ECONSTOR

WWW.ECONSTOR.EU

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationszentrum Wirtschaft
The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics

Krähmer, Daniel

Working Paper

Learning and self-confidence in contestsLearning and self-confidence in contests

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Mitglied der
Leibniz-Gemeinschaft

Daniel Krähmer *

Learning and Self-Confidence in Contests

* WZB – Wissenschaftszentrum Berlin

SP II 2003 – 10

September 2003

ABSTRACT

Learning and Self-Confidence in Contests

by Daniel Krähmer [*]

The paper studies a repeated contest when contestants are uncertain about their true abilities. A favourable belief about one's own ability (confidence) stimulates effort and increases the likelihood of success. Success, in turn, reinforces favourable beliefs. We consider a specific example in which this reinforcement mechanism implies that, with positive probability, players fail to learn their true abilities, and one player may eventually win the contest forever. As a consequence, persistent inequality arises, and the worse player may eventually prevail. Furthermore, confidence is self-serving in that it increases a player's utility and the likelihood to be the long-run winner.

Keywords: Contest, self-confidence, belief reinforcement, incomplete learning, dynamic programming

JEL Classification: C37, C61, D44, D83

ZUSAMMENFASSUNG

Lernen und Selbstvertrauen in Wettkämpfen

Das Papier betrachtet einen wiederholten Wettkampf, in dem die Wettkämpfer ihre wahren Fähigkeiten nicht kennen. Wettkämpfer mit einer hohen Einschätzung ihrer eigenen Fähigkeiten (Selbstvertrauen) zeigen eine höhere Einsatzbereitschaft und haben damit bessere Erfolgsaussichten. Umgekehrt verstärken Erfolge das Selbstvertrauen. Wir betrachten ein einfaches Beispiel, in dem dieser sich selbst verstärkende Effekt dazu führt, dass die Spieler mit positiver Wahrscheinlichkeit über ihre wahren Fähigkeiten im Ungewissen bleiben, und dass ein Spieler schließlich für immer als Sieger aus dem Wettkampf hervorgeht. Als Folge ergeben sich dauerhafte Ungleichheiten, wobei es der tatsächlich unfähigere Spieler sein kann, der langfristig überlegen ist. Darüber hinaus zeigt sich, dass Selbstvertrauen sowohl den Nutzen eines Spielers als auch die Erfolgswahrscheinlichkeit, der schließlich überlegene Spieler zu sein, erhöht.

---

1 Introduction

Consider a contest in which the probability of success is determined by a player's ability, effort, and luck. If ability and effort are complements, highly able players will exert more effort than less able ones. When ability is unknown, then effort is chosen according to perceived rather than to true ability: players who regard themselves as highly able will exert more effort than more pessimistic ones. Higher effort increases the probability of success for an optimistic player, and more successes, being indicative of high ability, will, in turn, reinforce an optimistic player's optimistic belief. At the same time as the winner becomes more optimistic, the loser becomes more pessimistic. This suggests that in a contest with unknown abilities beliefs tend to diverge, accompanied by long winning streaks for one player.

The paper illustrates this basic reinforcement mechanism by means of a simple example and explores its consequences for the long run. We show that belief reinforcement implies that, with positive probability, players fail to learn their true abilities in equilibrium, and that persistent inequality will arise in the long run in the sense that one player will eventually win the contest forever, whereas the other player will eventually be discouraged and reduce effort to a minimum. This is true irrespective of objective abilities. Therefore, the objectively better player may be discouraged, and at the same time, the objectively worse player may entertain a mistakenly favourable self-belief.

Besides complementarity of ability and effort, the driving force of these results is that we assume that the importance of ability declines if one player reduces effort. Only if both players invest sufficient effort, ability makes the difference. This reflects a particular form of strategic interaction in our contest: for example, if a player is the only one to apply for a job, he gets the job with certainty, irrespective of his ability.

Declining importance of ability implies that the contest outcome becomes uninformative if a player sufficiently reduces effort. At such a point, updating will stop, the optimistic player easily wins, and inequality becomes persistent. That such a point may be reached, is a result of belief reinforcement which will eventually drive beliefs sufficiently apart such that the discouraged player will reduce effort to an uninformative minimum.

Beliefs in our model may be interpreted in two ways. On the one hand, initial beliefs, i.e., priors, may represent the objective distribution from which nature draws actual ability.

A player with a higher belief then not only regards himself to be better than a player with a lower belief, but he is indeed objectively better (on average). On the other hand, beliefs may be seen as a player's internal, or subjective representation of the world which, in principle, may initially be entirely disconnected from external, or objective reality.[1] Seen as such, beliefs are a characteristic of a player and may be interpreted as self-confidence. The material consequences of (purely 'mental') self-confidence can then be studied by comparative statics with respect to beliefs. However, it is important to note that comparative statics is only meaningful for a given fixed realization of intrinsic abilities, i.e., ex post.[2]

With this second interpretation, we can study the impact of (initial) self-confidence on a player's performance. As a consequence of complementarity, higher self-confidence increases the probability to be the eventually prevailing player. In this sense, confidence is self-serving. Long term success may not necessarily be the result of ability but rather of confidence in combination with an initial streak of good luck.

The latter is consistent with evidence from social psychology suggesting that positive illusions promote motivation and persistence and thereby increase the likelihood to succeed in a given task (Taylor/Brown (1988), Colvin/Block (1994)). Conversely, there is evidence that individuals with low self-esteem exert less effort as a result of not expecting to succeed, triggering the self-fulfilling prophecy described above (Brockner (1984)).

As for contests, in a recent economic experiment Gneezy et al. (2001) find that women's absolute performance tends to be worse than men's in mixed- but not in single-gender tournaments. They also find that men are typically more confident than women. In line with our results, they argue that the difference in the two treatments might be due to a confidence effect. That confidence is self-serving in contests, is also consistent with anecdotal statements of top athletes that suggest that what makes the difference in top competition is not primarily ability but rather mental strength.

In a very stylized way, our model may also be relevant for other domains, particularly labour markets. Labour market success—getting accepted to a good university, getting a job,

---

[1]See Van den Steen (2002) for an elaboration of that view.

[2]To study the material consequences as objectively expected ex ante, one has to evaluate each possible ability scenario by some given objective distribution which is known only to some ideal outside observer but completely unknown to the players of the game. In what follows, all results are such that they hold for every possible ability scenario. Therefore, we do not need to distinguish between subjective and objective distribution.

or getting promoted—is often determined by relative performance, i.e., contests.

For example, our results are in line with research that suggests that long-term unemployment might, to a large extent, be the result of discouragement and loss of self-esteem and not only of missing intrinsic ability (Darity/Goldsmith (1996), Dunifon/Duncan (1998)).

Our model may also contribute to the discussion on intergenerational mobility in labour markets. Seen as an intergenerational model, it predicts a strong correlation between the labour market success of parents and of their children. 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 confidence, or self-esteem transmitted to children by parental upbringing may possibly be one of the missing elements in explaining the intergenerational stability of labour market outcomes.

Literature

Several studies look at the impact of beliefs on economic performance, but their focus is mostly different from ours. Our notion of confidence is similar to the one in Benabou/Tirole (2002a, 2002b), who examine the incentives to manipulate information concerning the self. For instance, by delegating tasks, a principal may foster an agent's confidence and thereby stimulate effort. There is evidence that overconfidence and optimism is prominent among CEOs and entrepreneurs which may give rise to distorted investment decisions (see Malmendier/Tate (2002) for an empirical paper, and Manove/Padilla (1999), de Meza/Southey (1996) for theoretical papers). Confidence-dependent performance may also explain attribution biases and judgement errors (Compte/Postlewaite (2001)).

As for learning, two recent studies show how misperceptions in form of overconfidence and optimism may survive in the long run. Heifetz/Spiegel (2000) consider a class of stage games that includes contests, but in contrast to us, they consider an evolutionary dynamics and formalize optimism as biased perceptions of the game rather than as beliefs. Optimism may also survive in a perfect competition environment (Manove (2000)).

Our no-learning result bears some similarity to no-learning results in bandit problems (Roth-

schild (1974), Berry/Fristedt (1985)).[3] In fact, in our model, a player's problem can be viewed as a two-armed bandit problem where an arm's payoff characteristic depends directly on which arm the other player chooses. We are not aware of any paper that has considered a bandit problem of that type.[4]

Finally, in the contest literature, only few papers consider players who do not know their own abilities. One example is Rosen (1986) who studies the optimal design of prizes in an elimination tournament. In an elimination tournament however, the reinforcement mechanism we study does not obtain because after each period winners are matched with similarily confident winners.

The paper is organized as follows. Section 2 presents the model. Section 3 illustrates the basic belief reinforcement mechanism. Section 4 analyses an infinite horizon model. To provide intuition, section 4.1 considers the case with myopic players. Section 4.2 deals with the dynamcally optimizing players. Section 5 concludes.

2 The Model

Time $t = 0, 1, 2, ..., \overline{t}$ is discrete and may be finite or infinite. In each period, there are two players, $i = 1, 2$, who engage in a contest. The winner prize is normalized to 1 and the loser prize to 0. The winning probability in period t is determined by players' types (*ability*) and efforts spent.

Players' types are unknown, also to players themselves. Let α_i be player i's type, where $\alpha_i \in \{\alpha_L, \alpha_H\}, \alpha_L < \alpha_H$, and denote by γ_i^t player i's belief in period t to be of type α_H. We assume that there is no asymmetric information, that is, beliefs are common knowledge. γ_i is referred to as player i's *confidence*. Let $\gamma = (\gamma_1, \gamma_2)$, and define $\Delta\gamma = \gamma_1 - \gamma_2$, and $\Delta\alpha = \alpha_H - \alpha_L$.

Efforts e_i are in some (ordered) set D, and the cost of spending effort e is $c(e)$. D and c are specified in the sections below.

