DEPARTMENT OF ECONOMICS WORKING PAPER SERIES

Updating Valuations and Stochastic Dynamics on Coordination

Hideki Fujiyama
Department of Economics
Dokkyo University

Working Paper 2007-06
http://www.bus.lsu.edu/economics/papers/pap07_06.pdf

Department of Economics
Louisiana State University
Baton Rouge, LA 70803-6306
http://www.bus.lsu.edu/economics/

Updating Valuations and Stochastic Dynamics on Coordination

Hideki FUJIYAMA
Department of Economics
Dokkyo University
1-1 Gakuen-cho, Soka-shi
Saitama 340-0042 Japan
e-mail: fujiyama@dokkyo.ac.jp

May 2007

Abstract

This work studies an equilibrium selection of infinitely repeated symmetric 2×2 coordination games that show a tension between Pareto efficiency and risk dominance, in which bounded rational agents adopt the following simple behavior rule: each agent has a valuation of actions, and chooses the highest one. Valuations are updated according to the sign of the difference between the current valuation and the realized payoff. By applying techniques from stochastic stable states (Kandori et al. 1993 and Young 1993), it is shown that the risk dominant outcome is selected; that is, it is realized more frequently in the long run.

JEL classification: C72, D83

Keywords: Satisficing behavior, Coordination game, Stochastic evolutionary game, Valuation of action

1 Introduction

It is not possible to know everything relevant when we face decisions. Generally, the limited knowledge we have is based on personal experience. If someone succeeds in gaining relatively higher payoff, he/she would retain that choice of action. Such subjective valuations of actions are critical factors in maintaining stable situations. Of course, the outcome of an action depends on the other person's action, and vice versa.

In this paper, under such patterns of behavior, we examine the equilibrium selection in a symmetric coordination game. We assume that each player has valuations of possible actions. If an action is chosen, the valuation of the action is updated according to the sign of the difference between current payoff and current valuation of the action. If the realized payoff of the action exceeds the valuation then the valuation increases; otherwise it decreases. If both are equal, it does not change. If the valuation of the current action exceeds the valuation of the alternative, the player chooses the same action next time. Also, because of the perturbation, the player chooses the undesirable action with small probability.[1]

This behavior rule shares basic features with satisficing behavior (Karandikar, Mookherjee, Ray and Vega-Redond 1998, In-Koo and Matsui 2005, Kim 1999, Pazgal 1997). The features of satisficing behavior are summarized as follows: (1) there is *one* endogenous parameter (aspiration) that is updated, based on realized payoffs, and converges to the realized payoffs.[2] (2) players use the parameter to trigger changes of action. In the present model we have *two* endogenous parameters (valuations for actions), of which one converges to the realized payoffs. Regarding (2), two endogenous parameters (valuations

[1]This trembling for choice is also assumed by In-Koo and Matsui (2005) and Pazgal (1997). Conversely, Karandikar et al (1998) assume that aspiration is perturbed directly.

[2]In Pazgal (1997), the endogenous parameter (aspiration) converges to the maximal payoff in past experience.

for actions) are used in turn as a reference point.

In the present model, the Risk dominant outcomes are selected in the long run in the coordination game irrespective of the initial conditions. This is in contrast to the situation in (Karandikar et al. 1998, Pazgal 1997, Kim 1999), in which the Pareto efficient outcome is selected under certain conditions.[3]

This difference stems from the number of endogenous reference points. In the other papers, the endogenous reference point is the aspiration level. Based on this reference point, the action is chosen. In the simple satisficing model (Karandikar et al. 1998), if the equilibrium outcome is Pareto efficient, the aspiration is higher than the Risk dominant outcome. It is therefore difficult to move to the Risk dominant outcome. On the other hand, if the equilibrium outcome is Pareto inefficient (and Risk dominant), the aspiration is lower than the Pareto efficient outcome and it is easy to move to the Pareto efficient outcome.

In case-based decision theory (Pazgal 1997, Kim 1999), valuations of actions are calculated at every time. The aspiration level is used as a reference point for the valuations, and there is a device to maintain the aspiration level high. In Pazgal (1997), it is assumed that aspiration converges to the maximum payoff in the past. Also, if the initial aspiration level is high enough, players are not satisfied with their current situation and make many experiments in the early periods. The Pareto efficient outcome then occurs with probability close to 1. In this situation the player is not satisfied with the Pareto inefficient outcome, generating an inclination to Pareto efficient outcomes.

In Kim (1999), let us suppose that the aspiration level is close to the efficient payoff and the adjustment speed is slow. If the efficient outcome is realized once, the outcome

[3] In Pazgal (1997), Kim (1999), an initial high aspiration level is needed. The results in Karandikar et al. (1998) do not depend on the initial conditions, but require sufficiently slow updating of aspiration.

then continues. This is because, by the way of making valuations, positive values are accumulated to the valuation of the efficient action. On the other hand, even if the risk dominant outcome is realized, the valuation of the risk dominant action becomes less than that of the efficient action after a finite period. This is because, with a high aspiration level, the negative values accumulate to the valuation of the risk dominant action. As a result of this negative accumulation and slow adjustment, the valuation of risk dominant action decreases.

In contrast to previous research, players in our model use two reference points in turn. This weakens commitment to the efficient outcome. This weakening is because the reference point is a valuation of the other actions, and it cannot maintain a high value. Hence, the risk dominant outcome comes to be realized in the long run.

When we focus more closely on behavior patterns, the differences are as follows. When a player quits the current action and chooses a new action, in Karandikar et al. (1998) and In-Koo and Matsui (2005), a very simple manner is assumed: all actions can be chosen with equal probability. In Kim (1999) and Pazgal (1997), a more sophisticated manner is assumed: players adopt an action that has the largest cumulative payoff rescaled by the current aspiration level. In that case, players have to remember all past experience in the past.[4] On the other hand, in the present situation, players update one of the valuations directly at every time, so that players need to remember only two numbers (the valuations of two actions). Of course, if the situation is more complex, it is too difficult to consider situations carefully at each choice of action. Karandikar et al. (1993) and In-Koo and Matsui (2005) are then helpful guides. If the situation is important for the players, they would adopt the methodology of Kim (1999) and Pazgal (1997). But when the situation

[4]There is a simple formula for levels of valuations. However, players have to remember the number of times players choose a particular action.

is not overly complex and not overly important, the present methodology applies.

The rule of thumb in the present model shares common features with the reinforcement dynamics of Roth and Erev (1995) and Erev and Roth (1998). Under reinforcement dynamics there are two endogenous parameters, as here. In the present model, however, players use a pure strategy except for tie breaking. Under reinforcement dynamics, players use a mixed strategy and valuations of actions are used as the weight for mixing. Hence, in our model, players stick firmly to the current action so long as the valuations indicate that it is better. In other words, the inertia is stronger than the reinforcement dynamics. In the coordination game, it is possible that players afraid of frequent changes of action would send complex signals to the other player. Consequently, a pure strategy assumption in this model may also be relevant.

The relationship among models are summarized in Table1.

	One Endogenous Parameter	Two Endogenous Parameters
Pure Strategy	Aspiration Model	Present Model
Mixed Strategy		Reinforcement Model

Table 1: The present model, aspiration model and reinforcement model

In the next section we describe the model. Section 3 presents results. Section 4 sketches the proofs of the results. Section 5 gives concluding comments.

2 The Model

Consider the following 2×2 game:

	A	B
A	a, a	d, c
B	c, d	b, b

in which we assume that $a > c$, $b > d$ and $a + d < b + c$. Both (A, A) and (B, B) are Nash equilibria, and (A, A) is a Pareto efficient outcome and (B, B) is a risk dominant outcome.

We also assume that players divided the continuous domain for evaluation into a same sized grid, and recognize only which sector the payoffs or valuations fall in. If we label all grids with positive integers, then $a, b, c, d \in Z^+$ where Z^+ denotes the set of positive integers. The players are indexed by $i = 1, 2$. Player i's state at time $t \geq 0$ is defined by two *valuations*, for two actions, as (v^i_{At}, v^i_{Bt}). We take all valuations as positive without loss of generality.[5] Initial states are $H, L \in Z^+$ such that $H > a$ and $L < d$, and for $i \in \{1, 2\}$, $v^i_{A,0}, v^i_{B,0} \in [L, H]$. A (social) *state* s is the pair of states of Players 1 and 2. Thus, at time t, it follows that $s \equiv [(v^1_{At} v^1_{Bt}); (v^2_{At}, v^2_{Bt})]$.

Given Player i's state at period t, Player i adopts the action that has higher valuation:

$$x^i_t = A \quad \text{if } v_{A,t} > v^i_{B,t} \tag{1}$$

$$x^i_t = B \quad \text{if } v_{A,t} < v^i_{B,t}. \tag{2}$$

If $v^i_{A,t} = v^i_{B,t}$ then player i randomizes with mixed probabilities over A and B; that is, $x^i_t =$ with probability p, and $x^i_t = B$ with probability $1 - p$, where $p \in (0, 1)$. This pair of actions determines the payoffs, $\pi^1_t(x^1_t, x^2_t)$ and $\pi^2_t(x^2_t, x^1_t)$.

Player i updates valuations according to the sign of the difference between π and v.

[5] If necessary, we can transform payoffs and initial valuations by adding large positive numbers to them.

Whatever the size of the difference, the extent of revision for the valuation is constant. Let $i \in \{1, 2\}$, $i \neq j$ and $\alpha \in \{A, B\}$,

$$v^i_{\alpha,t+1} = \begin{cases} v^i_{\alpha,t} + 1 & \text{if } x^i_t = \alpha \text{ and } \pi^i_t(x^i_t, x^j_t) > v^i_{\alpha,t} \\ v^i_{\alpha,t} & \text{if } x^i_t = \alpha \text{ and } \pi^i_t(x^i_t, x^j_t) = v^i_{\alpha,t} \\ v^i_{\alpha,t} - 1 & \text{if } x^i_t = \alpha \text{ and } \pi^i_t(x^i_t, x^j_t) < v^i_{\alpha,t}. \end{cases} \quad (3)$$

If $x^i_t \neq \alpha$ then $v^i_{\alpha,t+1}$ retains the same value. This updating rule expresses the fact that players change their valuations grid by grid; we assume the constant change of valuations.

These updating rules define a Markov process over the (social) state space defined by the set $[L, H] \times [L, H] \times [L, H] \times [L, H]$. Let S be the state space. The process will be denoted P and be referred to as the *untrembled process*.

Under the untrembled process there are in general many stationary states. If we introduce trembles in choosing actions, we can select the most robust outcome against small perturbations. Let x^* be the action chosen under the untrembled process, and let x' be another action. We assume that each player chooses x^* with probability $1 - \varepsilon$ and x' with ε, where ε is a small positive number less than 1. The process will be denoted P^ε and is referred to as the *trembled process*.

3 Results

We first define the stable states in the untrembled process.

A stable state is such that, for all $i, j \in \{1, 2\}$ and $i \neq j$: (1) it induces a pair of actions (x^1, x^2); (2) the valuation of the action is exactly equal to the achieved payoff, so that $v^i_{x^i,t} = \pi^i(x^i, x^j)$; and (3) the valuation of the action exceeds that of another one: $v^i_{x^i,t} > v^j_{y^i,t}$ where $x^i \neq j^i$. We refer to the state as the pure strategy state (PSS), following (Karandikar et al. 1998). Of course, every PSS is a recurrent class of the untrembled process by definition.

For convergence to PSS, we have the following result. To keep the proof simple, suppose that for any $i \in \{1,2\}$, $d \leq v^i_{A,0} \leq a$ and $c \leq v^i_{B,0} \leq b$ or $b \leq v^i_{B,0} \leq c$.[6] ("PSSs" will be the plural form of "PSS.")

Proposition 1

Assume that for any $i \in \{1,2\}$, $d \leq v^i_{A,0} \leq a$ and $c \leq v^i_{B,0} \leq b$ or $b \leq v^i_{B,0} \leq c$. The untrembled process converges to particular PSSs. PSSs are constituted only of Pareto efficient outcomes and risk dominant outcomes.

