Equilibrium Learning in Simple Contests*

Daniel Krähmer†

November 28, 2005

Abstract

The paper studies a repeated contest when contestants are uncertain about their true relative abilities. When ability and effort are complements, a favourable belief about one's own ability stimulates effort and increases the likelihood of success. Success, in turn, reinforces favourable beliefs. We show that this implies that with positive probability players fail to learn their true relative abilities in equilibrium, and one player wins the contest with high probability forever. In this case, the prevailing player may be the actually worse player, and persistent inequality arises. We discuss some features of the model when the complementarity assumption is dropped.

Keywords: Contest, Belief Reinforcement, Incomplete Learning, Dynamic Programming.
JEL Classification: C61, C73, D44, D83

*Acknowledgements: I would like to thank Helmut Bester, Tilman Börgers, Jürgen Bierbaum, Kai Konrad, Sven Rady, Francesco Squintani, and Roland Strausz for helpful comments and discussions. Financial support from the Deutsche Forschungsgemeinschaft (DFG) through SFB/TR 15 is gratefully acknoledged.

†Freie Universität Berlin, Institut für Wirtschaftstheorie, Boltzmannstr. 20, 14195 Berlin, Germany, ++49-(0)30-83855223, kraehmer@wiwiss.fu-berlin.de. Support by the German Science Foundation (DFG) through SFB/TR 15 is gratefully acknowlegded.

1 Introduction

This paper studies a repeated contest when the contestants initially do not know their intrinsic relative abilities but can learn about them over time. Examples where contestants encounter each other repeatedly include sports contests, promotion tournaments in small firms, research contests in specialized areas, or court trials, but also conflicts between states or family dynasties. In many of these cases, the players' relative ability, for example, physical or mental strength, is not known a priori.

If a contest is played repeatedly, by experiencing successes and failures, the players learn about their relative abilities over time and by varying their choices, they can generate more or less information that promotes or inhibits learning. This raises the issue whether equilibrium play will identify the players' relative abilities and whether the contest will, eventually, select the relatively more able player.

To address such questions, we consider a simple contest with two players who can exert either high or low effort. A player's win probability depends on relative ability, effort, and luck. We shall make the crucial assumption that a higher relative ability of one player translates into a higher success probability only if both players exert the same level of effort. Whenever the effort levels chosen by the players differ, relative ability does not matter for success in the contest.

This reflects, for a example, a situation where low effort stands for non-participation in the contest. Then the win probability of the participating player is not affected by how able or unable his non-participating rival might be. In an environment with two players, two effort levels and perfect monitoring of the contest outcome, this is a necessary requirement to make the question at study interesting. Otherwise, whatever actions might be played, each observation of the contest outcome would be informative and learning would obtain with certainty.

Whether learning obtains or not, therefore, depends on whether players will eventually choose different effort levels in equilibrium. To study this question, we shall first look at the case where relative ability and efforts are complements. This means that it is never myopically optimal for a player to reduce his effort when his belief about his relative ability increases.

We derive our main result for this case and show that there is an equilibrium such that whenever the players' belief that one of them is the less able player is sufficiently large, this player chooses low effort and the other player, who is believed to be the more able player, chooses high effort. We show that such a state where one player resigns is reached with positive probability irrespective of true abilities. Hence, beliefs become persistent, and the players will not learn their true relative abilities. Whenever this is the case, one player wins the contest with a high probability forever and persistent inequalities arise in the long run. Moreover, the prevailing player might be the "wrong", that is, the actually relatively less able player.

The basic idea rests on a simple belief reinforcement argument. A player who is optimistic about his relative ability will, due to complementarity, tend to choose high effort, resulting in a higher number of actual successes. Increased successes, in turn, will promote the player's optimism further. So beliefs become self-fulfilling. At the same time, a pessimistic player will thereby become discouraged and, as a result, ever more pessimistic. If the belief revision proceeds fast, the pessimistic player will resign at some point. From then on, no further information is revealed, and beliefs remain uncontradicted.[1]

We shall then discuss some features of the model when complementarity is dropped. In this case, it is myopically optimal for a sufficiently optimistic player to reduce effort against a rival who does not exert effort. This is because he believes that he will win anyway. We call this case complacency. For simplicity, we shall only look at the repeated game with myopic players. For moderate levels of beliefs, the same reinforcement mechanism as above will drive beliefs apart. At some point, the pessimistic player becomes discouraged and reduces his effort, given the optimistic player still exerts high effort. Whether learning obtains or not, depends then on the optimal reaction of the optimistic player at this point. If complacency sets in at very high levels of optimism only, it is optimal to stick to high effort, and therefore the belief process stops and learning does not obtain. This is no longer true, if already a moderately optimistic player is complacent. Such a player wants to reduce his effort against a resigning rival. However, the rival's optimal response against such a move is to exert high effort again, and no equilibrium in pure strategies exists. Therefore, with positive probability each observation of the contest outcome is informative. Learning is then complete.

Related work

[1] This sort of self-fulfilling prophecy is at the heart of social psychlogical accounts of the dynamics of optimism and pessimism (see Aspinwall et al. (2001), Brocker (1984)).

Our work is closely related to the literature on rational learning in two-armed bandit problems (Rothschild (1974), Berry/Fristedt (1985)). Indeed, if a player's choice of effort is exogenously fixed to be either high or low forever, then the problem the other player faces in our setup is a classical two-armed bandit problem with one safe and one risky arm. It is well known that in this case the optimal strategy does not lead to the optimal action with certainty.[2] This is because after sufficiently many failures with the risky arm, the player switches to the safe arm forever and his pessimistic belief about the risky arm remains uncontradicted even if the risky action is optimal. By contrast, optimistic beliefs about the risky arm, when in fact the safe action is optimal, will be contradicted.

The game setup is different because whether beliefs will be contradicted or not, depends on what the other player does. In the complementarity case, when the pessimistic player resigns, it becomes impossible in equilibrium that the optimistic player's beliefs will be contradicted. In this sense, learning is inhibited through the strategic interaction.[3]

Our model is a simple example of a multi-armed bandit game where an arm's payoff depends on the choices of a player's opponents.[4] Such games are studied in some generality in the work of Kalai/Lehrer (1995).[5] Their focus however is different from ours. They consider players with only limited knowledge about their environment. Players then form subjective beliefs about their environment, including their opponents' behaviour, and are assumed to optimize with respect to these beliefs. Kalai/Lehrer (1995) examine conditions such that the players' behaviour converges to equilibrium play of the underlying incomplete information game. We instead ask whether the players learn the underlying game, given they play an equilibrium of the incomplete information game.

As for contests, learning issues are addressed by Squintani/Välimäki (2002) who consider

[2] For a characterization of incomplete learning type results in general single player bandit problems see Easley/Kiefer (1988).

[3] We shall comment in more detail on the relation of our model to the single player bandit problem at the end of Section 3.

[4] Multi-armed bandit games are also studied in the strategic experimenation literature (Bolton/Harris (1999), Keller et al. (2005)). In these approaches, a player's payoff is not affected *directly* by other players' choices but by the possibility to observe their payoffs.

[5] A bandit game of this type is also examined by Fishman/Rob (1998). They consider a dynamic Bertrand game with two sellers who can learn about demand by experimenting with high or low prices. A seller's payoff from an experiment with a high price then directly depends on which price his rival chooses.

a model where players are randomly matched to play a contest in which the dominant, i.e., winning action is unknown and in which the dominant action changes at random times. They show that in contrast to myopic players, farsighted players optimally engage in experimentation in equilibrium and therefore play the optimal action almost all of the time. This is different to our results where experimentation stops and the optimal action is not identified with positive probability for all discount factors. The reason is that in a setup with random matching and exogenous state changes, our reinforcement mechanism is not at work, and occasional experimentation might be optimal for non-myopic players.

Our work also contributes to the tournament literature. Only few papers consider tournaments with players who do not know their abilities. One example is Rosen (1986) who studies the optimal design of prizes in an elimination tournament. In an elimination tournament however, beliefs about relative abilities are not reinforced because after each period winners are matched with similarily optimistic winners. A further example is Stone (2004). Stone shows that players who are concerned about their self-image might be led to exert either very high or very low effort so as to suppress self-relevant information that could threaten their ego. As in our case, although for a different reason, in equilibrium one player will then resign and the other player will exert high effort and win easily.

From a more applied side, our results suggest that contests might perform poorly as incentive and selection mechanisms when players do not know their abilities. They also suggest an explanation for the observed high degree of intergenerational mobility in labour markets. We postpone this discussion to section 3.2. The rest of the paper is organized as follows. Section 2 presents the model. Section 3 analyzes the complementarity case and presents the main results. Section 4 discusses the case with complacency. Section 5 concludes.

2 The Model

Time $t = 0, 1, 2, ...$ is discrete and infinite. There are two players, P1 and P2, who, in each period, engage in a contest. The winner prize is normalized to 1 and the loser prize to 0. After each period t, players observe the contest outcome and the action taken by the other player in t. The winning probability in period t is determined by the efforts spent by the players and by a state of nature realized in $t = 0$ which specifies players' relative abilities.

There are two states. In state $\theta = 1$, P1 is more able than P2, and in state $\theta = 2$, P2 is more able than P1. The realization of the state is unknown, also to players themselves. We assume that there is no asymmetric information, that is, the players' beliefs about states are the same, and this is common knowledge. Denote by γ_θ^t the players' common belief in period t that the state is θ. We shall also use plain γ to denote the probability γ_1 that the state is $\theta = 1$. Let $\Gamma = [0, 1]$ be the set of all possible beliefs γ.

There are two effort levels $e \in \{0, 1\}$, and the cost of spending effort e is ce, $c > 0$.

The true winning probability for P1 in a given period for given efforts $e_1 = i, e_2 = j$ in state θ is denoted by

$$p_{ij}^\theta = P\left[\text{P1 wins } | e_1 = i, e_2 = j; \theta\right].$$

The corresponding true winning probability for P2 is $1-p_{ij}^\theta$.[6]

For states to reflect abilities, we set $p_{ii}^1 > 1/2$ and $p_{ii}^2 < 1/2$ for $i = 0, 1$. We further assume that for given state, a player's winning likelihood increases in his own's and decreases in his rival's effort. Moreover, we make the following assumptions:

A1: $p_{10}^\theta = p$

A2: $p - c > 1/2$

A3: Either $p_{11}^2 < 1 - p + c$ or $p_{11}^1 > p - c$.

A1 says that if players exert unequal levels of effort, then the contest outcome is independent of θ. This reflects a specific form of strategic interaction in our contest. For example, if a player is the only candidate to apply for a job, he gets the job with high probability, irrespective of θ. In this sense, θ represents *relative* abilities.

A2 reflects the contest flavour of the game. It says that effort costs are moderate and implies that for some belief γ, it is myopically optimal for both players to exert effort in the current period, given the other player does so.[7]

A3 rules out the trivial case in which choosing high effort is a dominant strategy for both players.

[6] The fact that the probability of a draw is 0 is not essential for the analysis.

[7] Myopically, P1's is better of from choosing $e_1 = 1$ rather than $e_1 = 0$ if $p_{11}^1\gamma + p_{11}^2(1-\gamma) - c \geq 1 - p$ or, equivalently, if $p - c \geq 1 - p_{11}^1 + \left(p_{11}^1 - p_{11}^2\right)\gamma$. Because $p_{11}^1 > 1/2$, this inequality is implied by A2 for large γ. A similar statement holds for P2.

Notice also that because $p = p_{10}^1 > p_{11}^1 > 1/2$, A1 and A2 together imply that $p - c > 1 - p$. The following table summarizes the stage game payoffs in state θ where P1 is the row player, and P2 is the column player.

	$e_2 = 0$	$e_2 = 1$
$e_1 = 0$	$p_{00}^\theta, 1 - p_{00}^\theta$	$1 - p, p - c$
$e_1 = 1$	$p - c, 1 - p$	$p_{11}^\theta - c, 1 - p_{11}^\theta - c$

(1)

Table 1: Stage game payoffs.

Since the true winning probabilities are unknown, players have to form expectations. Denote by

$$\pi_{ij}(\gamma) = p_{ij}^1 \gamma + p_{ij}^2 (1 - \gamma)$$

P1's *expected* winning probability for efforts $e_1 = i$ and $e_2 = j$, given belief γ. We shall occasionally write $\pi(i, j; \gamma)$ for $\pi_{ij}(\gamma)$.

Players are assumed to be Bayesian rational and to discount future profits by a common discount factor $\delta \in [0, 1)$.

Remark: The specification can be seen as a reduced version of a more general model where also the win probabilities for unequal effort choices are unknown. For example, P1's win probability p_{10} could be either high or low, reflecting the absolute ability of P1. It would then appear reasonable to assume that higher absolute ability translates into higher relative ability, but relative abilities could still differ, conditional on absolute abilities being equal. Our specification looks at that latter contingency.

3 Complementarity

In this section we assume that $p_{00}^1 < p - c$ and $1 - p_{00}^2 < p - c$. That is, given a player exerts no effort, it is myopically optimal for the rival player to exert effort, irrespective of beliefs. Together with the other assumptions, this implies that, given player $-i$'s behaviour, it can never be myopically optimal for player i to reduce his effort as γ_i increases. We call this property *(effective) complementarity* between effort and relative ability.[8] We shall discuss the other case below.

3.1 Analysis

As common in repeated games, there may be many equilibria. To rule out equilibria with implicit agreements between the players, we restrict attention to Markov strategies and look for a Markov Perfect Bayesian equilibrium (MPBE). We state most of the definitions and results in terms of P1. Corresponding defnitions and results hold for P2.

Strategies

A Markov strategy depends only on the payoff relevant information in date t but not on the entire history up to t. In our setup, the only payoff relevant information in date t is players' beliefs in t. Denote by σ (and φ respectively) the event that P1 wins (loses) the contest in a given period. Let

$$H_t = \left\{ \left(e_1^1, e_2^1, \omega^1, ..., e_1^t, e_2^t, \omega^t \right) \mid e_i^s \in \{0,1\}, \omega^s \in \{\sigma, \varphi\}, 1 \leq s \leq t \right\} \qquad (2)$$

be the set of all possible histories up to t, and let $H = \cup H_t$ be the set of all possible histories.[9]

Definition 1 *1) A (pure) strategy* $\eta_i = (\eta_i^t)_{t=0,1,...}$ *for player* i *is a sequence of mappings* $\eta_i^t : H_t \to \{0,1\}$ *from histories into actions. The set of all strategies for player* i *is denoted by* Σ_i.