The true winning probability for player i in a given period for given efforts and given types is denoted by $p_i(e_i, e_{-i}; \alpha)$, and p_i is increasing in e_i and α_i and decreasing in e_{-i} and α_{-i}. Since

---

[3]For an interesting interpretation of a bandit problem as a self-confidence model see Bar-Isaac (2000).

[4]In the literature on multi-armed bandit games, an arm's payoff depends on what other players do (Bolton/Harris (1999), Keller/Rady (2002)). But in these approaches, a player's payoff is not affected *directly* by other players' choices but by the possibility to observe their payoffs.

true abilities are unknown, players have to form expectations about their winning chances. Denote by $\pi\left(e_1, e_2; \gamma\right)$ player 1's *expected* winning probability for efforts e_1 and e_2, given belief γ.

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

3   Belief Reinforcement in the Short Run

The purpose of this section is to illustrate the basic reinforcement mechanism when efforts and abilities are complements. We use a simple linear model with continuous effort levels and quadratic costs. That is,

$$p_i\left(e_i, e_{-i}; \alpha\right) = \frac{1}{2} + \alpha_i e_i - \alpha_{-i} e_{-i}, \tag{1}$$

$$c\left(e\right) = \frac{1}{2}e^2. \tag{2}$$

To guarantee that p_i is a probability, we restrict e to the interval $D = [0, 1/\left(2\alpha_H\right)]$.

The model is a score difference model where player i's score is given by $\alpha_i e_i$ plus some noise term with the noise difference uniformly distributed.

Consider first the one-shot contest ($\overline{t} = 1$). For given e_{-i}, the first order condition for player i yields the best reply

$$e_i^* = \overline{\alpha}_i, \tag{3}$$

where $\overline{\alpha}_i = \gamma_i \Delta \alpha + \alpha_L$ is the expected ability of player i. Since the best response does not depend on the rival's effort e_i, (3) also describes the equilibrium. Thus, in scenario (α_1, α_2) the objective winning probability for player 1 in equilibrium is

$$p_1^* = \frac{1}{2} + \alpha_1 \overline{\alpha}_1 - \alpha_2 \overline{\alpha}_2 \tag{4}$$

$$= \gamma_1 \Delta \alpha \cdot \alpha_1 - \gamma_2 \Delta \alpha \cdot \alpha_2 + \alpha_L \left(\alpha_1 - \alpha_2\right) + \frac{1}{2}. \tag{5}$$

This implies the following result.

Proposition 1 *For $\overline{t} = 1$ it holds:*
p_i^* is increasing in own confidence γ_i and decreasing in the rival's confidence γ_{-i}.

The proposition is a direct consequence of complementarity. It says that confidence increases a player's winning chance. In this sense, confidence is self-serving. Higher confidence increases the expected marginal return to effort. This increases the effort chosen which, in turn, raises the objective winning chance.

A further question is whether differences in abilities can be compensated by confidence. That is, whether the objectively worse player can have a higher objective winning chance if the former is sufficiently more confident than his rival. It turns out that in a two-type environment, this can never be the case. To see this, suppose player 1 is of type α_L and has the highest possible confidence $\overline{\alpha}_1 = \alpha_H$ while player 2 is of type α_H and has the lowest possible confidence $\overline{\alpha}_2 = \alpha_L$. Then player 1's objective winning chance is $1/2 + \alpha_L \alpha_H - \alpha_H \alpha_L = 1/2$ which is also the objective winning chance of player 2. Thus, differences in confidence cannot offset differences in abilities.

Note however that this is no longer true in an environment with three types. If there is a third, intermediate, type α_M with $\alpha_L < \alpha_M < \alpha_H$, then in state (α_L, α_M) with confidence $(\overline{\alpha}_1, \overline{\alpha}_2) = (\alpha_H, \alpha_L)$ the winning chance of player 1 is $1/2 + \alpha_L \alpha_H - \alpha_M \alpha_L$, and that of player 2 is $1/2 + \alpha_M \alpha_L - \alpha_L \alpha_H$. Hence, player 1 is more likely to win.

Suppose next that there are two time periods ($\overline{t} = 2$). One may think of two legs in a sports tournament or of two halfs in a match. Consider a symmetric situation with initially equally confident players, that is, $\gamma_1 = \gamma_2$. Due to symmetry, there is a symmetric perfect Bayesian Nash equilibrium in which both players choose the same effort in period 1. This equilibrium has the following two properties.

Proposition 2 *For $\overline{t} = 2$ and $\gamma_1 = \gamma_2$ it holds:*
(i) For all scenarios (α_1, α_2), the likelihood of winning in period 2 increases after a first period success and decreases after a first period failure.
(ii) If both players are equally able objectively, that is, $\alpha_1 = \alpha_2$, then one player winning twice is more likely than each player winning once.

The proposition results from Bayesian updating together with complementarity. By Bayesian updating, winning in period 1 is indicative of being highly able and raises confidence. By complementarity, this increases effort and thus the likelihood of winning in period 2.[5] This is stated

---

[5]The effort choice in period 1 has therefore two benefits. Besides the current winning probability, it affects

in (i). (ii) says that uneven outcomes, or inequalities are likely to obtain in the short run. With $\alpha_1 = \alpha_2$ and equal efforts in period 1, each player's chance of winning in period 1 is 1/2. Due to updating, the first period winner increases effort and the loser decreases effort in period 2, raising the winner's likelihood to also win in period 2 above 1/2. The total probability of two consecutive successes for a player is therefore larger than 1/2.

While these results illustrate the logic of our basic belief reinforcement mechanism for the short run, the next section explores its consequencesfor the long run.

4   Long-Run Implications and Learning

We now consider an infinite horizon model $(\bar{t} = \infty)$. For this, we assume that efforts are discrete, that is, a player can either spend effort $(e = 1)$ or not $(e = 0)$. The cost of spending effort is $c > 0$.

The winning probabilities in a given period are as follows. If exactly one player spends effort, this player will win with probability 1, irrespective of abilities. If both players spend the same effort $e_i = e_{-i}$, then the winning probability for player 1, conditional on types, is[6]

$$p(e_1, e_2; \alpha) = \frac{1}{2} + \alpha_1 e_1 - \alpha_2 e_2, \tag{6}$$

and for player 2 it is $1 - p(e_1, e_2; \alpha)$.

In summary, the matrix of winning probabilities of player 1 is

$$
\begin{array}{ccc}
 & e_2 = 1 & e_2 = 0 \\
e_1 = 1 & \frac{1}{2} + \alpha_i - \alpha_{-i} & 1 \\
e_1 = 0 & 0 & \frac{1}{2}.
\end{array}
\tag{7}
$$

Embodied in this contest success function are thus two crucial assumptions: first, as above, ability and effort are complements. Second, if one player does not exert effort, then the contest outcome is independent of abilities. In other words, only if both players invest effort, ability makes the difference. This is the most extreme form of a contest in which the importance of

the future winning chance via its influence on confidence. Period 1 effort is thus larger than period 2 effort. As for contest design, if the designer wants to induce even effort across periods, the simple repeated contest may thus be suboptimal.

[6]We shall now write plain p for p_1.

ability declines with declining effort of one player. This property reflects a specific form of strategic interaction in our contest. For example, if a player is the only one to apply for a job, he gets the job with certainty, irrespective of his ability.

The *expected* winning probability of player 1 for efforts e_1 and e_2, given belief γ, is thus

$$\pi(e_1, e_2; \gamma) = \begin{cases} 0 & \text{if} \quad (e_1, e_2) = (0, 1) \\ 1/2 & \text{if} \quad (e_1, e_2) = (0, 0) \\ 1/2 + \Delta\gamma\Delta\alpha & \text{if} \quad (e_1, e_2) = (1, 1) \\ 1 & \text{if} \quad (e_1, e_2) = (1, 0). \end{cases} \tag{8}$$

We make the following parameter restrictions:

$$0 < 1/2 - (\alpha_H - \alpha_L) < c < 1/2. \tag{9}$$

The first inequality says that p is indeed a probability. The second inequality says that if beliefs are sufficiently different ($\Delta\gamma$ close to $+1$ or -1), then, given the rival player chooses high effort, the current payoff for the pessimistic player from exerting effort is negative. The third inequality says that, given the rival player exerts no effort, the current payoff from choosing effort is higher than from not choosing effort.

4.1 Myopic Players

To build intuition, we shall first consider a sequence of one-shot contests played by myopic players ($\delta = 0$). For this, we first solve for the Bayesian Nash equilibrium of the one-shot contest.

The one-shot game

Best Response

Let $e_2 = 0$. Since $c < 1/2$, it is optimal to choose $e_1 = 1$, irrespective of beliefs. Let $e_2 = 1$, then $e_1 = 1$ is a best response for player 1 if, and only if,

$$\frac{1}{2} + \Delta\gamma\Delta\alpha - c \geq 0. \tag{10}$$

Hence, spending effort $e_1 = 1$ is a best response against $e_2 = 1$ if, and only if, $\gamma_1 \geq (c - 1/2)/\Delta\alpha + \gamma_2$. Let $\theta = (1/2 - c)/\Delta\alpha$, and let $r_1(\gamma_2) = -\theta + \gamma_2$. Then, player 1 exerts effort against $e_2 = 1$ if, and only if, his belief γ_1 exceeds the threshold $r_1(\gamma_2)$. By (9), $\theta \in (0, 1)$.

Because of symmetry, the best response of player 2 looks alike.

Equilibrium

We have the following equilibria in pure strategies:

- If $\Delta\gamma \leq -\theta$, then $e_1 = 0, e_2 = 1$ is the unique equilibrium.

- If $\Delta\gamma \geq \theta$, then $e_1 = 1, e_2 = 0$ is the unique equilibrium.

- If $\Delta\gamma \in (-\theta, \theta)$, then $e_1 = 1, e_2 = 1$ is the unique equilibrium.

Hence, if the difference in beliefs $\Delta\gamma$ is sufficiently large, one player is discouraged and the other player wins with certainty.[7] A typical equilibrium looks as follows.