From the construction of the trembled process, P^ε is irreducible and aperiodic for every $\varepsilon > 0$. Hence, by standard results for Markov processes, it has a unique stationary distribution. Denote this by μ^ε. Any initial distribution converges to it, so that for any initial distribution d, $dP^\varepsilon \rightarrow \mu^\varepsilon$ as $t \rightarrow \infty$. According to standard stochastic evolutionary game analysis (Young 1998, Young 1993, Kandori et al. 1993), $\lim_{\varepsilon\rightarrow 0} \mu^\varepsilon = \mu^*$ exists. If $\mu^*(s) > 0$, the state s is called stochastically stable, where $\mu^*(s)$ is the probability of state s given μ^*. For the probability of the set of states, let $\mu^\varepsilon(A) \equiv \sum_{s\in A} \mu^\varepsilon(s)$.

We derive the following result relating to stochastically stable states:

Proposition 2

Assume that for any $i \in \{1,2\}$, $d \leq v^i_{A,0} \leq a$ and $c \leq v^i_{B,0} \leq b$ or $b \leq v^i_{B,0} \leq c$. The risk dominant outcome corresponds to stochastically stable states. Formally, let s^R is the set of states such that $v^1_B = v^2_B = b$, $v^1_B > v^1_A$ and $v^2_B > v^2_A$. Then $\lim_{\varepsilon\rightarrow 0} \mu^\varepsilon(s^R) = 1$.

---

[6]This assumption is not used in Theorem 1. But it is reasonable if there are some test periods for deciding valuations based on the realized payoff.

Next, assume that for $i \in \{1, 2\}$, $v^i_{A,0}, v^i_{B,0} \in [L, H]$. Under this condition, for any $\epsilon > 0$, P^ε is not irreducible. The standard technique cannot therefore be applied in this case. However, the assertions of Proposition 2 do not change, so that we derive the following theorem:

Theorem 1 Assume that for any $i \in \{1, 2\}$, $v^i_{A,0}, v^i_{B,0} \in [L, H]$.

(1) The risk dominant outcome corresponds to stochastically stable states in the model. Formally, let s^R is the set of states such that $v^1_B = v^2_B = b$, $v^1_B > v^1_A$ and $v^2_B > v^2_A$. Then $\lim_{\varepsilon \to 0} \mu^\varepsilon(s^R) = 1$.

(2) Any initial state converges to a stochastically stable state.

4 Informal Sketches of Proofs

All proofs are straightforward but long, and are therefore set out in full in the Appendix. In this section we describe the intuitive basis of the proofs.

For Proposition 1, if a sequence of states does not enter in the PSS, the sequence circulates. However, from the way of updating valuations and the tie breaking rule, it is impossible. In the proof we show that, from arbitrary initial conditions, there is a path to PSS with positive probability.

For Proposition 2, the range of initial states is restricted. Hence it is easy to derive the minimum cost tree. The costs from (A, A)-type PSS to (B, B) can be shown to be less than the costs from (B, B)-PSS to (A, A), by comparing a cost tree whose root is in (B, B)-PSS and the minimum cost tree whose root is in (A, A)-PSS.

For Theorem 1, the range of initial states is extended. However, there is a unique recurrent class that is exactly the same as that of Proposition 2. In other words, if a state is PSS but is not included in the set of PSS in Proposition 2, then it is transient. According to standard results for Markov chains, this implies that there exists a unique limit distribution of P^ε and, for any transient state, as $t \to \infty$ then its limit distribution of P^ε is 0. From these facts it is easy to prove the statement of Proposition 2 under this assumption, and that there is good convergence from any initial state.

5 Concluding Remarks

This paper has examined stochastically stable states in the coordination game that show a tension between Pareto efficiency and risk dominance. In the present model, each agent has a valuation of the possible actions, which are updated according to a sign of the difference between the current valuation and the realized payoff. The agent chooses the action having higher valuation. In this situation, the risk dominant outcome is selected.

Let us now consider extensions to the analysis. First, the present results and proofs suppose that the domain of valuations is a positive integer, and that the extent of revision for the valuation is constant. It is therefore important to consider the case in which the extent of revision depends on the difference between the valuation and the realized payoff. Second, one might assume that the probability of trembling ε depends on the state; for example, it could depend on the difference between a valuation and the realized payoff. Third, there is a generalization to a finite-population model. Finally, we should consider the appropriateness of the present model in light of the data from experimental analysis.

Acknowledgment:

I thank Sarangi Sudipta for his comments, and LSU for an good research environment during 2006-2007.

Appendix

Proof of Proposition 1: There are two distinct cases, depending on the payoff structure. We examine them in turn.

Case 1: $a > b > c > d > 0$.

We divide all states into nine regions, depending on the valuations (see Table 1).

If the process enters Region 1 or 9, the state converges to PSS. Note that every PSS is a recurrent class. This is because there are no trembles, so that the relationship between two valuations does not change. Hence we can show that, from every region, there is a path to Region 1 or Region 9 with positive probability. Let (x^1, x^2) be a couple of actions, for Player 1 and 2 respectively.

	$v_A^2 > v_B^2$	$v_A^2 = v_B^2$	$v_A^2 < v_B^2$
$v_A^1 > v_B^1$	Region 1 (Convergence to PSS)	Region 2	Region 3
$v_A^1 = v_B^1$	Region 4	Region 5	Region 6
$v_A^1 < v_B^1$	Region 7	Region 8	Region 9 (Convergence to PSS)

Table 2: Regions of pairs of valuations

We first examine Region 3. In this region, (A, B) are chosen, and $v_{A,t}^1$ and $v_{B,t}^2$ decrease monotonically.

From the assumptions, for any $0 \le t < \infty$, $i \in \{1, 2\}$, then $d \le v_{A,t}^i \le a$ and $c \le v_{B,t}^i \le b$ and all values of valuations are integers. After a finite number of periods, the state enters Region 2, Region 5 or Region 6 with probability one.

In Region 2 and 5, by randomization due to tie breaking, (A, A) is chosen with positive probability. The state then moves into Region 1.

In Region 6, if $v^1_{B,t} < b$ then, with positive probability, (B, B) is chosen; after that, the state moves into Region 9. Otherwise, we can construct the following path with positive probability: first (A, B) is chosen, and $v^1_{A,t} \geq v^1_{B,t}$ and $v^2_{A,t} \geq v^2_{B,t}$ are realized. Next, (B, B) is chosen with positive probability, then the state moves into Region 9.

For Regions 4, 7, 8, a symmetry argument is applied. Hence, from every region, there is a path to the PSS in Region 1 or 9 with positive probability.

Case 2: $a > c > b > d > 0$.

We also use Table 1 in this case. By the same argument as in Case 1, every state in Region 2 and 5 moves into Region 1 with positive probability.

In Region 3, (A, B) is chosen with probability 1. Also, $v^1_{A,t}$ decreases by 1 and $v^2_{B,t}$ increases by 1 or is equal to c. From the assumption that for any $0 \leq t < \infty$, $i \in \{1, 2\}$, it follows that $d \leq v^i_A \leq a$ and $b \leq v^i_B \leq c$ and that all values of valuations are integers. After a finite number of periods, the state moves into Region 6 with probability one.

Below, there are many sub-cases, so that we describe a state via Figure 1 in which the values of $v^1_{A,\cdot}$, $v^1_{B,\cdot}$, $v^2_{A,\cdot}$ and $v^2_{B,\cdot}$ are denoted by dots on bars. Each bar shows the domain of each valuation.

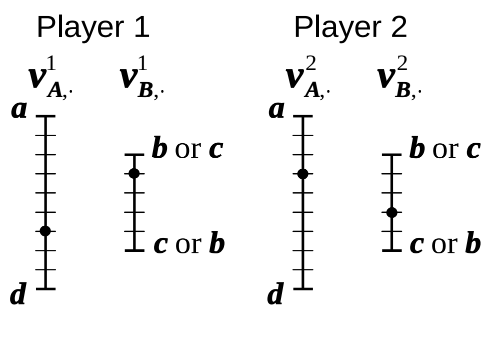

Figure 1: Description of a state

In Region 6, there are following sub-cases.

Sub-case 1: Suppose that $v^1_{A,t} = v^1_{B,t} = b$. We can construct the following sequence of pairs of actions. First, (B, B) is chosen with nonzero probability. After that, $v^1_{A,t+1} = v^1_{B,t+1}$ and $v^2_{A,t+1} \leq v^2_{B,t+1}$ are realized. If $v^2_{A,t+1} = v^2_{B,t+1}$, the state is in Region 5 (Figure 2). Otherwise, (B, B) is chosen repeatedly until $v^2_{A,\cdot} = v^2_{B,\cdot}$.

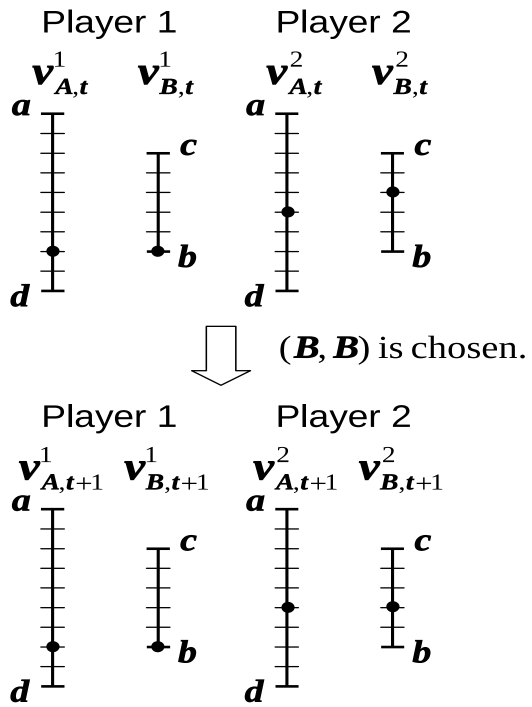

Figure 2: Transitions for states

Sub-case 2: Suppose that $v^1_{A,t} = v^1_{B,t} \neq b$. Suppose also that $v^2_{B,t} \leq c - 1$. First, players choose (A, B). After that, $v^1_{A,t+1} < v^1_{B,t+1}$ and $v^2_{A,t+1} < v^2_{B,t+1}$ are realized. Next, (B, B) is chosen. After that, $v^1_{A,t+2} = v^1_{A,t} - 1 = v^1_{B,t+2} = v^1_{B,t} - 1$ and $v^2_{A,t+2} = v^2_{A,t} < v^2_{B,t+2} = v^2_{B,t}$ are realized (Figure 3). This shows that Player 1's valuations decrease by 1 but Player 2's valuations do not change. If we repeat (A, B) and (B, B) enough times, then $v^1_{A,\cdot} = v^1_{B,\cdot} = b$ and $v^2_{A,\cdot} < v^2_{B,\cdot}$ are realized. Thereafter we can follow the argument of sub-case 1.