2) A (stationary) Markov strategy $\eta_i : \Gamma \to \{0,1\}$ *for player* i *maps beliefs into actions.*

Notice that we require a Markov strategy to be stationary, that is, to be the same for all periods.

Because players are Bayesian rational, beliefs are derived by Bayes' rule. Let γ_θ^σ be the updated belief that the state is θ upon P1 winning. That is, $\gamma_\theta^\sigma (e_1 = i, e_2 = j; \gamma) = p_{ij}^\theta \gamma / \pi_{ij} (\gamma)$. We shall also write plain γ_θ^σ for γ^σ. Define γ_θ^φ and γ^φ likewise.

[8] In contrast to proper complementarity, we do *not* have that, given player $-i$'s behaviour, i's incentives to choose high effort increases in γ_i. For example, if $e_2 = 0$, then P1's incentive to choose $e_1 = 1$ is $\pi_{10} - c - \pi_{00} = p - c - \left(p_{00}^1 - p_{00}^2\right) \gamma_1 - p_{00}^2$ which is declining in γ_1.

[9] Note, $H_0 = \{\emptyset\}$.

Utility

To define a player's expected utility, we need first to define the appropriate probability measure.[10]

A Markov strategy η_2 of P2 gives rise to a law of motion that can be controlled by P1. This results in transition probabilities that govern the evolution of beliefs. More formally, for $e_1 \in \{0, 1\}$,

$$q\left(\gamma' \,|\, e_1, \eta_2\left(\gamma\right), \gamma\right) = \begin{cases} \pi\left(e_1, \eta_2\left(\gamma\right); \gamma\right) & \text{if } \quad \gamma' = \gamma^{\sigma} \\ \\ 1 - \pi\left(e_1, \eta_2\left(\gamma\right); \gamma\right) & \text{if } \quad \gamma' = \gamma^{\varphi} \end{cases} \tag{3}$$

defines a transition kernel from current beliefs γ into next period beliefs γ'. q may be viewed as the (expected) conditional probability that the next period belief is γ', conditional on the current belief being γ.

Suppose that P1 plays a strategy $\eta_1 \in \Sigma_1$, and let the initial belief be γ^0. Then the probability of a finite sequence $\overline{\omega}_t = \left(\omega^1, ..., \omega^t\right), \omega^s \in \{\sigma, \varphi\}$ of successes and failures for P1 is

$$P^t\left[\overline{\omega}_t\,;\, \eta_1, \eta_2, \gamma^0\right] = \prod_{s=1}^{t} q\left(\gamma^s \,|\, \eta_1^{s-1}\left(h_{s-1}\right), \eta_2\left(\gamma^{s-1}\right), \gamma^{s-1}\right), \tag{4}$$

where h_s is the (unique) history induced by $\overline{\omega}_s, \eta_1, \eta_2$, and γ^s is the (unique) belief induced by h_s. It is well known, that the measures P^t thus defined constitute a consistent family of probability measures. Hence, by Kolmogorov's consistency theorem, there is a unique probability measure $P\left[\cdot\,;\, \eta_1, \eta_2, \gamma^0\right]$ on the set of infinite sequences $\omega = \left(\omega^1, ..., \omega^t, ...\right), \omega^t \in \{\sigma, \varphi\}$ such that P and P^t coincide on the set of finite sequences of length t.

With this P1's expected utility from strategy $\eta_1 \in \Sigma_1$ against a Markov strategy η_2 at initial belief γ is given by

$$U_1\left(\eta_1, \eta_2; \gamma\right) = \int \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t\left(h_t\right), \eta_2\left(\gamma^t\right); \gamma^t\right) - c\eta_1^t\left(h_t\right)\right]\, dP\left(\omega\,; \eta_1, \eta_2, \gamma\right), \tag{5}$$

where the integration is over all infinite sequences ω, and h_t and γ^t are the (unique) histories and beliefs induced by $\omega, \eta_1, \eta_2, \gamma$.

Best Response

[10] Our construction of the probability space is somewhat sloppy. We neglect measurability issues and the proper definition of σ-algebras.

P1's best response η_1^* against a Markov η_2 is thus given as a solution to the maximization problem

$$\sup_{\eta_1 \in \Sigma_1} U_1(\eta_1, \eta_2; \gamma). \tag{MP}$$

We derive the best response by dynamic programming. The Bellman equation for problem (MP) is

$$V_1(\gamma) = \max_{e_1 \in \{0,1\}} \{\pi(e_1, \eta_2(\gamma); \gamma) - ce_1 + \tag{BE}$$
$$+\delta [\pi(e_1, \eta_2(\gamma); \gamma) V_1(\gamma^\sigma) + (1 - \pi(e_1, \eta_2(\gamma); \gamma)) V_1(\gamma^\varphi)]\}.$$

It is well known that a solution V_1 to (BE) coincides with $\sup U_1$. Furthermore, if a solution to (BE) exists, then the maximizer in (BE) at γ coincides with the best response in (MP) at γ. In particular, the best response is (stationary) Markov (see Blackwell (1965)).

Hence, to establish existence of a Markov best response, it remains to show that (BE) has a solution V_1. We do this in the usual way by showing that V_1 is the limit of iterated applications of a contraction mapping. We have the following proposition. Details are in Appendix A.

Proposition 1 *For all η_2, V_1 exists and is bounded.*

Threshold Strategies

The previous considerations show that a best response against a Markov strategy exists and is again Markov. This could be used to establish the existence of a MPBE by an abstract existence theorem (see Maskin/Tirole (2001)). However, it would not tell much about how the equilibrium looks like, and how the system evolves over time in this equilibrium. Therefore, we shall further restrict the strategy space so as to derive a more specific equilibrium. We shall look for an equilibrium in threshold strategies. A threshold strategy is a strategy where a player chooses high effort only when the belief about his relative strength is sufficiently large. More formally:

Definition 2 *A Markov strategy η_i for player i is called threshold strategy if there is a number $r_i \in \Gamma$ such that*

$$\eta_1(\gamma) = \begin{cases} 1 & \text{if } \gamma > r_1 \\ \in \{0,1\} & \text{if } \gamma = r_1 \\ 0 & \text{if } \gamma < r_1 \end{cases}, \qquad \eta_2(\gamma) = \begin{cases} 0 & \text{if } \gamma > r_2 \\ \in \{0,1\} & \text{if } \gamma = r_2 \\ 1 & \text{if } \gamma < r_2. \end{cases}$$

The reason why the strategy is not specified at the threshold r_i is technical.[11] We impose the tie-breaking rule that if the player is indifferent at r_i, he chooses $e_i = 0$. Notice that, by definition, every threshold strategy is Markov.

A threshold strategy divides the belief space Γ into two connected sets: one in which effort is chosen, and one in which no effort is chosen. We call the latter *no-effort set.*

Definition 3 *Let η_i be a threshold strategy. Define by*

$$N(\eta_i) = \{\gamma \,|\, \eta_i(\gamma) = 0\} \tag{6}$$

the set of beliefs where player i chooses no effort under η_i.[12]

Alternatively, the no-effort set can be described by intervals in Γ. The size of the no-effort set measures how aggressive a strategy is. The smaller a player's no-effort set, the more aggressive this player.

Equilibrium

Our aim is to show that there is a MPBE in threshold strategies. For this, we proceed as follows. We first show that a best response against a threshold strategy is again a threshold strategy. We then establish a monotonicity property that says that a player optimally responds less aggressive against a more aggressive rival. We then use this monotonicity property to show that we can define a function $\widehat{r}_i : \Gamma \to \Gamma$ such that $\widehat{r}_i(r_{-i})$ is the threshold of player i's best response against a threshold strategy of player $-i$ with threshold r_{-i}. We then establish that $\widehat{r}_i$ is continuous on the compactum Γ which implies that there are thresholds (r_1^*, r_2^*) at which $\widehat{r}_1$ and $\widehat{r}_2$ intersect. The threshold strategies corresponding to (r_1^*, r_2^*) serve then as equilibrium candidates. We relegate the formal proof to Appendix B and state the result directly.

Theorem 1 *There are threshold strategies (η_1^*, η_2^*) with corresponding thresholds (r_1^*, r_2^*) that constitute a MPBE, and it holds:*

[11] Since we want to establish an equilibrium in threshold strategies, we need to make sure, first of all, that a best response against a threshold strategy is again a threshold strategy. If we prescribe a player, for example, to play $e = 0$ at the threshold, the best response of the rival player is generally *not* to play $e = 0$ at the threshold as well.

[12] If no cause for confusion, we shall omit the argument and write N_i instead of $N(\eta_i)$.

(i) $r_1^* > 0 \Leftrightarrow p_{11}^2 < 1 - p + c$.

(ii) $r_2^* < 1 \Leftrightarrow p_{11}^1 > p - c$.

(iii) $N_1^* \cap N_2^* = \emptyset$.

Property (i) characterizes the case in which there are some beliefs γ in which P1 chooses low effort. If the condition on the right hand side of (i) does not hold, then choosing high effort is a dominant strategy for P1 in all states. Property (ii) is the corresponding condition for P2. Notice that assumption A3 guarantees that either (i) or (ii) holds. Property (iii) says that there are some beliefs γ at which both players choose high effort, and that at least one player exerts effort in equilibrium.

Learning

We can now state our main result that players will with positive probability fail to learn their true abilities. Let

$$T(\omega) = \min \left\{ t \geq 0 \,|\, \gamma^t(\omega) \in N_1^* \cup N_2^* \right\} \tag{7}$$

be the first period in which one player ceases to spend effort. If this period is reached, the process stops and no information is generated any more. Intuitively, $N_1^* \cup N_2^*$ might be reached in finite time because it contains small neighbourhoods around $\gamma = 0$ or $\gamma = 1$. Therefore, if, say $r_2^* > 0$, and if P1 experiences a long sequence of consecutive successes, the belief tends to $\gamma = 1$ and P2's no-effort set is reached. More precisely, we have the following result.

Theorem 2 *(i) Let $r_1^* > 0$ and $r_2^* < 1$. Then for all states θ,*

$$P[T < \infty \,|\, \eta_1^*, \eta_2^*; \theta] = 1. \tag{8}$$

(ii) Let either $r_1^* > 0$ or $r_2^* < 1$ and let $\gamma^0 \notin \{0, 1\}$. Then for all states θ,

$$P[T < \infty \,|\, \eta_1^*, \eta_2^*; \theta] > 0. \tag{9}$$

Thus, learning will be incomplete, and with positive probability one player will be discouraged in the long run. If both no-effort sets contain a small neighbourhood, this happens with probability 1. We relegate the proof to Appendix C.

Theorem 2 implies that with positive probability the no-effort set of the actually more able player is reached such that, eventually, the actually weaker player wins the contest forever.

This naturally directs interest to properties of the distribution of the first entry time into N_i^*:

$$T_i(\omega) = \min\left\{t \geq 0 \,|\, \gamma^t(\omega) \in N_i^*\right\}. \tag{10}$$

The following result is straightforward.

Proposition 2 *For all states θ, the likelihood to reach the rival player's no effort set in finite time is increasing in one's own and decreasing in the rival's belief.*

Proof: The higher γ_i, the closer the belief γ_i is at N_{-i}^*. Therefore, less successes are required to move into N_{-i}^*. □

In Appendix D, we illustrate how in principle a quantitative expression for the distribution of the first entry time can be obtained. The basic idea is to transform the belief process γ^t into a random walk with (linear) drift, Y^t, such that γ crossing through a threshold r_i^* is equivalent to Y crossing through corresponding threshold α. The first passage time of Y through these corresponding thresholds can then be computed by the use of martingale techniques.

3.2 Discussion

Relation to single decision maker bandit problems

The problem of a player in our game is closely related to the problem of a single decision maker who faces a two-armed bandit with a safe and a risky arm. In fact, if P2 always plays action $e_2 = 1$,[13] then P1's problem *is* a two-armed single decision maker bandit problem where the "safe" action $e_1 = 0$ yields per period payoff $1-p$ and the "risky" action $e_1 = 1$ yields (average) per period payoff $p_{11}^\theta - c$ with probability γ_θ.

It is well known that the optimal strategy in this case is to choose the risky action as long as P1's belief γ_1 is above some threshold level and to choose the safe action otherwise.[14] Therefore, the optimal strategy does not always lead to the optimal action in state $\theta = 1$. This is because a sufficiently long series of failures leads P1 to choose the safe action. From this point on, no new information is revealed and P1's "overly" pessimistic beliefs remain uncontradicted. By contrast, in state $\theta = 2$, the optimal strategy always leads to the optimal action. Otherwise, P1

[13] By Theorem 1 (ii), this is the case if $p_{11}^1 \leq p - c$.

[14] Unless $e_1 = 1$ is a dominant strategy; that is, unless $p_{11}^2 > 1 - p + c$.

would hold overly optimistic beliefs in the long run and play the risky action infinitely often. But since the risky action reveals information in each period, such optimistic beliefs cannot, eventually, remain uncontradicted.

The latter changes in our game setup. Here, not only overly pessimistic but also overly optmistic beliefs of P1 might remain uncontradicted in the long run because optimistic beliefs of P1 go along with pessimistic beliefs of P2, triggering the latter to resign and the contest outcome to be uninformative. In this sense, the presence of P2 inhibits learning.

Implications of our results

Contests are often used as mechanisms to either induce effort or select able players.[15] Our contest, however, does not necessarily achieve these aims. With positive probability, one player will reduce his effort to 0, and in addition, the "wrong", that is, the less able player might eventually be selected by the contest. (For example, in state $\theta = 1$, P2 has a positive likelihood to be the long term winner.) This suggests that contests might perform poorly when contestants have limited knowledge about their ability, because the latter makes them vulnerable to self-fulfilling prophecies. To draw this conclusion in general, however, one would need to specify more precisely the options and constraints a designer has in setting up the contest and the mechanisms he could use other than a contest.