Evolution

Consider now a sequence of one-shot games which are linked only by Bayesian updating but not by intertemporal utility maximization. That is, players in period t play a Bayesian Nash equilibrium of the one-shot contest given beliefs γ_i^t. According to the outcome of the contest, beliefs are updated and transmitted to players in period $t + 1$. Hence, the model can be interpreted as an intergenerational game in which the successor generation inherits the updated beliefs of the predecessor generation.

Denote by $\Delta\gamma_t$ the difference between player 1's and player 2's beliefs in period t. Notice that once $|\Delta\gamma_t| \geq \theta$, the more confident player will win with certainty in period t. Accordingly, winning is not informative about players' abilities, and beliefs stay the same. That is, $|\Delta\gamma_t| \geq \theta$ implies $\Delta\gamma_{t+1} = \Delta\gamma_t$.

Let T be the first period in which one player ceases to spend effort, that is, $T = \min\{t : |\Delta\gamma_t| \geq \theta\}$. It follows from the previous paragraph that $\Delta\gamma_t = \Delta\gamma_s$ for all $t, s \geq T$. Accordingly, in T one player is discouraged and stops exerting effort forever. This implies the following result.

Proposition 3 *For all scenarios* (α_1, α_2), *with positive probability, one player will be discouraged, that is,*

$$\Pr[T < \infty | (\alpha_1, \alpha_2)] > 0. \tag{11}$$

---

[7] As in the previous section, player i's equilibrium winning chance increases in γ_i. Because efforts are now discrete however, it does not increase smoothly but jumps at the threshold $r_1(\gamma_2)$.

Proof: Since $\theta < 1$, it requires an at most finite number of consecutive successes of one player such that Bayesian updating leads to $|\Delta\gamma_t| \geq \theta$. But this event obtains with positive probability in all scenarios (α_1, α_2), that is, $\Pr[T < \infty | (\alpha_1, \alpha_2)] > 0$. $\square$

As mentioned in the introduction, confidence, or self-esteem, passed on to children in the course of their upbringing may play an important role in determining labour market success and help explain the strong correlation between parents' and childrens' earnings. To be sure, our model is, of course, too simple and stylized to provide a definite explanation. For instance, in reality, parents' and childrens' abilities are certainly not perfectly correlated which weakens the confidence effect. Nevertheless, our model may point into a direction of future research, exactly because it generates inequality without appealing to factors like education, wealth, or neighbourhood effects.

4.2　Farsighted Players

Consider now two infinitely lived players who, in each period, engage in the contest described above. The main difference to myopic players is that exerting effort yields now not only a return for the current period but, depending on what the rival player does, may also generate information that can be used in the future. In other words, by exerting effort a player can experiment.

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). Because of symmetry, it is sufficient to state most of the definitions and results in terms of player 1.

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. To define a Markov strategy properly, denote by σ (and φ respectively) the event that player 1 wins (loses) the contest in a given period. Let

$$H_t = \left\{ \left(e_1^1, e_2^1, \omega^1, ..., e_1^t, e_2^t, \omega^t \right) \mid e_i^s \in \{0, 1\}, \omega^s \in \{\sigma, \varphi\}, 1 \leq s \leq t \right\} \tag{12}$$

be the set of all possible histories up to t, and let $H = \cup H_t$ be the set of all possible histories.[8]

Definition 1 *1) A (pure) strategy $\eta_i = (\eta_i^t)_{t=0,1,\dots}$ for player i is a sequence of mappings where each η_i^t maps histories into actions, that is,*

$$\eta_i^t : H_t \to \{0,1\}. \tag{13}$$

The set of all strategies for player i is denoted by Σ_i.

2) A (stationary) Markov strategy η_i for player i maps beliefs into actions, that is,

$$\eta_i : [0,1]^2 \to \{0,1\}. \tag{14}$$

Notice that we require a Markov strategy to be stationary, that is, to be the same for all periods t. Denote by γ^t the *state* in t, and let $\Gamma = [0,1]^2$ be the set of all possible states. Because players are Bayesian rational, beliefs are derived by Bayes rule. Let $\gamma^\sigma = (\gamma_1^\sigma, \gamma_2^\varphi)$ be the updated beliefs upon player 1 winning and player 2 losing. That is, for (e_1, e_2),

$$\gamma_1^\sigma (e_1, e_2; \gamma) = \frac{\Pr[\sigma \,|\, \alpha_H]\, \gamma_1}{\Pr[\sigma]} \tag{15}$$

$$= \frac{[p(e_1, e_2; (\alpha_H, \alpha_H))\, \gamma_2 + p(e_1, e_2; (\alpha_H, \alpha_L))\,(1 - \gamma_2)]\, \gamma_1}{\pi(e_1, e_2; \gamma)}. \tag{16}$$

Define γ^φ likewise. Notice, by definition of our contest success function, the belief is not updated if one player chooses 0 effort because in this case the contest outcome does not depend on abilities and is therefore not informative.

Utility

Before defining a player's expected utility, we need first to define the appropriate probability measure.[9]

A Markov strategy η_2 of player 2 gives rise to a law of motion that can be controlled by player 1. The resulting transition probabilities govern the evolution of states accordingly. More formally, for $e_1 \in \{0,1\}$,

$$q(\gamma' \,|\, e_1, \eta_2(\gamma), \gamma) = \begin{cases} \pi(e_1, \eta_2(\gamma); \gamma) & \text{if } \gamma' = \gamma^\sigma \\[2ex] 1 - \pi(e_1, \eta_2(\gamma); \gamma) & \text{if } \gamma' = \gamma^\varphi \end{cases} \tag{17}$$

---

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

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

defines a transition kernel from current states γ into next period states γ'. $q\left(\gamma' \,|\, e_1, \eta_2\left(\gamma\right), \gamma\right)$ may be viewed as the (expected) conditional probability that the next period state is γ', conditional on the current state being γ, when player 2 plays η_2, and player 1 plays e_1 in the current period.

Suppose that player 1 plays a strategy $\eta_1 \in \Sigma_1$, and let the initial state 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 player 1 is given by

$$P^t\left[\overline{\omega}_t \,;\, \eta_1, \eta_2, \gamma^0\right] = \prod_{s=1}^{t} q\left(\gamma^s \,|\, \eta_1^{s-1}\left(h_{s-1}\right), \eta_2\left(\gamma^{s-1}\right), \gamma^{s-1}\right), \qquad (18)$$

where h_s is the (unique) history induced by $\overline{\omega}_t, \eta_1, \eta_2$, 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 finite sequences of length t. That is, for all finite sequences $\overline{\omega}_t$ it holds

$$P\left[\overline{\omega}_t \,;\, \eta_1, \eta_2, \gamma^0\right] = P^t\left[\overline{\omega}_t \,;\, \eta_1, \eta_2, \gamma^0\right]. \qquad (19)$$

With this we can define a player's expected utility. Player 1's expected current payoff from playing e_1 at a given state is $\pi\left(e_1, \eta_2\left(\gamma\right); \gamma\right) - ce_1$. Player 1's expected utility from strategy $\eta_1 \in \Sigma_1$ against a Markov strategy η_2 at initial state γ is thus given by

$$U_1\left(\eta_1, \eta_2; \gamma\right) = \int \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t\left(h_t\right), \eta_2\left(\gamma^t\right); \gamma^t\right) - c\eta_1^t\left(h_t\right)\right] dP\left(\omega \,;\, \eta_1, \eta_2, \gamma\right), \qquad (20)$$

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

Player 1's best response η_1^* against a Markov η_2 is thus given as a solution to the maximization problem

$$\sup_{\eta_1 \in \Sigma_1} U_1\left(\eta_1, \eta_2; \gamma\right). \qquad (\text{MP})$$

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 state γ coincides with the best response in (MP) at state 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. To define the contraction, let S be the space of all continuous functions on Γ equipped with the supremums norm. Let $\psi \in S$, and define the operator $F : S \to S$ by

$$F\psi(\gamma) = \max_{e_1 \in \{0,1\}} \{\pi(e_1, \eta_2(\gamma)) - ce_1 + \tag{21}$$

$$+ \delta[\pi(e_1, \eta_2(\gamma)) \psi(\gamma^\sigma) + (1 - \pi(e_1, \eta_2(\gamma))) \psi(\gamma^\varphi)].\}$$

Hence, V_1 is the solution to the fixed point problem $FV = V$.[10] We first show that F is a contraction.

Proposition 4 F *is a contraction.*

The proof, which is in the appendix, shows that Blackwell's sufficiency conditions are satisfied. Because F is a contraction, a fixed point exists and is continuous.

Proposition 5 *For all η_2, V_1 exists and is continuous.*

Proof: *Existence*: Since F is a contraction, it follows by Banach's fixed-point theorem that the fixed-point problem $FV = V$ has a unique solution. This establishes existence of V_1.
Continuity: 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 function on the right hand side of (21) are continuous. Thus, all elements of the sequence $(F^n\psi)_{n=1,2,...}$ are in S. Since S is complete, the

---

[10]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.

limit of this sequence is a member of S, thus continuous. □

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. The case with myopic players suggests to look for an equilibrium in threshold strategies. A threshold strategy is a strategy where a player chooses high effort only when his own belief is sufficiently large and his rival's belief is sufficiently small. More formally:

Definition 2 *A Markov strategy η_i is called threshold strategy if there is an increasing function $r : [0,1] \to [-1,1]$ such that*

$$\eta_1 (\gamma_i, \gamma_{-i}) = 1 \iff \gamma_i > r\left(\gamma_{-i}^+\right), \tag{22}$$

where $r(\gamma^+) = \lim_{\gamma' \downdownarrows \gamma} r(\gamma')$ is the right-hand limit.

Notice that the right-hand limit exists because $r(\gamma')$ is decreasing for $\gamma' \downdownarrows \gamma$ and bounded from below.