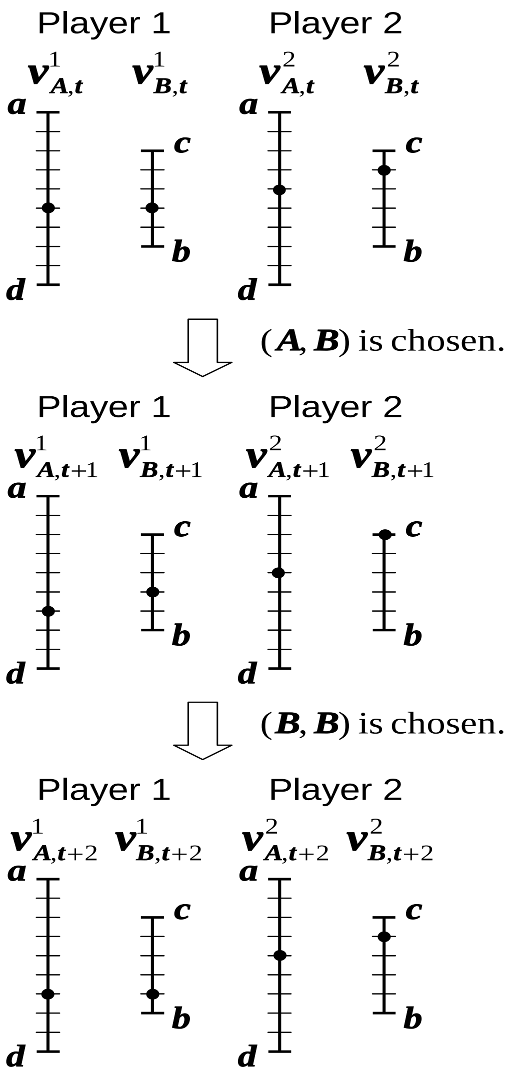

Figure 3: Transitions for states

Sub-case 3: Suppose that $v^1_{A,t} = v^1_{B,t} \neq b$. Suppose also that $v^2_{A,t} \leq c-2$ and $v^2_{B,t} = c$. First, players choose (A, B). After that, $v^1_{A,t+1} < v^1_{B,t+1}$ and $v^2_{A,t+1} < v^2_{B,t+1}$ are realized. Next, (B, B) is chosen. After that, $v^1_{A,t+2} = v^1_{A,t} - 1 = v^1_{B,t+2} = v^1_{B,t} - 1$ and $v^2_{A,t+2} = v^2_{A,t} < v^2_{B,t+2} = c-1$ are realized (Figure 4). Thereafter we can follow the argument of sub-case 2.

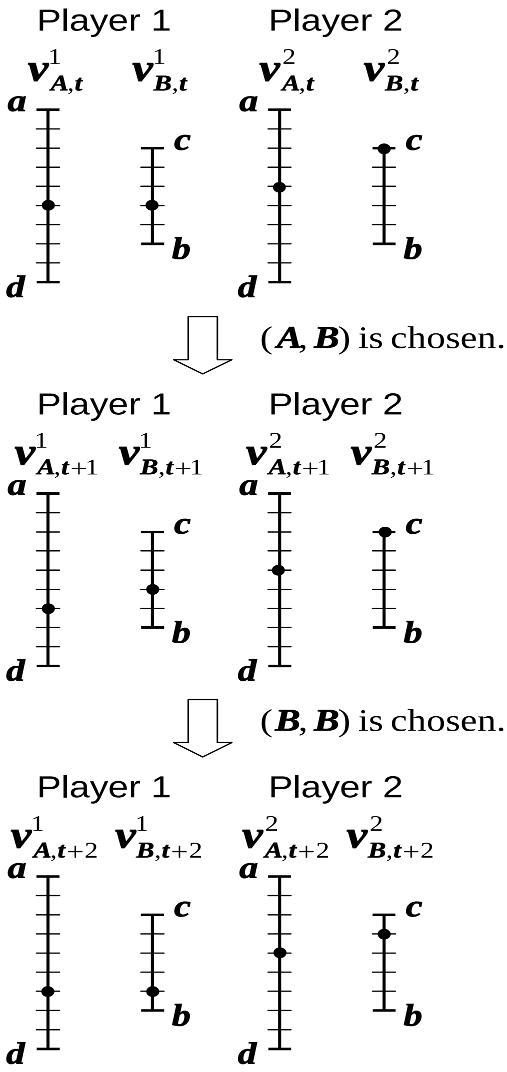

Figure 4: Transitions for states

Sub-case 4: Suppose that $v_A^1 = v_B^1 \neq b$. Assume also that $v_B^2 = c$ and $v_A^2 = c-1$. First, players choose (A, B). After that $v_{A,t}^1 < v_{B,t}^1$ and $v_{A,t}^2 < v_{B,t}^2$ are realized. Next, (B, B) is chosen. After that, $v_{A,t+2}^1 = v_{A,t}^1 - 1 = v_{B,t+2}^1 = v_{B,t}^1 - 1$ and $v_{A,t+2}^2 = v_{B,t+2}^2 = c - 1$ are realized (Figure 5). This state is in Region 5.

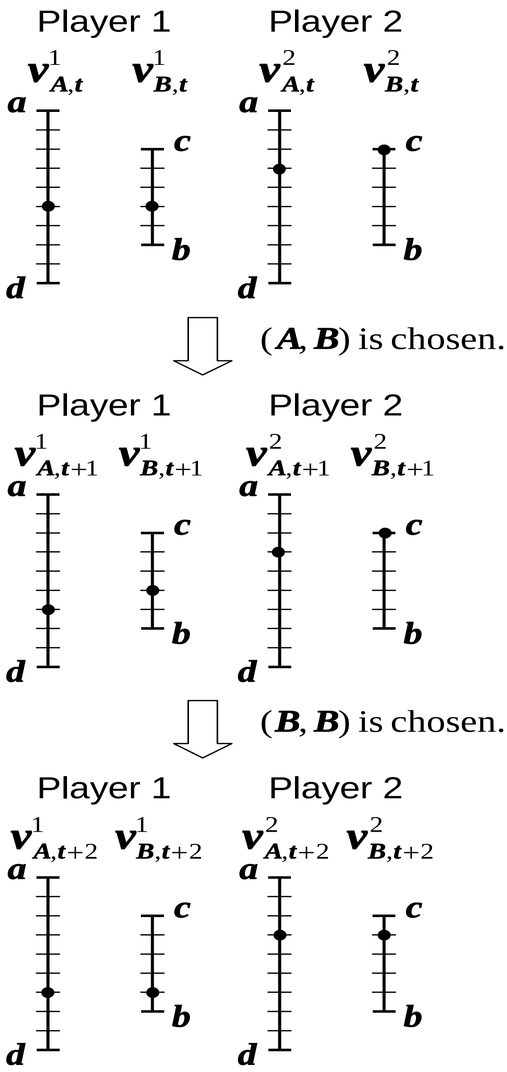

Figure 5: Transitions for states

For Regions 4, 7, 8, the symmetry argument is used. Hence, in every region, there is a path to the PSS in Region 1 or 9 with positive probability. The proposition in this case is proved.

Note that (B, B) is a PSS, as in the previous case. This is because we can consider the state in which $v^1_{A,0} = v^2_{A,0} = d$ and $v^1_{B,0} = v^2_{B,0} = b$. □

Proof of Proposition 2:

By irreducibility of the trembled process, we can use the standard technique (Young 1998). In this technique, cost trees are defined in the set of PSSs, and we find the smallest cost tree. Let (x^1, x^2)-PSS be a state that it is a PSS, and (x^1, x^2) is chosen. If a root of a cost tree is in (x^1, x^2)-PSSs, then we refer to it as a "(x^1, x^2)-PSS cost tree."

By definition, a cost tree includes all PSSs. If only a sub set of all states are included, we refer to it as "cost sub-tree". If only a subset of all (x^1, x^2)-PSSs is included, we refer to it as a "(x^1, x^2)-PSS cost sub-tree."

States are defined by the valuations made by the two players. If actions are denoted together it is easier to understand a state, so that we denote a state not as $[(v^1_{A,t}, v^1_{B,t}); ((v^2_{A,t}, v^2_{B,t}))]$ but as $[x^2_t, (v^1_{A,t}, v^1_{B,t}); x^2_t, ((v^2_{A,t}, v^2_{B,t}))]$ in this proof. If time is not important, we omit t and write this as $[x^2, (v^1_A, v^1_B); x^1((v^2_A, v^2_B))]$. Finally, let $(s \rightarrow s')$ be a path from state s to s'.

The proof is divided into two different cases, depending on the payoff structure.

Case 1: $a > b > c > d > 0$.

We first exhibit two claims used in this proof.

Claim 1: If, for Player $i = 1$ (or 2), $v^i_B \geq c + 1$ and the state is a (A, A)-PSS then, with one perturbation, v^i_B decreases by one.

This is because the sequence of action choices, (A, B) (or (B, A)) with one perturbation, and (A, A) with no perturbation, gives rise to a situation in which only v^i_B decreases by one.

Claim 2: If, for Player $i = 1$ (or 2), $v_A^i \geq d+1$ and the state is a (B,B)-PSS then, with one perturbation, v_A^i decreases by one.

This is because the sequence of action choices, (B,A) (or (A,B)) with one perturbation, and (B,B) with no perturbation, gives rise to the situation.

Lemma 1: There is a (B,B)-PSS cost tree having cost $\{(b-c+1)^2-1\}+(a-c)+\{(b-d)^2-1\}$.

Proof of Lemma 1:

Step 1: In this step we construct a (A,A)-PSS cost sub-tree such that the root is $[A,(a,c);A,(a,c)]$; it includes all (A,A)-PSSs but not (B,B)-PSSs, and its cost is $(b-c+1)^2-1$.

Every state in (A,A)-PSSs satisfies the following condition: for all $i,j \in \{1,2\}$, $v_A^1 = v_A^2 = a$ and $v_B^1, v_B^2 \in \{c, c+1, \ldots, b\}$.

Consider the following (A,A)-PSS cost sub-tree (Figure 6):

$$\bigcup_{v_B^1 \in \{b,b-1,\ldots,c\}} \bigcup_{v_B^2 \in \{b,b-1,\ldots,c+1\}} ([A,(a,v_B^1);A,(a,v_B^2)] \to [A,(a,v_B^1);A,(a,v_B^2-1)]) + \bigcup_{v_B^1 \in \{b,b-1,\ldots,c+1\}} ([A,(a,v_B^1);A,(a,c)] \to [A,(a,v_B^1-1);A,(a,c)]) \quad (4)$$

Since every link has cost one by Claim 1, its cost is $(b-c+1)^2-1 (= (b-c)(b-c+1)+(b-c))$.

Step 2: In this step we construct a (B,B)-PSS cost sub-tree such that the root is $[B,(d,b);B,(d,b)]$; it includes all (B,B)-PSSs but not (A,A)-PSSs, and its cost is $(b-d)^2-1$.

Every state in (B,B)-PSSs satisfies the following condition: for all $i,j \in \{1,2\}$, $v_B^1 = v_B^2 = b$ and $v_A^1, v_A^2 \in \{d, d+1, \ldots, b-1\}$.

Consider the following (B,B)-PSS cost sub-tree (Figure 7):

$$\bigcup_{v_A^1\in\{b-1,b-2,\ldots,d\}}\ \bigcup_{v_A^2\in\{b-1,\ldots,d+1\}}([B,(v_A^1,b);A,(v_A^2,b)]\to[B,(v_A^1,b);B,(v_A^2-1,b)])$$
$$+\bigcup_{v_A^1\in\{b-1,\ldots,d+1\}}([B,(v_A^1,b);B,(d,b)]\to[B,(v_A^1-1,b);B,(d,b)]) \quad (5)$$

Since every link has cost one by Claim 2, its cost is $(b-d)^2-1(=(b-d-1)(b-d)+(b-d-1))$.

Step 3: In this step we construct a path from $[A,(a,c);A,(a,c)]$ to $[B,(c,b);B,(c,b)]$ and show that its cost is $a-c$.

Consider the following transition. First, (A,B) is repeated $a-c$ times. This transition takes cost $a-c$. After this, $v_A^1=v_B^1=c$, $v_A^2=a$ and $v_B^2=c$ are realized (Figure 8). With no cost, (B,A) is repeated $a-c$ times (Figure 9). Finally (B,B) is repeated $b-c$ times with no cost, and $[B,(c,b);B,(c,b)]$ is reached (Figure 10). By summing all costs, a total cost of $a-c$ is required.