Our results are also relevant to the discussion on the intergenerational mobility in labour markets. Labour market success—getting accepted to a good university, getting a job, or getting promoted—is often determined by relative performance, i.e., contests. Seen as an intergenerational model, it predicts a strong correlation between the labour market success of parents and of their children. In fact, in our model persistent inequality arises in the long run, and one player may eventually win the contest with probability p forever and become rich whereas the other player will be discouraged and stay poor. This is consistent with substantial empirical evidence that parental earnings are a reliable predictor for childrens' earnings (see Solon (1999)) for a review). Yet, what appears to be puzzling is that the labour market success of successful parents' children cannot be accounted for by their superior education, or the inheritance of wealth, or cognitive ability (Bowles et al. (2001)). Our model suggests that optimistic or pessimistic beliefs transmitted to children by parental upbringing may possibly

[15] See, for example, Lazear/Rosen (1981), Meyer (1991, 1991).

be one of the missing elements in explaining the intergenerational stability of labour market outcomes.[16]

In general, the model points to the importance of beliefs in the determination of success. If beliefs are more broadly interpreted in terms of psychological attitudes such as optimism, pessimism, or self-confidence, then our results such as this in Proposition 2 illustrate how such attitudes might have real consequences.

4 Complacency

We shall now discuss some features of the model when the complementarity assumption is dropped. For simplicity, we focus on the symmetric case in which the relative advantage of P1 in state 1 is the same as that of P2 in state 2; that is, $p_{ii}^{\theta} = 1 - p_{ii}^{\theta}$. In this case, dropping complementarity means $p_{00}^{1} > p - c$. We shall only look at the case of myopic players ($\delta = 0$). The analysis with farsighted players is considerably more complicated and beyond the scope of the paper.

The main difference to the complementarity specification is that if a player is convinced that he is the better player, he optimally stops exerting effort when his rival does so, hence the wording *complacency*. To illustrate, suppose $\gamma = 1$. Then, it is easily seen that P1's best response against $e_2 = 1$ is $e_1 = 1$, and that against $e_2 = 0$ is $e_1 = 0$. By contrast, P2's best response against $e_1 = 0$ is $e_2 = 1$, and that against $e_1 = 1$ is $e_2 = 0$. So, while P1 wants to match the action of P2, P2 wants to chose an action opposed to that of P1. Hence, there is no equilibrium in pure strategies.

More precisely, P1's best response, BR_1, is now determined by two thresholds, $r_1, \widetilde{r}_1$ such that

$$BR_1(1) = 1 \Longleftrightarrow \gamma \geq r_1, \tag{11}$$

$$BR_1(0) = 0 \Longleftrightarrow \gamma \geq \widetilde{r}_1. \tag{12}$$

[16] See also Picketty (1995) who argues that persistent intergenerational income inequality can be explained as a result of incomplete learning in a rational experimentation setup.

A little bit of algebra yields that

$$r_1 = \frac{p_{11}^1 + c - p}{2p_{11}^1 - 1} \in \left(0, \frac{1}{2}\right), \tag{13}$$

$$\widetilde{r}_1 = \frac{p_{00}^1 + p - c - 1}{2p_{00}^1 - 1} \in \left(\frac{1}{2}, 1\right). \tag{14}$$

Note that $\widetilde{r}_1 < 1$ if and only if the complacency assumption holds. By symmetry, we obtain the analogous thresholds for P2 as $r_2 = 1 - r_1$ and $\widetilde{r}_2 = 1 - \widetilde{r}_1$.

The equilibrium can now be characterized by the following four sets:

$$M_\alpha = [0, \widetilde{r}_2] \cup (\widetilde{r}_1, 1], \quad M_{01} = (\widetilde{r}_2, r_1], \quad M_{11} = (r_1, r_2], \quad M_{10} = (r_2, \widetilde{r}_1]. \tag{15}$$

For $i, j \in \{0, 1\}$, if $\gamma \in M_{ij}$, the equilibrium is given by $(e_1^*, e_2^*) = (i, j)$. If $\gamma \in M_\alpha$ there is only an equilibrium in mixed strategies. The mixing probabilities of P1 and P2, respectively, can be computed to

$$\alpha_1(\gamma) = \frac{p - c - p_{00}^1 - (1 - 2p_{00}^1)\gamma}{(2\gamma - 1)(1 - p_{11}^1 - p_{00}^1)}, \quad \alpha_2(\gamma) = \alpha_1(1 - \gamma). \tag{16}$$

The following picture illustrates P1's equilibrium strategy if $\widetilde{r}_2 < r_1$.

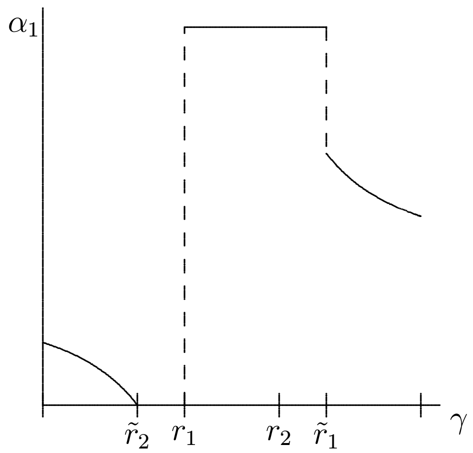

Figure 1: Equilibrium strategy of P1

Learning

We consider now the repeated game for myopic players. Whether learning obtains or not, depends on whether M_{01} and M_{10} are empty or not.[17] If they are *not* empty, then updating stops as soon as the belief process reaches M_{01} or M_{10}. By an argument analogous to that in the proof of Theorem 2, it follows that learning will be incomplete. If M_{01} and M_{10} are empty, however, in each period with positive probability an action pair is played for which the contest

[17] Because of symmetry: $M_{01} = \emptyset \Leftrightarrow M_{10} = \emptyset$.

outcome is informative. Therefore, with probability 1, players will make an infinite number of informative observations, and thus the true state will eventually be revealed.

Notice, M_{01} is not empty if and only if $r_1 \geq \widetilde{r}_2$. A little bit of algebra yields that the latter is equivalent to $p^1_{11} \geq p^1_{00}$. As $p^1_{11} - p^1_{00}$ declines to 0, M_{01} shrinks to $\emptyset$. We summarize these observations in the following proposition.

Proposition 3 *For all states θ learning is complete if and only if $p^1_{00} \geq p^1_{11}$.*

The condition $p^1_{00} \geq p^1_{11}$ says that the relative advantage for P1, conditional on being more able, is smaller when both players exert effort than when both players do not exert effort. It implies that whenever it is optimal for P1 to respond with $e_1 = 1$ against $e_2 = 0$, it cannot be optimal for P2 to respond with $e_2 = 0$ against $e_1 = 1$. In this case, as opposed to the complementarity case, learning is promoted by the presence of the other player.

These considerations illustrate for the myopic case what changes when the complementarity assumption is abandoned. The main difficulty with farsighted players is that one has to allow for mixed strategies. Suppose, for example, that the game has two periods. Then in period 2 the myopic equilibrium is played. This implies that a player may not wish to learn between period 1 and 2 that he is more able, because when doing so he might be led to reduce his effort in period 2 and thereby triggering a more aggressive response by his rival in period 2. We do not know how this twist affects the learning incentives in the long run.

5 Conclusion

The paper considers a dynamic contest when the players do initially not know their intrinsic ablities. If relative ability and effort are complements, a belief reinforcement effect encourages optimistic players and discourages pessimistic players. This implies that players may fail to learn their true abilities in the long run, and one player may eventually win with high probability. As a consequence, persistent inequality arises, and the actually worse player may prevail in the long run.

The specific form of our contest limits the generality of our results. However, as discussed in the previous section, some qualitative characteristics of the belief reinforcement mechanism might, under some conditions, carry over when the complementarity assumption is dropped.

But we do not know what happens with continuous effort levels or a richer state space. We therefore see our work as a first step to explore issues of learning in dynamic contests.

Generalizing our framework could also provide a better understanding of the extent to which the possibility that players become subject to self-fulfilling prophecies impairs the performance of contests as incentive and selection mechanisms. Such self-fulfilling prophecies can be expected to become particularly important when one also allows players to accumulate wealth over time, a possibility set aside in our setup. The wealthier player can then invest more resources in the contest, thus, increasing his likelihood to succeed, and, in turn, to become even wealthier.

Appendix A: Existence of Value Function

Proof of Proposition 1: We show that V_1 is the limit of iterated applications of a contraction mapping. Let S be the space of all bounded functions on Γ equipped with the supremum norm. Let $\psi \in S$, and define the mapping $F : S \to S$ by

$$F\psi(\gamma) = \max_{e_1 \in \{0,1\}} \{\pi(e_1, \eta_2(\gamma)) - ce_1 + \\ + \delta[\pi(e_1, \eta_2(\gamma))\psi(\gamma^\sigma) + (1 - \pi(e_1, \eta_2(\gamma)))\psi(\gamma^\varphi)]\}. \tag{17}$$

Hence, V_1 is the solution to the fixed point problem $FV = V$.[18]

We show in the next paragraph that F is a contraction. Since F is a contraction, it follows by Banach's fixed-point theorem that $FV = V$ has a unique solution. This establishes existence of V_1. Moreover, it is well known that V_1 is the limit of iterated applications of F on an arbitrary starting point $\psi \in S$. Notice that $F\psi \in S$ since all functions on the right hand side of (17) are bounded. Thus, all elements of the sequence $(F^n\psi)_{n=1,2,\ldots}$ are in S. Since S is complete, the limit is in S, thus bounded.

To establish that F is a contraction, we check Blackwell's sufficiency conditions:

(i) For all $\psi, \phi \in S$ with $\psi \geq \phi$ on Γ, it holds that $F\psi \geq F\phi$ on Γ.

(ii) For all $\psi \in S$ and $\xi > 0$, there is a $\beta \in (0,1)$ such that $F(\psi + \xi) \leq F\psi + \beta\xi$.

As for (i): Let $\psi, \phi \in S$ with $\psi \geq \phi$ on Γ. Let $\left(e_1^\psi, e_1^\phi\right) \in \{0,1\}^2$ be the maximizers of $F\psi$ and

[18] Note that both F and the corresponding fixed point V_1 depend on a specific Markov strategy η_2. Provided it does not cause confusion, we shall suppress this dependency.

$F\phi$, respectively. Then[19]

$$F\psi = \pi\left(e_1^{\psi}, \eta_2\right) - ce_1^{\psi} + \delta\left[\pi\left(e_1^{\psi}, \eta_2\right)\psi^{\sigma} + \left(1 - \pi\left(e_1^{\psi}, \eta_2\right)\right)\psi^{\varphi}\right] \quad (18)$$

$$\geq \pi\left(e_1^{\phi}, \eta_2\right) - ce_1^{\phi} + \delta\left[\pi\left(e_1^{\phi}, \eta_2\right)\psi^{\sigma} + \left(1 - \pi\left(e_1^{\phi}, \eta_2\right)\right)\psi^{\varphi}\right] \quad (19)$$

$$\geq \pi\left(e_1^{\phi}, \eta_2\right) - ce_1^{\phi} + \delta\left[\pi\left(e_1^{\phi}, \eta_2\right)\phi^{\sigma} + \left(1 - \pi\left(e_1^{\phi}, \eta_2\right)\right)\phi^{\varphi}\right] \quad (20)$$

$$= F\phi. \quad (21)$$

The first inequality follows by definition of the maximum, and the second inequality follows by the assumption that $\psi \geq \phi$.

As for (ii): Let $\psi \in S$ and $\xi > 0$, and let $\beta = \delta$. Then

$$F(\psi + \xi) = \max_{e_1 \in \{0,1\}} \{\pi(e_1, \eta_2) - ce_1 + \delta[\pi(e_1, \eta_2)\psi^{\sigma} + (1 - \pi(e_1, \eta_2))\psi^{\varphi}] + \quad (22)$$

$$+\delta[\pi(e_1, \eta_2)\xi + (1 - \pi(e_1, \eta_2))\xi]\} \quad (23)$$

$$= F\psi + \beta\xi. \quad (24)$$

This shows that F is a contraction. □

Appendix B: Proof of Theorem 1

The proof makes use of several propostions that we shall show below. We first characterize a best response against a threshold strategy.

Proposition 4 *Let η_2 be a threshold strategy. Then it holds:*

(i) V_1 is (weakly) increasing in γ and constant on N_2 with

$$V_1(\gamma) = \frac{p - c}{1 - \delta} \quad \text{for all } \gamma \in N_2.$$

(ii) A best response for P1 against η_2 is given by

$$\eta_1^*(\gamma) = \begin{cases} 1 & \text{if} \quad V_1(\gamma) > (1-p)/(1-\delta) \\ \\ 0 & \quad V_1(\gamma) = (1-p)/(1-\delta). \end{cases} \quad (25)$$

(iii) η_1^* is a threshold strategy with threshold $r_1^* = \sup\{\gamma \in \Gamma | V_1(\gamma) = (1-p)/(1-\delta)\}$.[20]

[19] In what follows, we shall occasionally suppress the argument γ and write ψ^{σ} instead of $\psi(\gamma^{\sigma})$ etc.

[20] If $V_1(\gamma) > (1-p)/(1-\delta)$ for all γ, then we define $r_1^* = 0$.

The next proposition states that the best response correspondence is monotone.

Proposition 5 *Let $\eta_2, \widetilde{\eta}_2$ be threshold strategies. Denote by BR_i player i's best response correspondence on the set of threshold strategies of player $-i$. Then it holds:*
(i) $\eta_1 = BR_1(\eta_2) \Longrightarrow N(\eta_1) \cap N(\eta_2) = \emptyset$.
(ii) $\eta_2 \geq \widetilde{\eta}_2 \Longrightarrow BR_1(\eta_2) \leq BR_1(\widetilde{\eta}_2)$.[21]

Our next aim is to define a function that maps thresholds of player $-i$'s threshold strategies on the thresholds of player i's corresponding best response threshold strategies. To do so, we need the following result.

Proposition 6 *Let $\eta_2, \widetilde{\eta}_2$ be threshold strategies with corresponding thresholds $r_2, \widetilde{r}_2$. Let $\eta_1 = BR_1(\eta_2), \widetilde{\eta}_1 = BR_1(\widetilde{\eta}_2)$ be best responses with corresponding thresholds $r_1, \widetilde{r}_1$. Then it holds:*

$$r_2 = \widetilde{r}_2 \Longrightarrow r_1 = \widetilde{r}_1. \tag{26}$$

Denote by $TH_i(\eta_i)$ the threshold of threshold strategy η_i. Proposition 6 allows us to make the following definition.