If player i plays a threshold strategy, then, for fixed γ_{-i}, he chooses high effort only if γ_i is (strictly) above the threshold $r\left(\gamma_{-i}^+\right)$. Equivalently, for fixed γ_i player i chooses high effort only if γ_{-i} falls below the threshold given by $r^{-1}(\gamma_{-i})$.[11]

The threshold need not be continuous. We shall construct an equilibrium in threshold strategies as the limit of some sequence of threshold strategies. For this we need that the limit of this sequence is again a threshold strategy. If we require a threshold strategy to be continuous, we would need the sequence to converge uniformly. As it turns out however, this is too strong a requirement.

Notice also that, by definition, every threshold strategy is Markov.

It will be convenient to work with an equivalent notion where the threshold r is the 0-level set of some function $\psi : \Gamma \to \mathbb{R}$ that increases if γ moves to the south-east.

---

[11]If r is not strictly increasing or discountinuous, r^{-1} is to be understood as the generalized inverse defined by $r^{-1}(\gamma_i) = \inf \{\gamma_{-i} | r(\gamma_{-i}) \geq \gamma_i\}$.

Definition 3 $\psi : \Gamma \to \mathbb{R}$ *is called increasing towards the south-east, or SE-increasing, if*

(i) for all γ_2 it holds: $\gamma_1 \geq \widetilde{\gamma}_1 \Rightarrow \psi(\gamma_1, \gamma_2) \geq \psi(\widetilde{\gamma}_1, \gamma_2)$,

(ii) for all γ_1 it holds: $\gamma_2 \geq \widetilde{\gamma}_2 \Rightarrow \psi(\gamma_1, \gamma_2) \leq \psi(\gamma_1, \widetilde{\gamma}_2)$.

Proposition 6 *A strategy η_i is a threshold strategy if, and only if, there is a continuous SE-increasing function $\psi \in S$ such that*

$$\eta_i(\gamma_1, \gamma_2) = \begin{cases} 1 & if \quad \psi(\gamma_i, \gamma_{-i}) > 0 \\ \\ 0 & if \quad \psi(\gamma_i, \gamma_{-i}) \leq 0. \end{cases} \tag{23}$$

The proof is in the appendix. A threshold strategy divides the state 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 4 *Let η_i be a threshold strategy. Define by*

$$N(\eta_i) = \{(\gamma_1, \gamma_2) \mid \eta_i(\gamma_1, \gamma_2) = 0\} \tag{24}$$

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

The no-effort set measures how aggressive a strategy is. The smaller a player's no-effort set, the more aggressive this player.

With these definitions, there are four equivalent ways of describing a threshold strategy: η, r, ψ, N. In the sequel, we shall use the notion which is most convenient for the problem at hand.

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. Then we show that best responses against threshold strategies have some monotonicity properties, and that the best response correspondence is continuous. Monotonicity says that a player responds to more aggressive strategies with more defensive strategies. We then consider the sequence of mutual best responses which is induced when one player starts with the most

---

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

aggressive strategy $\eta \equiv 1$. Monotonicity will imply that this sequence is monotone and therefore converges. By construction, and by continuity of the best response correspondence, the limit will then be an equilibrium in threshold strategies.

Our first result is that the value function against a threshold strategy is SE-increasing. Because it is continuous, it induces a threshold strategy. It turns out that this threshold strategy is a best response. We have the following proposition.

Proposition 7 *Let η_2 be a threshold strategy. Then it holds:*

(i) V_1 is SE-increasing.

(ii) The threshold strategy induced by V_1 is a best response for player 1 against η_2, that is,

$$
\eta_1^* (\gamma) = \begin{cases} 1 & if \quad V_1 (\gamma) > 0 \\ \\ 0 & if \quad V_1 (\gamma) = 0. \end{cases} \tag{25}
$$

The intuition is as follows. Because player 2 plays a threshold strategy, player 2 does not exert effort, if γ_2 is small and γ_1 is large. Player 1 should then exert effort, because this guarantees him a success at relatively small cost. Indeed, since updating stops if player 2 ceases to spend effort, player 1 obtains the highest possible value $(1 - c) / (1 - \delta)$ in this case. As soon as player 2 starts to exert effort, player 1's utility from exerting effort consists of the current value $\pi - c$ and the future value from learning. As γ_1 tends to 0, and γ_2 tends to 1, the current value monotonically declines and becomes eventually negative, and the future value monotonically declines to 0. Hence, there is threshold beyond which utility becomes negative. Thus, beyond this threshold, player 1 optimally stops exerting effort, and his value is $V_1 = 0$.

To show that an equilibrium in threshold strategies exists, we shall consider a particular sequence of mutual best responses. To show that the sequence converges, we need the following monotonicity result.

Proposition 8 *Let $\eta_2, \widetilde{\eta}_2$ be threshold strategies. Then it holds:*

(i) $\eta_1 = BR_1 (\eta_2) \implies N (\eta_1) \cap N (\eta_2) = \emptyset$.

(ii) $\eta_2 \geq \widetilde{\eta}_2 \implies BR_1 (\eta_2) \leq BR_1 (\widetilde{\eta}_2)$.[13]

---

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

Here, BR_i denotes player i's best response correspondence on the set of threshold strategies of player $-i$. Property (i) says that, whenever player 2 chooses no effort, player 1 optimally exerts effort. This follows by the same argument as in the previous paragraph.

Property (ii) says that a player responds to more aggressive strategies by more defensive ones. Intuitively, the logic is as follows. Suppose, player 2 exerts effort. If player 1 exerts effort, the process continues, and there are essentially two events: in the favourable event, the process reaches player 2's no-effort set in finite time from which on player 1 will win forever. In the unfavourable event, player 2's no-effort set is not reached in finite time in which case player 1 gets less than what he would, had player 2's no-effort set been reached. Now, player 2 becoming less aggressive means that his no-effort set becomes larger. But this makes the favourable event more likely which raises player 1's utility from exerting effort and thereby the incentives to exert effort.

We now come to convergence. We shall use the notion of pointwise (statewise) convergence. A sequence of threshold strategies $(\eta^n)_n$ converges *pointwisely* to a strategy η if the sequence $(\eta^n(\gamma))_n$ converges to $\eta(\gamma)$ for all $\gamma \in \Gamma$. This implies that the corresponding sequence of thresholds (r^n) converges pointwisely. Accordingly, there is a function $r : [0,1] \to [0,1]$ such that

$$r^n(\gamma) \underset{n\to\infty}{\to} r(\gamma) \quad \text{for all } \gamma \in [0,1]. \tag{26}$$

It is easily seen that r is increasing. Thus, r gives rise to a threshold strategy $\widetilde{\eta}$. But it is straighforward to show that η coincides with $\widetilde{\eta}$. Thus, η is a threshold strategy, and the set of threshold strategies is complete with respect to pointwise convergence.

We now define the following sequence $\eta = (\eta_1^n, \eta_2^n)_{n=1,2,\dots}$ of mutual best responses by

$$\eta_2^1 \equiv 1, \tag{27}$$

$$\eta_1^n = BR_1(\eta_2^n) \quad \text{for } n = 1, 2, \dots, \tag{28}$$

$$\eta_2^n = BR_2(\eta_1^{n-1}) \quad \text{for } n = 2, 3, \dots . \tag{29}$$

The sequence thus defined is monotone and converges:

Proposition 9 *The sequence η is increasing in the first, and decreasing in the second argument, that is,*

17

(i) $\eta_1^n \geq \eta_1^{n+1}$ *for* $n = 1, 2, ...$

(ii) $\eta_2^n \leq \eta_2^{n+1}$ *for* $n = 1, 2, ...$

Moreover, η converges to a limit of threshold strategies (η_1^*, η_2^*).

Results (i) and (ii), which follow by induction, are a consequence of Proposition 8. As a bounded, monotone sequence, η converges, and the limit is a pair of threshold strategies because the space of threshold strategies is complete. A proof is in the appendix.

To show that the limit (η_1^*, η_2^*) is an equilibrium, we need that BR is continuous. This is shown in the next proposition. The proof, which is technical, is in the appendix.

Proposition 10 *Let $\left(\eta_{-i}^n\right)_n$ be a sequence of threshold strategies with $\eta_{-i}^n \underset{n\to\infty}{\to} \eta_{-i}$. Then*

$$BR_i\left(\eta_{-i}^n\right) \underset{n\to\infty}{\to} BR_i\left(\eta_{-i}\right). \tag{30}$$

We are now in the position to state that an equilibrium in threshold strategies exists.

Theorem 1 (η_1^*, η_2^*) *is a MPBE, and it holds:*

(i) There is a $\rho > 0$ such that

$$\{\gamma \in \Gamma \mid \|\gamma - (1,0)\| < \rho\} \subset N_2^*, \tag{31}$$

$$\{\gamma \in \Gamma \mid \|\gamma - (0,1)\| < \rho\} \subset N_1^*. \tag{32}$$

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

That (η_1^*, η_2^*) is an equilibrium, follows by construction of η as a sequence of mutual best responses and by continuity of BR. Properties (i) and (ii) are general properties of any equilibrium in threshold strategies. Property (i), which is a consequence of continuity of V_1, says that player i's no-effort contains a small neighbourhood around the 'worst' state where $\gamma_i = 0$ and $\gamma_{-i} = 1$. Effectively, this implies that a player's no-effort set is reached in finite time if this player loses sufficiently often. Property (ii) says that at least one player exerts effort in equilibrium. This is direct consequence of Proposition 8 (i). A typical equilibrium may thus look as follows.[14]

---

[14]We have not looked at whether the equilibrium is unique or symmetric, because this is not needed to make our main point.

Learning

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

$$T(\omega) = \min\left\{t \geq 0 \mid \gamma^t(\omega) \in N_1^* \cup N_2^*\right\} \tag{33}$$

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. Also, by Theorem 1 (ii), one player will win and the other player will lose forever from that point on. Intuitively, $N_1^* \cup N_2^*$ will be reached in finite time, because it contains small neighbourhoods around the most discrepant beliefs. For, one player will experience a very long sequence of consecutive successes, thus, beliefs increasingly diverge, and the other player's no-effort set is reached. More precisely, we have the following result.