By Step1, Step2 and Step 3, we construct a cost tree that satisfies the condition of lemma 1 (Figure11). □

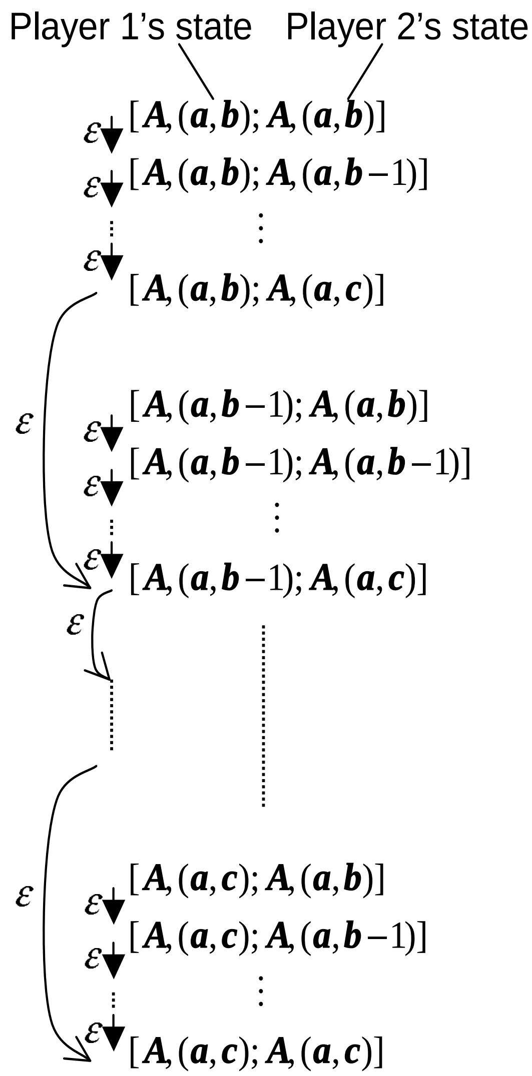

Figure 6: (A, A)-PSS cost sub-tree

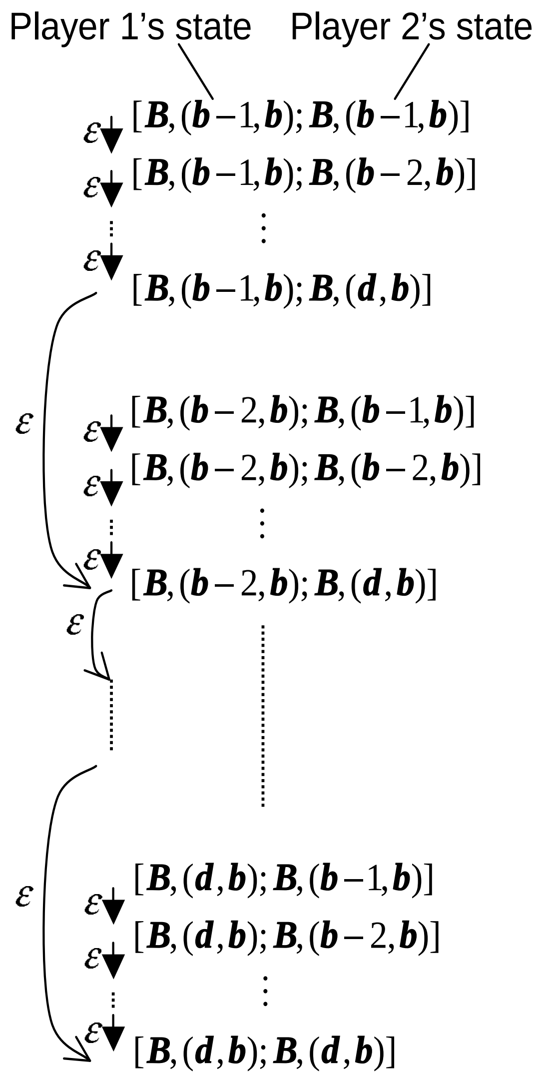

Figure 7: (B, B)-PSS cost sub-tree

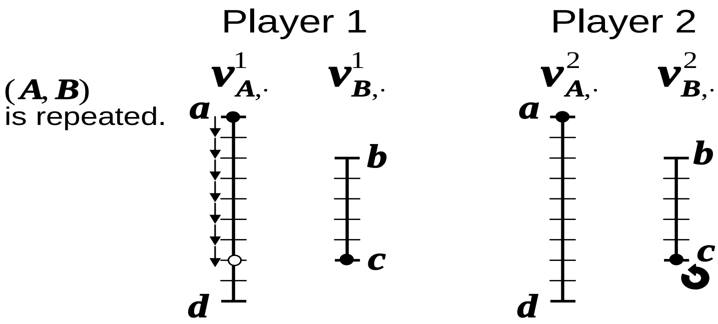

Figure 8: Transitions

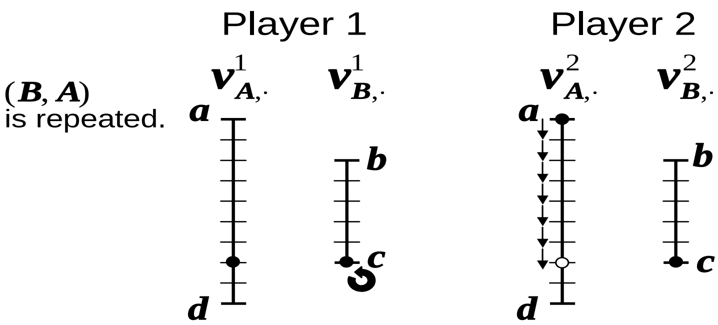

Figure 9: Transitions

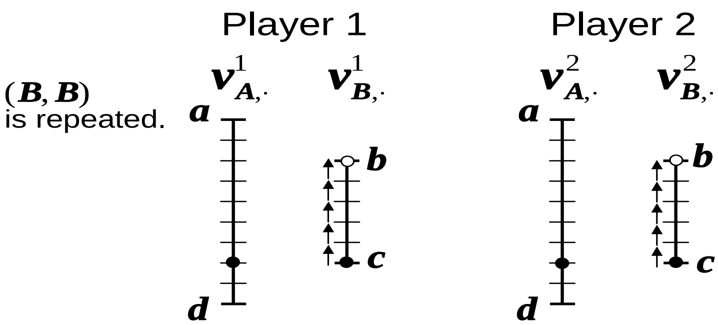

Figure 10: Transitions

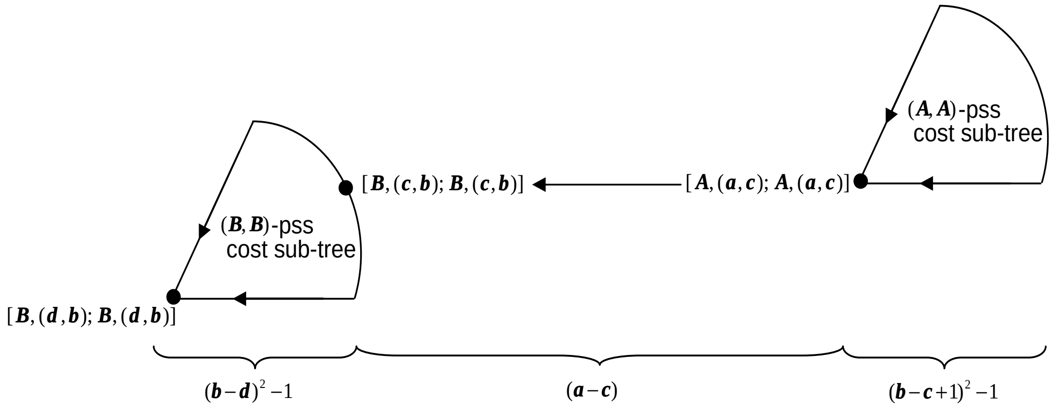

Figure 11: a cost tree

Before stating Lemma 2, we exhibit the following claim, similar to the previous claim.

Claim 3: If v_A^i ($i \in \{1, 2\}$) increases in (B, B)-PSS, then at least two perturbations are needed. More precisely, both v_B^1 and v_B^2 increase by one.

This is because the only pair of actions that increases v_A^i ($i \in \{1, 2\}$) is (A, A).

Lemma 2: The minimum cost in the (A, A)-PSS cost trees is $\{(b - c + 2)^2 - 1\} + 2 + \{(b - d - 1)(b - d + 2)\}$.

Proof of Lemma 2:

Step 1: We show the path that the minimum cost tree includes. From the definition of a cost tree, there is a path to the root from any state. Consider paths from $[B, (d, b); B, (d, b)]$ to (A, A)-PSS. We can construct the following path:

$$\left\{ \bigcup_{v_A \in \{d, d+1, \ldots, b-1\}} ([B, (v_A, b); A, (v_A, b)] \rightarrow [B, (v_A + 1, b); B, (v_A + 1, b)]) \right\} \\ + \{([B, (b, b); B, (b, b)] \rightarrow [B, (b + 1, b); A, (b + 1, b)])\} \tag{6}$$

The first term expresses transitions in (B, B)-PSSs. In the transitions, two perturbations are needed at each transition. In the second term there is no perturbation. Hence the total cost of this path is $2(b-d)$.

This path has the minimum cost among paths from $[B, (d, b); B, (d, b)]$ to (A, A)-PSS. This is because, by Claim 3, there is no redundant transition in this path.

Step 2: In this step we examine the minimum cost tree that includes the path defined in Step 1.

In the PSSs, at least one perturbation is needed, by definition. The path in Step 1 is a transition from (B, B)-PSS to (A, A)-PSS. Hence, if the cost tree whose root is (A, A)-PSS and includes the path in Step 1 is the minimum cost tree, then the other links need exactly one perturbation.

Step 3: In this step we construct a (A, A)-PSS cost sub-tree such that the root is $[A, (a, c); A, (a, c)]$; it includes all (A, A)-PSSs but not (B, B)-PSSs, and its cost is $(b-c+1)^2 - 1$.

This result follows from Step 1 of Lemma 1. In this tree, each one-step link needs exactly one perturbation.

Step 4: In this step we construct a (B, B)-PSS cost sub-tree such that the root is $(B, (b-1); B, (b-1, b))$; it includes only (B, B)-PSSs but not (A, A)-PSSs, and includes part of the path in Step 1. Its cost is $(b-d-1)(b-d+2)$.

Consider the following (B, B)-PSS cost sub-tree (Figure 12):

$$\bigcup_{i\in\{1,2,\dots,b-d-1\}} \left[\left\{ \bigcup_{v_A^2\in\{d+1,d+2,\dots,b-1\}} ([B,(b-i,b);B,(v_A^2,b)] \to [B,(b-i,b);B,(v_A^2-1,b)]) \right\} \right. \tag{7}$$

$$\left. - ([B,(b-i,b);B,(b-i,b)] \to [B,(b-i,b);B,(b-i-1,b)]) \right]$$

$$+ \bigcup_{i\in\{1,2,\dots,b-d-1\}} ([B,(b-i,b);B,(d,b)] \to [B,(b-i-1,b);B,(d,b)])$$

$$+ \bigcup_{v_A^2\in\{d+1,d+2,\dots,b-1\}} ([B,(b-i,b);B,(v_A^2,b)] \to [B,(b-i-1,b);B,(v_A^2-1,b)])$$

$$+ \bigcup_{v_A\in\{d,d+1,\dots,b-2\}} ([B,(v_A,b);B,(v_A,b)] \to [B,(v_A+1,b);B,(v_A+1,b)])$$

Since every link in the first, second and third terms has one perturbation by Claim 1, its cost is $(b-d-1)(b-d-1-1)+(b-d-1)+(b-d-1)$. Every link in the last terms has two perturbations by Claim 3, so that its cost is $2(b-d-1)$. Hence the total cost is $(b-d-1)(b-d+2)$.

Step 5: In this step we construct a path from $[B,(b-1,b);B,(b-1,b)]$ to $[A,(a,b);A,(a,b)]$ and show its cost is 2.