Definition 4 *Let r_{-i} be a threshold for player $-i$ induced by a threshold strategy η_{-i}. Then the function $\widehat{r}_i : \Gamma \to \Gamma$ is defined by*

$$\widehat{r}_i(r_{-i}) = TH_i(BR_i(\eta_{-i})).$$

We shall show existence of equilibrium by using that $\widehat{r}_1$ and $\widehat{r}_2$ have an intersection (r_1^*, r_2^*). To establish that an intersection exists, it is sufficient to show that $\widehat{r}_i$ is continuous on Γ. The existence of an intersection follows then by a standard fixed-point theorem.

Proposition 7 *The function $\widehat{r}_i$ is continuous for all $r_{-i} \in \Gamma$.*

We are now in the position to prove Theorem 1.

Proof of Theorem 1: *Existence*: Because $\widehat{r}_i$ is continuous on Γ, and Γ is compact, there is an intersection (r_1^*, r_2^*) such that $\widehat{r}_1(r_2^*) = \widehat{r}_2(r_1^*)$. Now, define the strategies $\mu = 1_{(r_1^*,1]}$,

[21] $\eta_i \geq \widetilde{\eta}_i$ if and only if $\eta_i(\gamma) \geq \widetilde{\eta}_i(\gamma)$ for all $\gamma \in \Gamma$. Equivalently: $\eta_i \geq \widetilde{\eta}_i$ if and only if $N(\eta_i) \subseteq N(\widetilde{\eta}_i)$.

$\nu = 1_{[r_1^*,1]}$ for P1, and $\xi = 1_{[0,r_2^*)}$, $\zeta = 1_{[0,r_2^*]}$ for P2. Then, by definition of $\widehat{r}_i$, $BR_1(\eta_2) \in \{\mu, \nu\}$ for $\eta_2 \in \{\xi, \zeta\}$, and $BR_2(\eta_1) \in \{\xi, \zeta\}$ for $\eta_1 \in \{\mu, \nu\}$. Thus, for an equilibrium to exist we have to show that there is a pair $(\eta_1^*, \eta_2^*) \in \{\mu, \nu\} \times \{\xi, \zeta\}$ such that $BR_1(\eta_2^*) = \eta_1^*$ and $BR_2(\eta_1^*) = \eta_2^*$.

To do so, we go through all possible combinations $(BR_1(\xi), BR_1(\zeta), BR_2(\mu), BR_2(\nu))$ of best responses and show that only combinations can arise in which there is a pair (η_1^*, η_2^*) of mutual best responses. We can represent the possible cases by arrow diagrams such as

$$\begin{array}{cc} \mu \searrow & \leftarrow \xi \\ \nu \rightarrow & \nwarrow \zeta \end{array} \tag{27}$$

where the arrow, for example, at μ points to $BR_2(\mu)$; that is, the diagram indicated corresponds to the best response combination (μ, μ, ζ, ξ). A pair (η_1^*, η_2^*) of mutual best responses exists if two arrows point directly towards each other. For example, in the diagram indicated, (μ, ζ) is such a pair.

There are 16 such diagrams, and it is tedious but straightforward to see that there are only two diagrams in which there is no pair of mutual best responses. These diagrams are

$$\text{A)} \quad \begin{array}{cc} \mu \searrow & \leftarrow \xi \\ \nu \nearrow & \leftarrow \zeta \end{array}, \qquad \text{and} \quad \text{B)} \quad \begin{array}{cc} \mu \rightarrow & \swarrow \xi \\ \nu \rightarrow & \nwarrow \zeta \end{array}. \tag{28}$$

We shall now show that these two diagrams are incompatible with the monotonicity of BR. To rule out case A), notice that $\mu < \nu$ and $\xi < \zeta$. Because $\xi < \zeta$, Proposition 5 implies that $BR_1(\xi) \geq BR_1(\zeta)$. Hence, it cannot be that $BR_1(\xi) = \mu$ and $BR_1(\zeta) = \nu$. Case B) is ruled out with an identical argument. This establishes existence of an equilibrium.

As for (i): "$\Rightarrow$": Let $r_1^* > 0$. By Proposition 5 (i), $\eta_2^*(\gamma) = 1$ for all $\gamma < r_1^*$. Notice that P1's utility from $e_1 = 1$ against $\eta_2(\gamma) = 1$ is at least

$$\lambda = p_{11}^1 \gamma + p_{11}^2 (1 - \gamma) - c + \delta \frac{1 - p}{1 - \delta}. \tag{29}$$

This is so because P1 gets $p_{11}^1 \gamma + p_{11}^2 (1 - \gamma) - c$ in the current period and then more than $1 - p$ in all future periods. Therefore for all $\gamma < r_1^*$: $(1 - p) / (1 - \delta) = V_1(\gamma) \geq \lambda$. That is, $p_{11}^1 \gamma + p_{11}^2 (1 - \gamma) - c \leq 1 - p$ for all $\gamma < r_1^*$. Because the weak inequality holds for all $\gamma < r_1^*$ and its left hand side is increasing, it must hold with strict inequality for $\gamma = 0$. This implies that $p_{11}^2 < 1 - p + c$. This shows "$\Rightarrow$".

"$\Leftarrow$": Let $p_{11}^2 < 1 - p + c$. Suppose to the contrary that $r_1^* = 0$. Because η_1^* is a threshold strategy, this implies that $\eta_1^* = 1$ for all γ except possibly for $\gamma = 0$. There are now two cases:

(a) $\Gamma \backslash N_2^*$ contains an open neighbourhood, or (b) it does not. In case (a), we show that there is a γ such that $\eta_1^*(\gamma) = 1$ and $V_1(\gamma) < (1-p)/(1-\delta)$, a contradiction to Proposition 4 (ii).

Because of (a) and because η_2^* is a threshold strategy, $\Gamma \backslash N_2^*$ contains $\gamma = 0$. Thus, for each $k \in \mathbb{N}$, we can find a γ_k close to 0 such that, first, the action profile $(1,1)$ is played for at least k periods even if P1 were to experience k successes in a row and, second, $\pi(1,1;\gamma^t) - c$ remains smaller than $1-p$ in these k periods, because by assumption $p_{11}^2 - c < 1 - p$. Hence, starting at γ_k, P1's per period payoff is less than $1-p$ for at least k periods and, trivially, less than $p - c$ from then on. Thus,

$$V_1(\gamma_k) < \sum_{t=0}^{k} \delta^t (1-p) + \sum_{t=k}^{\infty} \delta^t (p-c)$$

Hence, $V_1(\gamma_k) < (1-p)/(1-\delta)$ for sufficiently large k.

In case (b), $\eta_2^* = 0$ for all $\gamma > 0$. We show that for sufficiently small γ, P2 is better off by playing $e_2 = 1$ rather than $e_2 = 0$. Indeed, since for all $\gamma > 0$ the action profile $(1,0)$ is played, P2's value is $V_2(\gamma) = (1-p)/(1-\delta)$. However, by an argument similar to that in part (i), if P2 plays $e_2 = 1$ he obtains at least

$$\left(1 - p_{11}^1\right)\gamma + \left(1 - p_{11}^2\right)(1-\gamma) - c + \delta \frac{1-p}{1-\delta}.$$

But since $p_{11}^2 < 1 - p + c$, this is strictly larger than $(1-p)/(1-\delta)$ for sufficiently small γ, a contradiction to $V_2(\gamma) = (1-p)/(1-\delta)$. This completes "$\Leftarrow$".

As for (ii): (ii) follows from identical arguments as in (i).

As for (iii): This is an immediate consequence of Proposition 5 (i).

Remaining proofs of propositions 4 to 7:

Proof of Proposition 4: To show (i), we show that the fixed point mapping F preserves the property "*(weakly) increasing in* γ *and constant on* N_2". Denote by $S^{+,c}$ the set of all functions in S that are increasing in γ and constant on N_2. Notice that $S^{+,c}$ is complete.[22] The following proposition states that F maps $S^{+,c}$ in $S^{+,c}$.

Proposition 8 *Let* η_2 *be a threshold strategy, and let* $\psi \in S^{+,c}$*. Then* $F\psi \in S^{+,c}$*.*

[22] This is because the functions in $S^{+,c}$ are only *weakly* increasing.

Now, since F preserves $S^{+,c}$, all elements of a sequence $(F^n\psi)_{n=1,2,\ldots}$ with $\psi \in S^{+,c}$ are in $S^{+,c}$. Since $S^{+,c}$ is complete, the limit V_1 is in $S^{+,c}$.

To determine V_1 on N_2, compare the value of $e_1 = 1$ and the value of $e_1 = 0$. Since $\eta_2(\gamma) = 0$, there is no updating when P1 plays $e_1 = 1$, and $e_1 = 1$ gives

$$p - c + \delta V_1(\gamma).$$

Choosing $e_1 = 0$ gives

$$\pi_{00}(\gamma) + \delta\left[\pi_{00}(\gamma) V_1(\gamma^{\sigma}) + (1 - \pi_{00}(\gamma)) V_1(\gamma^{\varphi})\right]$$

Since V_1 is increasing, it is maximal on N_2, thus

$$V_1(\gamma) \geq \pi_{00}(\gamma) V_1(\gamma^{\sigma}) + (1 - \pi_{00}(\gamma)) V_1(\gamma^{\varphi}).$$

Moreover, by complementarity, $p - c > \pi_{00}(\gamma)$. Hence, $e_1 = 1$ is a maximizer against $\eta_2(\gamma)$ and

$$V_1(\gamma) = p - c + \delta V_1(\gamma) \quad \text{for all } \gamma \in N_2. \tag{30}$$

Thus, $V_1(\gamma) = (p - c) / (1 - \delta)$. This shows (i).

Proof of Proposition 8: We first show that $F\psi$ is constant on N_2. Let $\gamma \in N_2$. $e_1 = 0$ yields

$$\pi_{00}(\gamma) + \delta\left[\pi_{00}(\gamma)\psi(\gamma^{\sigma}) + (1 - \pi_{00}(\gamma))\psi(\gamma^{\varphi})\right],$$

and $e_1 = 1$ yields $p - c + \delta\psi(\gamma)$. Since ψ is increasing and constant on N_2, it is maximal on N_2. Thus,

$$\psi(\gamma) \geq \pi_{00}(\gamma)\psi(\gamma^{\sigma}) + (1 - \pi_{00}(\gamma))\psi(\gamma^{\varphi}).$$

Moreover, by complementarity, $p - c > \pi_{00}(\gamma)$. Hence, $e_1 = 1$ is a maximizer of $F\psi(\gamma)$. Thus,

$$F\psi(\gamma) = p - c + \delta\psi(\gamma) \quad \text{for all } \gamma \in N_2. \tag{31}$$

Since, by assumption, ψ is constant on N_2, it follows that $F\psi$ is constant on N_2.

It remains to show that $F\psi$ is increasing in γ. Let $\gamma \geq \widetilde{\gamma}$. We have to show that $F\psi(\gamma) \geq F\psi(\widetilde{\gamma})$. Let $e_1, \tilde{e}_1 \in \{0, 1\}$ be the maximizers of $F\psi(\gamma)$ and $F\psi(\widetilde{\gamma})$, respectively. We begin by observing that since η_2 is a threshold strategy, $\eta_2(\gamma) \leq \eta_2(\widetilde{\gamma})$. Hence, there are the following cases: A): $\eta_2(\gamma) = 0$ and $\eta_2(\widetilde{\gamma}) = 1$, and B): $\eta_2(\gamma) = \eta_2(\widetilde{\gamma})$.

Consider first case A). Since $\eta_2(\gamma) = 0$, $\gamma \in N_2$. Thus, it follows from the first part of the proof that $e_1 = 1$ is a maximizer of $F\psi(\gamma)$ and $F\psi(\gamma) = p - c + \delta\psi(\gamma)$. We want to show that this is larger than

$$F\psi(\widetilde{\gamma}) = \pi(\tilde{e}_1, \eta_2(\widetilde{\gamma})) - c\tilde{e}_1 + \delta[\pi(\tilde{e}_1, \eta_2(\widetilde{\gamma}))\psi(\widetilde{\gamma}^\sigma) + (1 - \pi(\tilde{e}_1, \eta_2(\widetilde{\gamma})))\psi(\widetilde{\gamma}^\varphi)]. \quad (32)$$

Now, because P1's win probability increases in P1's and decreases in P2's effort, it follows that $p - c \geq \pi(\tilde{e}_1, 1)$. Moreover, because ψ is increasing and constant on $N_2, \psi(\gamma)$ is maximal. Thus, $\psi(\gamma) \geq \psi(\widetilde{\gamma}^\sigma)$ and $\psi(\gamma) \geq \psi(\widetilde{\gamma}^\varphi)$. These two observations imply that $F\psi(\gamma) \geq F\psi(\widetilde{\gamma})$.

Consider next case B): $\eta_2(\gamma) = \eta_2(\widetilde{\gamma})$. We begin by noting that because e_1 is a maximizer of $F\psi(\gamma)$ it follows that

$$F\psi(\gamma) \geq \pi(\tilde{e}_1, \eta_2(\gamma)) - c\tilde{e}_1 + \delta[\pi(\tilde{e}_1, \eta_2(\gamma))\psi(\gamma^\sigma) + (1 - \pi(\tilde{e}_1, \eta_2(\gamma)))\psi(\gamma^\varphi)], \quad (33)$$

where γ^σ and γ^φ are computed on the basis of $\tilde{e}_1$. Denote the right hand side of (33) by ζ. We shall now show that $\zeta \geq F\psi(\widetilde{\gamma})$.