Theorem 2 *For all scenarios (α_1, α_2), learning will be incomplete, and with positive probability one player will be discouraged in the long run, that is,*

$$P[T < \infty \mid \eta_1^*, \eta_2^*; (\alpha_1, \alpha_2)] > 0. \tag{34}$$

Proof: The proof is essentially the same as in Proposition 1. Due to Bayesian updating, it requires a finite number of consecutive successes for one player for the state to move arbitrarily closely to either $(1, 0)$ or $(0, 1)$. Therefore, because of property (i) of Theorem 1, it takes a finite number of successes of player i for the state to reach N_{-i}^*. But conditional on (α_1, α_2), sequences are independent. Therefore, with positive probability in all scenarios (α_1, α_2), one player will have a sufficiently long sequence of successes such that the other player's no effort set is reached. In this case, because of property (ii) of Theorem 1, the more confident player exerts effort whereas the other player resigns. □

As a consequence, persistent inequality arises in the long run. One player may eventually win the contest forever and become rich whereas the other player will be discouraged and stay poor. In particular, for example in scenario (α_H, α_L), where player 1 is the better and player 2 the worse player, player 2 has a positive chance to be the long term winner. In this sense, the 'wrong' player may be prevailing in the long run.

To see the difference between a contest and a single decision maker problem, consider two individuals who, independently from each other, have to perform an infinite sequence of tasks with success probability αe at cost ce. It is known from the literature on bandit problems that

in this case *both* individuals may eventually stop exerting effort. This is because both players may fail and thus be discouraged at the same time. In contrast, in a contest one player getting less confident implies that the other player gets more confident.

In our specific contest, there is an additional element of strategic interaction in that a player's winning chance does not depend on ability if the other player exerts no effort. In a Markov equilibrium, this implies that the belief process stops if one player stops exerting effort. If, instead, a player's winning chance did always depend on ability, then even a highly confident and highly able player would at some point experience a long streak of failures and lose confidence. At the same time, the other player would be encouraged. We therefore conjecture that at least in the 'medium term' we would observe swings between long winning sequences of either player. Whether and under what conditions this would lead to learning or to cyclical behaviour in the long run, is an open question.

Confidence

We consider now the impact of confidence on the likelihood to be the long-run winner. For this, define by

$$T_i\left(\omega\right) = \min\left\{t \geq 0 \,\middle|\, \gamma^t\left(\omega\right) \in N_i^*\right\} \tag{35}$$

the first entry time into N_i^*. The following result says that the more confident a player and the less confident his rival, the higher the likelihood to reach the rival's no-effort set.

Theorem 3 *For all scenarios (α_1, α_2), the likelihood to reach the rival player's no effort set in finite time is increasing in own and decreasing in the rival's confidence, that is,*

$$P\left[T_{-i} < T_i \,\middle|\, \eta_1^*, \eta_2^*; (\alpha_1, \alpha_2)\right] \text{ is increasing in } \gamma_i \text{ and decreasing in } \gamma_{-i}. \tag{36}$$

Proof: The proof is straightforward. The higher γ_i, and the smaller γ_{-i}, the closer the state is at N_{-i}^*. Therefore, less successes are required to move into N_{-i}^*. $\square$

Furthermore, because V_i is SE-increasing, player i's expected utility is increasing in γ_i and decreasing in γ_{-i}. In this sense, confidence is self-serving. Admittedly, given complementarity, this result is not too surprising. Nevertheless, it illustrates that perception might be an important factor in contests, and that long term success may not necessarily be the result of ability but rather of confidence in combination with an initial streak of good luck.

5   Conclusion

The paper discusses the role of confidence in a repeated contest when players do not know their intrinsic abilities. The basic idea is that complementarity of ability and effort gives rise to a belief reinforcment effect which encourages confident players and discourages less confident players. In our example, this implis that players may fail to learn their true abilities in the long run, and one player may eventually win forever. As a consequence, persistent inequality arises, and the actually worse player may be the long run winner. Furthermore, confidence is self-serving in that it increases a player's utility and the likelihood to be the eventually prevailing player.

The specific form of our contest somewhat limits the generality of our results. However, as already discussed in the previous section, some qualitative characteristics of the belief rein-forcement mechanism can be expected to carry over to more general contests.

More importantly, we have neglected the fact that high confidence often leads to excessive risk taking, countervailing the self-serving effect of confidence (see Baumeister (1998) for evidence and Benabou/Tirole (2002a, 2002b) for a model). Recent contest literature suggests that at least in some contests risk taking may be beneficial (Hvide (2000), Kräkel (2002)). To what extent this would strengthen the self-serving effect of confidence is a question for future research.

Finally, we have also set aside the possibility that players may accumulate wealth over time. This should lead to an even more pronounced reinforcement effect, as the wealthier player can invest more resources in the contest, thus, increasing his chance to succeed, and, in turn, to become even wealthier.

Appendix

Proof of Proposition 4: We use Blackwell's sufficiency condition, i.e., we have to show:

(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 $F(\psi + \xi) \leq F\psi + \beta\xi$.

Ad (i): Let $\psi, \phi \in S$ with $\psi \geq \phi$ on Γ. Let $\left(e_1^\psi, e_1^\phi\right) \in \{0,1\}$ be the maximizers of $F\psi$ and $F\phi$, respectively. There are four possible cases: $\left(e_1^\psi, e_1^\phi\right) = (0,0)$, $\left(e_1^\psi, e_1^\phi\right) = (1,1)$, $\left(e_1^\psi, e_1^\phi\right) =$

$(1,0)$, $\left(e_1^\psi, e_1^\phi\right) = (0,1)$. For $e_1^\psi = e_1^\phi$ the claim follows immediately. If $\left(e_1^\psi, e_1^\phi\right) = (1,0)$, then[15]

$$F\psi = \pi\left(1, \eta_2\right) - c + \delta\left[\pi\left(1, \eta_2\right)\psi^\sigma + \left(1 - \pi\left(1, \eta_2\right)\right)\psi^\varphi\right] \tag{37}$$

$$\geq \pi\left(0, \eta_2\right) + \delta\left[\pi\left(0, \eta_2\right)\psi^\sigma + \left(1 - \pi\left(0, \eta_2\right)\right)\psi^\varphi\right] \tag{38}$$

$$\geq \pi\left(0, \eta_2\right) + \delta\left[\pi\left(0, \eta_2\right)\phi^\sigma + \left(1 - \pi\left(0, \eta_2\right)\right)\phi^\varphi\right] \tag{39}$$

$$= F\phi. \tag{40}$$

The first inequality follows by definition of the maximum, and the second inequality follows by the assumption that $\psi \geq \phi$. The case $\left(e_1^\psi, e_1^\phi\right) = (0,1)$ is shown in the same way.

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

$$F\left(\psi + \xi\right) = \max_{e_1 \in \{0,1\}} \left\{\pi\left(e_1, \eta_2\right) - ce_1 + \delta\left[\pi\left(e_1, \eta_2\right)\psi^\sigma + \left(1 - \pi\left(e_1, \eta_2\right)\right)\psi^\varphi\right] + \tag{41}$$

$$+ \delta\left[\pi\left(e_1, \eta_2\right)\xi + \left(1 - \pi\left(e_1, \eta_2\right)\right)\xi\right] \tag{42}$$

$$= F\psi + \beta\xi. \tag{43}$$

This shows that F is a contraction. $\square$

Proof of Proposition 6: Let ψ be given. Define $r\left(\gamma_{-i}\right) = \sup\left\{r \in [0,1] \mid \psi\left(r, \gamma_{-i}\right) \leq 0\right\}$. Since ψ is SE-increasing, r is increasing. Since ψ is continuous, the supremum is assumed and $\psi\left(r\left(\gamma_{-i}^+\right), \gamma_{-i}\right) \leq 0$. Therefore, if $\gamma_i > r\left(\gamma_{-i}^+\right)$, then $\psi\left(r\left(\gamma_{-i}^+\right), \gamma_{-i}\right) > 0$, hence $\eta\left(\gamma_i, \gamma_{-i}\right) = 1$.

For the converse, let r be given. W.l.o.g., let r be not constantly equal to 1 or 0 (in this case, define ψ constantly equal to 1 or -1). Define $\psi\left(\gamma_i, \gamma_{-i}\right) = 0$ for $\gamma_i \leq r\left(\gamma_{-i}^+\right)$, and define $\psi(1, 0) = 1$. For all other $\left(\gamma_i, \gamma_{-i}\right)$ define ψ as follows. Connect each point in $\left\{\left(\gamma_i, \gamma_{-i}\right) \mid \gamma_i \in \left[r\left(\gamma_{-i}^-\right), r\left(\gamma_{-i}^+\right)\right]\right\}$ and the point $\psi(1, 0)$ by a straight line and let ψ be the resulting function. ($r\left(\gamma_{-i}^-\right)$ is the left-hand limit.) ψ is obviously continuous, and since r is increasing, ψ is SE-increasing. Finally, by construction, $\psi\left(\gamma_i, \gamma_{-i}\right) > 0$ if, and only if, $\gamma_i > r\left(\gamma_{-i}^+\right)$. $\square$

Proof of Proposition 7: To show (i), we need some preliminary results. It is convenient to work with the following definition. Let η_2 be a threshold strategy. Denote by $S^{+,c}$ the set of all functions in S that are SE-increasing and constant on N_2.

---

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

Notice that $S^{+,c}$ is complete.[16] The following proposition states that F maps $S^{+,c}$ in $S^{+,c}$.

Proposition 11 *Let η_2 be a threshold strategy, and let $\psi \in S^{+,c}$. Then $F\psi \in S^{+,c}$.*

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 of such a sequence is a member of $S^{+,c}$ with respect to the supremum norm, in particular, it is SE-increasing. This shows (i).