Consider the following path:

$$[B,(b-1,b);B,(b-1,b)] \to [A,(b,b);A,(b,b)] \to [A,(b+1,b);A,(b+1,b)] \to \cdots \tag{8}$$
$$\to [A,(a-1,b);A,(a-1,b)] \to [A,(a,b);A,(a,b)]$$

In the first transition, two perturbations are needed. In the other transitions there is no cost, because there is a nonzero probability that both players choose A. Hence the

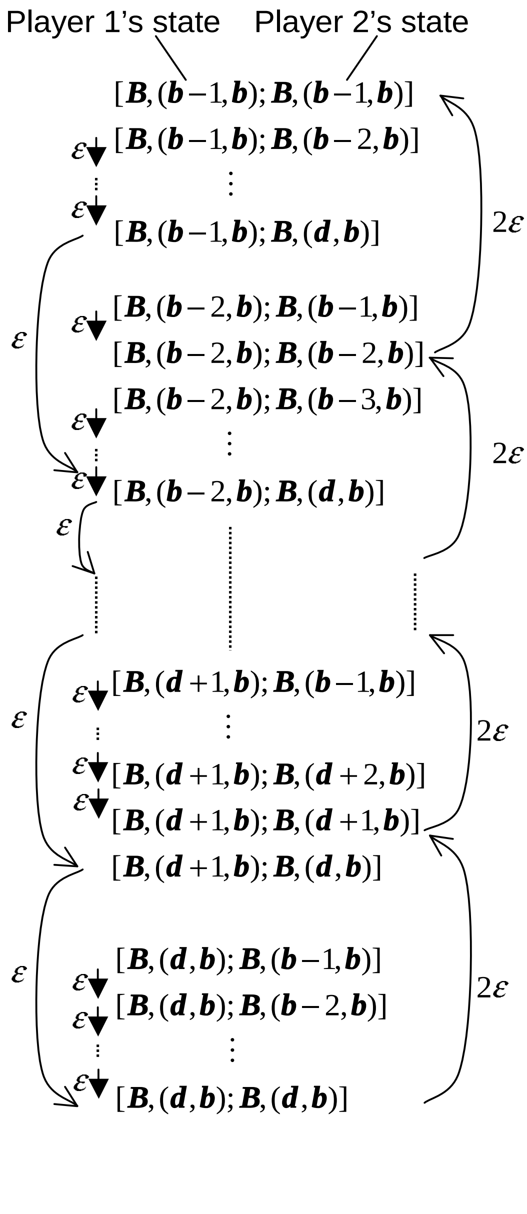

Figure 12: (B, B)-PSS cost sub-tree

total cost of this path is 2.

Step 6: We finally sum over each of the steps in this argument to prove the lemma. By collecting all of the links in Step 3, 4 and 5, we generate the cost tree that includes the path in Step 1 (Figure 13). Each link has exactly unit cost (perturbation) except the path in Step 1. By Step 2, this is the minimum cost tree in the (A, A)-PSS cost trees.

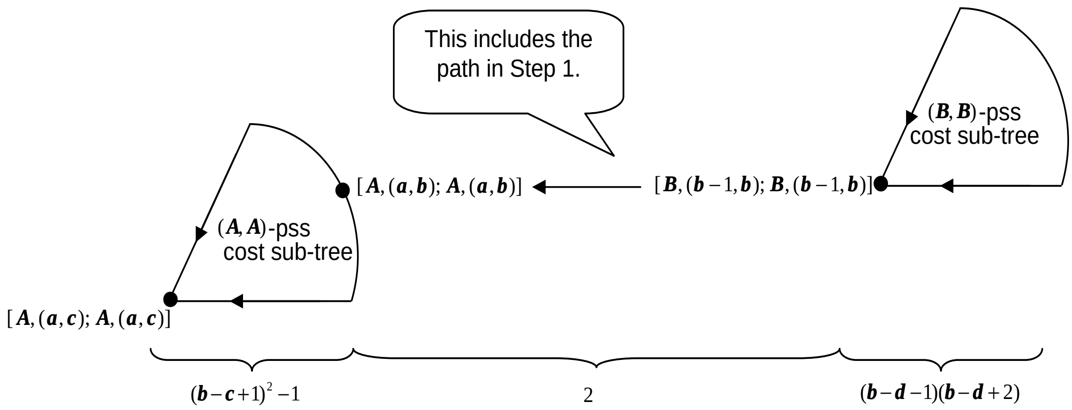

Figure 13: a cost tree

From steps 3, 4 and 5, the total cost is $\{(b-c+1)^2-1\}+\{2\}+\{(b-d-1)(b-d+2)\}$.

The lemma is proved. □

Finally, we prove the statement in Proposition 2 given the condition $a > b > c > d > 0$. By Lemmas 1 and 2,

$$\begin{aligned} &[\{(b-c+1)^2-1\}+\{a-c\}+\{(b-d)^2-1\}] \\ &\quad -[\{(b-c+1)^2-1\}+\{2\}+\{(b-d-1)(b-d+2)\}] \\ &=(a-c)-1-(b-d) \\ &=\{(a+d)-(b+c)\}-1<0. \end{aligned} \tag{9}$$

The last inequality follows from the assumption for the tension between efficiency and risk dominance. This inequality shows that the minimum cost tree has a root of (B,B)-PSS. That is, a risk dominant outcome is the stochastically stable state. □

Case 2: $a > c > b > d > 0$.

We first exhibit two claims used in this proof.

Claim 4: If, for Player $i = 1$ (or 2), $v^i_B < c - 1$ and the state is one of the (A,A)-PSSs then, with one perturbation, v^i_B increases by one.

This is because the sequence of action choices, (A,B) (or (B,A)) with one perturbation and (A,A) with no perturbation, leads to the situation in which only v^i_B increases by one.

Claim 5: If, for Player $i = 1$ (or 2), $v^i_A < d+1$ and the state is in (B,B)-PSS then, with one perturbation, v^i_A increases by one.

This is because the sequence of action choices, (B,A) (or (A,B)) with one perturbation and (B,B) with no perturbation, leads this situation.

Lemma 3: There is a (B,B)-PSS cost tree having cost $\{(c-b+1)^2\} + \{a-c\} + \{(b-d)^2 - 1\}$.

Proof of Lemma 3:

Step 1: In this step, we construct a (A,A)-PSS cost sub-tree such that the root is $[A,(a,c);A,(a,c)]$, it includes all (A,A)-PSSs but not (B,B)-PSSs, and its cost is $(c-b+1)^2 - 1$.

Every state in (A,A)-PSSs satisfies the following condition: for all $i,j \in \{1,2\}$, $v^1_A = v^2_A = a$ and $v^1_B, v^2_B \in [b, b+1, \ldots, c-1, c]$.

Consider the following (A, A)-PSS cost sub-tree (Figure 14):

$$\bigcup_{v_B^1 \in \{b,b+1,\ldots,c\}} \bigcup_{v_B^2 \in \{b,b+1,\ldots,c-1\}} ([A,(a,v_B^1);A,(a,v_B^2)] \to [A,(a,v_B^1);A,(a,v_B^2+1)])$$
$$+ \bigcup_{v_B^1 \in \{b,b+1,\ldots,c-1\}} ([A,(a,v_B^1);A,(a,c)] \to [A,(a,v_B^1+1);A,(a,c)]) \qquad (10)$$

Since every link has cost one by Claim 4, its overall cost is $(c-b+1)^2 - 1 (= (c-b)(c-b+1) + (c-b))$.

Step 2: In this step we construct a (B, B)-PSS cost sub-tree such that the root is $[B,(d,b);B,(d,b)]$; it includes all (B, B)-PSSs but not (A, A)-PSSs, and its cost is $(b-d)^2 - 1$.

Every state in (B, B)-PSSs satisfies the following condition: for all $i, j \in \{1, 2\}$, $v_B^1 = v_B^2 = b$ and $v_A^1, v_A^2 \in \{d, d+1, \ldots, b-1\}$.

Consider the following (B, B)-PSS cost sub-tree (Figure 15):

$$\bigcup_{v_A^1 \in \{b-1,b-2,\ldots,d\}} \bigcup_{v_A^2 \in \{b-1,\ldots,d+1\}} ([B,(v_A^1,b);A,(v_A^2,b)] \to [B,(v_A^1,b);B,(v_A^2-1,b)])$$
$$+ \bigcup_{v_A^1 \in \{b-1,\ldots,d+1\}} ([B,(v_A^1,b);B,(d,b)] \to [B,(v_A^1-1,b);B,(d,b)]) \qquad (11)$$

Since every link has cost one by Claim 5, its overall cost is $(b-d)^2 - 1 (= (b-d-1)(b-d) + (b-d-1))$.

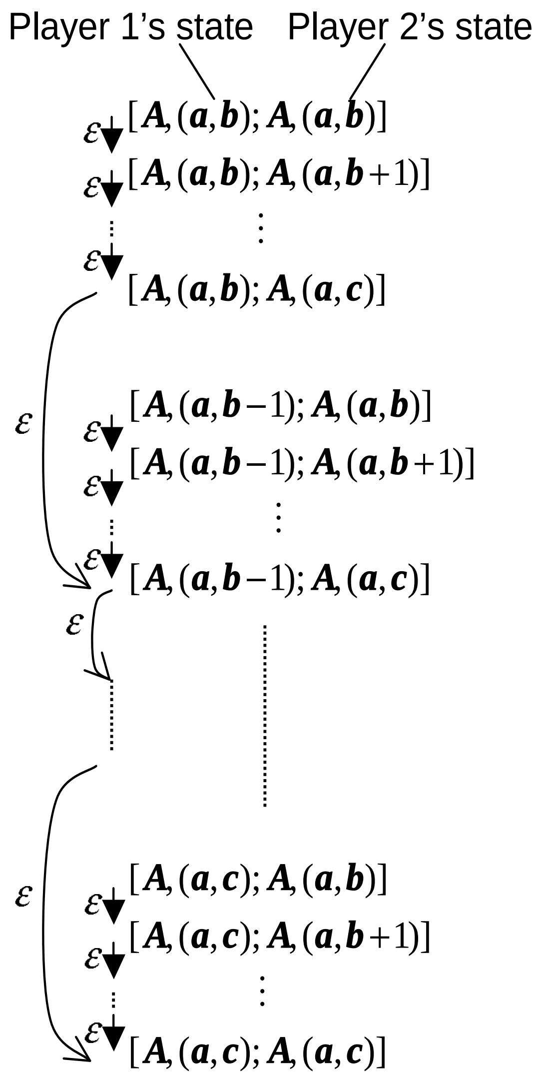

Figure 14: (A, A)-PSS cost sub-tree

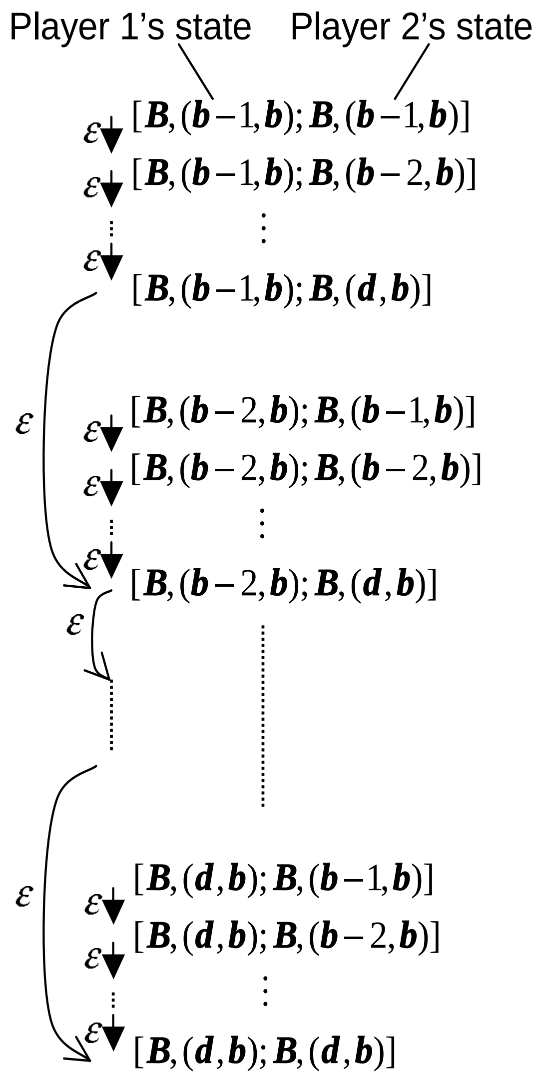

Figure 15: (B,B)-PSS cost sub-tree

Step 3: In this step, we construct a path from $[A,(a,c);A,(a,c)]$ to $[B,(b-1,b);B,(b-1,b)]$ and show that its cost is $(a-c)$.

