \geq a$. Since a is arbitrary, one gets $\lim_{N \rightarrow +\infty} \pi_2^{N,1}(\vec{k}^N) = 1$, as desired. ■