To see this, observe first that because $\eta_2(\gamma) = \eta_2(\widetilde{\gamma})$, it follows that $\pi(\tilde{e}_1, \eta_2(\gamma)) = \pi(\tilde{e}_1, \eta_2(\widetilde{\gamma}))$. Observe second that if $\tilde{e}_1 \neq \eta_2(\gamma)$, then updating stops, and $\gamma^\sigma = \gamma^\varphi = \gamma$. Likewise, $\widetilde{\gamma}^\sigma = \widetilde{\gamma}^\varphi = \widetilde{\gamma}$. If $\tilde{e}_1 = \eta_2(\gamma)$, then updating continues, and monotonicity of Bayes' rule implies that $\gamma^\sigma \geq \widetilde{\gamma}^\sigma$ and $\gamma^\varphi \leq \widetilde{\gamma}^\varphi$. Therefore, in either case we have that $\gamma^\sigma \geq \widetilde{\gamma}^\sigma$ and $\gamma^\varphi \leq \widetilde{\gamma}^\varphi$, and because ψ is increasing, it follows that $\psi(\gamma^\sigma) \geq \psi(\widetilde{\gamma}^\sigma)$ and $\psi(\gamma^\varphi) \geq \psi(\widetilde{\gamma}^\varphi)$. These two observations imply that $\zeta \geq F\psi(\widetilde{\gamma})$, and this completes the proof. □

We now show part (ii) of Proposition 4. Let η_2 be a threshold strategy. It may be that for some $\gamma \in \Gamma$, P1 is indifferent between $e_1 = 1$ and $e_1 = 0$. In these cases, we impose the tie-breaking rule that P1 chooses $e_1 = 0$. The best response gets thus single-valued at each $\gamma \in \Gamma$. Let η_1 be such a single-valued best response against η_2. We show:

(a) If $\eta_1(\gamma) = 0$, then $V_1(\gamma) = (1 - p)/(1 - \delta)$.

(b) If $\eta_1(\gamma) = 1$, then $V_1(\gamma) > (1 - p)/(1 - \delta)$.

In other words: $\eta_1(\gamma) = 1$ if and only if $V_1(\gamma) > (1 - p)/(1 - \delta)$.

As for (a): Let $\eta_1(\gamma) = 0$. Suppose, $\eta_2(\gamma) = 0$. The first step in the proof of Proposition 8 implies that P1 optimally chooses $e_1 = 1$ if $\eta_2(\gamma) = 0$, contradicting $\eta_1(\gamma) = 0$. Therefore, $\eta_2(\gamma) = 1$. In this case, since $\eta_1(\gamma) = 0$, there is no updating, and P1's value is given by

$$V_1(\gamma) = 1 - p + \delta V_1(\gamma). \quad (34)$$

This implies that $V_1(\gamma) = (1-p)/(1-\delta)$.

As for (b): Let $\eta_1(\gamma) = 1$. Suppose, $\eta_2(\gamma) = 0$. Then, by part (i),

$$V_1(\gamma) = (p-c)/(1-\delta) > (1-p)/(1-\delta).$$

Suppose, $\eta_2(\gamma) = 1$. We have seen under (a) that the value from $e_1 = 0$ against $\eta_2(\gamma) = 1$, conditional on optimal continuation, is equal to $(1-p)/(1-\delta)$. Our tie-breaking rule implies that P1 strictly prefers $e_1 = 1$ to $e_1 = 0$, if $\eta_1(\gamma) = 1$. Thus, if $\eta_1(\gamma) = 1$, it must be that $V_1(\gamma) > (1-p)/(1-\delta)$.

It remains to show part (iii) of Proposition 4. Due to part (ii),
$N_1 = \{\gamma \mid V_1(\gamma) = (1-p)/(1-\delta)\}$ is the no-effort set of P1's best response against a threshold stratey η_2. Since V_1 is increasing, it follows that $r_1 = \sup N_1$ is a threshold according to Definition 2. □

We now turn to the proofs of Propositions 5 to 7. We start by providing a useful formula to compute utility differences. To do so, we introduce some notation.

Let η_2, $\widetilde{\eta}_2$ be threshold strategies for P2 with $\eta_2 \geq \widetilde{\eta}_2$. Let η_1 be an abritrary threshold strategy for P1 with $N(\eta_1) \cap N(\eta_2) = \emptyset$. Divide $N(\widetilde{\eta}_2)$ into the following sets:

$$A = N(\eta_1) \cap N(\widetilde{\eta}_2), \quad B = N(\widetilde{\eta}_2) \setminus (A \cup N(\eta_2)). \tag{35}$$

The following picture illustrates these sets.

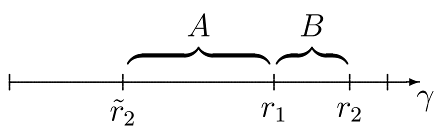

Figure 2: The sets A and B

Notice that B is well defined and non-empty because, by assumption, $\eta_2 \geq \widetilde{\eta}_2$. We shall now define first entry times into these sets. Define for $j = A, B$

$$T_j(\omega) = \min\left\{t \geq 0 \mid \gamma^t(\omega) \in j\right\}. \tag{36}$$

Furthermore, define by

$$\{T_A < T_B\} = \{\omega \mid T_A(\omega) < \infty, T_A(\omega) < T_B(\omega)\} \tag{37}$$

the event that the belief reaches A in finite time and that A is reached before B.[23] Define likewise $\{T_B < T_A\}$. Then we have the following result.

Proposition 9 *The utility difference from playing η_1 against η_2 rather than against $\widetilde{\eta}_2$ is given by*

$$U_1(\eta_1, \eta_2) - U_1(\eta_1, \widetilde{\eta}_2) \tag{38}$$

$$= \int_{\{T_A<T_B\}} \delta^{T_A(\omega)} \left[\frac{1-p}{1-\delta} - U_1(\eta_1, \widetilde{\eta}_2)\left(\gamma^{T_A(\omega)}\right)\right] dP(\omega\,; \eta_1, \eta_2, \gamma) \tag{39}$$

$$+ \int_{\{T_B<T_A\}} \delta^{T_B(\omega)} \left[U_1(\eta_1, \eta_2)\left(\gamma^{T_B(\omega)}\right) - \frac{p-c}{1-\delta}\right] dP(\omega\,; \eta_1, \eta_2, \gamma). \tag{40}$$

Proof: Let $\gamma^0 = \gamma \in \Gamma$, and let ω be an infinite sequence $\omega = (\omega^1, ..., \omega^t, ...), \omega^t \in \{\sigma, \varphi\}$ of successes and failures of player 1. Let $(\gamma^t)_{t=1,2,...}$ be the unique sequence of beliefs induced by ω under strategies η_1, η_2, and initial state γ^0. Let $\eta_i^t = \eta_i(\gamma^t)$. P1's utility of η_1 against η_2 is

$$U_1(\eta_1, \eta_2) = \int \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t\right] dP(\omega\,; \eta_1, \eta_2, \gamma) \tag{41}$$

$$= \int_{\{T_A<T_B\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t\right] dP(\omega\,; \eta_1, \eta_2, \gamma) \tag{42}$$

$$+ \int_{\{T_B<T_A\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t\right] dP(\omega\,; \eta_1, \eta_2, \gamma) \tag{43}$$

$$+ \int_{\{T_A=T_B=\infty\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t\right] dP(\omega\,; \eta_1, \eta_2, \gamma) \tag{44}$$

We can split the sum in (42) into dates before and after A is reached, that is, the integral in (42) is equal to

$$\int_{\{T_A<T_B\}} \sum_{t=0}^{T_A-1} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t\right] + \delta^{T_A(\omega)} U_1(\eta_1, \eta_2)\left(\gamma^{T_A(\omega)}\right) dP(\omega\,; \eta_1, \eta_2, \gamma). \tag{45}$$

For (43) we obtain the corresponding expression. Now notice that the two profiles (η_1, η_2) and $(\eta_1, \widetilde{\eta}_2)$ coincide as long as A or B is not reached. Hence, before A or B is reached both the transition probabilities and the current payoff are the same under (η_1, η_2) and under $(\eta_1, \widetilde{\eta}_2)$.

[23] Again, we neglect measurability issues. If we had defined the probability space properly, it would however be easy to show that T_j is a stopping time and that $\{T_A < T_B\}$ is measurable.

Therefore,

$$\int_{\{T_A<T_B\}} \sum_{t=0}^{T_A-1} \delta^t \left[\pi\left(\eta_1^t,\eta_2^t;\gamma^t\right)-c\eta_1^t\right] dP\left(\omega\,;\eta_1,\eta_2,\gamma\right) \tag{46}$$

$$+\int_{\{T_B<T_A\}} \sum_{t=0}^{T_B-1} \delta^t \left[\pi\left(\eta_1^t,\eta_2^t;\gamma^t\right)-c\eta_1^t\right] dP\left(\omega\,;\eta_1,\eta_2,\gamma\right) \tag{47}$$

$$=\int_{\{T_A<T_B\}} \sum_{t=0}^{T_A-1} \delta^t \left[\pi\left(\eta_1^t,\widetilde{\eta}_2^t;\gamma^t\right)-c\eta_1^t\right] dP\left(\omega\,;\eta_1,\widetilde{\eta}_2,\gamma\right) \tag{48}$$

$$+\int_{\{T_B<T_A\}} \sum_{t=0}^{T_B-1} \delta^t \left[\pi\left(\eta_1^t,\widetilde{\eta}_2^t;\gamma^t\right)-c\eta_1^t\right] dP\left(\omega\,;\eta_1,\widetilde{\eta}_2,\gamma\right). \tag{49}$$

By the same argument, since (η_1,η_2) and $(\eta_1,\widetilde{\eta}_2)$ coincide on the set $\{T_A=T_B=\infty\}$, the integral in (44) is the same under (η_1,η_2) and $(\eta_1,\widetilde{\eta}_2)$. Therefore, we can write the utility difference from playing η_1 against η_2 rather than against $\widetilde{\eta}_2$ as

$$U_1\left(\eta_1,\eta_2\right)-U_1\left(\eta_1,\widetilde{\eta}_2\right) \tag{50}$$

$$=\int_{\{T_A<T_B\}} \delta^{T_A(\omega)} U_1\left(\eta_1,\eta_2\right)\left(\gamma^{T_A(\omega)}\right)\, dP\left(\omega\,;\eta_1,\eta_2,\gamma\right) \tag{51}$$

$$-\int_{\{T_A<T_B\}} \delta^{T_A(\omega)} U_1\left(\eta_1,\widetilde{\eta}_2\right)\left(\gamma^{T_A(\omega)}\right)\, dP\left(\omega\,;\eta_1,\widetilde{\eta}_2,\gamma\right) \tag{52}$$

$$+\int_{\{T_B<T_A\}} \delta^{T_B(\omega)} U_1\left(\eta_1,\eta_2\right)\left(\gamma^{T_B(\omega)}\right)\, dP\left(\omega\,;\eta_1,\eta_2,\gamma\right) \tag{53}$$

$$-\int_{\{T_B<T_A\}} \delta^{T_B(\omega)} U_1\left(\eta_1,\widetilde{\eta}_2\right)\left(\gamma^{T_B(\omega)}\right)\, dP\left(\omega\,;\eta_1,\widetilde{\eta}_2,\gamma\right). \tag{54}$$

To complete the proof, we need two further arguments. First, since (η_1,η_2) and $(\eta_1,\widetilde{\eta}_2)$ coincide as long as A or B is not reached, a path ω reaches A or B in finite time under (η_1,η_2) if and only if it does so under $(\eta_1,\widetilde{\eta}_2)$. Thus, we can replace $dP\left(\omega\,;\eta_1,\widetilde{\eta}_2,\gamma\right)$ by $dP\left(\omega\,;\eta_1,\eta_2,\gamma\right)$ in (52) and (54).

Second, when A is reached under (η_1,η_2), P1 stops exerting effort while P2 does not, hence the process stops, and thus

$$U_1\left(\eta_1,\eta_2\right)\left(\gamma^{T_A(\omega)}\right)=\frac{1-p}{1-\delta}. \tag{55}$$

Similarily, when B is reached under $(\eta_1,\widetilde{\eta}_2)$, P2 stops exerting effort while P1 does not, hence the process stops, and thus

$$U_1\left(\eta_1,\widetilde{\eta}_2\right)\left(\gamma^{T_B(\omega)}\right)=\frac{p-c}{1-\delta}. \tag{56}$$

Collecting terms yields now the desired expression. □

Proof of Proposition 5: As for (i): Proposition 4 (i) and (ii) imply that P1 optimally chooses $e_1 = 1$ on N_2 . This implies the claim.

As for (ii): Let $\eta_2 \geq \widetilde{\eta}_2$. Denote by $\eta_1, \widetilde{\eta}_1$ the corresponding best responses, and by $V_1, \widetilde{V}_1$ the corresponding value functions of P1. By Proposition 4, (ii), $\eta_1(\gamma) = 1$ if and only if $V_1(\gamma) > (1-p)/(1-\delta)$. Hence, it is enough to show that $\widetilde{V}_1(\gamma) \geq V_1(\gamma)$ for all $\gamma \in \Gamma$. We shall show that η_1 gives a higher utility against $\widetilde{\eta}_2$, than η_1 gives against η_2. That is,

$$U_1(\eta_1, \widetilde{\eta}_2) \geq U_1(\eta_1, \eta_2). \tag{57}$$

By definition of a best response, this implies $\widetilde{V}_1 \geq U_1(\eta_1, \widetilde{\eta}_2) \geq U_1(\eta_1, \eta_2) = V_1$.