Proof of Proposition 11: We show first that $F\psi$ is SE-increasing. Let $\gamma_1 \geq \widetilde{\gamma}_1$. We have to show that $F\psi(\gamma_1, \gamma_2) \geq F\psi(\widetilde{\gamma}_1, \gamma_2)$ for all γ_2. Let $e_1, \tilde{e}_1 \in \{0, 1\}$ be the maximizers of $F\psi(\gamma_1, \gamma_2)$ and $F\psi(\widetilde{\gamma}_1, \gamma_2)$, respectively. Suppose first—case I)—that $e_1 = 1$ and $\tilde{e}_1 = 1$. Then

$$F\psi(\gamma_1, \gamma_2) = \pi(1, \eta_2(\gamma)) - c + \delta\left[\pi(1, \eta_2(\gamma))\psi(\gamma^\sigma) + (1 - \pi(1, \eta_2(\gamma)))\psi(\gamma^\varphi)\right], \qquad (44)$$

and

$$F\psi(\widetilde{\gamma}_1, \gamma_2) = \pi(1, \eta_2(\widetilde{\gamma})) - c + \delta\left[\pi(1, \eta_2(\widetilde{\gamma}))\psi(\widetilde{\gamma}^\sigma) + (1 - \pi(1, \eta_2(\widetilde{\gamma})))\psi(\widetilde{\gamma}^\varphi)\right]. \qquad (45)$$

Since η_2 is a threshold strategy, $\eta_2(\gamma_1, \gamma_2) \leq \eta_2(\widetilde{\gamma}_1, \gamma_2)$. Hence, there are three possible cases: A): $\eta_2(\gamma_1, \gamma_2) = \eta_2(\widetilde{\gamma}_1, \gamma_2) = 1$, B): $\eta_2(\gamma_1, \gamma_2) = \eta_2(\widetilde{\gamma}_1, \gamma_2) = 0$, and C): $\eta_2(\gamma_1, \gamma_2) = 0$ and $\eta_2(\widetilde{\gamma}_1, \gamma_2) = 1$. Consider first case A). In this case,

$$\pi(1, \eta_2(\gamma)) = 1/2 + (\gamma_1 - \gamma_2)\Delta\alpha \geq 1/2 + (\widetilde{\gamma}_1 - \gamma_2)\Delta\alpha = \pi(1, \eta_2(\widetilde{\gamma})). \qquad (46)$$

Furthermore, monotonicity of Bayes' rule implies that $\gamma_1^\sigma \geq \widetilde{\gamma}_1^\sigma$ and $\gamma_1^\varphi \leq \widetilde{\gamma}_1^\varphi$, thus, because ψ is SE-increasing, $\psi(\gamma^\sigma) \geq \psi(\widetilde{\gamma}^\sigma)$ and $\psi(\gamma^\varphi) \geq \psi(\widetilde{\gamma}^\varphi)$. This implies that $F\psi(\gamma_1, \gamma_2) \geq F\psi(\widetilde{\gamma}_1, \gamma_2)$.

Consider next case B): $\eta_2(\gamma_1, \gamma_2) = \eta_2(\widetilde{\gamma}_1, \gamma_2) = 0$. Then, because beliefs are not updated if no effort is chosen,

$$F\psi(\gamma_1, \gamma_2) = 1 - c + \delta\psi(\gamma), \qquad (47)$$

and

$$F\psi(\widetilde{\gamma}_1, \gamma_2) = 1 - c + \delta\psi(\widetilde{\gamma}). \qquad (48)$$

Hence, since ψ is SE-increasing, $F\psi(\gamma_1, \gamma_2) \geq F\psi(\widetilde{\gamma}_1, \gamma_2)$.

---

[16]This is because the SE-increasing concept involves only *weak* inequalities.

Consider finally case C): $\eta_2(\gamma_1, \gamma_2) = 0$ and $\eta_2(\widetilde{\gamma}_1, \gamma_2) = 1$. Then

$$F\psi(\gamma_1, \gamma_2) = 1 - c + \delta\psi(\gamma), \tag{49}$$

and

$$F\psi(\widetilde{\gamma}_1, \gamma_2) = \pi(1, \eta_2(\widetilde{\gamma})) - c + \delta\left[\pi(1, \eta_2(\widetilde{\gamma}))\psi(\widetilde{\gamma}^\sigma) + (1 - \pi(1, \eta_2(\widetilde{\gamma})))\psi(\widetilde{\gamma}^\varphi)\right]. \tag{50}$$

Notice first that $1 \geq \pi(1, \eta_2(\widetilde{\gamma}))$. Moreover, $\psi(\gamma) \geq \psi(\widetilde{\gamma}^\varphi)$, because ψ is SE-increasing. Furthermore, ψ is SE-increasing and constant on N_2. Hence, ψ is maximal on N_2. But, by assumption $\eta_2(\gamma_1, \gamma_2) = 0$, thus $(\gamma_1, \gamma_2) \in N_2$. Hence, $\psi(\gamma) \geq \psi(\widetilde{\gamma}^\sigma)$. All three arguments together imply that $T\psi(\gamma_1, \gamma_2) \geq T\psi(\widetilde{\gamma}_1, \gamma_2)$.

This shows the claim for case I). Cases II): $e_1 = 0, \tilde{e}_1 = 0$, III): $e_1 = 1, \tilde{e}_1 = 0$, and IV): $e_1 = 0, \tilde{e}_1 = 1$ follow analogously.

Likewise, it follows that $F\psi(\gamma_1, \widetilde{\gamma}_2) \geq F\psi(\gamma_1, \gamma_2)$ for $\gamma_2 \geq \widetilde{\gamma}_2$ and for all γ_1. This shows that $F\psi$ is SE-increasing.

It remains to show that $F\psi$ is constant on N_2. Let $\gamma \in N_2$. Hence, whatever player 1 chooses, there will be no updating, since player 2 chooses $e_2 = 0$. Therefore, $e_1 = 0$ yields $1/2 + \delta\psi(\gamma)$, and $e_1 = 1$ yields $1 - c + \delta\psi(\gamma)$. Since, by assumption, $1 - c > 1/2$, $e_1 = 1$ is a maximizer of $F\psi(\gamma)$. Thus,

$$F\psi(\gamma) = 1 - c + \delta\psi(\gamma) \quad \text{for all } \gamma \in N_2. \tag{51}$$

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

It remains to show part (ii) of Proposition 7. For this, let η_2 be given. Consider a best response of player 1 against η_2. It may be that in some states player 1 is indifferent between $e_1 = 1$ and $e_1 = 0$. To break ties we assume that in these cases player 1 chooses $e_1 = 0$. The best response gets thus single-valued. Let η_1 be such a statewise single-valued best response against η_2. We show

(a) If $\eta_1(\gamma) = 0$, then $V_1(\gamma) = 0$

(b) If $\eta_1(\gamma) = 1$, then $V_1(\gamma) > 0$.

In other words: $\eta_1(\gamma) = 1$ if, and only if, $V_1(\gamma) > 0$.

Ad (a): Let $\eta_1(\gamma) = 0$. Suppose, $\eta_2(\gamma) = 0$. Compare player 1's value of playing $e_1 = 1$ and of playing $e_1 = 0$, conditional on optimal continuation. Since $\eta_2(\gamma) = 0$, there is no learning

and the future value $\delta V_1(\gamma)$ is the same for both $e_1 = 1$ and $e_1 = 0$. The current value of $e_1 = 1$ is $1 - c$, whereas the current value of $e_1 = 0$ is $1/2 < 1 - c$. Thus $\eta_1(\gamma) = 1$, a contradiction. Therefore, $\eta_2(\gamma) = 1$. In this case, since $\eta_1(\gamma) = 0$, there is again no updating, and player 1's value is given by

$$V_1(\gamma) = 0 + \delta V_1(\gamma). \tag{52}$$

This implies that $V_1(\gamma) = 0$.

Ad (b): Let $\eta_1(\gamma) = 1$. Suppose, $\eta_2(\gamma) = 0$. Then, player 1 wins with certainty and again there is no updating. Player 1's value is thus given by

$$V_1(\gamma) = 1 - c + \delta V_1(\gamma). \tag{53}$$

Hence $V_1(\gamma) = (1 - c) / (1 - \delta) > 0$.

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 0. Our tie-breaking rule implies that player 1 strictly prefers $e_1 = 1$ to $e_1 = 0$, if $\eta_1(\gamma) = 1$. Thus, if $\eta_1(\gamma) = 1$, it must be that $V_1(\gamma) > 0$. $\square$

Proof of Proposition 8: Ad (i): Suppose, there is a state in which player 1 chooses $e_1 = 0$ against $e_2 = 0$. Then he gets 0. Yet, if player 1 deviates and chooses $e_1 = 1$ instead, he gets $(1 - c) / (1 - \delta) > 0$. This implies the claim.

Ad (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 player 1. By Proposition 7, $\eta_1(\gamma) = 1$ if, and only if, $V_1(\gamma) > 0$. Hence, it is enough to show that

$$\widetilde{V}_1(\gamma) \geq V_1(\gamma) \quad \text{for all } \gamma \in \Gamma. \tag{54}$$

We shall show that η_1 gives a higher utility against $\widetilde{\eta}_2$, than η_1 gives against η_2. That is,

$$U_1(\eta_1, \widetilde{\eta}_2) \geq U_1(\eta_1, \eta_2). \tag{55}$$

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 (55), 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. Likewise, let $(\widetilde{\gamma}^t)_{t=1,2,...}$ the beliefs induced

under $\eta_1, \widetilde{\eta}_2$, and initial state γ^0. We divide now $N\left(\widetilde{\eta}_2\right)$ into several sets. Define

$$A = N\left(\eta_1\right) \cap N\left(\widetilde{\eta}_2\right), \tag{56}$$

$$B = N\left(\widetilde{\eta}_2\right) \backslash N\left(\eta_2\right), \tag{57}$$

$$C = N\left(\widetilde{\eta}_2\right) \backslash \left(A \cup B\right). \tag{58}$$

The following picture illustrates these sets.

Notice that C is well defined because, by assumption, $\eta_2 \geq \widetilde{\eta}_2$.