Consider the following transition. First (A,A) is repeated $a-c$ times. This transition involves $a-c$ costs. After this, $v_A^1 = v_B^1 = c$, $v_A^2 = a$ and $v_B^2 = c$ are realized (Figure 16). Second, with no cost, (B,A) is repeated $a-c$ times and $v_A^1 = v_B^1 = v_A^2 = v_B^2 = c$ is realized (Figure 17).

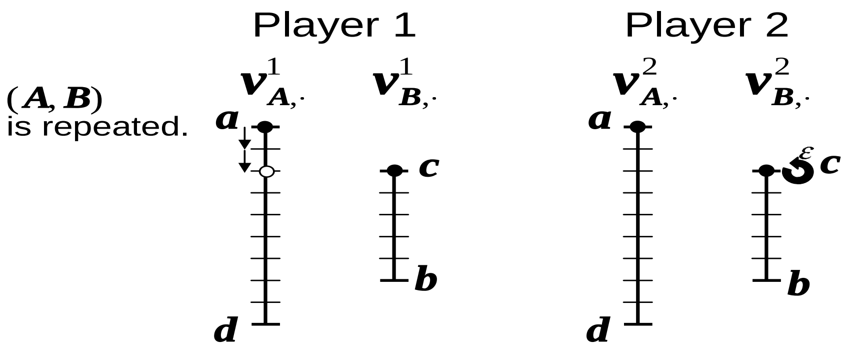

Figure 16: transitions

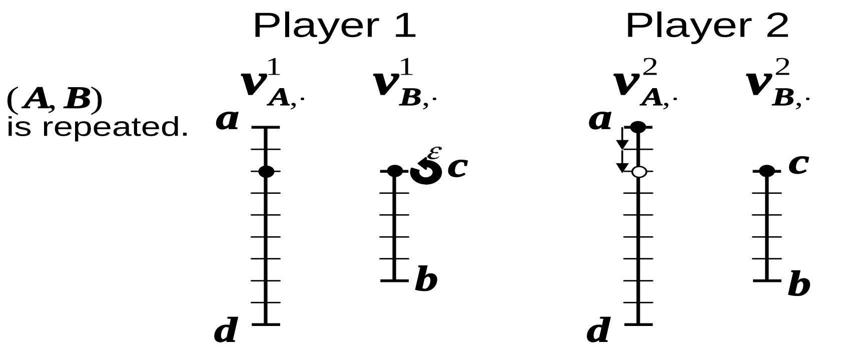

Figure 17: Transitions

Third, the sequence of actions, (A, B), (B, A), and (B, B), is chosen. All $v^1_A, v^1_B, v^2_A, v^2_B$ decrease by one (Figure 18). If $b = c - 1$, skip the fourth procedure and go to the fifth one. Fourth, repeat the sequence of action, (A, B), (B, B), (B, A), and (B, B) until $v^1_A = v^1_B = v^2_A = v^2_B = b$ is realized (Figure 19). Fifth, once again (A, B), (B, B), (B, A), and (B, B) are chosen (Figure 20). Consequently $v^1_A = v^2_A = b - 1 < v^1_B = v^2_B = b$ is realized. The costs occur only in the first procedure and are $a - c$. Hence, the total cost of this path is $a - c$.

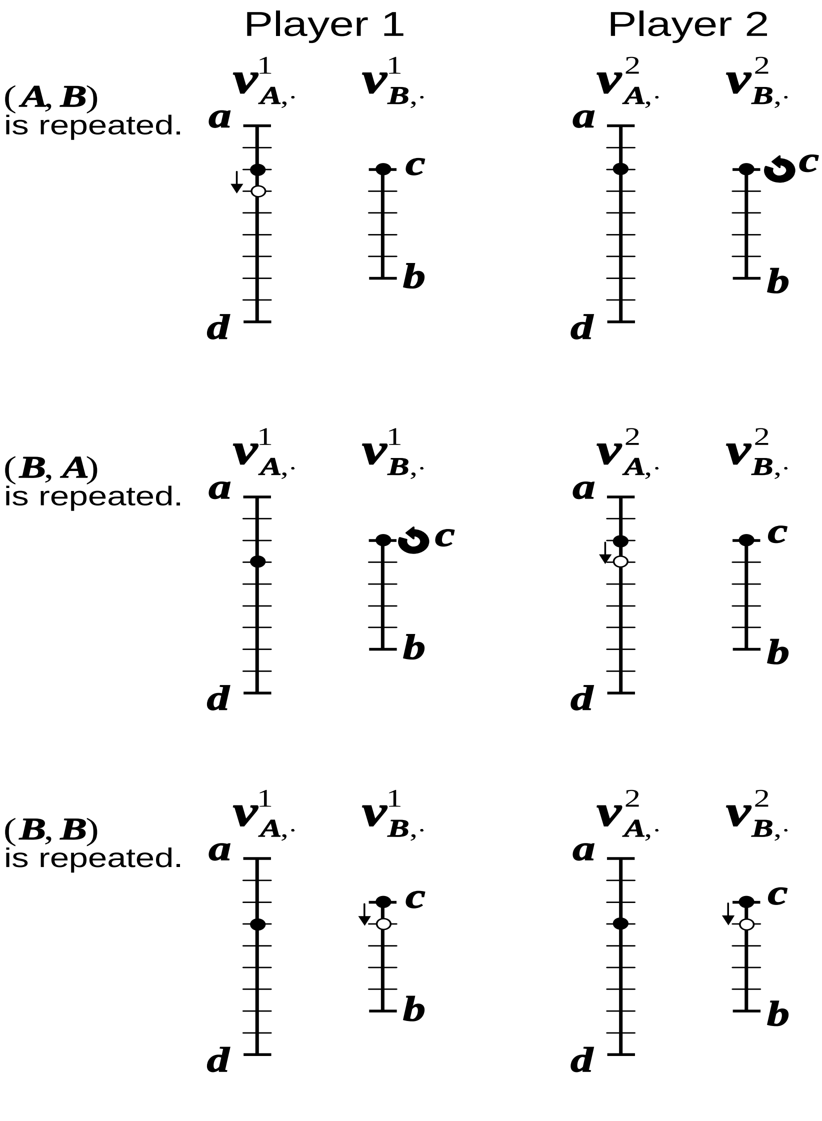

Figure 18: Transitions

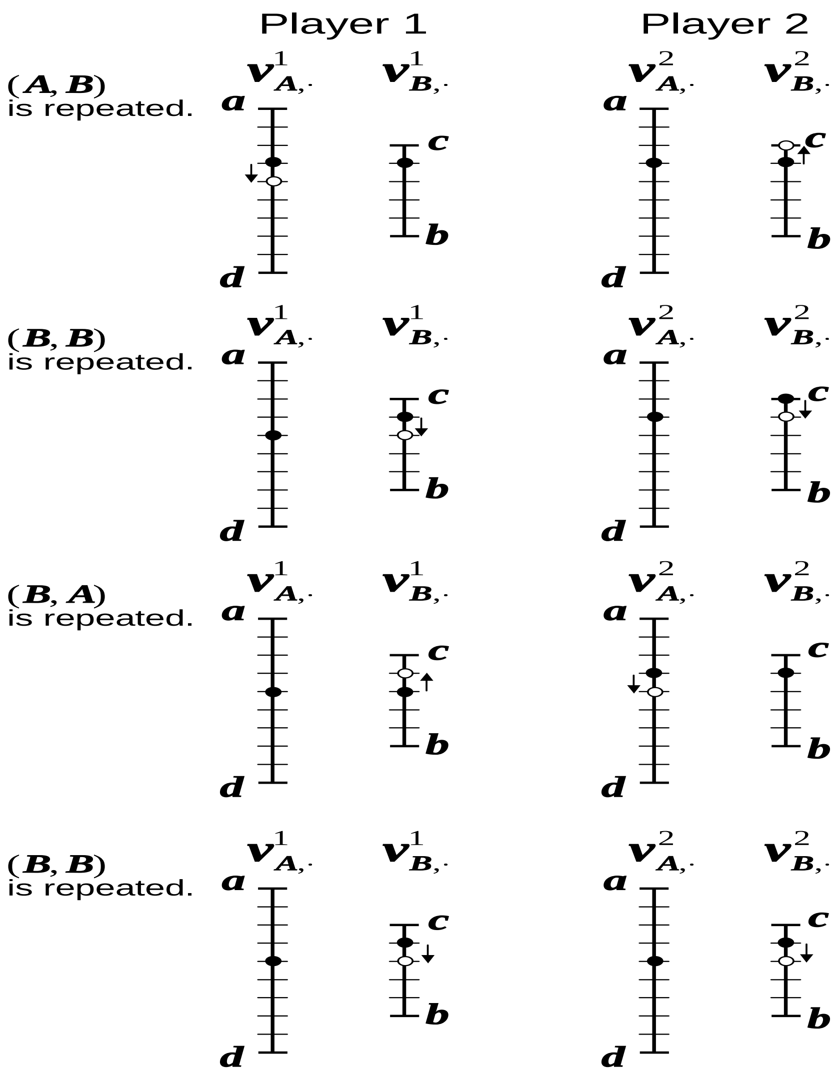

Figure 19: Transitions

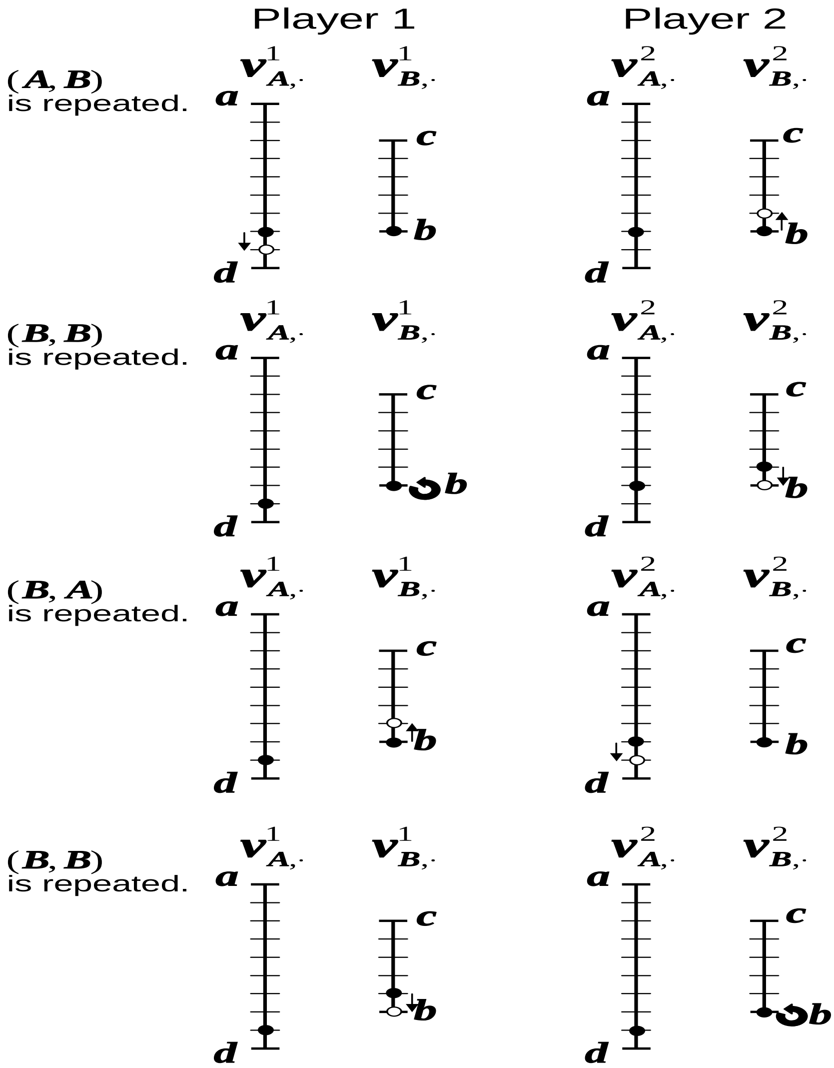

Figure 20: Transitions

Using Step 1, Step 2 and Step 3, we construct a cost tree that satisfies the condition of lemma 3 (Figure 21). □

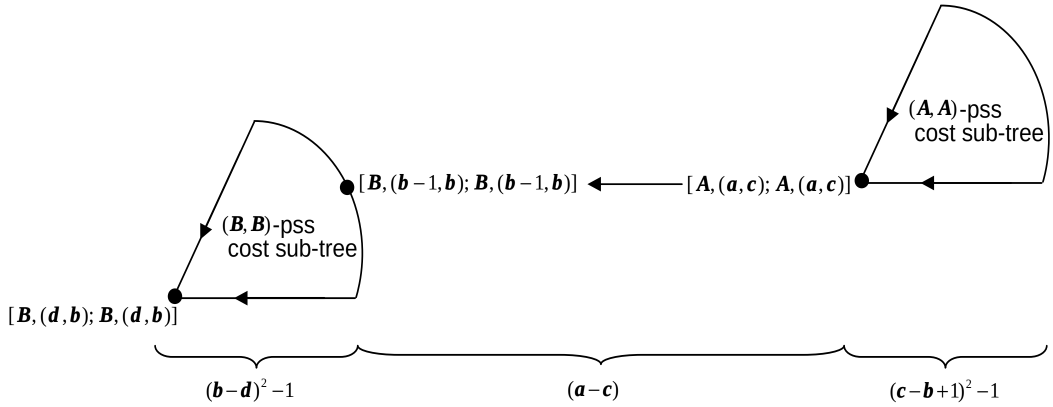

Figure 21: A cost tree

Before stating Lemma 4, we exhibit the following claim, similar to previous claims.