To show (57), notice that by (i), $N(\eta_1) \cap N(\eta_2) = \emptyset$. Hence we can apply Proposition 9 to obtain

$$U_1(\eta_1, \eta_2) - U_1(\eta_1, \widetilde{\eta}_2) \tag{58}$$

$$= \int_{\{T_A < T_B\}} \delta^{T_A(\omega)} \left[\frac{1-p}{1-\delta} - U_1(\eta_1, \widetilde{\eta}_2)\left(\gamma^{T_A(\omega)}\right)\right] dP(\omega\,; \eta_1, \eta_2, \gamma) \tag{59}$$

$$+ \int_{\{T_B < T_A\}} \delta^{T_B(\omega)} \left[U_1(\eta_1, \eta_2)\left(\gamma^{T_B(\omega)}\right) - \frac{p-c}{1-\delta}\right] dP(\omega\,; \eta_1, \eta_2, \gamma). \tag{60}$$

We shall now show that this expression is non-positive. To see this, observe first that when A is reached under $(\eta_1, \widetilde{\eta}_2)$, both players stop exerting effort. The worst thing that can happen to P1 in this case is that he loses in the current period and the next period belief is such that $(0, 1)$ is played from then on and P1 would receive a period payoff of $1-p$ forever (note that this follows from complementarity: $p_{00}^2 > 1 - p + c > 1 - p$). Hence,

$$U_1(\eta_1, \widetilde{\eta}_2)\left(\gamma^{T_A(\omega)}\right) \geq \pi_{00}(\gamma) + \delta\frac{1-p}{1-\delta} \geq \frac{1-p}{1-\delta}. \tag{61}$$

Second, notice that because η_1 is a best response against η_2, $U_1(\eta_1, \eta_2)\left(\gamma^{T_B(\omega)}\right) = V_1\left(\gamma^{T_B(\omega)}\right)$. Thus, $U_1(\eta_1, \eta_2)\left(\gamma^{T_B(\omega)}\right) \leq (p-c)/(1-\delta)$ according to Proposition 4 (i). This completes the proof. □

Proof of Proposition 6: If $\eta_2 = \widetilde{\eta}_2$, the claim is trivially true. If $\eta_2 \neq \widetilde{\eta}_2$, then there are only two threshold strategies with equal thresholds, $\eta_2 = 1_{[0,r_2]}$ and $\widetilde{\eta}_2 = 1_{[0,r_2)}$. Notice that $\eta_2 \geq \widetilde{\eta}_2$. Let V_1 and $\widetilde{V}_1$ be the value functions against η_2 and $\widetilde{\eta}_2$, respectively. Suppose now to the contrary that $r_1 \neq \widetilde{r}_1$. Then Proposition 5 implies that $\widetilde{r}_1 < r_1$. Thus, there is a non-empty interval $I \subset (\widetilde{r}_1, r_1)$, and it follows that $V_1(\gamma) = (1-p)/(1-\delta)$ for all $\gamma \in I$. We shall show that there is a $\gamma \in I$ such that

$$U_1(\widetilde{\eta}_1, \eta_2)(\gamma) > \frac{1-p}{1-\delta}, \tag{62}$$

which yields a contradiction, because by definition of a best response $V_1(\gamma) \geq U_1(\widetilde{\eta}_1, \eta_2)(\gamma)$.

To prove (62), notice first that since $I \cap N(\widetilde{\eta}_1) = \emptyset$, it holds that $\widetilde{V}_1 > (1-p)/(1-\delta)$ on I. Hence, (62) is implied if there is a $\gamma \in I$ such that

$$U_1(\widetilde{\eta}_1, \eta_2)(\gamma) = \widetilde{V}_1(\gamma). \tag{63}$$

We shall now show (63). To do so, we apply Proposition 9. (Notice: the strategy denoted η_1 in the Proposition, is now $\widetilde{\eta}_1$). Due to Proposition 5 (i), $A = N(\widetilde{\eta}_1) \cap N(\widetilde{\eta}_2) = \emptyset$. Moreover, $B = N(\widetilde{\eta}_2) \setminus (A \cup N(\eta_2)) = \{r_2\}$. Thus,

$$U_1(\widetilde{\eta}_1, \eta_2)(\gamma) - \widetilde{V}_1(\gamma) \tag{64}$$

$$= U_1(\widetilde{\eta}_1, \eta_2)(\gamma) - U_1(\widetilde{\eta}_1, \widetilde{\eta}_2)(\gamma) \tag{65}$$

$$= \int_{\{T_{\{r_2\}}<\infty\}} \delta^{T_{\{r_2\}}(\omega)} \left[U_1(\widetilde{\eta}_1, \eta_2)\left(\gamma^{T_{\{r_2\}}(\omega)}\right) - \frac{p-c}{1-\delta}\right] dP(\omega\,; \widetilde{\eta}_1, \eta_2, \gamma). \tag{66}$$

We now show that there is a $\gamma \in I$ such that $P\left[T_{\{r_2\}} < \infty\,; \widetilde{\eta}_1, \eta_2, \gamma\right] = 0$, hence the integral is 0. To do so, let

$$\Gamma_{r_2}^t = \left\{\gamma \mid \exists\omega : \gamma^t(\omega) = r_2, \text{ given } \gamma^0 = \gamma \text{ and } (\widetilde{\eta}_1, \eta_2) \text{ is played}\right\} \tag{67}$$

the set of all possible initial beliefs γ such that $\{r_2\}$ can be possibly hit after t periods. Define by $\Gamma_{r_2} = \cup_t \Gamma^t(r_2)$ the set of all initial beliefs such that $\{r_2\}$ can be hit in finite time. Since in each period, there are only two events, success and failure, $\Gamma_{r_2}^t$ is finite. Accordingly, $\Gamma_{r_2} = \cup_t \Gamma^t(r_2)$ is countable. Hence the intersection $I \cap (\neg\Gamma_{r_2})$ is non-empty and for all initial beliefs $\gamma \in I \cap (\neg\Gamma_{r_2})$, $\{r_2\}$ is not reached in finite time, that is, $P\left[T_{\{r_2\}} < \infty\,; \widetilde{\eta}_1, \eta_2, \gamma\right] = 0$. This completes the proof. □

Proof of Proposition 7: Let $\left(r_{-i}^n\right)_n$ be a sequence with limit r_{-i}. We show that $\lim \widehat{r}_i\left(r_{-i}^n\right) = r_{-i}$. To do so, we make use of the following Proposition.

Proposition 10 *Let $\left(\eta_{-i}^n\right)_n$ be a sequence of threshold strategies with $\eta_{-i}^n \underset{n\to\infty}{\longrightarrow} \eta_{-i}$ (pointwisely). Let $r_i^n = \widehat{r}_i\left(TH_{-i}\left(\eta_{-i}^n\right)\right)$ and $r_i = \widehat{r}_i\left(TH_{-i}(\eta_{-i})\right)$. Then*

$$r_i^n \underset{n\to\infty}{\longrightarrow} r_i \tag{68}$$

To complete the proof, it therefore suffices to construct a sequence $\left(\eta_{-i}^{n}\right)_{n}$,$\eta_{-i}$ of threshold strategies with $TH_{-i}\left(\eta_{-i}^{n}\right)=r_{-i}^{n}$ and $TH_{-i}\left(\eta_{-i}\right)=r_{-i}$ such that $\left(\eta_{-i}^{n}\right)$ converges to η_{-i} (pointwisely). We demonstrate the construction for $i=1$ (the other case is analogous).

Define $\eta_{2}^{n}=1_{\left[0,r_{2}^{n}\right]}$. There are two cases. (I): there are infinitely many sequence members $r_{2}^{n}\geq r_{2}$ and (II): otherwise. Now define $\eta_{2}=1_{[0,r_{2}]}$ in case (I), and $\eta_{2}=1_{[0,r_{2})}$ in case (II). It then follows that η_{2}^{n} converges to η_{2} pointwisely and this completes the proof. □

Proof of Proposition 10: We only show left-continuity, that is, if $\eta_{2}^{n}\upuparrows\eta_{2}$, then $r_{1}^{n}\upuparrows r_{1}$. Right-continuity is shown with identical arguments.

Let V_{1}^{n} and V_{1} be the value functions associated with the best responses η_{1}^{n} and η_{1} against η_{2}^{n} and η_{2}. The proof proceeds in two steps. In STEP 1, we show that $\lim V_{1}^{n}=V_{1}$. In STEP 2, we show that $\lim V_{1}^{n}=V_{1}$ implies that $r_{1}^{n}\upuparrows r_{1}$.

STEP 1: To show the pointwise convergence of V_{1}^{n} to V_{1}, notice that since, $\eta_{2}^{n}\leq\eta_{2}^{n+1}\leq\eta_{2}$ by assumption, it follows from the proof of Proposition 5 that $V_{1}^{n}\geq V_{1}^{n+1}\geq V_{1}$. Thus, since V_{1}^{n} is bounded from below, the (pointwise) limit exists and $\lim_{n}V_{1}^{n}\geq V_{1}$. We shall show that

$$U_{1}\left(\eta_{1}^{n},\eta_{2}^{n}\right)-U_{1}\left(\eta_{1}^{n},\eta_{2}\right)\rightarrow 0. \tag{69}$$

Thus, since $\lim U_{1}\left(\eta_{1}^{n},\eta_{2}^{n}\right)$ exists, it exists $\lim U_{1}\left(\eta_{1}^{n},\eta_{2}\right)$, and the two limits coincide. This implies that $\lim V_{1}^{n}=\lim U_{1}\left(\eta_{1}^{n},\eta_{2}^{n}\right)=\lim U_{1}\left(\eta_{1}^{n},\eta_{2}\right)\leq U_{1}\left(\eta_{1},\eta_{2}\right)=V_{1}$, where the inequality holds because η_{1} is a best response against η_{2}. Because we also have that $\lim V_{1}^{n}\geq V_{1}$, it follows that $\lim V_{1}^{n}=V_{1}$.

To show (69), let $\varepsilon>0$. Let $\tau\in\mathbb{N}$ be such that

$$\delta^{\tau}\left(\frac{p}{1-\delta}\right)<\varepsilon. \tag{70}$$

Define by

$$\Gamma^{\tau}=\left\{\gamma\in\Gamma\,|\,\text{there is an }\omega\text{ such that }\gamma^{t}\left(\omega\right)=\gamma\text{ for some }t\in\{0,...,\tau\}\right\} \tag{71}$$

the set of all possible beliefs up to time τ. Because there are only two events, success or failure, in each period, Γ^{τ} contains at most $2^{\tau}+1<\infty$ elements. Therefore it exists

$$\xi_{\tau}=\min_{\gamma\in\Gamma^{\tau}\backslash N(\eta_{2})}dist\left(\gamma,N\left(\eta_{2}\right)\right), \tag{72}$$

where $dist\left(\gamma, N\left(\eta_{2}\right)\right)$ is the smallest (Euclidean) distance between γ and the closure of $N\left(\eta_{2}\right)$. In particular, $\xi_{\tau}>0$.[24]

Define by $C^{n}=N\left(\eta_{2}^{n}\right) \backslash N\left(\eta_{2}\right)$ the set where η_{2}^{n} and η_{2} differ. Let r_{2}^{n} and r_{2} be the thresholds associated with η_{2}^{n} and η_{2}. By Proposition 5 $\eta_{2}^{n} \upuparrows \eta_{2}$ implies $r_{2}^{n} \upuparrows r_{2}$. Thus, there is a number n^{τ} such that

$$r_{2}-r_{2}^{n}<\xi_{\tau} / 2 \quad \text{for all } n>n^{\tau}, \tag{73}$$

and it follows by construction that the belief process can not reach $C^{n^{\tau}}$ before time τ; that is, $C^{n^{\tau}} \cap \Gamma^{\tau}=\emptyset$. The following picture illustrates the set C^{n}.

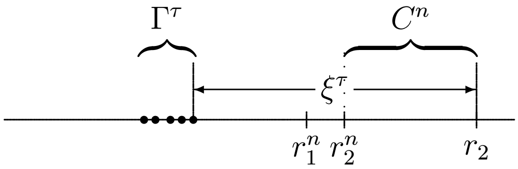

Figure 3: The set C^{n} for $n>n^{\tau}$

To determine $\left|U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)-U_{1}\left(\eta_{1}^{n}, \eta_{2}^{n}\right)\right|$, notice first that it follows from the proof of Proposition 5 that $U_{1}\left(\eta_{1}^{n}, \eta_{2}^{n}\right) \geq U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)$, and we can delete the absolute value operator. To determine $U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)-U_{1}\left(\eta_{1}^{n}, \eta_{2}^{n}\right)$, we shall apply Proposition 9 (where now η_{2} plays the role of $\widetilde{\eta}_{2}$, and $\eta_{1}^{n}, \eta_{2}^{n}$ play the role of η_{1}, η_{2} respectively). Notice first that by Proposition 5 (i), $A=N\left(\eta_{2}^{n}\right) \cap N\left(\eta_{1}^{n}\right)=\emptyset$. Thus, $B=N\left(\eta_{2}^{n}\right) \backslash\left(A \cup N\left(\eta_{1}\right)\right)=C^{n}$. Hence, we obtain that

$$U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)-U_{1}\left(\eta_{1}^{n}, \eta_{2}^{n}\right)=\int_{\left\{T_{C^{n}}<\infty\right\}} \delta^{T_{C^{n}}(\omega)}\left[\frac{p-c}{1-\delta}-U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)\left(\gamma^{T_{C^{n}}(\omega)}\right)\right] d P\left(\omega ; \eta_{1}, \eta_{2}, \gamma\right). \tag{74}$$

Because $U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)\left(\gamma^{T_{C^{n}}(\omega)}\right)$ is certainly larger than $-c /(1-\delta)$, it follows that

$$U_{1}\left(\eta_{1}^{n}, \eta_{2}\right)-U_{1}\left(\eta_{1}^{n}, \eta_{2}^{n}\right) \leq \frac{p}{1-\delta} \int_{\left\{T_{C^{n}}<\infty\right\}} \delta^{T_{C^{n}}(\omega)} d P. \tag{75}$$

Now we can write the integral as

$$\int_{\left\{T_{C^{n}}<\tau\right\}} \delta^{T_{C^{n}}(\omega)} d P+\int_{\left\{\tau \leq T_{C^{n}}<\infty\right\}} \delta^{T_{C^{n}}(\omega)} d P. \tag{76}$$

The first term is smaller than $P\left[T_{C^{n}}<\tau\right]$ which is 0, because, by construction, the belief process cannot reach C^{n} before time $t=\tau$. Moreover, the second term is smaller than

[24] If $\Gamma^{\tau} \subset N\left(\eta_{2}\right)$, then $\Gamma^{\tau} \backslash N\left(\eta_{2}\right)=\emptyset$, and $\xi_{\tau}=0$. In this case, $\gamma^{0} \in N\left(\eta_{2}\right)$. Therefore, $V_{1}\left(\gamma^{0}\right)=(p-c) /(1-\delta)$, and it follows directly that $V_{1} \geq V_{1}^{n}$.

$\delta^\tau P\left[\tau \leq T_{C^n} < \infty\right] \leq \delta^\tau$. Hence, we obtain

$$U_1\left(\eta_1^n, \eta_2\right) - U_1\left(\eta_1^n, \eta_2^n\right) \leq \delta^\tau \left(\frac{p}{1-\delta}\right) < \varepsilon. \tag{77}$$

This shows step 1.