Clearly, $N\left(\widetilde{\eta}_2\right) = A \cup B \cup C$. Hence, if player 2 plays $\widetilde{\eta}_2$, once the state has reached A, or B, or C, the process stops, and players receive a constant per period payoff from then on. In particular, suppose, $\left(\eta_1, \widetilde{\eta}_2\right)$ is played. Then, if the state reaches A, both players exert no effort from then on, and player 1 gets per period payoff $1/2$. If the state reaches B or C, then $e_1 = 1$ and $e_2 = 0$, and player 1 obtains per period payoff $(1 - c)$ from then on. Suppose instead, $\left(\eta_1, \eta_2\right)$ is played. Then player 1 receives the same as under $\left(\eta_1, \widetilde{\eta}_2\right)$ when the state reaches B. But, he receives less than under $\left(\eta_1, \widetilde{\eta}_2\right)$ when the state reaches A or C, because under η_2, player 2 chooses effort $\widetilde{e}_2 = 1$ on C.

This argument suggests that $\left(\eta_1, \eta_2\right)$ is better for player 1 than $\left(\eta_1, \widetilde{\eta}_2\right)$. To make it more precise, we write the overall utility of player 1 in terms of first entry times in these sets. Define for $j = A, C$

$$T_j\left(\omega\right) = \min\left\{t \geq 0 \,\middle|\, \gamma^t\left(\omega\right) \in j\right\}. \tag{59}$$

Furthermore, define by

$$\left\{T_A < T_C\right\} = \left\{\omega \,\middle|\, T_A\left(\omega\right) < \infty, T_A\left(\omega\right) < T_C\left(\omega\right)\right\} \tag{60}$$

the event that the state reaches A in finite time and that A is reached before B.[17] Define

---

[17]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 $\left\{T_A < T_B\right\}$ is measurable.

likewise $\{T_C < T_A\}$. Then, with $\eta_i^t = \eta_i(\gamma^t)$, the utility of player 1 is

$$U_1(\eta_1, \eta_2) = \int \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \eta_2, \gamma) \tag{61}$$

$$= \int_{\{T_A < T_C\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \eta_2, \gamma) \tag{62}$$

$$+ \int_{\{T_C < T_A\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \eta_2, \gamma) \tag{63}$$

$$+ \int_{\{T_A = T_C = \infty\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \eta_2, \gamma) \tag{64}$$

We can split the sum in (62) into dates before and after A is reached, that is, the integral in (62) is equal to

$$\int_{\{T_A < T_C\}} \sum_{t=1}^{T_A - 1} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] + \delta^{T_A(\omega)} V_1\left(\gamma^{T_A(\omega)}\right) dP(\omega; \eta_1, \eta_2, \gamma). \tag{65}$$

For (63) 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 C is not reached. Hence, before A or C is reached both the transition probabilities and the current payoff are the same under (η_1, η_2) and under $(\eta_1, \widetilde{\eta}_2)$. Therefore,

$$\int_{\{T_A < T_C\}} \sum_{t=0}^{T_A - 1} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \eta_2, \gamma) \tag{66}$$

$$+ \int_{\{T_C < T_A\}} \sum_{t=0}^{T_C - 1} \delta^t \left[\pi\left(\eta_1^t, \eta_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \eta_2, \gamma) \tag{67}$$

$$= \int_{\{T_A < T_C\}} \sum_{t=0}^{T_A - 1} \delta^t \left[\pi\left(\eta_1^t, \widetilde{\eta}_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \widetilde{\eta}_2, \gamma) \tag{68}$$

$$+ \int_{\{T_C < T_A\}} \sum_{t=0}^{T_C - 1} \delta^t \left[\pi\left(\eta_1^t, \widetilde{\eta}_2^t; \gamma^t\right) - c\eta_1^t \right] dP(\omega; \eta_1, \widetilde{\eta}_2, \gamma). \tag{69}$$

By the same argument, since (η_1, η_2) and $(\eta_1, \widetilde{\eta}_2)$ coincide on the set $\{T_A = T_C = \infty\}$, the integral in (64) is the same under (η_1, η_2) and $(\eta_1, \widetilde{\eta}_2)$.

Furthermore, when A is reached, player 1 stops exerting effort while player 2 does not under (η_1, η_2), thus

$$V_1\left(\gamma^{T_A(\omega)}\right) = 0. \tag{70}$$

However, when A is reached under $(\eta_1, \widetilde{\eta}_2)$, both players stop exerting effort, thus

$$U_1(\eta_1, \widetilde{\eta}_2)\left(\gamma^{T_A(\omega)}\right) = \frac{1}{2(1-\delta)}. \tag{71}$$

27

When C is reached under $(\eta_1, \widetilde{\eta}_2)$, then only player 2 stops exerting effort, thus

$$U_1 (\eta_1, \widetilde{\eta}_2) \left(\gamma^{T_C(\omega)} \right) = \frac{1-c}{1-\delta} \geq V_1 \left(\gamma^{T_C(\omega)} \right), \tag{72}$$

where the inequality holds because V_1 is bounded by $(1-c)/(1-\delta)$. Hence, we can estimate $U_1 (\eta_1, \eta_2)$ against

$$\int_{\{T_A < T_C\}} \sum_{t=0}^{T_A - 1} \delta^t \left[\pi \left(\eta_1^t, \widetilde{\eta}_2^t; \gamma^t \right) - c\eta_1^t \right] + \delta^{T_A(\omega)} U_1 (\eta_1, \widetilde{\eta}_2) \left(\gamma^{T_A(\omega)} \right) dP (\omega \, ; \eta_1, \widetilde{\eta}_2, \gamma) \tag{73}$$

$$+ \int_{\{T_C < T_A\}} \sum_{t=0}^{T_C - 1} \delta^t \left[\pi \left(\eta_1^t, \widetilde{\eta}_2^t; \gamma^t \right) - c\eta_1^t \right] + \delta^{T_C(\omega)} U_1 (\eta_1, \widetilde{\eta}_2) \left(\gamma^{T_C(\omega)} \right) dP (\omega \, ; \eta_1, \widetilde{\eta}_2, \gamma) \tag{74}$$

$$+ \int_{\{T_A = T_C = \infty\}} \sum_{t=0}^{\infty} \delta^t \left[\pi \left(\eta_1^t, \widetilde{\eta}_2^t; \gamma^t \right) - c\eta_1^t \right] dP (\omega \, ; \eta_1, \widetilde{\eta}_2, \gamma). \tag{75}$$

But this sum is just equal to $U_1 (\eta_1, \widetilde{\eta}_2)$. Thus, (55) is shown. $\square$

Proof of Proposition 9: Ad (i) and (ii): The proof is by induction. Let $n = 1$. Clearly $\eta_2^1 \equiv 1 \geq \eta_2^2$, hence (ii). Thus, by Proposition 8 (ii), $\eta_1^1 = BR_1 (\eta_2^1) \leq BR_1 (\eta_2^2) = \eta_1^2$, hence (i). Let (i) and (ii) hold for $i = 1, ..., n$. Then, by (ii) and Proposition 8 (ii), $\eta_1^n = BR_1 (\eta_2^n) \leq BR_1 (\eta_2^{n+1}) = \eta_1^{n+1}$, thus, (i) holds for $n + 1$. Likewise, $\eta_2^{n+1} = BR_2 (\eta_1^{n-1}) \geq BR_2 (\eta_1^n) = \eta_2^n$, thus, (ii) holds for $n + 1$.

Convergence: Let r_i^n be the threshold from Definition 2 associated with η_i^n. By the remark above, it is enough to show that the sequence $(r_i^n)_n$ converges pointwisely. Consider player 1. $\eta_1^n \leq \eta_1^{n+1}$ on Γ translates into $r_1^n \geq r_1^{n+1}$ on $[0, 1]$. Thus, for all γ_2, the sequence $(r_1^n (\gamma_2))_n$ is decreasing and bounded from below, hence, convergent. But, again by the remark above, the limit is a threshold strategy. $\square$

Proof of Proposition 10: We only show that BR_1 is left-continuous, that is $\lim_{\eta_2^n \Uparrow \eta_2} BR_i (\eta_2^n) = BR_i (\eta_2)$. Right-continuity can be shown in the same way. Notice first that convergence of a sequence $(\eta^n)_n$ of threshold strategies is implied by pointwise convergence of the corresponding SE-increasing functions $(\psi^n)_n$. If $(\psi^n)_n$ converges pointwisely to a *continuous* function ψ, then ψ is SE-increasing, and the corresponding sequence $(\eta^n)_n$ of threshold strategies converges to the threshold strategy induced by ψ.

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. By the remark above and by Proposition 7, it is enough to show that V_1^n converges

statewisely to V_1. Since, by assumption, $\eta_2^n \leq \eta_2^{n+1} \leq \eta_2$, it follows from the proof of Proposition 8 that $V_1^n \geq V_1^{n+1} \geq V_1$. Thus, since V_1^n is bounded from below, the (statewise) limit exists and $\lim_n V_1^n \geq V_1$. We shall show that

$$U_1\left(\eta_1^n, \eta_2^n\right) - U_1\left(\eta_1^n, \eta_2\right) \to 0. \tag{76}$$

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

$$\delta^\tau \frac{1-c}{1-\delta} < \varepsilon. \tag{77}$$

Define by

$$\Gamma^\tau = \left\{\gamma \in \Gamma \mid \text{there is an } \omega \text{ such that } \gamma^t\left(\omega\right) = \gamma \text{ for } t = 0, ..., \tau\right\} \tag{78}$$

the set of all possible states 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

$$\xi_\tau = \min_{\gamma \in \Gamma^\tau \backslash N(\eta_2)} dist\left(\gamma, N\left(\eta_2\right)\right), \tag{79}$$

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, $\xi_\tau > 0$.[18]

Let r_2^n and r_2 be the thresholds associated with η_2^n and η_2. Define for each $\gamma \in \Gamma$ the number $n^{\tau,\gamma}$ such that

$$r_2^n\left(\gamma_1\right) - r_2\left(\gamma_1\right) < \xi_\tau/2 \quad \text{for all } n > n^{\tau,\gamma}. \tag{80}$$

Because Γ^τ is finite, it exists

$$n^\tau = \max_{\gamma \in \Gamma^\tau} n^{\tau,\gamma}, \tag{81}$$

and it follows by construction that

$$\left(N\left(\eta_2^{n^\tau}\right) \backslash N\left(\eta_2\right)\right) \cap \Gamma^\tau = \emptyset. \tag{82}$$

---

[18]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, $V_1\left(\gamma^0\right) = (1-c)/(1-\delta)$, and it follows directly that $V_1 \geq V_1^n$.