Claim 6: If $v_A^i (i \in \{1, 2\})$ increases in (B, B)-PSS, then at least two perturbations are needed. More precisely, both v_B^1 and v_B^2 increase by one.

This is because the only pair of actions that increases $v_A^i (i \in \{1, 2\})$ is (A, A).

Lemma 4: The minimum cost in (A, A)-PSS cost trees is $\{(c - b + 2)^2 - 1\} + 2 + \{(b - d - 1)(b - d + 2)\}$.

Proof of Lemma 4:

Step 1: We show the path that the minimum cost tree can include. From the definition of cost trees, there is a path to the root from any state. Consider paths from

$[B,(d,b);B,(d,b)]$ to (A,A)-PSS. We can construct the following path:

$$\left\{ \bigcup_{v_A \in \{d,d+1,\ldots,b-1\}} ([B,(v_A,b);A,(v_A,b)] \to [B,(v_A+1,b);B,(v_A+1,b)]) \right\} + \{([B,(b,b);B,(b,b)] \to [B,(b+1,b);A,(b+1,b)])\} \tag{12}$$

The first term expresses transitions in (B,B)-PSSs. In the transitions, two perturbations are needed at each transition. In the second term there is no perturbation. Hence, the total cost of this path is $2(b-d)$.

This path has the minimum cost among paths from $[B,(d,b);B,(d,b)]$ to (A,A)-PSS. This is because, by Claim 6, there is no redundant transition in this path.

Step 2: In this step, we examine the minimum cost tree that includes the path defined in Step 1. In the PSSs, at least one perturbation is needed, by definition. The path via Step 1 is a transition from (B,B)-PSS to (A,A)-PSS. Hence, if the cost tree whose root is (A,A)-PSS and includes the path in Step 1 is the minimum cost tree, then the other links involve exactly one perturbation.

Step 3: In this step, we construct a (A,A)-PSS cost sub-tree such that the root is $[A,(a,c);A,(a,c)]$; it includes only (A,A)-PSS, and its cost is $(b-c+1)^2-1$.

This result follows from Step 1 of Lemma 1. In this tree, each one-step link needs exactly one perturbation.

Step 4: In this step, we construct a (B,B)-PSS cost sub-tree such that the root is $(B,(b-1);B,(b-1,b))$; it includes only (B,B)-PSS and a part of the path in Step 1, and its cost is $(b-d-1)(b-d+2)$.

Consider the following (B,B)-PSS cost sub-tree (Figure 22):

$$\bigcup_{i\in\{1,2,\ldots,b-d-1\}} \left[\left\{ \bigcup_{v_A^2\in\{d+1,d+2,\ldots,b-1\}} ([B,(b-i,b);B,(v_A^2,b)] \to [B,(b-i,b);B,(v_A^2-1,b)]) \right\} \right. \tag{13}$$

$$\left. - ([B,(b-i,b);B,(b-i,b)] \to [B,(b-i,b);B,(b-i-1,b)]) \right]$$

$$+ \bigcup_{i\in\{1,2,\ldots,b-d-1\}} ([B,(b-i,b);B,(d,b)] \to [B,(b-i-1,b);B,(d,b)])$$

$$+ \bigcup_{v_A^2\in\{d+1,d+2,\ldots,b-1\}} ([B,(b-i,b);B,(v_A^2,b)] \to [B,(b-i-1,b);B,(v_A^2-1,b)])$$

$$+ \bigcup_{v_A\in\{d,d+1,\ldots,b-2\}} ([B,(v_A,b);B,(v_A,b)] \to [B,(v_A+1,b);B,(v_A+1,b)])$$

Since every link in the first, second and third terms has one perturbation by Claim 4, its cost is $(b-d-1)(b-d-1-1)+(b-d-1)+(b-d-1)$. Every link in the last terms has two perturbations by Claim 6, and its cost is $2(b-d-1)$. Hence the total cost is $(b-d-1)(b-d+2)$.

Step 5: In this step, we construct a path from $[B,(b–1,b);B,(b–1,b)]$ to $[A,(a,b);A,(a,b)]$ and show that its cost is 2.

Consider the following path:

$$[B,(b-1,b);B,(b-1,b)] \to [A,(b,b);A,(b,b)] \to [A,(b+1,b);A,(b+1,b)] \to \cdots \tag{14}$$

$$\to [A,(a-1,b);A,(a-1,b)] \to [A,(a,b);A,(a,b)]$$

In the first transition, two perturbations are needed. In the other transitions there is no cost, because there is a nonzero probability with which both players choose A. Hence

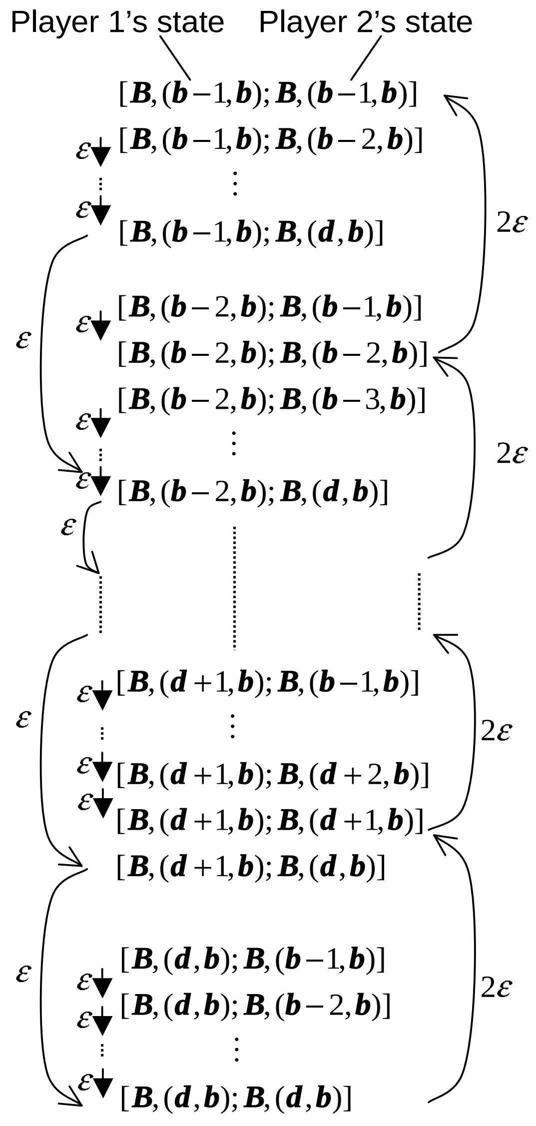

Figure 22: (B, B)-PSS cost sub-tree

the total cost of this path is 2.

Step 6: We sum over the steps to prove the lemma. If we collect all links in Step 3, 4 and 5, it constitutes the cost tree that includes the path in Step 1 (Figure 23). From the construction, all links have unit cost (perturbation) except the path in Step 1. By Step 2, this is the minimum cost tree in the (A, A)-PSS cost trees.

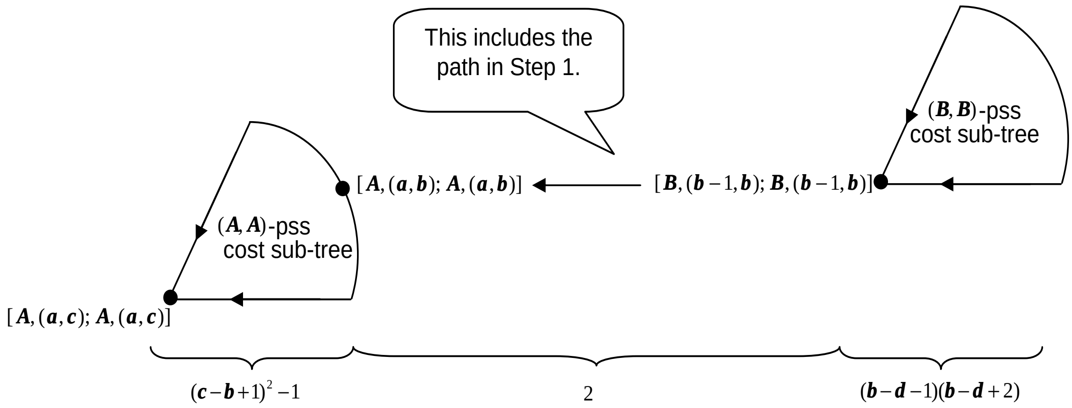

Figure 23: A cost tree

From steps 3, 4 and 5, the total cost is $\{(c-b+1)^2-1\}+\{2\}+\{(b-d-1)(b-d+2)\}$.

This proves the lemma. □

Finally, we prove the statement in Proposition 2 under the condition $a > c > b > d > 0$. By Lemmas 3 and 4,

$$\begin{aligned}
&[\{(c-b+1)^2-1\}+\{a-c\}+\{(b-d)^2-1\}] \qquad (15)\\
&\quad -[\{(c-b+1)^2-1\}+\{2\}+\{(b-d-1)(b-d+2)\}]\\
&=(a-c)-1-(b-d)\\
&=\{(a+d)-(b+c)\}-1<0.
\end{aligned}$$

The last inequality follows from the assumption for the tension between Pareto efficiency and risk dominance. This inequality shows that the minimum cost tree has a root of (B, B)-PSS. That is, a risk dominant outcome is the stochastically stable state. □

Proof of Theorem 1:

We can extend Proposition 2 directly by standard techniques (Young 1998).

First, we describe Young's definition and theorem.

Definition (Young 1998, p. 54)

Let P^ε be a Markov process on S. We call P^ε a regular perturbed Markov process if P^ε is irreducible for $\varepsilon \in (0, \varepsilon^*]$, and for every $s, s' \in S$, $P^\varepsilon_{ss'}$ approaches $P^0_{ss'}$ at an exponential rate; that is, $\lim_{\varepsilon \to 0} P^\varepsilon_{ss'} = P^0_{ss'}$, and if $P^\varepsilon_{ss'} > 0$ for some $\varepsilon > 0$, then some $\varepsilon \in (0, \varepsilon^*]$ and $0 < \lim_{\varepsilon \to 0} \frac{P^\varepsilon_{ss'}}{\varepsilon^{r(s,s')}}$ for some $r(s, s')$.

Young's Theorem (Young 1998, Theorem 3.1, p. 153)

Let P^ε be a regular perturbed Markov process. Assume that P^ε has a unique recurrent class.[7]

(1) There exists a unique stationary distribution (μ^ε) of P^ε for each $\varepsilon > 0$.[8] That is, $\mu^\varepsilon P^\varepsilon = \mu^\varepsilon$.