STEP 2: Suppose to the contrary that $\left(r_1^n\right)_n$ does not converge to r_1. Since $\eta_2^n \leq \eta_2^{n+1} \leq \eta_2$ by assumption, it follows from Proposition 5 that $r_1^n \leq r_1^{n+1} \leq r_1$. Hence, it exists $\overline{r}_1 = \lim r_1^n$ and there is an $\varepsilon > 0$ such that $r_1 - \overline{r}_1 > \varepsilon$. Moreover, by Proposition 5 (i), $r_1 \leq r_2$ and, by assumption, $r_2^n \upuparrows r_2$. Thus there is a $K \in \mathbb{N}$ such that $\left|r_1 - r_2^k\right| < \varepsilon/2$ for all $k > K$. Thus, $\min\left\{r_2^K, r_1\right\} - r_1^n > \varepsilon/2$ for all n. Hence, we can pick beliefs $\gamma, \widetilde{\gamma} \in \left(\overline{r}_1, \min\left\{r_2^K, r_1\right\}\right)$ with $\gamma - \widetilde{\gamma} > \varepsilon/4$. Notice that, by construction, $\eta_1^n\left(\widetilde{\gamma}\right) = \eta_1^n\left(\gamma\right) = 1$ and $\eta_2^n\left(\widetilde{\gamma}\right) = \eta_2^n\left(\gamma\right) = 1$ for all $n > K$. Therefore,

$$V_1^n\left(\gamma\right) - V_1^n\left(\widetilde{\gamma}\right) = \pi\left(1,1;\gamma\right) - \pi\left(1,1;\widetilde{\gamma}\right) \tag{78}$$

$$+ \delta\left[\pi\left(1,1;\gamma\right) V_1^n\left(\gamma^\sigma\right) + \left(1 - \pi\left(1,1;\gamma\right)\right) V_1^n\left(\gamma^\varphi\right)\right] \tag{79}$$

$$- \delta\left[\pi\left(1,1;\widetilde{\gamma}\right) V_1^n\left(\widetilde{\gamma}^\sigma\right) + \left(1 - \pi\left(1,1;\widetilde{\gamma}\right)\right) V_1^n\left(\widetilde{\gamma}^\varphi\right)\right]. \tag{80}$$

We shall now estimate this expression. First, the first line computes to $\left(p_{11}^1 - p_{11}^2\right)\left(\gamma - \widetilde{\gamma}\right)$ and is therefore strictly larger than $\left(p_{11}^1 - p_{11}^2\right)\left(\varepsilon/4\right)$. Next, because $V_1^n\left(\gamma^\sigma\right) \geq V_1^n\left(\gamma^\varphi\right)$ and because $\pi\left(1,1;\gamma\right) \geq \pi\left(1,1;\widetilde{\gamma}\right)$, the second line is larger than

$$\delta\left[\pi\left(1,1;\widetilde{\gamma}\right) V_1^n\left(\gamma^\sigma\right) + \left(1 - \pi\left(1,1;\widetilde{\gamma}\right)\right) V_1^n\left(\gamma^\varphi\right)\right]. \tag{81}$$

Moreover, monotonicity of Bayes' rule implies that $V_1^n\left(\gamma^\sigma\right) \geq V_1^n\left(\widetilde{\gamma}^\sigma\right)$ and $V_1^n\left(\gamma^\varphi\right) \geq V_1^n\left(\widetilde{\gamma}^\varphi\right)$. Therefore, the second and the third line taken together are larger than zero. Hence,

$$V_1^n\left(\gamma\right) - V_1^n\left(\widetilde{\gamma}\right) > \left(p_{11}^1 - p_{11}^2\right)\left(\varepsilon/4\right) \text{ for all } n > K. \tag{82}$$

Now, by construction, $\gamma, \widetilde{\gamma} < r_1 \leq r_2$. Thus, $\eta_1\left(\widetilde{\gamma}\right) = \eta_1\left(\gamma\right) = 0$ and $\eta_2\left(\widetilde{\gamma}\right) = \eta_2\left(\gamma\right) = 1$. Hence, at γ and $\widetilde{\gamma}$, learning stops under $\left(\eta_1, \eta_2\right)$. Consequently, $V_1\left(\gamma\right) = V_1\left(\widetilde{\gamma}\right)$. Now recall that, by hypothesis, $\lim V_1^n = V_1$, hence $V_1^n\left(\gamma\right) - V_1^n\left(\widetilde{\gamma}\right) \to V_1\left(\gamma\right) - V_1\left(\widetilde{\gamma}\right) = 0$, a contradiction to (82). This establishes step 2 and completes the proof. □

Appendix C: Proof of Theorem 2

Proof of Theorem 2: (i) Let λ be the the smallest number such that, starting from initial belief $\gamma^0 = r_1^*$, the belief process $\left(\gamma^t\right)_t$ moves to a belief $\gamma^{\sigma\ldots\sigma} \geq r_2^*$ after λ consecutive successes for P1, conditional on both players choosing $e_1 = e_2 = 1$. Notice that because (γ^t) is Markov, it is also true that if $\gamma^t = r_1^*$ then $\gamma^{t+\lambda} \geq r_2^*$ after λ consecutive successes for P1. Moreover, because γ^t converges to 1 if P1 wins infinitely often, λ is finite. Now, for $k = 0, 1, ...$ define by

$$A_k = \left\{\omega | \omega^{\lambda k+1} = ... = \omega^{(k+1)\lambda} = \sigma\right\} \tag{83}$$

the event that P1 wins λ times in a row beginning in period $t(k) = \lambda k + 1$. Define by $A_\infty = \cap_{m=0}^{\infty} \cup_{k \geq m} A_k$ the event that A_k occurs infinitely often. The Lemma of Borel-Cantelli (see Chow/Teicher 1978, p. 60, Thm 1) implies that $P[A_\infty | \theta] = 1$ for all θ. This is so because A_k is stochastically indepent of A_{k+1} and because $\sum P[A_k | \theta] = \sum \left(p_{11}^\theta\right)^\lambda = \infty$.

Now notice that A_∞ is included in the event $\{T < \infty\}$. To see this, let $\omega \in A_\infty$. Then there is a k such that $\omega \in A_k$. There are two possibilities: first, $\gamma^{t(k)}(\omega) \in N_1^* \cup N_2^*$ in which case $T(\omega) \leq t(k) < \infty$. Second, $\gamma^{t(k)}(\omega) \notin N_1^* \cup N_2^*$. In this case, monotonicity of Bayes' rule implies that $\gamma^{t(k+1)}(\omega) \in N_2^*$. This is so because, given that $(\gamma^t)_t(\omega)$ moves into N_2^* after λ success of P1 starting from $\gamma^0 = r_1^*$, then it does so a forteriori when it starts from $\gamma^0 > r_1^*$. Hence, $T(\omega) \leq t(k+1) < \infty$. This shows that $A_\infty \subseteq \{T < \infty\}$. Since, in addition, $P[A_\infty] = 1$, the assertion follows.

(ii) Suppose $r_1^* = 0$ and $r_2^* < 1$. Then by a similar argument as in (i) for each initial belief $\gamma \in (0, 1)$ there is a finite number $\lambda(\gamma)$ such that starting from γ the belief process moves into N_2^* after $\lambda(\gamma)$ consecutive successes of P1. The probability of this to happen is $\left(p_{11}^\theta\right)^{\lambda(\gamma)} > 0$ for all θ. This implies the assertion. The argument for the case $r_1^* > 0$ and $r_2^* = 1$ is identical. □

Appendix D: Derivation of Entry Time

In this appendix, we provide a quantitative expression for the first entry time. We shall look at the special case in which $p_{11}^2 = 1/2$ and $p_{11}^1 > p - c$. That is, in state $\theta = 2$ P1 is as able as P2. By Theorem 1 (i), (ii), $r_1^* = 0$ and $r_2^* < 1$. Thus, the action profile $(1, 1)$ is played if $\gamma < r_2^*$, and the action profile $(1, 0)$ is played if $\gamma \geq r_2^*$. We are interested in the likelihood that the belief process enters N_2^* in finite time although players are equally able, that is, in $P[T_2 < \infty \,|\, \theta = 2]$.

Theorem 3 *Let $\gamma^0 < r_2^*$ be the initial belief that the state is $\theta = 1$. Define*

$$\alpha = 2\left(\log\frac{p_{11}^1}{1 - p_{11}^1}\right)^{-1}\left(\log\frac{(1-\gamma^0)\, r_2^*}{\gamma^0(1 - r_2^*)}\right), \quad \beta = -\left(\log\frac{p_{11}^1}{1 - p_{11}^1}\right)^{-1}\log 4p_{11}^1\left(1 - p_{11}^1\right).$$

Let q_0 be the unique solution $q > 1$ to $q^2 - 2q^{1+\beta} + 1 = 0$. Then

$$P[T_2 < \infty \mid \theta = 2] \approx q_0^{-\alpha}.$$

The comparative statics properties with respect to p_{11}^1 are ambiguous. It can be shown that α decreases, and q_0 increases in p_{11}^1. Therefore, the sign of the derivative of $q_0^{-\alpha}$ with respect to p_{11}^1 generally depends on p_{11}^1. Intuitively, if p_{11}^1 is close to $p - c$, then r_2^* is close to 1 and the contest outcome is not very informative. Therefore, $P[T_2 < \infty \mid \theta = 2]$ should be exptected to be small. On the other extreme, if p_{11}^2 is close to 1, updating is strong, and only few successes for P1 might be enough to enter N_2^*. However, few failures of P1, which obtains with probablity 1/2 in $\theta = 2$, are enough to identify the state.

Proof of Theorem 3: Fix $\theta = 2$. For all sequences ω, define the random variable $X_s(\omega)$ as equal to $+1$ if P1 wins in period s, and equal to -1 if P1 loses in period s. The sum $L_t(\omega) = \sum_{s=1}^{t} X_s(\omega)$ is then P1's *lead* after t periods, and $(L_t + t)/2$ is the number of P1's successes. Conditional on $\theta = 2$, P1's win probability is $p_{11}^2 = 1/2$ as long as N_2 has not been reached by γ^t and so L_t is a symmetric random walk. For all ω such that $\gamma^t(\omega) \notin N_2$, iterated application of Bayes' rule yields that

$$\gamma^t = \frac{(p_{11}^1)^{(L_t+t)/2}(1 - p_{11}^1)^{t-(L_t+t)/2}\gamma_0}{(p_{11}^1)^{(L_t+t)/2}(1 - p_{11}^1)^{t-(L_t+t)/2}\gamma^0 + (1/2)^t(1 - \gamma^0)}. \tag{84}$$

A little bit of algebra yields that

$$\gamma^t \geq r \Longleftrightarrow L_t \geq \alpha + t\beta, \tag{85}$$

where $\alpha > 0$ and $\beta \in (0, 1)$ are as stated in the Theorem.

Define now the function $\varphi(q) = q^2 - 2q^{1+\beta} + 1$ and let q_0 be the solution $q > 1$ to $\varphi(q) = 0$. Notice that q_0 is unique since $\beta \in (0, 1)$. Define further the function $h(x) = q_0^x$, and the process $H_t(\omega) = h(L_t(\omega) - t\beta)$. It is easy to show that H_t is a martingale.

Let $\xi > 0$ and define by

$$T_\alpha(\omega) = \min\{t \geq 0 \mid L_t(\omega) - t\beta \geq \alpha\}, \qquad T_\xi^-(\omega) = \min\{t \geq 0 \mid L_t(\omega) - t\beta \leq -\xi\} \tag{86}$$

the first passage times of $L_t - t\beta$ through α and $-\xi$, respectively. Denote by $T_{\min} = \min(T_\alpha, T_\xi^-)$ the minimum of the two. By the same argument as in the proof of Theorem 2 (i), it follows that $T_{\min} < \infty$ almost surely. Thus, the stopped process $H_{t \wedge T}$ is a bounded martingale, and we can apply the optional sampling theorem (see Chung 1974, Thm 9.35). This says that $E[H_0] = E[H_{T_{\min}}]$. Hence,

$$1 = E[H_0] \tag{87}$$

$$= E[H_{T_{\min}}] \tag{88}$$

$$= E\left[H_{T_\alpha} \,|\, T_\alpha < T_\xi^-\right] P\left[T_\alpha < T_\xi^-\right] + E\left[H_{T_\xi^-} \,|\, T_\xi^- < T_\alpha\right] P\left[T_\xi^- < T_\alpha\right] \tag{89}$$

$$\approx \left(q_0^\alpha - q_0^{-\xi}\right) P\left[T_\alpha < T_\xi^-\right] + q_0^{-\xi}. \tag{90}$$

The approximation in the last line comes from the fact that $L_t - t\beta$ has only discrete values, so H_{T_α} is not exactly equal to q_0^α. Hence, $P\left[T_\alpha < T_\xi^-\right] \approx \left(1 - q_0^{-\xi}\right) / \left(q_0^\alpha - q_0^{-\xi}\right)$. With $\xi \to \infty$, we obtain

$$P[T_2 < \infty \,|\, \theta = 0] = P[T_\alpha < \infty] \approx \frac{1}{q_0^\alpha}. \tag{91}$$

This shows the claim. □

References

Aspinwall, L. G., L. Richter, R. R. Hoffman (2001). "Understanding how optimism "works": An examination of optimists' adaptive moderation of belief and behavior." In *Optimism and Pessimism: Theory, Research, and Practice*, edited by E. C. Chang, Washington, American Psychological Association.

Berry, D. A., B. Fristedt (1985): *Bandit Problems*, New York, Chapman and Hall.

Blackwell, D. (1965): "Discounted Dynamic Programming," *Annals of Mathematical Statistics*, 36, 226-235.

Bolton, P., C. Harris (1999): "Strategic Experimentation," *Econometrica*, 67, 349-374.

Bowles, S., H. Gintis, M. Osbourne (2001): "The Determinants of Earnings: A Behavioral Approach," *Journal of Economic Literature*, 39, 1137-1176.

Brockner, J. (1984): "Low Self-Esteem and Behavioral Plasticity," *Review of Personality and Social Psychology* (Vol. 4), edited by L. Wheeler, Beverly-Hills, CA, Sage.

Chow, Y., H. Teicher (1978): *Probability Theory*, New York, Springer.