Define $A^n = N(\eta_2^n) \setminus N(\eta_2)$, and let

$$T_{A^n}(\omega) = \left\{ t \geq 0 \,|\, \gamma^t(\omega) \in A^n \right\} \tag{83}$$

be the first entry time into A^n. Notice, that $A^n \subseteq N\left(\eta_2^{n^\tau}\right) \setminus N(\eta_2)$ for all $n > n^\tau$. This implies that

$$P\left[T_{A^n} \leq \tau\right] = 0 \quad \text{for all } n > n^\tau \tag{84}$$

since, by construction, no state can reach $N\left(\eta_2^{n^\tau}\right) \setminus N(\eta_2)$ before time τ. As an illustration, consider the following picture.

We shall now write U_1 in terms of the first entry time. Let $n > n^\tau$, and define $\eta^{n,t} = \eta^n(\gamma^t)$, then

$$U_1\left(\eta_1^n, \eta_2^n\right) = \int_{\{T_{A^n} < \infty\}} \sum_{t=0}^{T_{A^n}-1} \delta^t \left[\pi\left(\eta_1^{n,t}, \eta_2^{n,t}\right) - c\eta_1^{n,t}\right] \tag{85}$$

$$+ \delta^{T_{A^n}} U_1\left(\eta_1^n, \eta_2^n\right)\left(\gamma^{T_{A^n}}\right) \, dP\left(\omega; \eta_1^n, \eta_2^n, \gamma\right) \tag{86}$$

$$+ \int_{\{T_{A^n} = \infty\}} \sum_{t=0}^{\infty} \delta^t \left[\pi\left(\eta_1^{n,t}, \eta_2^{n,t}\right) - c\eta_1^{n,t}\right] \, dP\left(\omega; \eta_1^n, \eta_2^n, \gamma\right). \tag{87}$$

Now, because of Proposition 8 (i), the profiles (η_1^n, η_2^n) and (η_1^n, η_2) coincide on $\{T_{A^n} = \infty\}$ and as long as A^n is not reached. Hence, they give rise to the same transition probabilities and payoffs on $\{T_{A^n} = \infty\}$ and as long as A^n is not reached. Therefore, the second integral and the sum in the first integral coincide under (η_1^n, η_2^n) and (η_1^n, η_2). Accordingly

$$U_1\left(\eta_1^n, \eta_2^n\right) - U_1\left(\eta_1, \eta_2^n\right) \tag{88}$$

$$= \int_{\{T_{A^n} < \infty\}} \delta^{T_{A^n}} U_1\left(\eta_1^n, \eta_2^n\right)\left(\gamma^{T_{A^n}}\right) \, dP\left(\omega; \eta_1^n, \eta_2^n, \gamma\right) \tag{89}$$

$$- \int_{\{T_{A^n} < \infty\}} \delta^{T_{A^n}} U_1\left(\eta_1, \eta_2^n\right)\left(\gamma^{T_{A^n}}\right) \, dP\left(\omega; \eta_1, \eta_2^n, \gamma\right) \tag{90}$$

Moreover, the process stops under (η_1^n, η_2^n) when A^n is reached since player 2 then stops exerting effort. Player 1 either exerts effort, if $\gamma^{T_{A^n}} \in N(\eta_1^n) \cap A^n$. Or, if $\gamma^{T_{A^n}} \notin N(\eta_1^n) \cap A^n$, he does not exert effort. In either case he obtains utility

$$U_1\left(\eta_1^n, \eta_2^n\right)\left(\gamma^{T_{A^n}}\right) \leq \frac{1-c}{1-\delta}. \tag{91}$$

Furthermore, by definition of a best response, $U_1(\eta_1, \eta_2^n) \leq U_1(\eta_1^n, \eta_2^n)$. Hence,

$$|U_1(\eta_1^n, \eta_2^n) - U_1(\eta_1, \eta_2^n)| = U_1(\eta_1^n, \eta_2^n) - U_1(\eta_1, \eta_2^n) \tag{92}$$

$$\leq \int_{\{T_{A^n} < \infty\}} \delta^{T_{A^n}} \frac{1-c}{1-\delta} dP(\omega; \eta_1^n, \eta_2^n, \gamma) - 0 \tag{93}$$

$$\leq P[T_{A^n} \leq \tau] \frac{1-c}{1-\delta} \tag{94}$$

$$+ \int_{\{\tau < T_{A^n} < \infty\}} \delta^{T_{A^n}} \frac{1-c}{1-\delta} dP$$

$$\leq 0 + \delta^\tau \frac{1-c}{1-\delta} \tag{95}$$

$$< \varepsilon. \tag{96}$$

This shows the claim. $\square$

Proof of Theorem 1: *Equilibrium*: By Propositions 9 and 10 it follows that $\eta_1^* = \lim \eta_1^n = \lim BR_1(\eta_2^n) = BR_1(\lim \eta_2^n) = BR_1(\eta_2^*)$. Likewise, $\eta_2^* = BR_2(\eta_1^*)$. Thus, (η_1^*, η_2^*) is an equilibrium.

Ad (i): To prove the claim, we show that player 1's utility from choosing $e_1 = 1$ against η_2^* in state $(0,1)$ is strictly negative. Then, continuity of U_1 implies that there is a neighbourhood around $(0,1)$ such that $e_1 = 1$ gives strictly negative utility in this neighbourhood. Therefore, in this neighbourhood $V_1 = 0$, and $\eta_1^* = 0$.

Notice first that in state $\gamma = (0,1)$, the state is not updated any more. Thus, it is optimal for player 2 to choose $e_2 = 1$, irrespective of what player 1 does. Indeed, if player 1 chooses $e_1 = 0$, then player 2 gets $(1-c)/(1-\delta)$ from choosing $e_2 = 1$, and from choosing $e_2 = 0$ he gets $1/(2(1-\delta)) < (1-c)/(1-\delta)$. If player 1 chooses $e_1 = 1$, then player 2 gets $(1/2 + \Delta\alpha - c)/(1-\delta)$ from choosing $e_2 = 1$, and from choosing $e_2 = 0$ he gets $0 < (1/2 + \Delta\alpha - c)/(1-\delta)$. Hence, $\eta_2^*(0,1) = 1$. Likewise, $\eta_1^*(1,0) = 1$.

Suppose now that player 1 chooses $e_1 = 1$ against η_2^* in state $(0,1)$. Then player 1 gets $(1/2 + \Delta\alpha - c)/(1-\delta) < 0$. But this is what we wanted to show.

Ad (ii): This is an immediate consequence of Proposition 8 (ii). $\square$

6 References

Bar-Isaac, H. (2000): "Good Advisers Survive Despite Bad Luck—Self-Confidence in a Model of a Privately Informed and Strategic One-Armed Bandit," Mimeo, LSE.

Baumeister, R. (1998): "*The Self*," in The Handbook of Social Psychology, edited by D. Gilber, S. Fiske, and G. Lindsey, Boston, McGraw-Hill.

Benabou, R., J. Tirole (2002a):"Intrinsic and Extrinsic Motivation," Mimeo, Princeton University, http://www.princeton.edu/~rbenabou/.

Benabou, R., J. Tirole (2002b): "Self-Confidence and Personal Motivation," *Quarterly Journal of Economics*, 117, 871 - 915.

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.

Colvin, C. R., Block, J. (1994): "Do Positive Illusions Foster Mental Health? An Examination of the Taylor and Brown Formulation," *Psychological Bulletin*, 116, 3-20.

Compte, O., A. Postlewaite (2001): "Confidence Enhanced Performance," Mimeo, University of Pennsylvania.

Darity, W., A.H. Goldsmith (1996): "Social Psychology, Unemployment and Macroeconomics," *Journal of Economic Perspectives*, 10, 121-140.

De Meza, D., C. Southey (1996): "The Borrower's Curse: Optimism, Finance and Entrepreneurship," *Economic Journal*, 106, 375-386.

Dunifon, R., G. Duncan (1998): "Long-Run Effects of Motivation on Labor-Market Success," *Social Psychology Quarterly*, 61, 33-48.

Gneezy, U., M. Niederle, A. Rustichini (2001): "Performance in Competitive Environments: Gender Differences," Mimeo, www.cepr.org/meets/wkcn/6/683/Papers/rustichini1.pdf

Heifetz, A., Y. Spiegel (2000): "On the Evolutionary Emergence of Optimism," Mimeo, Tel Aviv University.

Hvide, H.K. (2000): "Tournament Rewards and Risk Taking," Mimeo, Norwegian School of Economics and Business.

Keller, G. S. Rady (2002): "Strategic Experimentation: The Case of Poisson Bandits," CESifo Working Paper No. 737.

Kräkel, M. (2002): "Risk Taking in Asymmetric Tournaments," Mimeo, University of Bonn.

Malmendier, U, G. Tate (2002): "CEO Overconfidence and Corporate Investment," Mimeo, Harvard University.

Manove, M. (2000): "Entrepreneurs, Optimism, and the Competitive Edge," Mimeo, Boston University.

Manove, M., J. Padilla (1999): "Banking (conservatively) with Optimists," *RAND Journal of Economics*, 30, 324-350.

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

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.

Taylor, S.E., Brown, J.D. (1988): "Illusion and Well-Being: A Social Psychological Perspective on Mental Health," *Psychological Bulletin*, 103, 193 - 210.

Van den Steen, E. (2002): "Skill or Luck. Biases of Rational Agents," Mimeo, MIT.