(2) $\lim_{\varepsilon \to 0} \mu^\varepsilon = \mu^0$ exists, and μ^0 is a stationary distribution of P^0.

(3) The stochastically stable states are precisely those states that are contained in the recurrent class(es) of P^0 having minimum stochastic potential.

The minimum stochastic potential means that the cost tree constructed among recurrent classes with a root of the state has minimum cost.

[7] In Young (1998), this is not assumed, but it is assumed that μ^* is the unique stationary distribution of P^ε.

[8] This is derived from that there is the unique stationary distribution if and only if P^ε has a unique recurrent class (Young 1998, p. 49).

In our model, the condition of irreducibility is violated. We therefore use the following definition and theorem.

Definition A.1

Let P^{ε} be a regular perturbed Markov process on S, and let $\tilde{P}^{\varepsilon}$ be a Markov process on $\tilde{S} = S \cup Z$ such that, for any $s \in Z$, s is transient and S is a unique recurrent class of $\tilde{P}^{\varepsilon}$, and $\tilde{P}^{\varepsilon}$ is constructed by P^{ε} as follows:

$$\tilde{P}^{\varepsilon} = \begin{pmatrix} P^{\varepsilon} & \mathbf{0} \\ Q_1^{\varepsilon} & Q_2^{\varepsilon} \end{pmatrix}. \tag{16}$$

We call this Markov process, $\tilde{P}^{\varepsilon}$, an enlarged regular perturbed Markov process. If $\varepsilon \to 0$, it approachs $\tilde{P}^{0}$ at an exponential rate, so that

$$\tilde{P}^{\varepsilon} \to \tilde{P}^{0} = \begin{pmatrix} P^{0} & \mathbf{0} \\ Q_1^{0} & Q_2^{0} \end{pmatrix}. \tag{17}$$

Theorem A.1

Let P^{ε} be a regular perturbed Markov process, and let μ^{*} be the unique stationary distribution of P^{ε} for each $\varepsilon > 0$. Let $\tilde{P}^{\varepsilon}$ be an enlarged regular perturbed Markov process.

(1) $\tilde{\mu}^{\varepsilon} = (\mu^{\varepsilon}, \mathbf{0})$ is a unique stationary distribution of $\tilde{P}^{\varepsilon}$ for each $\varepsilon > 0$, where $\mathbf{0}$ is a zero vector having size $|Z|$.

(2) $\lim_{\varepsilon \to 0} \tilde{\mu}^{\varepsilon} = \tilde{\mu}^{0}$ exists, where $\tilde{\mu}^{0} = (\mu^{0}, \mathbf{0})$. Here, $\tilde{\mu}^{0}$ is a stationary distribution of $\tilde{P}^{0}$.

(3) The stochastically stable states are precisely those states that are contained in the recurrent class(es) of P^0 having minimum stochastic potential. That is, stochastically stable states under the original regular perturbed Markov process P^ε on S are equal to stochastically stable states under the enlarged regular perturbed Markov process $\tilde{P}^\varepsilon$ on $\tilde{S}$.

Proof:

(1) From standard properties of Markov process, the stationary distribution of the transient state is 0. By Young's theorem, $\mu^\varepsilon P^\varepsilon = \mu^\varepsilon$. Hence we have

$$(\mu^\varepsilon, \mathbf{0}) \begin{pmatrix} P^\varepsilon & \mathbf{0} \\ Q_1^\varepsilon & Q_2^\varepsilon \end{pmatrix} = (\mu^\varepsilon, \mathbf{0}) \tag{18}$$

where $\mathbf{0}$ is a vector whose size is $|Z|$. From the assumption of an enlarged regular perturbed Markov process, there is unique recurrent class. This shows that there is a unique stationary distribution, $\tilde{\mu}^\varepsilon = (\mu^\varepsilon, \mathbf{0})$; that is, $\tilde{\mu}^\varepsilon \tilde{P}^\varepsilon = \tilde{\mu}^\varepsilon$.

(2) By Young's theorem, $\lim_{\varepsilon\to 0} \mu^\varepsilon$ exists and μ^0 is a stationary distribution of P^0. Hence, by (1), $\lim_{\varepsilon\to 0} \tilde{\mu}^0$ exists and $\tilde{\mu}^0$ is a stationary distribution of $\tilde{P}^0$, where $\tilde{\mu}^0 = (\mu^0, \mathbf{0})$.

(3) From (1) and (2), any state $s \in Z$ does not have strictly positive value in the stationary distribution. Note also that μ^0 is the same limit distribution in Young's Theorem. □

In Young's Theorem and Theorem A.1, aperiodicity is not assumed, because only stationary distribution is involved. It is well known that aperiodicity carries the good convergence property that, for any initial distribution d, $dP^\varepsilon \to \mu^\varepsilon$ as $t \to \infty$.

We next set out the corresponding result.

Proposition A.1

Let $\tilde{P}^{\varepsilon}$ be an enlarged regular perturbed Markov process. Then, for any initial distribution $\tilde{d}$,

(1) $\lim_{t\to\infty} \tilde{d}\tilde{P}^{\varepsilon} = \tilde{\mu}^{\varepsilon}$,

(2) $\lim_{\varepsilon\to\infty} \lim_{t\to\infty} \tilde{d}\tilde{P}^{\varepsilon} = \tilde{\mu}^{0}(\equiv (\mu^{0}, \mathbf{0}))$.

Proof:

From the properties of Markov processes and a form of the matrix, it follows that

$$\tilde{P}^{\varepsilon(n)} = \tilde{P}^{\varepsilon(n)} = \begin{pmatrix} P^{\varepsilon^n} & \mathbf{0} \\ Q_1^{\varepsilon(n)} & Q_2^{\varepsilon^n} \end{pmatrix}. \tag{19}$$

By the standard results of P^{ε},

$$\lim_{n\to\infty} P^{\varepsilon^n} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \cdot \mu^{\varepsilon}. \tag{20}$$

From the standard properties of Markov processes, if j is transient, then for any $i \in \tilde{S}$, $\lim_{n\to\infty} p_{ij}^{n} = 0$. Hence,

$$\lim_{n\to\infty} Q_2^{\varepsilon^n} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}. \tag{21}$$

Next, we examine $Q_1^{\varepsilon^{(n)}}$. Note first that

$$\tilde{P}^{\varepsilon^n} = \tilde{P}^{\varepsilon^m} \cdot \tilde{P}^{\varepsilon^k} \tag{22}$$

$$= \begin{pmatrix} P^{\varepsilon^m} & \mathbf{0} \\ Q_1^{\varepsilon^{(m)}} & Q_2^{\varepsilon^m} \end{pmatrix} \cdot \begin{pmatrix} P^{\varepsilon^k} & \mathbf{0} \\ Q_1^{\varepsilon^{(k)}} & Q_2^{\varepsilon^k} \end{pmatrix} \tag{23}$$

where $n = m + k$. For $|S| + 1 \leq i \leq |\tilde{S}|$ and $1 \leq j \leq |S|$,

$$\tilde{p}_{ij}^{\varepsilon^{(n)}} = \sum_{l=1}^{|\tilde{S}|} \tilde{p}_{il}^{\varepsilon^{(m)}} \tilde{p}_{lj}^{\varepsilon^{(k)}} \tag{24}$$

$$= \sum_{l=1}^{|S|} \tilde{p}_{il}^{\varepsilon^{(m)}} \tilde{p}_{lj}^{\varepsilon^{(k)}} + \sum_{l=|S|+1}^{|\tilde{S}|} \tilde{p}_{il}^{\varepsilon^{(m)}} \tilde{p}_{lj}^{\varepsilon^{(k)}}. \tag{25}$$

Hence,

$$\left| \mu_j - \tilde{p}_{ij}^{\varepsilon^{(n)}} \right| = \sum_{l=1}^{|S|} \tilde{p}_{il}^{\varepsilon^{(m)}} \left| \mu_j - \tilde{p}_{lj}^{\varepsilon^{(k)}} \right| + \sum_{l=|S|+1}^{|\tilde{S}|} \tilde{p}_{il}^{\varepsilon^{(m)}} \left| \mu_j - \tilde{p}_{lj}^{\varepsilon^{(k)}} \right|. \tag{26}$$

By equation (20), if $1 \leq l \leq |S|$ then $\lim_{k\to\infty} \tilde{p}_{lj}^{\varepsilon^{(k)}} = \lim_{n\to\infty} \tilde{p}_{lj}^{\varepsilon^k} = \mu_j$. From equation (21), if $|S| + 1 \leq l \leq |\tilde{S}|$ then $\lim_{m\to\infty} \tilde{p}_{il}^{\varepsilon^{(m)}} = 0$.

Therefore, for any $\delta > 0$, there exist k^* and m^* such that for any $m > m^*$ and $k > k^*$,

$$\left| \mu_j - \tilde{p}_{ij}^{\varepsilon^{(n)}} \right| = \sum_{l=1}^{|S|} \tilde{p}_{il}^{\varepsilon^{(m)}} \left| \mu_j - \tilde{p}_{lj}^{\varepsilon^{(k)}} \right| + \sum_{l=|S|+1}^{|\tilde{S}|} \tilde{p}_{il}^{\varepsilon^{(m)}} \left| \mu_j - \tilde{p}_{lj}^{\varepsilon^{(k)}} \right| < \delta. \tag{27}$$

That is, for $|S|+1 \leq i \leq |\tilde{S}|$ and $1 \leq j \leq |S|$, $\lim_{n\to\infty} \tilde{p}_{ij}^{\varepsilon^{(n)}} = \mu_j$. Hence

$$\lim_{n\to\infty} Q_1^{\varepsilon^{(n)}} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \cdot \mu^{\varepsilon}. \tag{28}$$

From equations (20), (21) and (28),

$$\lim_{n\to\infty} \tilde{P}^{\varepsilon^{(n)}} = \left(\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \cdot (\mu^{\varepsilon}, \mathbf{0}) \right). \tag{29}$$

From the definition of μ^0 and $\tilde{d}$,

$$\lim_{\varepsilon\to\infty}\lim_{t\to\infty}\tilde{d}\tilde{P}^{\varepsilon} = \tilde{d}\lim_{\varepsilon\to\infty}\lim_{t\to\infty}\tilde{P}^{\varepsilon} \tag{30}$$

$$= \tilde{d}\left(\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix}\cdot(\mu^0,\mathbf{0})\right) \tag{31}$$

$$= \tilde{d}\cdot\begin{pmatrix}(\mu^0,\mathbf{0})\\(\mu^0,\mathbf{0})\\\vdots\\(\mu^0,\mathbf{0})\end{pmatrix} \tag{32}$$

$$= (\mu^0,\mathbf{0}). \tag{33}$$

In the last equation, because $\tilde{d}$ is a distribution, the sum of all elements in $\tilde{d}$ is 1. □

Under the extended initial condition, the Markov process is an enlarged regular perturbed Markov process. From Theorem A.1 and Proposition 2, we have (1) of Theorem 1. By Proposition A.1, we have (2) of Theorem 1. □

References

[1] Erev, I. and A. E. Roth (1998) "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria," American Economic Review 88, pp. 848-881.

[2] In-Koo Cho and Akihiko Matsui (2005) "Learning aspiration in repeated games,"

Journal of Economic Theory 124, pp. 171-201.

[3] Karandikar, Rajeeva, Dilip Mookherjee, Debraj Ray, Fernando Vega-Redondo (1998) "Evolving Aspirations and Cooperation," Journal of Economic Theory 80, pp. 292-331.

[4] Kim, Youngse (1999) "Satisficing and optimality in 2×2 common interest games," Economic Theory 13, pp. 365-375.

[5] Pazgal, Amit (1997) "Satisficing Leads to Cooperation in Mutual Interests Games," International Journal of Game Theory 26, pp.439-453.

[6] Roth, A. E. and I. Erev (1995) "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," Games and Economic Behavior 8, pp. 164-212.