Chung, K. L. (1974): *A Course in Probability Theory*, New York, Academic Press.

Easley, D., N. Kiefer (1988): "Controlling Stochastic Processes with Unknown Parameters," *Econometrica*, 56, 1045-1064.

Fishman, A., R. Rob (1998): "Experimentation and Competition," *Journal of Economic Theory*, 78, 299-320.

Kalai, E., E. Lehrer (1995): "Subjective Games and Equilibria," *Games and Economic Behaivor*, 8, 123-163.

Keller, G., S. Rady, M. Cripps (2005): "Strategic Experimentation with Exponential Bandits," *Econometrica*, 73, 39-69.

Lazear, E., S. Rosen (1981): "Rank-Order Tournaments as Optimum Labour Contracts," *Journal of Political Economy*, 89, 841-864.

Maskin, E., J. Tirole (2001): "Markov Perfect Equilibrium 1. Observable Actions," *Journal of Economic Theory*, 100, 191-219.

Meyer, M. (1991): "Learning from Coarse Information: Biased Contests and Career Profiles," *Review of Economic Studies*, 58, 15-41.

Meyer, M. (1991): "Biased Contests and Moral Hazard: Implications for Career Profiles," *Annales d' Économie et de Statistique*, 25/26, 165-187.

Picketty, T. (1995): "Social Mobility and Redistributive Politics," *Quarterly Journal of Economics*, 110, 551-584.

Rothschild, M. (1974): "A Two-Armed Bandit Theory of Market Pricing", *Journal of Economic Theory*, 9, 185-202.

Rosen, S. (1986): "Prizes and Incentives in Elimination Tournaments," *American Economic Review*, 76, 701-715.

Solon, G. (1999): "Intergenerational Mobility in the Labor Market," in *The Handbook of Labor Economics*, Vol. 3a, edited by Ashenfelter, O., D. Card, Amsterdam, North-Holland.

Squintani, F, J. Välimäki (2002): "Imitation and Experimentation in Changing Contests," *Journal of Economic Theory*, 104, 376-404.

Stone, R. (2004): "Self-Handicapping, Tournaments, and Ego-Utility," mimeo.

Appendix E: Missing calculations—not for publication

Missing Calculations on page 16

Proof of

$$r_1 = \frac{p_{11}^1 + c - p}{2p_{11}^1 - 1} \in \left(0, \frac{1}{2}\right), \tag{92}$$

$$\widetilde{r}_1 = \frac{p_{00}^1 + p - c - 1}{2p_{00}^1 - 1} \in \left(\frac{1}{2}, 1\right). \tag{93}$$

To determine r_1, suppose $e_2 = 1$. We show that $e_1 = 1$ is a best response if and only if $\gamma \geq r_1$. Indeed, P1's utility from choosing $e_1 = 1$ is

$$p_{11}^1 \gamma + p_{11}^2 (1 - \gamma) - c = \left(2p_{11}^1 - 1\right) \gamma + 1 - p_{11}^1 - c, \tag{94}$$

where we have used the symmetry assumption $p_{11}^2 = 1 - p_{11}^1$. P1's utility from choosing $e_1 = 0$ is $1 - p$. Thus $e_1 = 1$ is a best response if and only if

$$\left(2p_{11}^1 - 1\right) \gamma + 1 - p_{11}^1 - c \geq 1 - p \Longleftrightarrow \tag{95}$$

$$\gamma \geq \frac{p_{11}^1 + c - p}{2p_{11}^1 - 1} = r_1. \tag{96}$$

Notice also that $r_1 < 1/2$ if and only if $p_{11}^1 + c - p < p_{11}^1 - 1/2$ or, equivalently, $p - c > 1/2$. But the latter holds by A2.

To determine $\widetilde{r}_1$, suppose $e_2 = 0$. We show that $e_1 = 0$ is a best response if and only if $\gamma \geq \widetilde{r}_1$. Indeed, P1's utility from choosing $e_1 = 0$ is

$$p_{00}^1 \gamma + p_{00}^2 (1 - \gamma) = \left(2p_{00}^1 - 1\right) \gamma + 1 - p_{00}^1, \tag{97}$$

where we have used the symmetry assumption $p_{00}^2 = 1 - p_{00}^1$. P1's utility from choosing $e_1 = 1$ is $p - c$. Thus $e_1 = 0$ is a best response if and only if

$$\left(2p_{00}^1 - 1\right) \gamma + 1 - p_{00}^1 \geq p - c \Longleftrightarrow \tag{98}$$

$$\gamma \geq \frac{p_{00}^1 + p - c - 1}{2p_{00}^1 - 1} = \widetilde{r}_1. \tag{99}$$

Notice also that $\widetilde{r}_1 < 1$ if and only if $p_{00}^1 + p - c - 1 < 2p_{00}^1 - 1$ or, equivalently, $p_{00}^1 > p - c$. But the latter is the complacency assumption. □

Derivation of mixing probabilities: Let $\alpha_1 \in [0,1]$ be the probability with which P1 chooses $e_1 = 1$. Then P2's utility from choosing $e_2 = 0$ is

$$\alpha_1 (1-p) + (1-\alpha_1) \left[\left(1 - p_{00}^1\right)\gamma + \left(1 - p_{00}^2\right)(1-\gamma)\right] \tag{100}$$

$$= \alpha_1 (1-p) + (1-\alpha_1) \left[\left(1 - 2p_{00}^1\right)\gamma + p_{00}^1\right], \tag{101}$$

where again we have used symmetry to get from the first to the second line. Similarly, P2's utility from choosing $e_2 = 1$ is

$$\alpha_1 \left[\left(1 - p_{11}^1\right)\gamma + \left(1 - p_{11}^2\right)(1-\gamma)\right] + (1-\alpha_1) p - c \tag{102}$$

$$= \alpha_1 \left[\left(1 - 2p_{11}^1\right)\gamma + p_{11}^1\right] + (1-\alpha_1) p - c. \tag{103}$$

Thus, P2 is indifferent between $e_2 = 0$ and $e_2 = 1$ if and only if

$$\alpha_1 (1-p) + (1-\alpha_1) \left[\left(1 - 2p_{00}^1\right)\gamma + p_{00}^1\right] \tag{104}$$

$$= \alpha_1 \left[\left(1 - 2p_{11}^1\right)\gamma + p_{11}^1\right] + (1-\alpha_1) p - c. \tag{105}$$

Rearranging terms gives

$$\alpha_1 \left[(1-p) + p - \left(1 - 2p_{00}^1\right)\gamma - p_{00}^1 - \left(1 - 2p_{11}^1\right)\gamma - p_{11}^1\right] \tag{106}$$

$$= p - c - \left(1 - 2p_{00}^1\right)\gamma - p_{00}^1. \tag{107}$$

The square bracket in the first line simplifies to $1 - p_{11}^1 - p_{00}^1 - 2\left[1 - p_{11}^1 - p_{00}^1\right]\gamma$, thus we obtain

$$\alpha_1(\gamma) = \frac{p - c - p_{00}^1 - \left(1 - 2p_{00}^1\right)\gamma}{(2\gamma - 1)\left(1 - p_{11}^1 p_{00}^1\right)}, \tag{108}$$

which is what is claimed on page 16. The mixing probability for P2 is $\alpha_1(1-\gamma)$ due to symmetry. □

Missing best response diagrams in the proof of Theorem 1: There are the following 16 best response diagrams

$$\begin{matrix} \mu \rightarrow & \leftarrow \xi \\ \nu \rightarrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \rightarrow & \leftarrow \xi \\ \nu \rightarrow & \nwarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \rightarrow & \swarrow \xi \\ \nu \rightarrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \rightarrow & \swarrow \xi \\ \nu \rightarrow & \nwarrow \zeta \end{matrix} \tag{109}$$

$$\begin{matrix} \mu \rightarrow & \leftarrow \xi \\ \nu \nearrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \rightarrow & \leftarrow \xi \\ \nu \nearrow & \nwarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \rightarrow & \swarrow \xi \\ \nu \nearrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \rightarrow & \swarrow \xi \\ \nu \nearrow & \nwarrow \zeta \end{matrix} \tag{110}$$

$$\begin{matrix} \mu \searrow & \leftarrow \xi \\ \nu \rightarrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \searrow & \leftarrow \xi \\ \nu \rightarrow & \nwarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \searrow & \swarrow \xi \\ \nu \rightarrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \searrow & \swarrow \xi \\ \nu \rightarrow & \nwarrow \zeta \end{matrix} \tag{111}$$

$$\begin{matrix} \mu \searrow & \leftarrow \xi \\ \nu \nearrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \searrow & \leftarrow \xi \\ \nu \nearrow & \nwarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \searrow & \swarrow \xi \\ \nu \nearrow & \leftarrow \zeta \end{matrix}, \quad \begin{matrix} \mu \searrow & \swarrow \xi \\ \nu \nearrow & \nwarrow \zeta \end{matrix} \tag{112}$$

In all diagrams there are two arrows pointing directly towards each other except in the top right and the bottom left diagram. These are diagrams A) and B). □

Missing calculations in the proof of Theorem 3 in Appendix C

Proof of inequality (85): Multiplying through by the denominator of γ^t gives that $\gamma^t \geq r$ is equivalent to

$$\left(p_{11}^1\right)^{(L_t+t)/2}\left(1-p_{11}^1\right)^{t-(L_t+t)/2}\gamma^0 \geq r_2^*\left[\left(p_{11}^1\right)^{(L_t+t)/2}\left(1-p_{11}^1\right)^{t-(L_t+t)/2}\gamma^0+(1/2)^t\left(1-\gamma^0\right)\right]. \tag{113}$$

Sorting for p_{11}^1 gives

$$\left(p_{11}^1\right)^{(L_t+t)/2}\left(1-p_{11}^1\right)^{t-(L_t+t)/2} \geq \frac{r_2^*(1/2)^t\left(1-\gamma^0\right)}{\gamma^0\left(1-r_2^*\right)}. \tag{114}$$

Taking log and multiplying through by 2 yields

$$(L_t+t)\log p_{11}^1+(t-L_t)\log\left(1-p_{11}^1\right) \geq 2\left[t\log(1/2)+\log\frac{r_2^*\left(1-\gamma^0\right)}{\gamma^0\left(1-r_2^*\right)}\right] \Leftrightarrow \tag{115}$$

$$L_t\left(\log p_{11}^1-\log\left(1-p_{11}^1\right)\right) \geq t\left(2\log(1/2)-\log p_{11}^1-\log\left(1-p_{11}^1\right)\right) \tag{116}$$

$$+2\log\frac{r_2^*\left(1-\gamma^0\right)}{\gamma^0\left(1-r_2^*\right)} \tag{117}$$

Rearranging the log expressions gives

$$L_t\log\frac{p_{11}^1}{\left(1-p_{11}^1\right)} \geq t\left(-\log 4p_{11}^1\left(1-p_{11}^1\right)\right)+2\log\frac{r_2^*\left(1-\gamma^0\right)}{\gamma^0\left(1-r_2^*\right)}. \tag{118}$$

Dividing by $\log\left(p^1_{11}/\left(1-p^1_{11}\right)\right)$ yields now the expressions for α and β. $\square$

Proof that $H = (H_t)_t$ **with** $H_t = q_0^{L_t - t\beta}$ **is a martingale**: For H to be a martingale, we have to show:

(i) $E[H_t] < \infty$ for all t.

(ii) $E[H_{t+1}|H_t] = H_t$ for all t.

As for (i): Since $(L_t)_t$ is a random walk, L_t is bounded. Morover, $t\beta$ is bounded for all t. This implies the claim.

As for (ii): Because $(L_t)_t$ is a random walk, $L_{t+1} = L_t + 1$ with probability $1/2$, and $L_{t+1} = L_t - 1$ with probability $1/2$. Thus,

$$E[H_{t+1}|H_t] = \frac{1}{q_0^{(t+1)\beta}}\frac{1}{2}\left[q_0^{L_t+1} + q_0^{L_t-1}\right] \tag{119}$$

$$= \frac{q_0^{L_t}}{q_0^{t\beta}}\frac{1}{q_0^{\beta}}\frac{1}{2}\left[q_0^{1} + q_0^{-1}\right] \tag{120}$$

$$= H_t\frac{1}{q_0^{\beta}}\frac{1}{2}\left[q_0^{1} + q_0^{-1}\right]. \tag{121}$$

Thus, we need to show that the expression following H_t in the last line equals 1. This is the case if and only if $\left[q_0^1 + q_0^{-1}\right] = 2q_0^{\beta}$. Multiplying through with q_0 (which is positive by definition) gives $q_0^2 - 2q_0^{\beta+1} + 1 = 0$. But this is true by definition of q_0. $\square$

Proof that $T_{\min} < \infty$ **almost surely**: Let α and ξ be integers. Define for $k = 0, 1, ...$the event

$$A_k = \left\{\omega | X_{k(\alpha+\xi)+1} = ... = X_{(k+1)(\alpha+\xi)} = -1\right\} \tag{122}$$

that P1 loses $\alpha + \xi$ times in a row beginning in period $t(k) = k(\alpha + \xi) + 1$. Define by $A_\infty = \cap_{m=0}^{\infty} \cup_{k \geq m} A_k$ the event that A_k occurs infinitely often. The Lemma of Borel-Cantelli (see Chow/Teicher 1978, p. 60, Thm 1) implies that $P[A_\infty] = 1$. This is so because A_k is stochastically indepent of A_{k+1} and because $\sum P[A_k] = \sum (1/2)^{\alpha+\xi} = \infty$.

Now notice that A_∞ is included in the event $\{T_{\min} < \infty\}$. To see this, let $\omega \in A_\infty$. Then there is a k such that $\omega \in A_k$. There are two possibilities: first, $L_{t(k)} - t(k)\beta \geq \alpha$ in which case $T_{\min}(\omega) \leq t(k) < \infty$. Second, $L_{t(k)} - t(k)\beta < \alpha$ in which case $L_{t(k+1)} - t(k+1)\beta < -\xi$. Hence, $T_{\min}(\omega) \leq t(k+1) < \infty$. This shows that $A_\infty \subseteq \{T_{\min} < \infty\}$. Since, in addition, $P[A_\infty] = 1$, it follows that $T_{\min} < \infty$ almost surely.